UNIVERSITY OF GRANADA

FACULTY OF SCIENCES

Department of Analytical Chemistry

Research Group FQM-297

"Environmental, Biochemical and Foodstuff Analytical Control"

Functional Food Research and Development Center (CIDAF)

DOCTORAL THESIS

CHARACTERIZATION OF PHENOLIC COMPOUNDS IN HIGHLY-CONSUMED VEGETABLE MATRICES BY USING ADVANCED ANALYTICAL TECHNIQUES

CARACTERIZACIÓN DE COMPUESTOS FENÓLICOS EN MATRICES VEGETALES MEDIANTE TÉCNICAS ANALTICALAS AVANZADAS

Presented by

IBRAHIM M. ABU REIDAH

Submitted for a Doctoral degree in Chemistry

GRANADA, 2013

Editor: Editorial de la Universidad de Granada
Autor: Ibrahim M. Abu Reidah
D.L.: GR 1899-2013
ISBN: 978-84-9028-591-6

This doctoral thesis has been conducted through financing from the Ministry of Foregin Affairs of Spain & The Spanish Agency Of International Cooperation for Development (MAEC-AECID) scholarship and funds from the Research Group FQM-297 “Environmental, Biochemical and Foodstuff Analytical Control” (Department of Analytical Chemistry, University of Granada) and Functional Food Research and Development Center (CIDAF) from different projects, contracts and grants from the central and autonomic administrations and research plan of the University of Granada.

CHARACTERIZATION OF PHENOLIC COMPOUNDS IN HIGHLY-CONSUMED VEGETABLE MATRICES BY USING ADVANCED ANALYTICAL TECHNIQUES

By

IBRAHIM M. ABU REIDAH

Granada, 2013

Signed by Dr. Alberto Fernández-Gutiérrez
Full Professor of the Department of Analytical Chemistry
Faculty of Sciences, University of Granada

Signed by Dr. Antonio Segura Carretero
Full Professor of the Department of Analytical Chemistry
Faculty of Sciences, University of Granada

Signed by Dr. David Arráez-Román
Assistant Professor of the Department of Analytical Chemistry
Faculty of Sciences, University of Granada

Submitted for a Doctoral Degree in Chemistry
Signed by Ibrahim M. Abu Reidah

ALBERTO FERNÁNDEZ GUTIÉRREZ, Full Professor of University of Granada, Department of Analytical Chemistry, Head of Research Group FQM-297 "Environmental, Biochemical, and Foodstuff Analytical Control" and Director of CIDAF,

Certifies:

That the work presented in this doctoral thesis entitled "**CHARACTERIZATION OF PHENOLIC COMPOUNDS IN HIGHLY-CONSUMED VEGETABLE MATRICES BY USING ADVANCED ANALYTICAL TECHNIQUES**" has been carried out under my direction and also under the supervision of both Dr. Antonio Segura Carretero and Dr. David Arráez-Román in the laboratories of the Department of Analytical Chemistry and Functional Food Research and Development Center (CIDAF). He demonstrates all requirements of eligibility to obtain the Doctoral Degree in Chemistry from the University of Granada.

Granada, March 2013

Acknowledgments

Acknowledgments

In the name of Allah, Praise be to him & Peace be upon his Prophet Mohammad

First, I would like to give a heartfelt, special thanks to my thesis directors: to Dr. Alberto Fernández Gutiérrez, Dr. Antonio Segura Carretero for having given me the opportunity to work within the FQM 297 and at CIDAF Groups, and to earning my Ph.D. In fact, you have been always motivating, encouraging, and enlightening. Also, I'd thank Dr. David Arráez Román for his continuous assistance and perspective remarks.

Dr. Antonio, thanks for everything, fatherly correcting me, advising, and also your hospitality in your house in summer vacation time with all CIDAF family members, is also unforgettable. Thank you for everything.

I am very grateful to the members of my dissertation committee. Their academic support and input, correcting and advising are greatly appreciated. Thank you. The friendly and supportive atmosphere inherent to the whole FQM 297 Group contributed essentially to the final outcome of my studies.

In this context I would like to alphabetically thank: Arantzazu, Ana, Ana Cobo, Cecilia, Celia, Cristina, Hakim, Ihsan, Isa, Javier, Jesús, Maria del Mar, Maria Luz, Nassima, Patricia, Rosa, and Salva, for their friendship, collaboration and the enjoyable and nice times we have passed together. Without their support the completion of this dissertation would not have been possible.

And particularly, I'd thank Ihsan, partnership and friendship. Isa, unforgettable moments with Dr. Leopoldo, registration, master classes, I won't forget your help in everything, many thanks. Jesús, you have correctly got my style of colors, thanks for your compliments. Maria Luz thanks for your humor giving to everybody. Maria del Mar, thanks for everything, collaboration, discussions, and for sweets offering.

Salva, thanks for the collaboration, guiding us to the Paris's congress. Javier; the manager, thank you. Patricia, I'd thank you for everything, cooperation, helping in samples' preparations, experiments.

Also form other members of FQM 297: I'd thank Alegria, Angel, Antonio-Luis Elena, Paco, Paulina, Maria, Marta, and Santi for their encouragement throughout.

Santi, Patchi and Marta thank you a lot for your technical help in installing the computer programs and the softwares, intersting discussions. Alegria, you look like your name, I hope you every success.

I'd thank in fact collectively, all people and (friends, teachers, relatives, professors) that have helped, encouraged, supported me during all my study periods.

Before and after all the thankfulness should be raised to Allah creator of everything the most merciful and Grateful, then to my parents, the factor of my success in this life, thanks for your continues support morally was or financially along all my life stations.

Saeeda, thanks for your infinite support throughout everything, in fact, I cannot thank you sufficiently. Kawthar and Mohammad; my lovely kids, your birth were paralleled with my academic progress and advancement.

Last but not least, I'd like to thank my family, relatives and friends who always ask, help, encourage also for their support to pursue this study.

Finally, I'd like to thank everybody who was part of my success during my entire study career at all levels, apologizing for not being able to mention all since they are lots.

Table of Contents

TABLE OF CONTENTS

Summary

SUMMARY

This doctoral thesis, entitled "Characterization of phenolic compounds in highly consumed vegetables matrices by using advanced analytical techniques", consists of two main sections: the Introduction and the Experimental Section.

The first section offers an overview of the Mediterranean diet and its key foods (mainly fruits and vegetables) and their importance as highly-consumed plant foods and also as good sources of bioactive compounds and phytochemicals, with a special focus on the phenolic compounds. Finally, the analytical tools used to characterize phenolic compounds are described, including high-performance liquid chromatography (HPLC) and ultra-high-performance liquid chromatography (UPLC) coupled to different detectors such as UV-visible (UV-vis) and electrospray ionization (ESI) mass spectrometry (MS) with time-of-flight (TOF) and quadrupole time-of-flight (QTOF) analyzers, are described.

The experimental section is divided in five chapters as follows:

Chapter 1 demonstrates a powerful analytical method to separate and characterize the phytochemical compounds in whole pods of three cultivars of green bean (*Phaseolus vulgaris*). Fresh green beans, the immature pods of any type of common beans, are very low in calories and contain no saturated fat. Nevertheless, this vegetable is a very good source of vitamins, minerals, and plant-derived phytochemicals. In this sense, an extraction method has been used to isolate the phytochemical compounds from the study sample using an aqueous-methanolic solvent, followed by HPLC separation coupled to TOF-MS. The proposed method characterized a total of 72 phytochemical compounds, 54 being reported in green beans for the first time, by using the retention time, the data generated by TOF-MS, and the literature. The cultivar Strike was found to be richer in flavonoids compared to other cultivars analyzed.

Chapter 2 Cucumber (*Cucumis sativus L.),* a highly consumed crop, is rich in phenolic compounds. Due to the low amount of data on cucumber, concerning to these compounds, a simple and rapid method to characterize the phenolic

compounds present in an extract of the edible part using HPLC–ESI-QTOF-MS was developed. Thus, based on the accuracy mass data together with the fragmentation pattern, the full scan run generated by QTOF-MS proved to be a useful tool for tentatively characterizing 73 phenolic compounds in the extract of the matrix studied, marking the first comprehensive characterization of phenolic compounds in cucumber.

Chapter 3 describes a comprehensive characterization of phenolic and other phytochemical compounds in an extract of the edible part of watermelon which is a highly cultivated and consumed fruit especially in summer times. Thus, the use of HPLC-ESI-QTOF-MS was useful to detect 71 phenolic compounds from different families such as phenolic acids, flavonoids, iridoids, coumarins and lignan using the MS and MS/MS data provided by the QTOF-MS, together with the data taken from the relevant literature on the same botanical family. In addition, watermelon flesh was shown to contain an array of diverse phytochemical components. This is the first report available to characterize individual phenolic and other phytochemical compounds in watermelon.

Chapter 4 concerns the characterization of an extract of artichoke (*Cynara scolymus)* using of HPLC-DAD-ESI-QTOF-MS as a separation and detection method in negative ionization mode. Artichoke is low in calories and fat, but is a rich source of dietary fiber, vitamins, minerals and antioxidant phenolic compounds. Thus, the proposed analytical technique has proven to be useful to characterize 61 phenolic and other polar compounds, 33 being phenolic compounds with their isomers, which have been tentatively characterized in artichoke for the first time, namely: 3 hydroxybenzoic acids, 17 hydroxycinnamic acids, 4 lignans, 7 flavones, 2 flavonols, and one phenol derivative. Furthermore, a total of 28 isomers of previously described phenolics have also been detected.

Finally, chapter 5 shows a deep comprehensive profiling of the "non-targeted" phytochemical compounds present in the hydro-methanolic extract of three cultivars of lettuce, namely, baby, romaine, and iceberg. Thus, an

UPLC-ESI-QTOF-MS method was used as a powerful tool to ensure the wide detection of phytochemical compounds from different families with highly deviating properties. In this sense, the proposed method tentatively characterized a total 171 compounds belonging to various structural classes, many of which are reported for the first time in lettuce, phenolics being the most abundant compounds among the phytochemicals characterized.

Resumen

RESUMEN

La presente tesis doctoral titulada "Caracterización de compuestos fenólicos en matrices vegetales de alto consumo mediante técnicas analíticas avanzadas", consta de dos secciones principales: introducción y parte experimental.

La primera sección describe un panorama general sobre la dieta mediterránea en cuanto a sus alimentos clave, particularmente frutas y verduras, así como su importancia como alimentos de origen vegetal altamente consumidos. Por otro lado se describe el interés de estos alimentos de origen vegetal como una buena fuente de compuestos bioactivos y fitoquímicos dando un enfoque especial a los compuestos fenólicos. Por último, se describen las herramientas analíticas utilizadas para caracterizar estos compuestos fenólicos de interés, como son la cromatografía líquida de alta resolución (HPLC) y la cromatografía líquida de ultra alta resolución (UPLC) acopladas a diferentes sistemas de detección como la espectrofotometría UV-visible (UV-vis) y la espectrometría de masas (MS) empleando para ello analizadores de tiempo de vuelo (TOF) y cuadrupolo-tiempo de vuelo (Q-TOF) acopladas a LC mediante ionización por electrospray (ESI).

La parte experimental se divide en cinco capítulos de la siguiente manera:

En el **capítulo 1** se propone un método para separar y caracterizar los compuestos fitoquímicos presentes en tres variedades de vaina de judía verde (*Phaseolus vulgaris L*). Esta verdura es muy baja en calorías y no contienen grasas saturadas y una muy buena fuente de vitaminas, minerales y fitoquímicos. En este sentido, se propuso un método de extracción con el objetivo de extraer los compuestos fitoquímicos presentes en las muestras estudiadas usando para este fin una mezcla polar de metanol/agua. Posteriormente se llevó a cabo el análisis de estos extractos mediante HPLC acoplada a TOF-MS. Así, el método utilizado permitió caracterizar un total de 72 compuestos fitoquímicos, 54 de ellos caracterizados por primera vez en este tipo de vegetales, haciendo uso del tiempo de retención, los datos generados por el TOF-MS y la literatura. Se encontró que la variedad Strike era la más rica en flavonoides en comparación con las otras variedades analizadas.

En el **capítulo 2** se llevó a cabo la caracterización de los compuestos fenólicos presentes en muestras de pepino (*Cucumis sativus L.*), el cual además de ser un cultivo muy importante y altamente consumido, es una fuente rica en compuestos fenólicos. Por ello y debido a la escasez de datos sobre el pepino en cuanto a compuestos fenólicos, en el presente capítulo se propuso un método sencillo y rápido para caracterizar los compuestos fenólicos en un extracto de pepino utilizando HPLC-ESI-QTOF-MS. Así, basándonos en la exactitud de los datos de masa y en los patrones de fragmentación generados por el QTOF-MS, se pudieron identificar de manera tentativa 73 compuestos fenólicos en la matriz estudiada, siendo esta investigación la primera caracterización exhaustiva de compuestos fenólicos presentes en pepino.

El **capítulo 3** describe una caracterización pormenorizada de compuestos fenólicos y otros compuestos fitoquímicos polares en un extracto de la porción comestible de sandía ya que es una fruta altamente cultivada y consumida sobre todo en épocas de verano. Así, el empleo de HPLC-ESI-QTOF-MS se convirtió en una poderosa herramienta analítica para la detección de 71 compuestos de diferentes familias entre las que destacan ácidos fenólicos, flavonoides, iridoides, cumarinas y lignanos, utilizando los datos de MS y MS/MS proporcionados por el QTOF-MS junto con los datos obtenidos a partir de la literatura pertinente sobre la misma familia botánica. Por lo que se conoce, esta investigación podría ser la primera en la que se lleva a cabo la caracterización de compuestos fenólicos y otros compuestos fotoquímicos polares en la sandía de forma exhaustiva.

El **capítulo 4** versa sobre la caracterización de un extracto de alcachofa (*Cynara scolymus*) utilizando para este fin HPLC-DAD-ESI-QTOF-MS como herramienta analítica. La alcachofa es una verdura baja en calorías y grasas, pero es una fuente rica en fibra dietética con grandes cantidades de vitaminas, minerales, antioxidantes y compuestos fenólicos. Así, la metodología analítica propuesta ha demostrado ser útil para caracterizar 61 compuestos fenólicos así como otros fitoquímicos polares, siendo 33 de ellos compuestos fenólicos, con sus isómeros, que han sido tentativamente caracterizados en muestras de alcachofa por primera vez.

Por último, en el **capítulo 5** se llevó a cabo un estudio, con un enfoque más metabolómico, para la caracterización de compuestos fitoquímicos polares presentes en extractos de diferentes variedades de lechuga (Baby, Romaine y Iceberg) mediante UPLC-ESI-QTOF-MS para garantizar una completa detección de sus fitoquímicos. En este sentido, con el método propuesto se pudieron identificar tentativamente un total de 171 compuestos pertenecientes a diferentes clases estructurales, muchos de los cuales son la primera vez que se caracterizan en muestras de lechuga, siendo de ellos los compuestos fenólicos los más abundantes entre los fitoquímicos caracterizados.

Objective

OBJECTIVE

There is substantial epidemiological evidence that a Mediterranean dietary pattern, a plant-based diet, can reduce risk of cardiovascular diseases, cancer, diabetes, and even aging-associated cognitive decline. The Mediterranean diet is consists mainly of high consumption of fruits and vegetables rich in phytochemical compounds. Over the last two decades, interest has been a steadily growing interest in the presence of phytochemical compounds in foods and also a great need for understanding how these compounds can benefit human health through the diet. The dietary intake of phenolic compounds has been a focus point of many recent publications and it has drawn much attention to the health benefits that these compounds may offer as antioxidant/anti-radical, anti-carcinogenic, anti-inflammatory, antiviral, and antimicrobial agents, etc. Thus, information on food composition is necessary to assess diet quality and the development and application of food-based dietary guidelines, providing a useful tool for the field of public-health nutrition. Although, most food composition tables focus on energy and macro- and micronutrients, interest in non-nutritive components, such as phenolic compounds, is increasing. In view of the beneficial effects of biologically active secondary plant-cell compounds, such as phenolic compounds, more data on these are needed.

In this context, the aim of this doctoral thesis has been to characterize phytochemical compounds in the edible fraction of plant matrices, sometimes from a metabolic point of view in order to characterize all the phytochemical compounds, such as green beans, cucumber, watermelon, artichoke, and lettuce by using liquid chromatography (LC) coupled with two different detection systems: UV-visible and mass spectrometry (MS) using time-of-flight (TOF) and quadrupole time-of-flight (QTOF) analyzers. The TOF analyzer provides excellent mass accuracy over a wide, dynamic range and enables measurements of the isotopic pattern, providing important additional information to determine the elemental composition. Lastly, the QTOF analyzer offers an extra advantage, since it can provide both MS and MS/MS

analyses, saving analysis time and having a less equipment use for identifying target compounds in highly complex matrices.

Objetivo

OBJETIVO

La dieta mediterránea es una dieta basada fundamentalmente en el alto consumo de frutas y verduras, a las cuales se les atribuyen numerosas propiedades beneficiosas, como reducir el riesgo de padecer enfermedades cardiovasculares, cáncer, diabetes e incluso el envejecimiento asociado al deterioro cognitivo. En las últimas dos décadas ha habido un creciente interés en estudiar los compuestos fitoquímicos presentes en los alimentos además de la necesidad de entender como estos compuestos pueden producir efectos beneficiosos para la salud humana a través de la dieta. Es por ello por lo que es necesario tener una mayor información sobre la composición de los alimentos que permitan una mejor evaluación de la calidad de la dieta, su desarrollo así como aplicar los alimentos a las directrices dietéticas, proporcionando con ello una herramienta útil para el campo de la nutrición y salud pública. Hoy día, aunque la mayoría de las tablas de composición de alimentos se centran en datos de energía, macro y micronutrientes, el interés en componentes no nutritivos, tales como compuestos fenólicos, está siendo de enorme interés. Es por ello por lo que, considerando los efectos beneficiosos de compuestos biológicamente activos tales como los compuestos fenólicos, se precisan más datos en cuanto a estos compuestos en los alimentos.

En este contexto, el objetivo principal de esta tesis doctoral ha sido caracterizar los compuestos fitoquímicos presentes, a veces introduciendo un punto de vista metabólico, en la parte comestible de matrices vegetales tales como judía verde, pepino, sandía, alcachofa y lechuga mediante el uso de cromatografía líquida (LC) acoplada a diferentes sistemas de detección: UV-visible y espectrometría de masas (MS) utilizando analizadores de tiempo de vuelo (TOF) y cuadrupolo tiempo de vuelo (QTOF). El empleo del analizador TOF se debe a que éste es capaz de proporcionar datos de masa exacta en un amplio intervalo además de realizar mediciones del patrón isotópico, proporcionando con ello una información adicional para la determinación de la composición elemental. En lo referente al analizador QTOF es capaz de ofrecer una ventaja adicional frente al TOF ya que puede proporcionar datos

tanto de MS como de MS/MS proporcionando información de patrones de fragmentación permitiendo con ello una mejor identificación de los compuestos de interés en matrices muy complejas.

Introduction

1. The Mediterranean diet: An overview

In the history of human nutrition, one of the most widespread alimentary regimens linked to the promotion of health is represented by the Mediterranean diet, which evolved naturally in the populations that settled along the Mediterranean Sea and used particular foods that grow well in this area[1].

Although the Mediterranean food culture is as diverse as the societies of the Mediterranean region, there are several key ingredients that unite the foods of the region. Historians frequently describe how the great civilizations of the past, such as the Phoenicians, Greeks, Romans, Carthaginians, and Arabs exerted a unifying influence on philosophy, art and architecture, and technology in the Mediterranean region. These civilizations also contributed to the food culture as well, and thus there are many similarities in the kinds of ingredients and preparation methods throughout the Mediterranean[2]. Today, the term "Mediterranean diet" refers to a combination of foods that include fresh fruits and vegetables, fish, pasta, bread, and a key ingredient, olive oil. The term "Mediterranean diet" was first used by Keys[3], referring to a general model of dietary behavior he detected in southern Italy and the island of Crete in the 1960s and 1970s[4].

Several definitions and particular versions (scores) of the Mediterranean diet have been proposed. These include, among others, the following several characteristics: (1) high consumption of vegetables and fruit; (2) low consumption of red meat; (3) low to moderate consumption of fish; (4) high consumption of cereals; and (5) last, and perhaps foremost, olive oil as the principal source of fat, which may be a unique common feature of Mediterranean diet[5].

The Mediterranean diet can be also defined as a modern nutritional recommendation inspired by the traditional dietary patterns of the populations bordering the Mediterranean. The Mediterranean basin includes three continents and more than 15 countries, including southern part of Europe (e.g. Portugal, Spain, Italy, and Greece),

[1] Keys, A. 1995. Mediterranean diet and public health: personal reflections. *The American Journal of Clinical Nutrition,* **61,** 1321S-1323S.
[2] Helstosky, C. 2009. *Food Culture in the Mediterranean* Greenwood Press.
[3] Keys, A. 1980 (Ed.), Seven Countries. *A Multivariate Analysis of Death and Coronary Heart Diseases*. Cambridge: Harvard University Press.
[4] Cloutier, M. and Adamson, E. 2004. *The Mediterranean Diet* HarperCollins.
[5] Gallus, S., Bosetti, C. and Vecchia, C. L. 2004. Mediterranean diet and cancer risk. *European Journal of Cancer Prevention,* **13,** 447-452.

North Africa (e.g. Morocco, Algeria, Tunisia, Libya, and Egypt), and part of Asia (Asia Minor and the Middle East, e.g. Turkey, Syria, Lebanon, Palestine) **(Figure 1)**. Traditional dietary habits in Mediterranean countries represent a food model that promotes health and fosters a high quality of life[6,7].

Figure 1. The area bordering the Mediterranean Sea

The virtues of the Mediterranean diet have been advocated ever since the Renaissance. More recently, since the 1980s, numerous epidemiological studies have shown that the people of Mediterranean countries have a long life expectancy and a low risk of suffering certain chronic diseases, including cardiovascular disease, metabolic disorders and certain types of cancer. Many authors have underlined the beneficial role of the Mediterranean diet with respect to lipid metabolism, blood pressure, body-mass index, inflammation, and blood coagulation[8,9].

The food pyramid **(Figure 2)** reflects Mediterranean dietary traditions, which historically have been associated with good health. This Mediterranean diet pyramid is based on food patterns typical of Crete, much of the rest of Greece, and southern Italy in the early 1960s, where the adult life expectancy was among the highest in the world and rates of coronary heart disease, certain cancers, and other diet-related chronic diseases were among the lowest.

[6] Cloutier, M. and Adamson, E. 2004. *The Mediterranean Diet* Harper Collins.
[7] Demarin, V., Lisak, M. and Morović, S. 2011. Mediterranean Diet in Healthy Lifestyle and Prevention of Stroke. *Acta Clinica Croatia*, **50**, 67-77.
[8] Puga, F. L. and Urquiaga, I. The Mediterranean Diets: Nutrition and Gastronomy. In *Functional Food Product Development* pp. 322-343, Wiley-Blackwell.
[9] Gallus, S., Bosetti, C. and Vecchia, C. L. 2004. Mediterranean diet and cancer risk. *European Journal of Cancer Prevention*, **13,** 447-452.

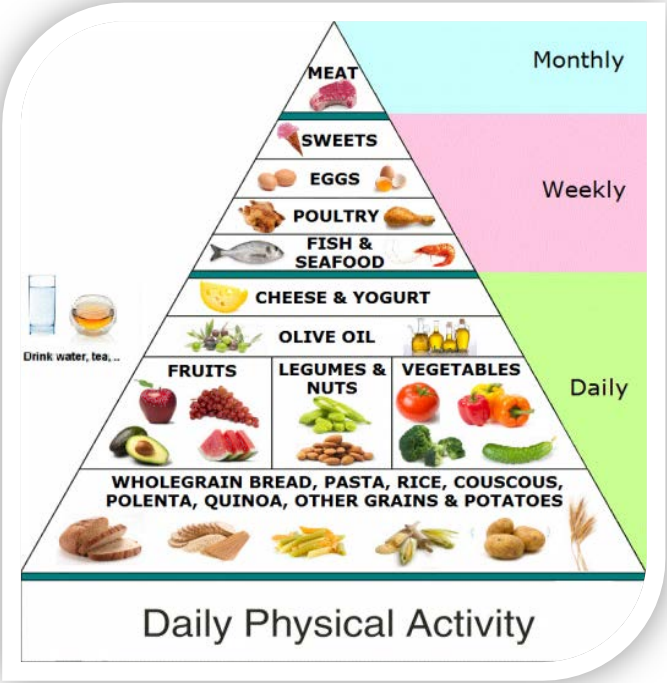


Figure 2. Mediterranean Diet Pyramid model

Work in the field or kitchen resulted in a lifestyle that included regular physical activity and was associated with low rates of obesity[10].

As reflected in this scheme, the Mediterranean diet is characterized mainly by abundant plant foods: fruit, vegetables, breads (and other forms of cereals), potatoes, beans, nuts, and seeds, with fresh fruit as the typical daily dessert, and olive oil as the principal source of fat. Meanwhile, dairy products (mainly cheese and yogurt), fish, and poultry are consumed in low to moderate amounts; zero to four eggs are consumed weekly; and red meat is consumed in low amounts. The pyramid represents a dietary pattern that is attractive for its famous palatability as well as for its health benefits[11].

A number of diets have received attention, but the strongest evidence for a beneficial health effect and decreased mortality after switching to a largely plant-based diet

[10] Willett, W. C., Sacks, F., Trichopoulou, A., Drescher, G., Ferro-Luzzi, A., Helsing, E. And Trichopoulos, D. 1995. Mediterranean diet pyramid: a cultural model for healthy eating. *The American Journal of Clinical Nutrition,* **61**, 1402S-1406S.

[11] Willett, W. C., Sacks, F., Trichopoulou, A., Drescher, G., Ferro-Luzzi, A., Helsing, E. and Trichopoulos, D. 1995. Mediterranean diet pyramid: a cultural model for healthy eating. *The American Journal of Clinical Nutrition,* **61,** 1402S-1406S.

comes from studies of Mediterranean diet[12]. Furthermore, this diet is often cited as beneficial for being low in saturated fat but high in monounsaturated fat and dietary fiber. One of the main explanations concerns the health effects of the abundant use of olive oil in this diet.

Dietary factors are only part of the reason for the health benefits enjoyed by certain Mediterranean cultures. A healthy lifestyle (notably a physically active lifestyle or physical labor) is also beneficial[13,14]. Therefore, the health effects of the Mediterranean diet are credited to factors such as small portions, daily exercise, and the emphasis on freshness, balance, and pleasure regarding food[15]. The environment may also be involved. However, on the population level, i.e. for the population of a whole country or a region, the influence of genetics is apparently rather minimal, given that it has been shown that significant increases in the risk of heart disease have resulted from the slowly changing habits of Mediterranean populations, from a healthy active lifestyle and Mediterranean diet to a not-so-healthy, less physically active lifestyle together with a diet influenced by predominant Western patterns[16-18]. There is an inverse association between adherence to the Mediterranean diet and the incidence of fatal and non-fatal heart disease in initially healthy middle-aged adults in the Mediterranean region[19]. A 10-year study published in the *Journal of American*

[12] Mitrou, P. N., Kipnis, V., Thiébaut, A. C. M., Reedy, J., Subar, A. F., Wirfält, E., Flood, A., Mouw, T., Hollenbeck, A. R., Leitzmann, M. F. and Schatzkin, A. 2007. Mediterranean dietary pattern and prediction of all-cause mortality in a US population: Results from the NIH-AARP diet and health study. *Archives of Internal Medicine,* **167,** 2461-2468.

[13] Dahlöf, B. R. 2010. Cardiovascular Disease Risk Factors: Epidemiology and Risk Assessment. *The American Journal of Cardiology,* **105,** 3A-9Al.

[14] Pollock, M. L., Franklin, B. A., Balady, G. J., Chaitman, B. L., Fleg, J. L., Fletcher, B., Limacher, M., Piña, I. L., Stein, R. A., Williams, M. and Bazzarre, T. 2000. Resistance Exercise in Individuals with and Without Cardiovascular Disease: Benefits, Rationale, Safety, and Prescription an Advisory from the Committee on Exercise, Rehabilitation, and Prevention, Council on Clinical Cardiology, American Heart Association. *Circulation,* **101,** 828-833.

[15] Guiliano, Mireille. *French Women Don't Get Fat*. Knopf.

[16] Vardavas, C. I., Linardakis, M. K., Hatzis, C. M., Saris, W. H. M. and Kafatos, A. G. Cardiovascular disease risk factors and dietary habits of farmers from Crete 45 years after the first description of the Mediterranean diet. *European Journal of Cardiovascular Prevention & Rehabilitation,* **17,** 440-446.

[17] Kafatos, A., Diacatou, A., Voukiklaris, G., Nikolakakis, N., Vlachonikolis, J., Kounali, D., Mamalakis, G. and Dontas, A. S. 1997. Heart disease risk-factor status and dietary changes in the Cretan population over the past 30 y: the Seven Countries Study. *The American Journal of Clinical Nutrition,* **65,** 1882-1886.

[18] Menotti, A., Keys, A., Kromhout, D., Blackburn, H., Aravanis, C., Bloemberg, B., Buzina, R., Dontas, A., Fidanza, F., Giampaoli, S., Karvonen, M., Lanti, M., Mohacek, I., Nedeljkovic, S., Nissinen, A., Pekkanen, J., Punsar, S., Seccareccia, F. and Toshima, H. 1993. Inter-cohort differences in coronary heart disease mortality in the 25-year follow-up of the seven countries study. *European Journal of Epidemiology,* **9,** 527-536.

[19] Martínez-González, M. A., García-López, M., Bes-Rastrollo, M., Toledo, E., Martínez-Lapiscina, E. H., Delgado-Rodriguez, M., Vazquez, Z., Benito, S. and Beunza, J. J. 2010. Mediterranean diet and the

Medical Association (JAMA) found that adherence to a Mediterranean diet and a healthful lifestyle was associated with more than a 50% lowering of early-death rates[20].

Moreover, several studies demonstrated the beneficial and preventive role of Mediterranean diet on the occurrence of cardiovascular diseases[21], chronic neurodegenerative diseases and neoplasm, diabetes, and obesity[22]. Therefore, this diet is frequently prescribed for patients having such illnesses, although it can also be used as a weight-reduction regimen, and it is thought by some to prevent certain forms of cancer. It is now firmly established as a model of healthful eating habits, and it has been adapted to suit many regional food tastes[23].

The presumed benefits of the Mediterranean diet for cardiovascular health are primarily correlative in nature; although they reflect a very real disparity in the geographic incidence of heart disease, identifying the causal determinant of this disparity has proven difficult. The most popular dietary candidate, olive oil, has been undermined by a body of experimental evidence that diets enriched in monounsaturated fats, such as olive oil, are not atheroprotective when compared to diets enriched in either polyunsaturated or even saturated fats[24,25]. A recently emerging alternative hypothesis to the Mediterranean diet is that differential exposure to solar ultraviolet radiation accounts for the disparity in cardiovascular health between residents of Mediterranean and more northerly countries. The proposed mechanism is solar UVB-induced synthesis of Vitamin D in the oils of the skin, which has been observed to reduce the incidence of coronary heart disease, and which rapidly diminishes with increasing latitude[26]. It bears noting that residents of the

incidence of cardiovascular disease: A Spanish cohort. *Nutrition, metabolism, and cardiovascular diseases: NMCD,* **21,** 237-244.

[20] Knoops, K. T., De Groot, L. C. P. G. M., Kromhout, D., Perrin, A.-E., Moreiras-Varela, O., Menotti, A. and Van Staveren, W. A. 2004. Mediterranean diet, lifestyle factors, and 10-year mortality in elderly european men and women: The hale project. *JAMA,* **292,** 1433-1439.

[21] Nuñez-Cordoba, J. M., Alonso, A., Beunza, J. J., Palma, S., Gomez-Gracia, E., and Martinez-Gonzalez, M. A. 2009. Role of vegetables and fruits in Mediterranean diets to prevent hypertension, *European Journal of Clinical Nutrition,* **63**, 605-612.

[22] Cloutier, M. and Adamson, E. 2004. *The Mediterranean Diet* HarperCollins.

[23] Demarin V., Lisak M. and Morović S. 2011. Mediterranean Diet in Healthy Lifestyle and Prevention of Stroke. *Acta Clinica Croatica,* **5**,67-77.

[24] Mark Brown, J., Shelness, G. and Rudel, L. 2007. Monounsaturated fatty acids and atherosclerosis: Opposing views from epidemiology and experimental animal models. *Current Atherosclerosis Reports,* **9,** 494-500.

[25] Rudel, L. L., Kelley, K., Sawyer, J. K., Shah, R. and Wilson, M. D. 1998. Dietary Monounsaturated Fatty Acids Promote Aortic Atherosclerosis in LDL Receptor-null, Human ApoB100-overexpressing Transgenic Mice. *Arteriosclerosis, Thrombosis, and Vascular Biology,* **18,** 1818-1827.

[26] Wong, A. 2008. Incident solar radiation and coronary heart disease mortality rates in Europe. *European Journal of Epidemiology,* **23,** 609-614.

Mediterranean also register very low rates of skin cancer (which is widely believed to be caused by over-exposure to solar UV radiation); and the incidence of melanomas is lower than in Northern Europe and significantly lower than in other hot countries such as Australia and New Zealand. It has been hypothesized that some components of the Mediterranean diet may provide protection against skin cancer.

On the other hand, a study published in the Archives of General Psychiatry shows that people who followed the Mediterranean diet were less likely to develop depression[27]. Also, a New Zealand study in 2012 and 2013 found the Mediterranean diet could reduce the risks of Alzheimer's disease after only six weeks[28].

2. Fruits and vegetables as a vital part of the Mediterranean diet

Fruits and vegetables are one of the key foods characterizing the Mediterranean diet. In fact, results found in the Mediterranean region support the contention that a primarily plant-based diet may indeed promote a longer life and better health[29]. Plant foods, such as fruits and vegetables, appear to be particularly rich in nutrient and non-nutrient antioxidants besides being the crucial items in the current Mediterranean diet, contributing to the persistence of health-promoting attributes[30].

In general, fruits and vegetables (relatively, all plant foods, with the possible exception of coconut and palm oils) have their highly recommended qualities. Moreover, each type of plant food has a unique combination of nutrients and phytochemicals. Therefore, a variety of plant foods is more important than choosing a few recommended plant foods and eating only those. The more variety in the diet, the more nutritionally complete the diet will be[31]. A large body of evidence indicates the positive effect of fruit and vegetable consumption on health, suggesting that the whole may be more than the sum of the parts. That is, individual components appear

[27] Sanchez-Villegas, A., Delgado-Rodríguez, M., Alonso, A., Schlatter, J., Lahortiga, F., Majem, L. S., Martínez-González, M. A. 2009. Association of the mediterranean dietary pattern with the incidence of depression: The seguimiento universidad de navarra/university of navarra follow-up (sun) cohort. *Archives of General Psychiatry,* **66,** 1090-1098.

[28] "Diet reduces illness risk in six weeks". 3 News NZ. Retrieved on 15 January 2013.

[29] Demarin, V., Lisak, M. and Morović, S. 2011. Mediterranean Diet in Healthy Lifestyle and Prevention of stroke. *Acta Clinica Croatia*, **50,** 67-77.

[30] Ferro-Luzzi, A., Cialfa, E., Leclercq, C. and Toti, E. 1994. The Mediterranean diet revisited. Focus on fruit and vegetables. *International Journal of Food Sciences and Nutrition,* **45,** 291-300.

[31] Helstosky, C. 2009. *Food Culture in the Mediterranean* Greenwood Press.

to act synergistically whereby the influence of at least some of these becomes additive[32].

2.1. Green beans (*Phaseolus vulgaris L.*)

Phaseolus vulgaris L. is the best known and most widely cultivated bean in the world and it has many common or local names. Archaeological remains of this legume date to about 5000 BC. In temperate areas, the bean is grown mainly for the young pods, under the general term green beans. Pods contain a large amount of water, about 2% protein, 0.5% fat, 3% carbohydrate with carotenes, vitamins B, C, and E. They are marketed fresh, canned, and frozen. The common bean is one of the most important food legumes, consumed worldwide as pods of green beans or seeds of dry beans. The dried seeds (pulses) contain 22% protein, 1.6% fat, and 50% carbohydrate, and only vitamins B and E[33]. On the other hand, green pods are a superior source of calcium, iron, and vitamin C.

Figure 3. Green beans (*Phaseolus vulgaris L.)*

Green beans **(Figure 3)**, also known as the French bean, haricot, kidney bean, snap bean, or string-beans, are the unripe fruit of any kind of bean. Green bean cultivars have been bred especially for the fleshiness, flavor, or sweetness of the pods.

Green bean is an excellent remedy against hydropsy, diseases related to the kidney, and bladder dysfunctions. Bean-pod tea can help in cases of spasms, difficulty with urination, and can help the body remove harmful toxins. It is useful as remedy against uric acid accumulation and the loss of albumin in the urine during pregnancy. Infusions made from the bean pods can help those suffering from hyperglycemia. In addition, it can help prevent diabetes. Used externally, bean flour can be beneficial in treatment of various skin conditions, such as acne, eczema, sores and ulcers[34]. In the prevention

[32] De La Rosa, L. A., Alvarez-Parrilla, E. and Gonzalez-Aguilar, G. A. 2009. *Fruit and Vegetable Phytochemicals: Chemistry, Nutritional Value and Stability* Wiley.
[33] Vaughan, J. and Geissler, C. 2009. *The New Oxford Book of Food Plants* OUP Oxford.
[34] http://health-from-nature.net/Bean-Common.html, retrieved on 12 Dec. 2013.

of atherosclerosis and heart disease (of diabetic origin) green beans are rich in antioxidant effect, helping to lower the level of free radicals in the body and prevent the oxidation of cholesterol. Also, bean help prevent cardiovascular disease are also of benefit in lowering blood pressure.

Moreover, green beans have powerful anti-inflammatory effects and valuable for a healthy immune system. Therefore, green beans are beneficial both to the diseases in which inflammation plays a major role in asthma, osteoarthritis, and rheumatoid arthritis, and can reduce frequency of migraines. It can also help combat infectious diseases such as the common cold, flu, tonsillitis, and otitis. For their content in phytochemicals, green beans have an important role in preventing colon cancer. They protect colon cells from the damaging action of the free radicals, prevent DNA damage, and discourage mutations in the colon-cell structure[35].

2.2. Cucumber (*Cucumis sativus L.*)

Cucumber (*Cucumis sativus L.*) belongs to the genus Cucumis of the family Cucurbitaceae. This vegetable is grown mostly in temperate to sub-tropical regions and is considered among the top ten vegetables in world production[36]. *Cucumis sativus* L. is thought to have originated in the foothills of the Himalayas, possibly from the wild *Cucumis hardwickii*. In India, the cucumber was cultivated some 3000 years ago and was known in ancient Egypt, Greece, and Rome. Today, cucumbers are found worldwide. The fruits are normally harvested before the seeds mature. They are eaten raw (usually peeled, depending on the cultivar), sliced, and served in salads, or they can be consumed in their common processed form (pickled).

Cucumbers contain about 96% water with 2% sugars (glucose and fructose). They also contain small amounts of protein, fat, and vitamin B complex, and has 2–8 mg/100 g vitamin C. Unpeeled cucumbers contain moderate amounts of carotenes[37]. *Cucumis Sativus L.* is the main species, extensively cultivated in many cultivars for its edible fruit, cucumber (**Figure 4**), which has been a major part of traditional Mediterranean diet since antiquity. Cucumbers are reportedly used in folk medicine as a treatment against diarrhea, gonorrhea, inflammation, hypertension, and diabetes mellitus, and as a serum-lipid regulator, an antioxidant, and an analgesic.

[35] http://www.liveandfeel.com/vegetables/Green_beans.html, retrieved on 21 Jan. 2013.
[36] Zhou, X.-H., Qian, C.-T., Lou, Q.-F. and Chen, J.-F. 2009. Molecular analysis of introgression lines from Cucumis hystrix Chakr. to C. sativus L. *Scientia Horticulturae,* **119,** 232-235.
[37] Vaughan, J. and Geissler, C. 2009. *The New Oxford Book of Food Plants* OUP Oxford.

Figure 4. Cucumber (*Cucumis Sativus* L.)

2.3. Watermelon (*Citrullus lanatus*)

Watermelon (*Citrullus lanatus or C. vulgaris*), which probably originated in Africa, belongs to the Cucurbitaceae family, which includes cucumbers, squash, and gourds. The first records of its cultivation in the Mediterranean region date to some 3000 years ago[38].

Watermelon, **(Figure 5)** is a popular fruit, which is usually consumed fresh directly as slices or diced and mixed with other fruits (fruit salad). Additionally, it can be processed as juice, jam, or pickled rind and the seeds (from seed cultivars) are also consumed. This highly consumed fruit, besides being a thirst quencher in the summertime, is also a fruit that is consumed year round. Moreover, C. lanatus, one of the most widely cultivated crops in the world, and constitutes the largest production of all cucurbits[39].

[38] Vaughan, J. and Geissler, C. 2009. *The New Oxford Book of Food Plants* OUP Oxford.

[39] Helstosky, C. 2009. *Food Culture in the Mediterranean* Greenwood Press.

Figure 5. Watermelon (*Citrullus lanatus*)

Watermelon flesh contains over 90% water, providing a source of liquid in arid regions, but little protein or fat. The total sugar content is around 7%, with some carotenes and a vitamin-C content of 8 mg/100 g[40].

Medicinally, watermelons are mildly diuretic, being effective in the treatment of dropsy and renal stones, as well as reducing hypertension[41]. It also helps to prevent erectile dysfunction, acts as an antioxidant, and is used to treat enlarged liver and jaundice. In addition, it can also be used as an natural alternative to chemical treatment for giardiasis[42,43].

2.4. Artichoke (*Cynara scolymus*)

Globe Artichoke (*Cynara scolymus*), referring to the unopen flower-heads or "chokes", have numerous large scales (bracts) with fleshy bases. The globe artichoke is a thistle-like plant belonging to the Asteraceae (Compositae) family. Historically, it was a food plant for the Greeks and Romans[44,45]. Artichokes are a prickly, round vegetable that are popular in Mediterranean cooking[46], where their commercial production represents a considerable proportion of the agricultural economy. The flower-head may be baked,

[40] Vaughan, J. and Geissler, C. 2009. *The New Oxford Book of Food Plants* OUP Oxford.
[41] Figueroa, A., Sanchez-Gonzalez, M. A., Wong, A., and Arjmandi, B. H. 2012. Watermelon Extract Supplementation Reduces Ankle Blood Pressure and Carotid Augmentation Index in Obese Adults With Prehypertension or Hypertension. *American Journal of Hypertension,* **25**, 640-643.
[42] Erukainure O.L., Oke O.V., Daramola A.O., Adenekan S.O. and Umanhonlen E.E. 2010. Improvement of the Biochemical Properties of Watermelon Rinds Subjected to Saccharomyces cerevisae Solid Media Fermentation. *Pakistan Journal of Nutrition,* **9**, 806-809.
[43] Hassan L. E. A., Koko W. S., Osman E.-B. E., Dahab M. M. and Sirat H. M. 2011. In vitro antigiardial activity of Citrullus lanatus Var. citroides extracts and cucurbitacins isolated Compounds, *Journal of Medicinal Plants Research,* **5**, 3338-3346.
[44] Helstosky, C. 2009. *Food Culture In The Mediterranean* Greenwood Press.
[45] Vaughan, J. and Geissler, C. 2009. *The New Oxford Book of Food Plants* Oup Oxford.
[46] Fao Statistical Database, 2010, http://faostat.fao.org/site/339/default.aspx, retreived on 11 Dec. 2012.

fried, boiled, stuffed, served hot with various sauces, used as an ingredient in other dishes, or canned.

Artichokes contain about 3% protein and 3% carbohydrate, little fat, and a small amount of vitamin C. The tender inner flesh constitutes nearly 35–55% of the fresh weight of the head, depending on the cultivar and the harvesting time[47].

Globe artichoke **(Figure 6)** is also a promising source of biopharmaceuticals, such as luteolin and mono-/di-caffeoylquinic acids, which are responsible for its therapeutic effects[48]. On the other hand, the leaves are an herbal medicine and have been recognized since ancient times for their beneficial and therapeutic properties on humans[49].

Figure 6. Artichoke (*Cynara scolymus*)

Many studies have demonstrated artichoke to possess valuable medicinal properties, having antioxidative, anti-carcinogenic, cholesterol-lowering, hyper-lipidemic, antifungal and antibacterial properties[50-52].

[47] Lattanzio, V., Kroon, P. A., Linsalata, V. and Cardinali, A. 2009. Globe artichoke: A functional food and source of nutraceutical ingredients. *Journal of Functional Foods,* **1,** 131-144.

[48] Jun N. J., Jang K. C., Kim S. C., Moon D. Y., Seong K. C., Kang K. H., Tandang L., Kim P. H., Cho S. K. and Park K. H. 2007. Radical Scavenging Activity and Content of Cynarin (1,3-dicaffeoylquinic acid) in Artichoke (Cynara scolymus L.). *Journal of Applied Biological Chemistry,* **50**, 244-248.

[49] Lattanzio, V., Kroon, P. A., Linsalata, V. and Cardinali, A. 2009. Globe artichoke: A functional food and source of nutraceutical ingredients. *Journal of Functional Foods,* **1,** 131-144.

[50] Zhu, X. F., Zhang, H. X. and Lo, R. 2005. Antifungal activity of Cynara scolymus L. extracts. *Fitoterapia,* **76,** 108-111.

[51] Brown, J. E. and Rice-Evans, C. A. 1998. Luteolin-rich artichoke extract protects low density lipoprotein from oxidation In vitro. *Free Radical Research,* **29,** 247-255.

[52] Shimoda, H., Ninomiya, K., Nishida, N., Yoshino, T., Morikawa, T., Matsuda, H. and Yoshikawa, M. 2003. Anti-Hyperlipidemic sesquiterpenes and new sesquiterpene glycosides from the leaves of artichoke (Cynara scolymus L.): structure requirement and mode of action. *Bioorganic & Medicinal Chemistry Letters,* **13,** 223-228.

2.5. Lettuce (*Lactuca sativa*)

Lettuce (*Lactuca sativa L.*), of the Asteraceae (Compositae) family, is one of the most widely consumed leafy salad vegetables **(Figure 7)**, with world production exceeding some 24 million tons[53]. Lettuce can be an annual or biennial cultivated in temperate, subtropical, and tropical lands. Lettuce is a widely used as a minimally processed food product in the preparation of salads and ready-to-eat foodstuffs because it has a long shelf life and is beneficial for human health[54]. While lettuce leaves are commonly used in salads, stem lettuces can be boiled as a vegetable. Lettuce is grown not only as a food but also as a medicinal herb[55]. The flowering stems contain latex which has soporific effects and can be used as a sedative[56]. In the diet, lettuce provides little protein, fat, starch, or sugars, but is useful for fiber, minerals (a large amount of potassium), carotenes, vitamin E, and vitamin C (5mg/100 g). It has been hypothesized that lettuce consumption improves both the lipoprotein profile and the antioxidant status in humans,

Figure 7. Lettuce (*Lactuca sativa L.*)

resulting in a cardiovascular benefits[57]. Several studies have reported the health effects of lettuce in preventing cardiovascular diseases in laboratory animals and humans[58].

[53] FAO Statistical Database. 2010. http://faostat.fao.org/site/339/default.aspx, retreived on 19 July 2012.

[54] Cano, A. and Arnao, M. B. 2005. Hydrophilic and lipophilic antioxidant activity in different leaves of three lettuce varieties. *International journal of food properties,* **8,** 521-528.

[55] http://www.naturallifemagazine.com/9704/lettuce.htm. Retrieved on 12 Dec. 2012.

[56] Vaughan, J. and Geissler, C. 2009. *The New Oxford Book of Food Plants* OUP Oxford.

[57] Nicolle, C., Cardinault, N., Gueux, E., Jaffrelo, L., Rock, E., Mazur, A., Amouroux, P. and Rémésy, C. 2004. Health effect of vegetable-based diet: lettuce consumption improves cholesterol metabolism and antioxidant status in the rat. *Clinical Nutrition,* **23,** 605-614.

[58] Llorach, R., Martinez-Sanchez, A. N., Tomas-Barberan, F. A., Gil, M. A. I. and Ferreres, F. 2008. Charact-erisation of polyphenols and antioxidant properties of five lettuce varieties and escarole. *Food Chemistry,* **108,** 1028-1038.

3. Fruits and vegetables as natural sources of bioactive compounds

Fruits and vegetables contain significantly high levels of biologically active components that impart health benefits beyond basic nutrition. They are a major source of dietary antioxidants that increase the plasma antioxidant capacity, reducing mortality by inhibiting atherosclerosis-related diseases[59], cardio- and cerebrovascular diseases. Also, they may lower high blood pressure.

Recently, the search for specific plant components that offer health benefits has broadened to encompass the vast range of "non-nutritive" compounds present in plant foods, and their potential to improve health. This scientific inquiry was triggered by many epidemiologic studies that have shown the effects of plant-based diets against non-communicable diseases such as stroke cardiovascular diseases, diabetes, and cancer. A multitude of bioactive compounds have been discovered, which, individually or in combination, can benefit health[60,61]. The beneficial effects of these non-nutrients or bioactive substances are often used to ascribe such effects to the foods containing them. Evidence is growing that such plant constituents, belonging to the group termed **"Bioactive compounds"**, which are non-nutritional constituents that typically occur in small quantities in foods, may promote optimal health and help to reduce the risk of chronic diseases such as cancer, coronary heart disease, stroke, and perhaps Alzheimer's disease[62].

Therefore, some researchers have conferred on fruits and vegetables the status of **"functional foods"**. There are many biologically plausible reasons for associating plants with this sort of protection, including the fact that many of the phytochemicals act as antioxidants.

Bioactive components in plant food (fruits and vegetables) are very diverse and include a wide range of chemical compounds with varying structures, such as vitamins,

[59] Cao, G., Booth, S. L., Sadowski, J. A. and Prior, R. L. 1998. Increases in human plasma antioxidant capacity after consumption of controlled diets high in fruit and vegetables. *The American Journal of Clinical Nutrition,* **68,** 1081-1087.

[60] Stavric, B. 1994. Role of chemopreventers in human diet. *Clinical Biochemistry,* **27,** 319-332.

[61] Kris-Etherton, P. M., Hecker, K. D., Bonanome, A., Coval, S. M., Binkoski, A. E., Hilpert, K. F., Griel, A. E. and Etherton, T. D. 2002. Bioactive compounds in foods: their role in the prevention of cardiovascular disease and cancer. *The American Journal of Medicine,* **113,** 71-88.

[62] Kris-Etherton, P. M., Hecker, K. D., Bonanome, A., Coval, S. M., Binkoski, A. E., Hilpert, K. F., Griel, A. E. and Etherton, T. D. 2002. Bioactive compounds in foods: their role in the prevention of cardiovascular disease and cancer. *The American Journal of Medicine,* **113,** 71-88.

phytochemicals (e.g. phenolic compounds), isopernoids (terpenoids), protein/amino acid, carbohydrates (and derivatives), lipidic compounds, minerals [63-66] **(Figure 8)**.

One of the predominant mechanisms of the protective action of these components, their antioxidant activity, has been demonstrated *in vitro* to have the capacity to scavenge free radicals. Evidence indicates that for the effect of fruit and vegetable consumption on health, the whole may be more than the sum of the parts, as individual components appear to act synergistically, the influence of at least some of these being additive[67].

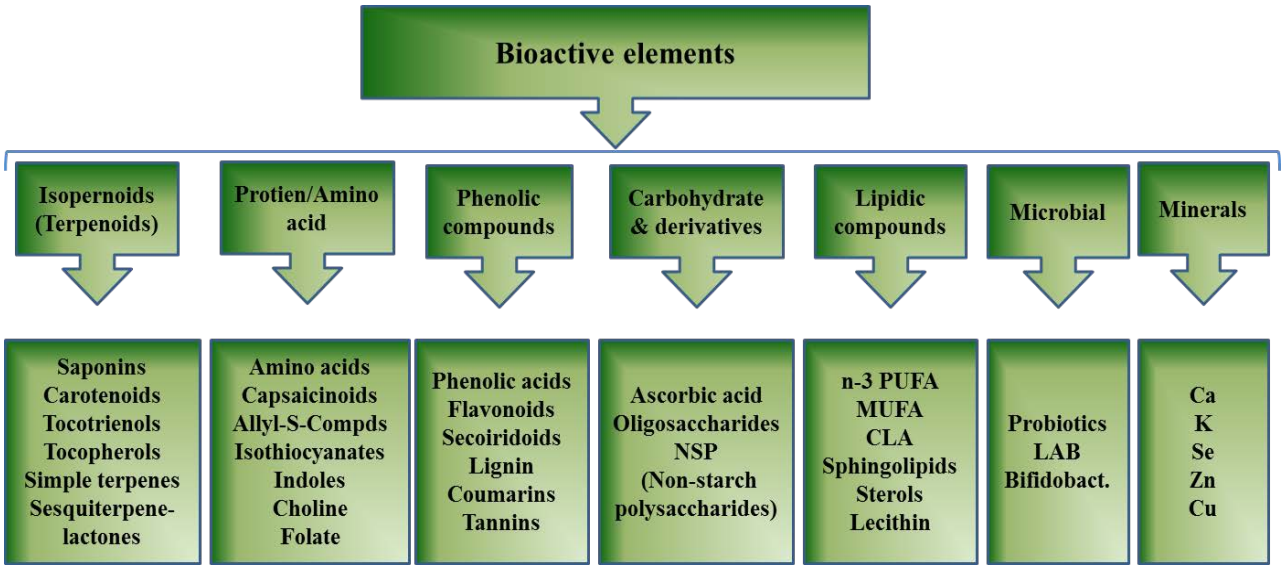


Figure 8. Bioactive phytochemicals in plant foods

The complexity of foods is well illustrated with reference to fruits and vegetables. In recent years, a growing amount of attention has been focused on fruits and vegetables and the role they might have in preventing a range of chronic diseases such as cancer, coronary heart disease, stroke, diabetes, cataracts, arthritis, Alzheimer's disease, and inflammatory bowel disease.

[63] Ajila, C. M., Brar, S. K., Verma, M., Tyagi, R. D., Godbout, S. and Valéro, J. R. Extraction and Analysis of Polyphenols: Recent trends. *Critical Reviews in Biotechnology,* **31,** 227-249.

[64] De La Rosa, L. A., Alvarez-Parrilla, E. and Gonzalez-Aguilar, G. A. 2009. *Fruit and Vegetable Phytochemicals: Chemistry, Nutritional Value and Stability* Wiley

[65] Baghurst, P., Beaumont-Smith, N., Baghurst, K., and Cox, D.1999.The relationship between the consumption of fruits and vegetables and health status, Report to Department of Health and Aged Care and the Strategic Intergovernmental Nutrition Alliance CSIRO Health Sciences & Nutrition.

[66] Liu, R. H. 2004. Potential Synergy of Phytochemicals in Cancer Prevention: Mechanism of Action. *The Journal of nutrition,* **134,** 3479S-3485S.

[67] De La Rosa, L. A., Alvarez-Parrilla, E. and Gonzalez-Aguilar, G. A. 2009. *Fruit and Vegetable Phytochemicals: Chemistry, Nutritional Value and Stability* Wiley.

Increased fruit and vegetable consumption has been advised by a number of international health authorities, partly in view of the abundance of protective components (bioactive compounds) in comparison to other harmful ones[68].

Although the natural combination of phytochemicals in fruit and vegetables is thought to be responsible for their potent antioxidant activity, the protective effects could be partly attributed to secondary phenolic metabolites[69].

4. Phenolic compounds: Structure and classification

Phenolic compounds are the most abundant antioxidants in human diets. It is estimated that more than 8,000 phenolic structures are currently known[70], with more than 500 phenolic compounds reported in foods[71]. Being secondary metabolites of plants which do not participate directly in growth and development, phenolic compounds are distributed among many types of fruits and vegetables and have specific functions, some influencing sensory properties (flavor and color)[72]. Furthermore, these compounds play key roles in protecting plants from herbivores, microorganisms, and UV radiation, in attracting pollinators or seed-dispersing animals, and in acting as stress-signaling molecules, in addition to other important functions[73].

The term "phenolic" or "polyphenol" can be defined chemically as a substance which possesses an aromatic ring bearing one hydroxy substituent or more, including functional derivatives (esters, methyl ethers, glycosides, etc.)[74]. Phenolic compounds can be divided into at least 10 different classes based upon their chemical structure, ranging from simple molecules, such as phenolic acids, to highly polymerized

[68] Baghurst, P., Beaumont-Smith, N., Baghurst, K., and Cox, D.1999.The relationship between the consumption of fruits and vegetables and health status, Report to Department of Health and Aged Care and the Strategic Intergovernmental Nutrition Alliance CSIRO Health Sciences & Nutrition.

[69] Crozier, A., Jaganath, I. B. and Clifford, M. N. 2009. Dietary phenolics: chemistry, bioavailability and effects on health. *Natural Product Reports,* **26,** 1001-1043.

[70] Pietta, P., Minoggio, M., Bramati, L. and ATTA-UR, R. 2003. Plant polyphenols: Structure, occurrence and bioactivity. In *Studies in Natural Products Chemistry* pp. 257-312, Elsevier.

[71] Smith, E. 2007. Plant secondary metabolites: occurrence, structure and role in the human diet. *Phytotherapy Research,* **21,** 904-904.

[72] Hounsome, N., Hounsome, B., Tomos, D. and Edwards-Jones, G. 2008. Plant Metabolites and Nutritional Quality of Vegetables. *Journal of Food Science,* **73,** R48-R65.

[73] Gomez-Romero, M. A., Segura-Carretero, A. and Fernandez-Gutierrez, A. Metabolite profiling and quantification of phenolic compounds in methanol extracts of tomato fruit. *Phytochemistry,* **71,** 1848-1864.

[74] Ho, C. T., Lee, C. and Huang, M. T. 1992. *Phenolic Compounds in Food and Their Effects on Health: Volume I: Analysis, Occurrence, and Chemistry* American Chemical Society.

compounds, such as tannins[70,75] **(Figure 9)**. They are commonly found conjugated to sugars and organic acids.

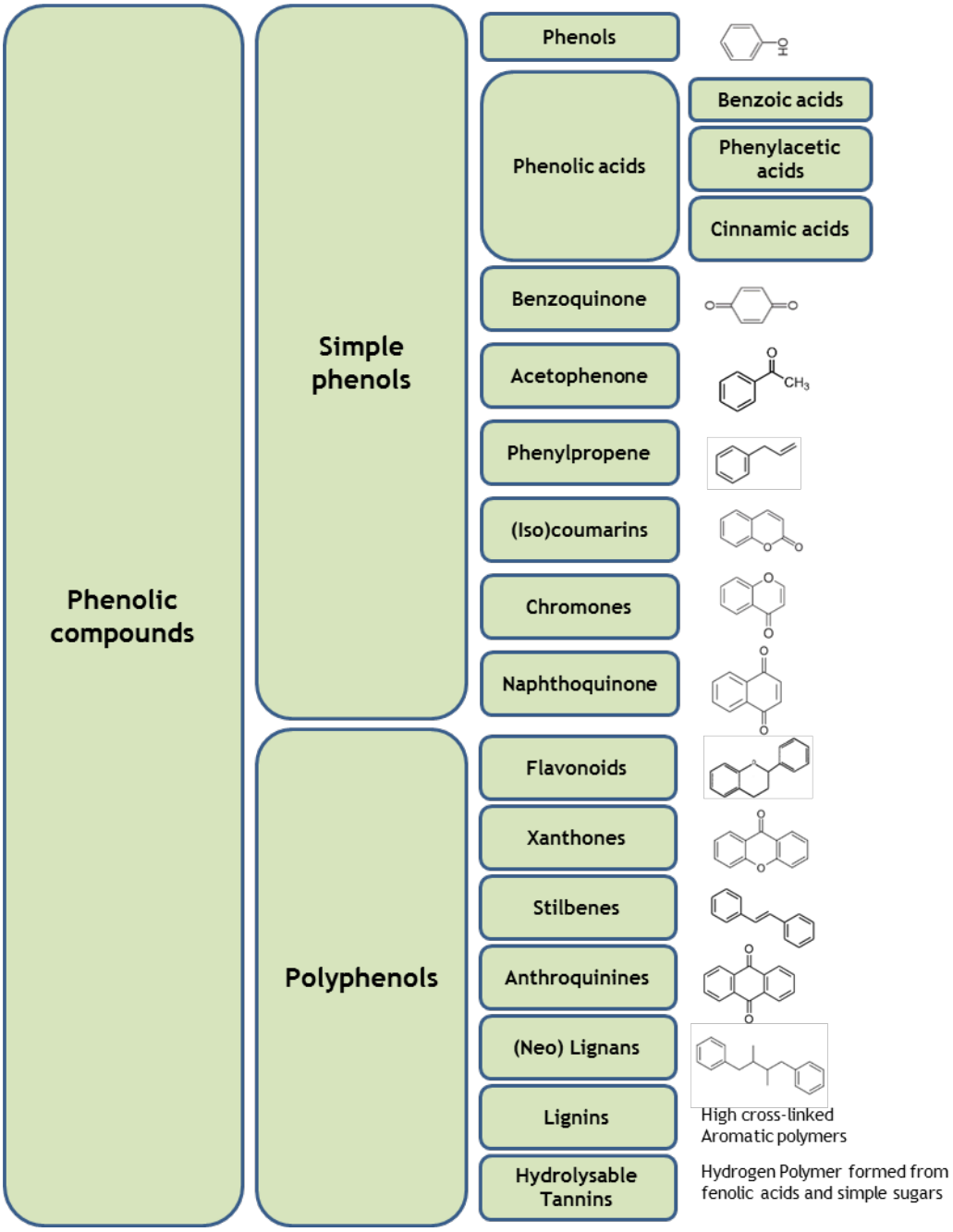


Figure 9. Classification of the principle families of phenolic compounds.

Phenolic compounds are not uniformly distributed in plants at the tissue at the cellular or subcellular levels. Insoluble phenolics are the components of cell walls, while soluble phenolics are compartmentalized within the plant-cell vacuoles. At the tissue

[75] Manach, C., Williamson, G., Morand, C., Scalbert, A. and Rémésy, C. 2005. Bioavailability and bioefficacy of polyphenols in humans. I. Review of 97 bioavailability studies. *The American Journal of Clinical Nutrition,* **81,** 230S-242S.

level, the outer layers of plants contain higher levels of phenolics than do the inner layers[76].

The two main subgroups of phenolic acids are derivatives of benzoic acid and derivatives of cinnamic acid **(Figure 10),** A third subgroup, which is less common, includes phenylacetic acid derivatives. Natural phenolic acids occur in fruits and vegetables, either in free or conjugated forms, usually as esters or amides[77]. The hydroxybenzoic acid present in edible plants is well known, with the exception of certain red fruits, black radish, and onions, which can have concentrations of several tens of milligrams per kilogram fresh weight[78]. Additionally, hydroxybenzoic acids are components of complex structures such as hydrolyzable tannins (gallotannins in mangoes, and ellagitannins in red fruit such as strawberries, raspberries, and blackberries)[79].

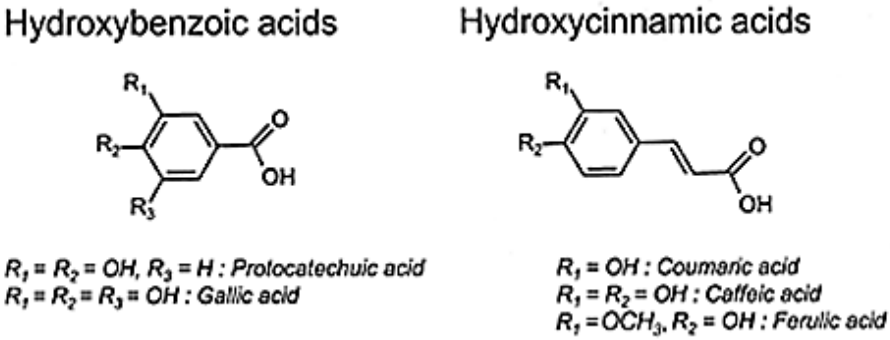


Figure 10. Chemical structure of the main phenolic acids

Hydroxycinnamic acids and their derivatives are almost exclusively derived from p-coumaric, caffeic, and ferulic acid, whereas sinapic acid is comparatively rare. Hydroxycinnamic acids usually occur in various conjugated forms, more frequently as esters than glycosides. The occurrences of hydroxycinnamic acids in human food are more common than hydroxybenzoic acids. These latter acids are rarely found in their free form, except in processed food that has undergone freezing, sterilization, or fermentation. The most important member of this group in food material is

[76] Naczk, M. and Shahidi, F. 2006. Phenolics in cereals, fruits and vegetables: Occurrence, extraction and analysis. *Journal of Pharmaceutical and Biomedical Analysis,* **41,** 1523-1542.
[77] Watson, R. R. and Preedy, V. R. 2009. *Bioactive Foods in Promoting Health: Fruits and Vegetables* Elsevier Science.
[78] Shahidi, F. and Naczk, M. 1995. *Food Phenolics: Sources, Chemistry, Effects, Applications* Technomic Publishing Company.
[79] *Clifford, M. N. and Scalbert, A. 2000. Ellagitannins – nature, occurrence and dietary burden. Journal of the Science of Food and Agriculture,* ***80,*** *1118-1125.*

chlorogenic acid, which is the key substrate for enzymatic browning, particularly in apples and pears[80].

The types of fruit having the highest concentrations (blueberries, kiwis, plums, cherries, apples) contain 0.5–2 g hydroxycinnamic acids per kg of fresh weight[81]. In this sense, p-coumaric acid can be found in a wide variety of edible plants such of as peanuts, tomatoes, carrots, and garlic. This acid has demonstrated antioxidant properties and is believed to lower the risk of stomach cancer by reducing the formation of carcinogenic nitrosamines[82,83]. Caffeic acid frequently occurs in fruits, grains and vegetables as simple esters with quinic acid (forming chlorogenic acid) or saccharides, and is also found in traditional Chinese herbs[84]. Finally, chlorogenic acid is found in particularly high concentrations in coffee, also isomers of this mono-caffeic acid are widely reported in Asteraceae plant family members. This compound, long known as an antioxidant, also slows the release of glucose into the blood stream after a meal[85].

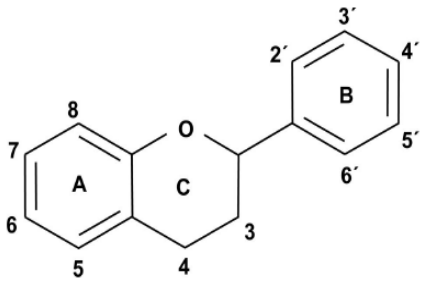


Figure 11. Basic structure of flavonoids.

The most important single group of phenolic compounds in food is the flavonoids, a chemically defined family of polyphenols that includes several thousand compounds. These have a C_6—C_3—C_6 basic structure **(Figure 11)**, and several subclasses of flavonoids which share a common structure consisting of 2 aromatic rings (A and B)

[80] Ho, C. T., Lee, C. and Huang, M. T. 1992. *Phenolic Compounds in Food and Their Effects on Health: Volume I: Analysis, Occurrence, and Chemistry* American Chemical Society.

[81] Macheix, J. J., Fleuriet, A. and Billot, J. 1990. *Fruit Phenolics* CRC Press.

[82] Ferguson, L. R., Zhu, S.-T. and Harris, P. J. 2005. Antioxidant and antigenotoxic effects of plant cell wall hydroxycinnamic acids in cultured HT-29 cells. *Molecular Nutrition & Food Research,* **49,** 585-593.

[83] Kikugawa, K., Hakamada,T., Hasunuma, M. and Kurechi, T. 1983. Reaction of p-hydroxycinnamic acid derivatives with nitrite and its relevance to nitrosamine formation. *Journal of Agricultural and Food Chemistry,* **31,** 780-785.

[84] Jiang, R. W., Lau, K. M., Hon, P. M., Mak, T. C., Woo, K. S. and Fung, K. P. 2005. Chemistry and biological activities of caffeic acid derivatives from Salvia miltiorrhiza. *Current Medical Chemistry,* **12,** 237-246.

[85] Clifford, M. N. 2000. Chlorogenic acids and other cinnamates – nature, occurrence, dietary burden, absorption and metabolism. *Journal of the Science of Food and Agriculture,* **80,** 1033-1043.

bound together by 3 carbon atoms that form an oxygenated heterocycle (ring C) may themselves be divided into different subclasses including: flavonols, flavones, isoflavones, flavanones, anthocyanidins, and flavanols (catechins and proanthocyanidins) and their glycosides[86,87] **(Figure 12)**.

Flavonoids are often present in glycoside forms, which have effects to render the flavonoid less reactive and more water soluble and allowing their storage safely in the cell vacuole[88].

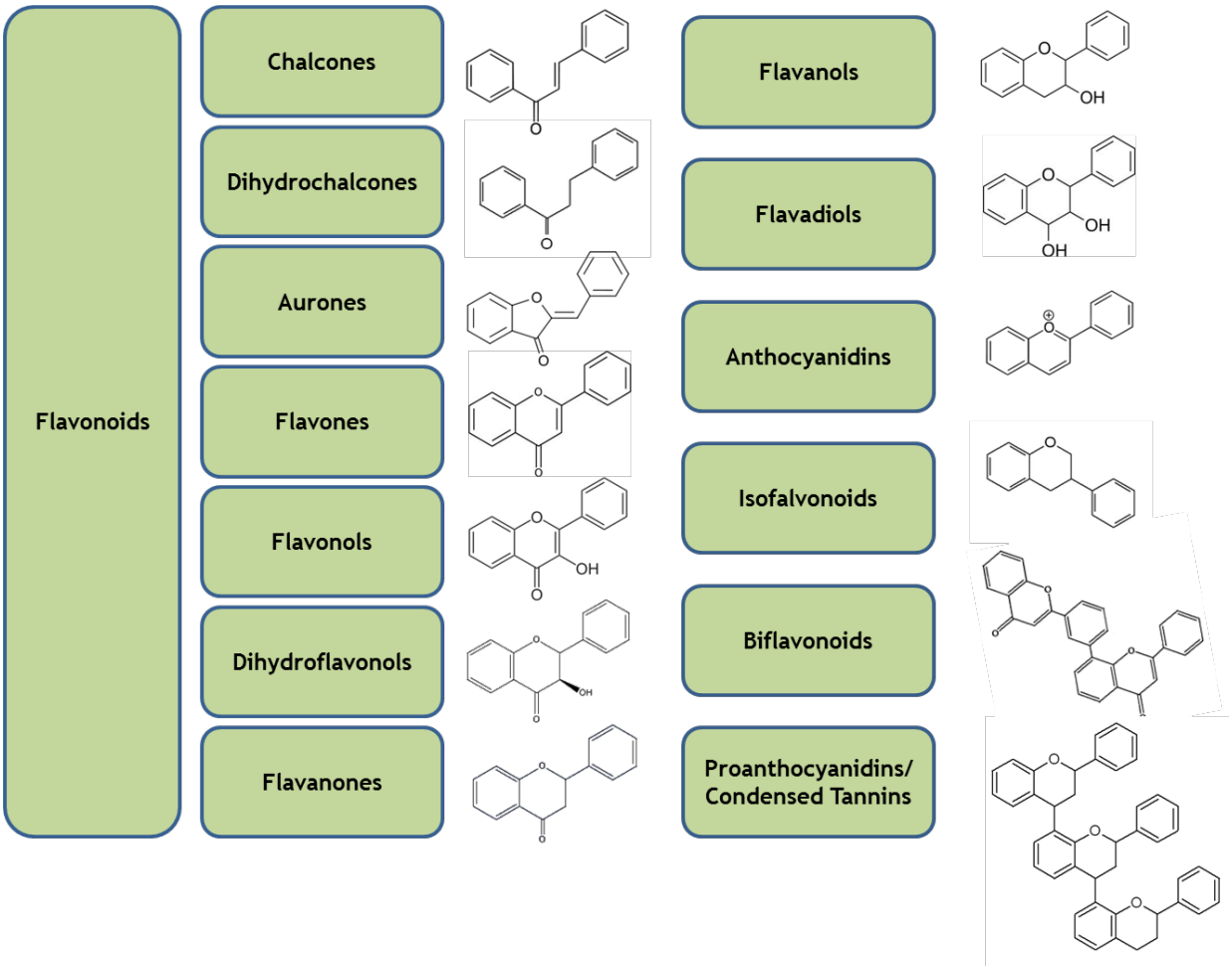


Figure 12. Classification of flavonoids

There are two types of glycosidic bonds i.e. *O*- and *C*-glycosides. The former site is often found in the phenolic hydroxyl group of the aglycone with formation of a glycosidic O—C bond. In theory, any of the hydroxyl groups can be glycosylated but certain positions are favored, such as the 7-hydroxyl group in flavones, flavanones, and isoflavones, the 3- and 7-hydroxyls in flavonols and flavanols and the 3- and 5-hydroxyls

[86] Manach, C., Scalbert, A., Morand, C., Rémésy, C. and Jimenez, L. 2004. Polyphenols: food sources and bioavailability. *The American Journal of Clinical Nutrition,* **79,** 727-747.

[87] Iwashina, T. 2000. The Structure and Distribution of the Flavonoids in Plants. *Journal of Plant Research,* **113,** 287-299.

[88] Harborne, J. B. and Mabry, T. J. 1982. *The Flavonoids: advances in research* Chapman and Hall.

in anthocyanidins. The latter site, which takes place by direct linkage of sugar to the flavonoid's basic nucleus via an acid-resistant C—C bond, has been found only at the *C*-6 or *C*-8 position of the aglycone to date[89,90]. Flavonoid *C*-glycosides are commonly further divided into mono-*C*-glycosylflavonoids, di-*C*-glycosylflavonoids and *C*-lycosylflavonoid- *O*-glycosides. In fruits and vegetables, flavonoids are frequently present as glycosides, since glycosylation makes the molecule less reactive but more soluble. Glucose is the sugar that is most frequently involved in glycoside formation, but we can also find galactose, rhamnose, xylose, and arabinose, as well as disaccharides such as rutinose[91].

Flavonols are arguably the most widespread of the flavonoids, being dispersed throughout the plant kingdom especially in fruits and vegetables. The distribution and structural variations of flavonols are extensive and have been well documented. Flavonols such as myricetin, quercetin, isorhamnetin, and kaempferol are most commonly found as *O*-glycosides. Conjugation occurs most frequently at the 3 position of the C-ring but substitutions can also occur at the 5, 7, 4', 3', and 5' positions of the carbon ring. Although the number of aglycones is limited, there are numerous flavonol conjugates with more than 200 different sugar conjugates of kaempferol alone[92].

On the other hand, flavones have a very close structural relationship to flavonols **(Figure 12)**. Although flavones such as luteolin and apigenin have A- and C-ring substitutions, they lack oxygenation at C3. A wide range of substitutions is also possible with flavones, including hydroxylation, methylation, *O*- and *C*-alkylation, and glycosylation. Most flavones occur as 7-*O*-glycosides. Flavones are reportedly contained in celery, parsley, and some herbs. In addition, polymethoxylated flavones, such as nobiletin and tangeretin, have been found in citrus species. Flavones, flavonols and their glycosides are widely distributed in the plant kingdom. It has been estimated that humans consuming high fruit and vegetable diets ingest up to 1 g of these

[89] Becchi, M. and Fraisse, D. 1989. Fast atom bombardment and fast atom bombardment collision-activated dissociation/mass-analysed ion kinetic energy analysis of C-glycosidic flavonoids. *Biological Mass Spectrometry,* **18,** 122-130.

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

[91] Heim, K. E., Tagliaferro, A. R., and Bobilya, D. G. 2002. Flavonoid antioxidants: Chemistry, metabolism and structure-activity relationships. *The Journal of Nutritional Biochemistry*, **13**, 572-584.

[92] Crozier, A., Clifford, M. N. and Ashihara, H. 2006. *Plant Secondary Metabolites: Occurrence, Structure and Role in the Human Diet* Blackwell.

compounds daily. The most common and biologically active dietary flavonol is quercetin[93].

Stilbenes are phenolic compounds containing two benzenic rings linked by an ethane or ethene bridge with C6–C2–C6 structure **(Figure 9)**. They are widely distributed in higher plants, acting as phytoalexins and growth regulators. Stilbenes are phytoalexins, compounds produced by plants in response to attack by fungal, bacterial, and viral pathogens. Resveratrol (3,4,5-trihydroxystilbene) is the most common stilbene, reputed to help prevent heart disease[94]. It occurs as both the *cis* and the *trans* isomers and is present in plant tissues primarily as *trans*-resveratrol-3-*O*-glucoside. The major dietary sources of stilbenes include grape, soy, and peanuts, and their respective products.

The lignans family of phenolic compounds is derived from the combination of two phenylpropanoids (C6-C3 units) **(Figure 9)**. Lignans are basically dimers of the cinnamic alcohol, which cyclizes in different ways, generating a wide range of molecules. Lignans are contained in woody tissues, cereals, and vegetables such as carrots, broccoli, and berries. Together with isoflavones, lignans belong to the class of phytoestrogens, which are protective factors of cardiovascular and immune systems[95]. Lignans are a diverse group of plant-derived compounds that form the building blocks for plant-cell walls. The richest source of lignans is flaxseed. Even though, lignans can also be found in *Brassica* vegetables, such as kale, and fruits such as apricot, strawberry, peach, and pear[96].

Finally, Tannins are polymers of phenolic acids or flavonoids, present in nature as hydrolyzable and non-hydrolyzable (condensed) tannins. The basic units of hydrolyzable tannins are gallic and ellagic acids, esterified to a core molecule, commonly glucose or a polyphenol such as catechin. Condensed tannins, also called proanthocyanidins, are mainly flavonoid polymers[97]. These compounds are astringent

[93] Ho, C. T., Lee, C. and Huang, M. T. 1992. *Phenolic Compounds in Food and Their Effects on Health: Volume I: Analysis, Occurrence, and Chemistry* American Chemical Society.

[94] Vidavalur, R., Otani, H., Singal, P. K., and Maulik, N. 2006. Significance of wine and resveratrol in cardiovascular disease: French paradox revisited. *Experimental and Clinical Cardiology*, **11**, 217–225.

[95] Adlercreutz, H. 1995. Phytoestrogenes: Epidemiology and a possible role in cancer protection. *Environmental Health Perspectives*, **103**, 103–112.

[96] De La Rosa, L. A., Alvarez-Parrilla, E. and Gonzalez-Aguilar, G. A. 2009. *Fruit and Vegetable Phytochemicals: Chemistry, Nutritional Value and Stability* Wiley.

[97] Watson, R. R. and Preedy, V. R. 2009. *Bioactive Foods in Promoting Health: Fruits and Vegetables* Elsevier Science.

and bitter plant phenolic compounds. They are widely distributed in many species of plants, where they play a role in protection from predation (perhaps even as pesticides) and in plant-growth regulation. The most abundant polyphenols are condensed tannins, found in virtually all plant families.

5. Phenolic compound bioactivity

Interest in phenolic compounds has grown over the last few decades, as these agents have become popular among the public, who attribute them with several medicinal properties, mostly related to their antioxidant activity, which can have important implications for health[98].

A number of phenolic compounds, particularly flavonoids, are efficient anti-proliferative agents, being able to inhibit tumor-cell proliferation by interfering with cell-cycle proteins or inducing apoptosis. Notably, flavonoids comprise the most widely studied class of polyphenols with respect to their antioxidant and biological properties. They display powerful antioxidant activities *in vitro*, being able to scavenge a wide range of reactive oxygen species (ROS), reactive nitrogen species (RNS), and chlorine species, including superoxide, hydroxyl, and peroxyl radicals, as well as hypochlorous and peroxynitrous acids. Some flavonoids can also chelate metal ions, such as copper and iron[99].

The antioxidant properties of phenolic compounds are owed mainly to their redox properties, which allow them to act as reducing agents, hydrogen donators, and singlet oxygen quenchers[100]. A growing body of chemical, biochemical, clinical, and epidemiological evidence supports contentions of the chemoprotective effects of phenolic antioxidants against oxidative stress-mediated disorders. The pharmacological actions of phenolic antioxidants stem mainly from their free-radical-scavenging and

[98] Duthie, G. G., Duthie, S. J. and Kyle J. A. M. 2000. Plant polyphenols in cancer and heart disease: implications as nutritional antioxidants. *Nutrtional Research Reviews,* **13**, 79-106.

[99] Yang, C. S., Landau, J. M., Huang, M. T., and Newmark, H. L. 2001. Inhibition of carcinogenesis by dietary polyphenolic compounds. *Annual Review of Nutrition*, **21**, 381–406.

[100] Kaur, C. and Kapoor, H. C. 2001. Antioxidants in fruits and vegetables – the millennium's health. *International Journal of Food Science & Technology,* **36,** 703-725.

metal-chelating properties as well as their effects on cell-signaling pathways and on gene expression[101].

Moreover, phenolic compounds have been shown to have the capacity to delay or prevent diseases associated with oxidative stress (OS), e.g. Alzheimer's, Parkinson's, atherosclerosis, cancer, diabetes, and aging. It is extremely difficult to estimate the food content of phenolic compounds and thus the daily average intake, because of the diversity of their chemical structure.

Mechanisms by which phenolic compounds act as anti-tumoral agents are manifold and have been shown to be connected with the functions of radical scavengers, detoxification agents, cell-signaling modulators, inhibitors of cell-cycle phases, and activators of apoptosis. Some flavonoids are able to achieve this effect by inhibiting the enzyme DNA topoisomerase II, which is necessary for the survival and spread of cancer cell[102]. Additionally, phenolic compounds show anti-inflammatory activity, which is mediated by inhibiting the formation of transcription factors closely linked to inflammation, such as NF-κB and enzymes such as xanthine oxidase, cytochrome oxidase, and lipoxygenase, which mediate the inflammatory process[103].

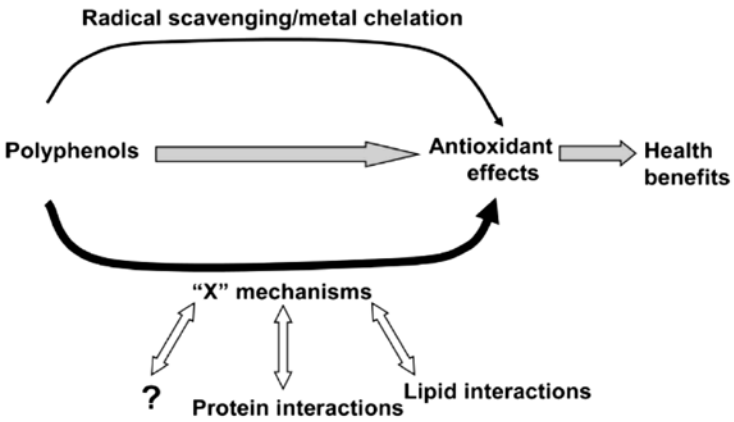


Figure 13. Scheme the relationship of polyphenols with health benefits

[101] Soobrattee, M. A., Neergheen, V. S., Luximon-Ramma, A., Aruoma, O. I. and Bahorun, T. 2005. Phenolics as potential antioxidant therapeutic agents: Mechanism and actions. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis,* **579,** 200-213.

[102] Fresco, P., Borges, F., Diniz, C., and Marques, M. P. M. (2006). New insights on the anticancer properties of dietary poliphenols. *Medicinal Research Reviews*, **26**, 747–766.

[103] Watson, R. R. and Preedy, V. R. 2009. *Bioactive Foods in Promoting Health: Fruits and Vegetables* Elsevier Science.

Beyond the capacity of phenolic compounds to scavenge free radicals and/or to chelate metals, other biochemical mechanisms can explain their *in vitro* and *in vivo* antioxidant effects **(Figure 13)**.

These mechanisms are related to polyphenol-lipid and polyphenol-protein interactions, and are based on the presence of hydroxyl groups that give phenolic compounds high possibilities for physical and chemical interactions. These mechanisms do not preclude a direct free-radical or metal-chelating action, but are more consistent with levels of *in vivo* phenolic compounds detected in human and animals. Furthermore, the actual occurrence of these potential actions *in vivo* should be considered when linking an antioxidant action to a phenolic compound, or to any other molecule, for which the metabolic fate and biological actions are not firmly established[104].

Several studies have shown that phenolic compounds are extensively metabolized *in vivo*, mainly during transfer across the small intestine, by colonic microflora and in the liver, resulting in significant alteration in their redox potentials[105].

6. Polyphenols in highly consumed fruits and vegetables in relation to health

Evidence suggests that dietary antioxidants can reduce cancer risk. The relationship between fruit and vegetable intake and cancers of the lung, colon, breast, cervix, esophagus, oral cavity, stomach, bladder, pancreas, and ovary has been established[106]. Fruit and vegetable consumption has been found to have a significant protective effect. The risk of cancer for most cancer sites is twice as high in persons whose intake of fruits and vegetables was low compared with those with high intake. Significant protection was also found in most studies on lung cancer. Fruits proved to be significantly protective in cancers of the esophagus, oral cavity, and larynx. Overall, protective effects of fruit and vegetable intake have been found in different cancer[107]. Oxidative stress induced by free radicals cause DNA damage, which, when left unrepaired, can lead to base mutation, single- and double-strand breaks, DNA cross-linking, and chromosomal breakage and rearrangement. This potentially cancer-

[104] Fraga, C. G. 2007. Plant polyphenols: How to translate their in vitro antioxidant actions to in vivo conditions. *IUBMB Life,* **59,** 308-315.

[105] Soobrattee, M. A., Neergheen, V. S., Luximon-Ramma, A., Aruoma, O. I. and Bahorun, T. 2005. Phenolics as potential antioxidant therapeutic agents: Mechanism and actions. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis,* **579,** 200-213.

[106] Block, G., Patterson, B. and Subar, A. 1992. Fruit, vegetables, and cancer prevention: A review of the epidemiological evidence. *Nutrition and Cancer,* **18,** 1-29.

[107] Liu, R. H. 2003. Health benefits of fruit and vegetables are from additive and synergistic combinations of phytochemicals. *The American Journal of Clinical Nutrition,* **78,** 517S-520S.

inducing oxidative damage might be prevented or limited by dietary antioxidants found in fruit and vegetables.

The most thoroughly investigated dietary components in fruits and vegetables acting as antioxidants are fiber, polyphenols, flavonoids, conjugated isomers of linoleic acid, limolene, epigallocatechin, gallate, some peptides, isoflavanones, vitamins A, B, C, E, tocopherols, calcium, selenium, chlorophyllin, alipharin, sulfides, catechin, tetrahydrocurecumin, seasaminol, glutathione, uric acid, indoles, thiocyanates, and protease inhibitors[108]. These compounds may act independently or in combination as anticancer or cardio-protective agents by a variety of mechanisms **(Table 1)**.

Table 1. Mechanism of actions of phenolic compounds in various pathophysiological conditions (adapted from [109])

Phenolic compounds	Pathology	Mechanism of actions
Quercetin, Kaempferol, Genistein, Resveratrol	Colon cancer	Suppression of COX-2 expression by inhibiting tyrosine kinases important for induction of COX-2 gene expression
Catechins	Neurodegenerative diseases	Increase in the activity of SOD and catalase
(−)-EGCG	Neurodegenerative conditions	Decrease in the expression of proapototic genes (bax, bad, caspase-1 and -6, cyclin dependent kinase inhibitor) thus maintaining the integrity of the mitochondrial membrane
(−)-EGCG	Cancer, diabetic retinopathy,chronic inflammation	Suppression of angiogenesis by inhibiting growth factor, activating receptors and PKC. Down regulation of VEGF production in tumor cells. Repression of AP-1, NF-KB and STAT-1 transcription factor pathways. Inhibition of capillary endothelial-cell proliferation and blood-vessel formation.
Proanthocyanidin (GSPE)	Cardiovascular disorders	Inhibitory effects on proapoptotic and cardioregulatory genes. Modulation of apoptotic regulatory bcl-XL , p53 and c-myc genes
Ferulic acid	Diabetes	Decrease in lipid peroxidation and rise in the level of glutathione and antioxidant enzymes

[108] Kaur, C. and Kapoor, H. C. 2001. Antioxidants in fruits and vegetables – the millennium's health. *International Journal of Food Science & Technology,* **36,** 703-725.

[109] Soobrattee, M. A., Neergheen, V. S., Luximon-Ramma, A., Aruoma, O. I. and Bahorun, T. 2005. Phenolics as potential antioxidant therapeutic agents: Mechanism and actions. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis,* **579,** 200-213.

Numerous studies suggest a strong link between dietary intake of phytochemicals and a reduced risk of cardiovascular disease. Dietary flavonoid intake was inversely associated with mortality from coronary-artery disease and inversely related (more weakly but still significantly) with the incidence of myocardial infarction. In a study in Finland, the intake of apples and onions, both high in quercetin, was inversely correlated with total mortality and coronary mortality [110]. In a recent Japanese study, the total intake of flavonoids (quercetin, myricetin, kaempferol, luteolin, and ficetin) was inversely correlated with the plasma total cholesterol and low-density lipoprotein (LDL) cholesterol concentrations[111].

The intake of quercetin alone was inversely related to total cholesterol and LDL plasma levels. Joshipura et al. [112] reported that total fruit intake and total vegetable intake were both individually associated with a decreased risk of coronary-artery disease; the inverse association between total consumption of fruit and vegetables and coronary-artery disease was observed when the dietary intake was > 4 servings/d. Mechanisms for the prevention of arteriosclerosis by antioxidants have been proposed. Dietary antioxidants that are incorporated in LDL are themselves oxidized when these LDL are exposed to pro-oxidative conditions before any extensive oxidation can occur in the sterol or polyunsaturated fatty acids[113].

Many kinds of beans containing natural phenolic compounds have antimutagenic effects and can function as health-promoting foods[114]. And so the green beans which represent the immature pods of any kind of beans are considered a rich source of phytochemicals such as flavonoids (flavonols). Several reports link bean consumption to reduced risk of cardiovascular disease, diabetes mellitus, obesity, cancer, and diseases of digestive tract, attributing these potential health benefits of beans to the presence of secondary metabolites such as phenolic compounds that possess

[110] Paul, K., Ritva, J., Antti, R. And Jouni, M. 1996. Flavonoid intake and coronary mortality in Finland: a cohort study. *BMJ,* **312,** 478-481.

[111] Arai, Y., Watanabe, S., Kimira, M., Shimoi, K., Mochizuki, R. and Kinae, N. 2000. Dietary Intakes of Flavonols, Flavones and Isoflavones by Japanese Women and the Inverse Correlation between Quercetin Intake and Plasma LDL Cholesterol Concentration. *The Journal of nutrition,* **130,** 2243-2250.

[112] Joshipura, K. J., Hu, F. B., Manson, J. E., Stampfer, M. J., Rimm, E. B., Speizer, F. E., Colditz, G., Ascherio, A., Rosner, B., Spiegelman, D. and Willett, W. C. 2001. The Effect of Fruit and Vegetable Intake on Risk for Coronary Heart Disease. *Annals of Internal Medicine,* **134,** 1106-1114.

[113] Sanchez-Moreno, C. N., Jimenez-Escrig, A. and Saura-Calixto, F. 2000. Study of low density lipoprotein oxidizability indexes to measure the antioxidant activity of dietary polyphenols. *Nutrition Research,* **20,** 941-953.

[114] De Mejía ¤A, E. G. L., Castaño -Tostado, E. and Loarca- Piña, G. 1999. Antimutagenic effects of natural phenolic compounds in beans. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis,* **441,** 1-9.

antioxidant properties[115]. The biological activity of flavonoid compounds from beans has been reported *in vitro* as well as *in vivo*, e.g. the antioxidant activity has been evaluated using different methods and different common bean (*Phaseolus vulgaris* L) cultivars[116].

Green beans could be considered a functional food with health properties which have been discussed extensively in the literature. For example, the antiglycemic, antioxidant, and hypolipidemic effects of green bean pods in diabetic rats have previously been reported[117]. These positive effects have been linked to the antioxidant activity supported by the flavonoids present in green bean pod extracts[118]. The specific consumption of green beans in relation to reducing cancer risk has also been studied elsewhere[119]. Furthermore, isoflavones, a major subclass of flavonoids which are the most abundant in Fabaceae, have been proposed to play a role in learning and memory during aging, bone metabolism, thrombogenicity, and the prevention of climacteric syndrome as well as several types of cancer[120].

Watermelons have many medicinal uses and it is considered mildly diuretic, being effective in the treatment of dropsy and renal stones, hypertension reducing, prevent erectile dysfunction, act as antioxidant and used to treat enlarged liver and jaundice, also it can be used as an alternative natural and chemical treatment for giardiasis[121-124]. Recently, it was reported that phenolics may be the principal hydrophilic

[115] Luthria, D. L. and Pastor-Corrales, M. A. 2006. Phenolic acids content of fifteen dry edible bean (Phaseolus vulgaris L.) varieties. *Journal of food composition and analysis: an official publication of the United Nations University, International Network of Food Data Systems,* **19,** 205-211.

[116] Guevara-González, R. G. and Torres-Pacheco, I. 2006. *Advances in Agricultural and Food Biotechnology: 2006* Research Signpost.

[117] Pari, L. and Venkateswaran, S. 2003. Effect of an aqueous extract of Phaseolus vulgaris on the properties of tail tendon collagen of rats with streptozotocin-induced diabetes. *Brazilian Journal of Medical and Biological Research,* **36,** 861-870.

[118] Roman-Ramos, R., Flores-Saenz, J. L. and Alarcon-Aguilar, F. J. 1995. Anti-hyperglycemic effect of some edible plants. *Journal of Ethnopharmacology,* **48,** 25-32.

[119] Takezaki, T., Gao, C. M., Wu, J. Z., Ding, J. H., Liu, Y. T., Zhang, Y., Li, S. P., Su, P., Liu, T. K. and Tajima K. 2001. Dietary protective and risk factors for esophageal and stomach cancers in a lowepidemic area for stomach cancer in Jiangsu Province, China: comparison with those in a high-epidemic area. *Japanese Journal of Cancer Research* **92,**1157-65

[120] Marin, F. R., Perez-Alvarez, J. A., Soler-Rivas, C. and Atta Ur, R. 2005. Isoflavones as functional food components. In *Studies in Natural Products Chemistry* pp. 1177-1207, Elsevier.

[121] Erukainure, O. L., Oke, O. V., Daramola, A. O., Adenekan, S. O., and Umanhonlen, E. E. 2010. Improvement of the Biochemical Properties of Watermelon Rinds Subjected to Saccharomyces cerevisae Solid Media Fermentation. *Pakistan Journal of Nutrition,* **9,** 806-809.

[122] Elsir, L., Hassan, A., Koko, W. S., Osman, E.-B. E., Dahab, M. M., Sirat, M. H. 2011. In vitro antigiardial activity of Citrullus lanatus Var. citroides extracts and cucurbitacins isolated Compounds. *Journal of Medicinal Plants Research,* **5,** 3338-3346.

[123] Sharma S., paliwal S., Dwivedi J., and Tilak. A. 2011. First report on laxative activity of Citrullus lanatus, *Pharmacology online,* **2**, 790-797.

compounds contributing to the hydrophilic antioxidant activity in watermelon[125]. On the other hand, cucumbers are reportedly used in traditional medicine as anti-diarrheal, anti-gonorrheal, anti-inflammatory, hypertension reducing agent, diabetes mellitus, and analgesic for their wide array of metabolites containing[126-128].

Asteraceae (Compositae) vegetables also have a remarkable content of antioxidants like sesquiterpenes, alkaloids, flavonoids, tetraterpenes, triterpenes, diterpenes, amines, cyanogenic glycosides, non-protein amino acids, monoterpenes, phenylpropanes, ascorbic acid and carotenoids, and phenolic acids such as mono- and di-caffeoylquinic acid derivatives[129]. Lettuce and artichoke, belonging to the same family (Asteraceae), are also of interest for chemoprevention. In regard to artichoke, there are many studies demonstrating that this vegetable has major medicinal properties, including antioxidative, anticarcinogenic, antigenotoxic, cholesterol-lowering, hepatoprotective, bile-expelling, diuretic, and anti-inflammatory, as well as antifungal, anti-HIV, and antibacterial[130-135].

---

[124] Figueroa A., Sanchez-Gonzalez M. A., Wong A., and Arjmandi, B. H. 2012. Watermelon Extract Supplementation Reduces Ankle Blood Pressure and Carotid Augmentation Index in Obese Adults With Prehypertension or Hypertension. *American Journal of Hypertension*, **25**,640-643.

[125] Tlili, I., Hdider, C., Lenucci, M. S., Riadh, I., Jebari, H. and Dalessandro, G. Bioactive compounds and antioxidant activities of different watermelon (Citrullus lanatus (Thunb.) Mansfeld) cultivars as affected by fruit sampling area. *Journal of Food Composition and Analysis,* **24**, 307-314.

[126] Gill N. S., Sood S., Muthuraman A., Garg M., Kumar R., Bali M., and Sharma P. D. 2010. Antioxidant, Anti-inflammatory and Analgesic Potential of Cucumis sativus Seed Extract. *Latin America Journal of Pharmacy,* **29,** 927-932.

[127] Kai, H., Baba, M., and Okuyama, T. 2008. Inhibitory Effect of Cucumis sativus on Melanin Production in Melanoma B16 Cells by Down regulation of Tyrosinase Expression. *Planta Medicina*, **74**, 1785-1788.

[128] Mukherjee S., Das J. K., and Roy C. K. 2009. Screening of Antioxidant Activity and Phenolic Content in Cucumis Sativus L. Seeds. *Pharmacology online*, **2**, 711-714.

[129] Watson, R. R. and Preedy, V. R. 2009. *Bioactive Foods in Promoting Health: Fruits and Vegetables* Elsevier Science.

[130] Agarwal, R. and Mukhtar, H. 1996. Cancer chemoprevention by polyphenols in green tea and artichoke. *Advances in experimental medicine and biology,* **401,** 35-50.

[131] Gebhardt, R. 1997. Antioxidative and Protective Properties of Extracts from Leaves of the Artichoke (Cynara scolymus L.) against Hydroperoxide-Induced Oxidative Stress in Cultured Rat Hepatocytes. *Toxicology and Applied Pharmacology,* **144,** 279-286.

[132] Brown, J. E. and Rice-Evans, C. A. 1998. Luteolin-rich artichoke extract protects low density lipoprotein from oxidation In vitro. *Free Radical Research,* **29,** 247-255.

[133] Crosby, D. C. and Robinson, W. E. Dicaffeoyltartaric Acid and Dicaffeoylquinic Acid HIV Integrase Inhibitors. In *HIV-1 Integrase* pp. 341-362, John Wiley & Sons, Inc..

[134] Shimoda, H., Ninomiya, K., Nishida, N., Yoshino, T., Morikawa, T., Matsuda, H. And Yoshikawa, M. 2003. Anti-Hyperlipidemic sesquiterpenes and new sesquiterpene glycosides from the leaves of artichoke (Cynara scolymus L.): structure requirement and mode of action. *Bioorganic & Medicinal Chemistry Letters,* **13,** 223-228.

[135] Zhua, X. F., Zhanga, H. X., and Lob, R. 2005. Antifungal activity of Cynara scolymus L. extracts. *Fitoterapia*, **76**, 108–111.

Nutritionally, the high levels of bioactive phenolic compounds (caffeoylquinic acids and flavonoids) in the inner portion represent an added value for artichoke flower heads. Moreover, many chemical studies of the artichoke have revealed it to be a rich source of the polyphenol compounds, with mono-, di-caffeoylquinic acids and flavonoids, as the major chemical components[136]. It is thought that the antioxidant properties of artichoke are related to its abundant phenolic composition[137]. In comparison to other vegetables, artichoke flower heads contain high levels of total polyphenols[138,139].

On the other hand, several recent studies have reported the health effects of lettuce in preventing cardiovascular diseases in laboratory animals and humans[140].
Nevertheless, lettuce's latex possesses soporific effects and may be utilized as a sedative[141]. Serafini et al. [142] have reported in their work that the ingestion of fresh lettuce is able to raise plasma antioxidant defense.

7. Analytical techniques used for separating phenolic compounds

Providing the necessary data for the estimation of the intake of nutrients and other food components is still one of the major applications of food composition data tables. Although energy and total amounts of macronutrients constitute the basic components, data on some compounds are less readily available. This is true for some fatty acids and trace elements, folate, as well as for particularly bioactive compounds e.g. phenolic compounds.

The phenolic fraction of food sample is very complex and, despite having been studied for decades and excellent progress having been made, it must be admitted that a

[136] Lattanzio, V., Kroon, P. A., Linsalata, V. and Cardinali, A. 2009. Globe artichoke: A functional food and source of nutraceutical ingredients. *Journal of Functional Foods,* **1,** 131-144.
[137] Mulinacci, N., Prucher, D., Peruzzi, M., Romani, A., Pinelli, P., Giaccherini, C. and Vincieri, F. F. 2004. Commercial and laboratory extracts from artichoke leaves: estimation of caffeoyl esters and flavonoidic compounds content. *Journal of Pharmaceutical and Biomedical Analysis,* **34,** 349-357.
[138] Brat, P., Georgé, S., Bellamy, A., Du Chaffaut, L., Scalbert, A., Mennen, L., Arnault, N. and Amiot, M. J. 2006. Daily polyphenol intake in France from fruit and vegetables. *The Journal of nutrition,* **136,** 2368-2373.
[139] Lombardo, S., Pandino, G., Mauromicale, G., Knödler, M., Carle, R. and Schieber, A. Influence of genotype, harvest time and plant part on polyphenolic composition of globe artichoke [Cynara cardunculus L. var. scolymus (L.) Fiori]. *Food Chemistry,* **119,** 1175-1181.
[140] Llorach, R., Martínez-Sánchez, A. N., Tomas -Barberan, F. A., Gil, M. A. I. and Ferreres, F. 2008. Characterisation of polyphenols and antioxidant properties of five lettuce varieties and escarole. *Food Chemistry,* **108,** 1028-1038.
[141] http://www.naturallifemagazine.com/9704/lettuce.htm, retrieved on 10 Jan. 2013.
[142] Serafini, M., Bugianesi, R., Salucci, M., Azzini, E., Raguzzini, A. and Maiani, G. 2002. Effect of acute ingestion of fresh and stored lettuce (*Lactuca sativa*) on plasma total antioxidant capacity and antioxidant levels in human subjects. *British Journal of Nutrition,* **88,** 615-623.

considerable number of compounds present in it have still not been completely characterized and many problems remain to be resolved. The reason behind these difficulties is the complexity of the chemical nature of these compounds and the similar complexity of the matrix in which they are found. One of the current difficulties hindering rapid and reproducible analyses of phenolic compounds is the scarcity of suitable pure standards, in particular of secoiridoids and lignan compounds. Phenolic acids of natural origin are weak acids and, owing to their phenolic hydroxyl groups, flavonoids and tannins also have a slightly acidic nature.

The analysis of phenolic compounds is challenging due to the great variety and reactivity of these compounds. On the other hand, modern separation and detection methods, such as hyphenated techniques of liquid chromatography (LC) with mass spectrometry (MS), ultraviolet-visible light (UV/Vis), or nuclear magnetic resonance (NMR) spectroscopy, have made it possible to analyze phenolic compounds.

7.1. Liquid chromatography (LC)

Many chromatographic analytical techniques based on differences in physical-chemical properties of the various components of a sample are used for separation. Continuous techniques are generally used, i.e. in which the analytes are detected continuously (on-line) after separation. One of the most widely used of these techniques is liquid chromatography (LC).

LC is a separation technique by which the components of the sample are distributed between a mobile phase (liquid) and a stationary phase (typically consisting of small porous particles with a large surface area). Sample molecules travel through the chromatography system and interact with the surface of the stationary phase. Sample molecules that interact strongly with the stationary phase will take a longer time to travel through the chromatography system, i.e. will have longer retention times.

In LC separation, the mobile phase is aspirated by the pump through the chromatographic column. Upon entering the sample in the column, its components interact with variable length in the stationary phase and it is distributed differently between the two phases. On being eluted from the end of the column, each analyte enters the detector and produces a signal that is measurable in some way. The intensity and duration of the signal is related to the amount or nature of the analyte.

Generally, the signal is amplified and recorded by an electronic integrator, a computer or other means of producing the chromatogram in order to identify and quantify the analyte.

LC equipment consists of a series of building blocks that are common to all models: pump, injector, chromatographic column, thermostatically controlled oven, detector, and data-acquisition system. **Figure 14** presents the basic scheme of commercial LC equipment, showing its commercial appearance.

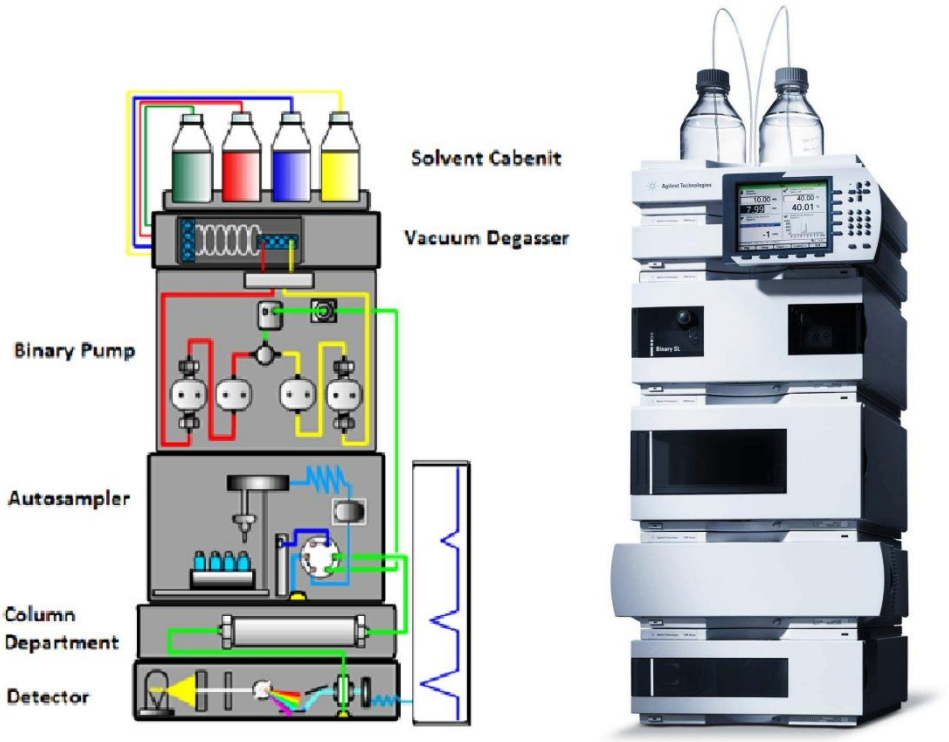


Figure 14. A simplified model of HPLC

Traditionally, LC columns (the stationary phase) are highly polar, e.g. silica, and the mobile phases relatively nonpolar, e.g. hexane. This mode of LC is called “normal-phase” liquid chromatography. The retention time of the analyte lengthens with increasing polarity. If the polarity of the mobile phase is increased, the retention time can be decreased. This mode of chromatography can be useful for separation on nonpolar or weakly polar flavonoid aglycones.

More common today is “reversed-phase” liquid chromatography for the analytical and preparative separations of analytes. This term derives from the fact that the mobile phase is more polar than the stationary phase and that it reverses the order of elution of the analytes.

The retention time, which is longer for more hydrophobic molecules, can be decreased by the addition of nonpolar solvent to the mobile phase. The hydrophobicity of the sample molecules is pH dependent, and therefore an organic acid such as formic acid, acetic acid, or trifluoroacetic acid, is often added to the mobile phase. The effect varies depending on application but generally improves the chromatography[143-145].

The quality of any chromatography separation depends on the characteristics of the compounds to be separated and their interactions with the column and the solvent. A profiling method for phenolic compounds requires uniform separation across the range of polarities; from the most polar (hydroxybenzoic and hydroxycinnamic acids) to the least polar (aglycones and poly-methoxylated flavonoids) compounds[146].

Phenolic compounds, because of their polar character, are separated mainly by partition chromatography on reversed phase, using columns filled with hydrocarbons such as C_8 (n-octyl) or more commonly, C_{18} (n-octadecyl). The elution was performed with a mobile phase of high polarity such as aqueous solutions which may contain a certain percentage of polar organic solvents such as acetonitrile or methanol.

Figure 15. "UPLC", a commercial model.

Only reversed-phase columns were used, since phenolic compounds are weak acids that can be separated as neutral, relatively hydrophobic compounds in a weak acid matrix. Normal-phase columns have been used for pro-anthocyanidins in order to

[143] Lough, W. J. and Wainer, I. W. 1995. *High Performance Liquid Chromatography: Fundamental Principles and Practice* Blackie Academic & Professional.
[144] Mcmahon, G. 2008. *Analytical Instrumentation: A Guide to Laboratory, Portable and Miniaturized Instruments* Wiley.
[145] Meyer, V. R. 2010. *Practical High-Performance Liquid Chromatography* Wiley.
[146] Harnly, J., Bhagwat, S. and Lin, L.-Z. 2007. Profiling methods for the determination of phenolic compounds in foods and dietary supplements. *Analytical and Bioanalytical Chemistry,* **389,** 47-61.

provide separation of the monomers through the decamers and a broad peak for the higher polymers[147].

In LC, formerly *high-performance liquid chromatography* (HPLC), a combination of packing materials with smaller and more uniform particles sizes, greater selectivity, and increased flow pressure yields both higher resolving power and faster analyses. Martin and Synge (1941), the inventors of modern chromatography, were aware that in theory the stationary phase requires very small particles, and therefore a high pressure was used to generate the flow required for LC in packed columns. Hence, HPLC was sometimes referred to as *high-pressure liquid chromatography*[148].

Recently, there have been improvements in the LC technique with the development of ultra-high performance liquid chromatography (UPLC) **(Figure 15)**. This technique takes full advantage of chromatographic principles to perform separations using columns packed with smaller-sized particles, i.e. less than 2 μm, and a chromatographic system operating at very high pressures. This leads to a shorter analysis time, higher peak efficiency, and higher resolution[149].

7.2. Detectors coupled to LC

Different types of detectors can be coupled to liquid chromatography. *Hyphenated* LC detectors refer to the coupling of an independent analytical instrument (e.g. MS, NMR, FTIR) to the LC system to provide detection. The mass spectrometry (MS) detector is the most popular hyphenated LC detector in use today **(Figure 16)**.

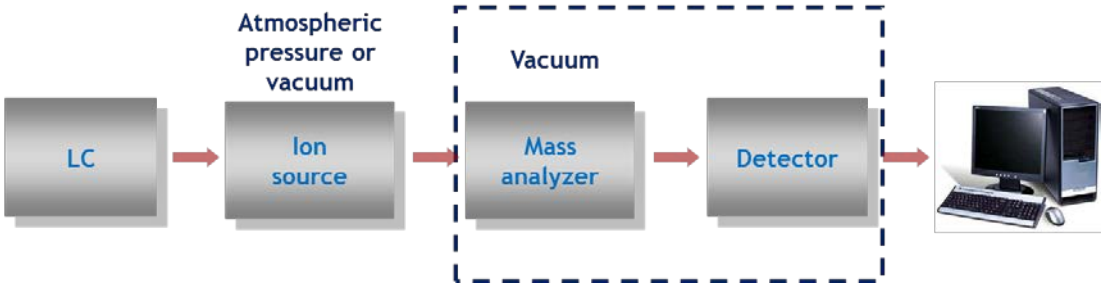


Figure 16. Scheme of hyphenated LC with MS

[147] Prior, R. L. and Gu, L. 2005. Occurrence and biological significance of proanthocyanidins in the American diet. *Phytochemistry,* **66,** 2264-2280.
[148] Martin A. J. P. and Synge R. L. M.1941. A new form of chromatogram employing two liquid phases. *Biochemical Journal,* **35**,1358-1368.
[149] Wilson, I. D., Nicholson, J. K., Castro-Perez, J., Granger, J. H., Johnson, K. A., Smith, B. W. and Plumb, R. S. 2005. High Resolution "Ultra Performance" Liquid Chromatography Coupled to Oa-TOF Mass Spectrometry as a Tool for Differential Metabolic Pathway Profiling in Functional Genomic Studies. *Journal of Proteome Research,* **4,** 591-598.

The choice of the ideal detector for each application is based on the nature and properties of the analytes to be determined and the sensitivity required and what information is sought (structural, quantitative, etc.).

For the detection of phenolic compounds, two detection systems are widely used, i.e. spectroscopic UV-visible and mass spectrometry. Additionally, for the identification of compounds for which the standard is not available, nuclear magnetic resonance is very useful.

7.2.1. UV-visible absorbance

Polyphenols absorb within the ultraviolet (UV) region. Two absorption bands are characteristic of flavonoids. Band II, with a maximum in the 240-285 nm range, is believed to arise from the A-ring. Band I, with a maximum in the 300-550 nm range, presumably arises from the B-ring[150,151] **(Figure 17)**.

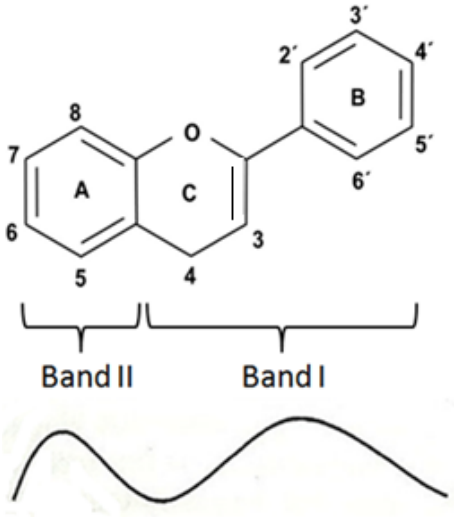


Figure 17. Basic structure of flavonoids and spectrum.

Among the flavonoids, anthocyanins are the most unique subclasses because they absorb visible light near 520 nm when the molecules are in the flavylium cation status (when pH is low).

[150] Merken, H. M. and Beecher, G. R. 2000. Measurement of Food Flavonoids by High-Performance Liquid Chromatography A Review. *Journal of Agricultural and Food Chemistry,* **48,** 577-599.
[151] Robards, K. and Antolovich, M. 1997. Analytical Chemistry of Fruit Bioflavonoids, A Review. *Analyst,* **122,** 11R-34R.

Table 2. Absorption bands of each family of polyphenols.

Polyphenol classes	UV max (nm)	
Phenolic acids	270-280	
Hydroxycinamic acids	305-325	
Flavonoids		
Anthocyanidins	240-280	450-560
Aurones	240-270	340-370
Chalcones	220-270	340-390
Dihydrochalcones	~ 220	~ 280
Flavones	250-270	330-350
Flavonols	250-270	350-380
Flavanones	270-295	
Flavanols	270-280	
Proanthocyanidins	~ 280	
Isoflavonoids	245-270	300-340
Other polyphenols		
Coumarins	220-230	310-350

The detector itself consists of a small liquid flow cell through which the eluent from the column flows. UV light passes through the cell and hits the UV photo-detector. UV absorbance detectors are not destructive and respond only to substances that absorb radiation at the wavelength of the source light. There are two kinds of wavelength detectors: fixed and multiple wavelength detectors. The former does not allow changing the wavelength of the radiation; meanwhile, the latter utilizes a narrow range of wavelengths to detect the solute. Its sensitivity is not as good as the fixed wavelength but it has advantage of flexibility. There are two main types of multiple wavelength detectors: the dispersion detector (variable wavelength detectors) that monitors the eluent at one wavelength only, and the diode array detector (DAD) that monitors the eluted solute over a range of wavelengths simultaneously. The second

type has two major advantages, firstly it allows for the best wavelength(s) to be selected for analysis and secondly it allows all the spectra under a peak to be determined and its purity to be established[144,152].

There is also combination of UV-vis spectrometric detector which is the most commonly used detector for HPLC. In terms of the electromagnetic spectrum, UV radiation covers the region from 190-350 nm, visible radiation covers the region 350-800 nm. Detection system based on the absorption of UV-visible is the most common in commercial HPLC equipment. The detection of analytes is based on the interaction between the UV-visible and the material that gives rise to absorption of certain wavelengths of radiation from such compounds. The multiple links combined in the compounds into the phenolic chromosphere showing characteristic absorption bands in the UV region and visible (see **Table 2**).

Many compounds in solution can absorb radiation in this region of the electromagnetic spectrum and, therefore, this detector can be considered universal. This behavior has the advantage that it can be used to solve a large number of analytical problems but may also be a drawback in cases requiring selectivity. The detection of UV-visible is very robust, making it one of the best ways to perform quantifications.

While UV-visible spectroscopy gives no structural information and cannot unequivocally identify compounds when there is no pattern, it can be useful in determining phenolic compounds in order to limit the family to which the analytes belong, as every family has certain characteristic absorption bands. Therefore, other techniques such as mass spectrometry are often necessary.

7.2.2. Mass spectrometry (MS)

Mass spectrometry is a detection system commonly used for the structural elucidation or confirmation of organic, bioorganic and organometallic compounds, and for quantitative analysis in environmental, pharmaceutical, forensic, food, and other sciences.

MS detectors are also widely used to provide structural information or confirmation of unknowns. MS detectors come in two popular configurations:

[152] Phillips, T. M. and Corradini, D. 2010. *Handbook of HPLC* CRC Press.

1- *Single-stage detector*, sometimes called an MSD (*mass selective detector*), is used to measure a single ionic species for each analyte, often the deprotonated (M - H) or protonated molecular ions (M + H). (Within a given run, more than one analyte ion can be monitored by switching back and forth between different *m/z* values or scanning between ions.) Instruments using this type of detection are referred to as LC-MS technique.

2- A more complex detector design isolates the primary ionic species (parent or precursor ion), fragments it into additional ions (daughter or product ions), and monitors one or more of these product ions. This process, gives added selectivity when the transition from precursor to product ion is used as a "signature" of a specific analyte. Such systems are referred to as the LC-MS/MS technique.

The first step in MS measurement is the conversion of neutral molecules to charged species (i.e. ions), which are then separated according to their mass-to-charge (*m/z*) ratio in a mass analyzer. It is a fundamental requirement of mass spectrometry that the ions be in the gas phase before they can be separated according to their individual *m/z* values and detected. Any particles that are not ionic (molecules or radicals) are removed from the mass spectrometer by the continuous pumping that maintains the vacuum. The relative abundances of individual *m/z* values are recorded by a suitable detector to produce what is known as a mass spectrum (MS). MS can be coupled to both gas-phase and liquid-phase separation techniques, enabling the structural analysis of complex mixtures after their chromatographic separation without time-consuming off-line isolation[153-155]. The unit of measure has become the Dalton (Da), displacing other terms such as amu. 1 Da = 1/12 of the mass of a single atom of the isotope of carbon-12 (^{12}C) [156].

A mass spectrometry should always perform the following processes:

1. Produce ions from the sample in the ionization source.
2. Separate these ions according to their mass-to-charge ratio in the mass analyzer.

[153] Ekman, R., Silberring, J., Westman-Brinkmalm, A. M., Kraj, A., Desiderio, D. M. and Nibbering, N. M. 2008. *Mass Spectrometry: Instrumentation, Interpretation, and Applications* Wiley.

[154] Miller, J. M. 2005. *Chromatography: Concepts and Contrasts* Wiley.

[155] Steehler, J. K. 2009. Introduction to Mass Spectrometry: Instrumentation, Applications, and Strategies for Data Interpretation, 4th Edition (by J. Throck Watson and O. David Sparkman). *Journal of Chemical Education,* **86,** 810.

[156] Balogh M. P. 2008. A Mass Spectrometry Primer: Part I. URL http://www.chromatographyonline.com/lcgc/article/articleDetail.jsp?id=542753. Accessed 10 July 2012.

3. Eventually fragment the selected ions and analyze the fragments in a second analyzer.
4. Detect the ions emerging from the last analyzer and measure their abundance with the detector that converts the ions into electrical signals.
5. Process the signals from detector that are transmitted to the computer and control the instrument through feedback[157].

Generally, the mass spectrometer consists of three main parts: the ion source, the mass analyzer, and the detector.

7.2.2.1. Ion source

In this part, analyte atoms, molecules, or clusters are transferred into gas phase and ionized neutral species into charged particles. The sample solution is introduced to the source by a syringe pump (direct infusion) or as eluent from a separation technique such as LC. The main ion sources used to separate and determine phenolic compounds are electrospray ionization (ESI)[158-161], atmospheric-pressure chemical ionization (APCI)[162,163], fast-atom bombardment (FAB)[164,165], and matrix-assisted laser-desorption ionization (MALDI) [166]. These interphases supply energy to the solid sample or liquid in different ways, so that gaseous ions are directly formed while avoiding volatilization, thereby allowing the subsequent ionization of the compounds. APCI is commonly used

[157] Gross, J. H. and Roepstorff, P. *Mass Spectrometry: A Textbook* Springer.

[158] Waridel, P., Wolfender, J.-L., Ndjoko, K., Hobby, K. R., Major, H. J. and Hostettmann, K. 2001. Evaluation of quadrupole time-of-flight tandem mass spectrometry and ion-trap multiple-stage mass spectrometry for the differentiation of C-glycosidic flavonoid isomers. *Journal of Chromatography A,* **926,** 29-41.

[159] Ferreres, F., Silva, B. M., Andrade, P. B., Seabra, R. M. and Ferreira, M. A. 2003. Approach to the study of C-glycosyl flavones by ion trap HPLC-PAD-ESI/MS/MS: application to seeds of quince (Cydonia oblonga). *Phytochemical Analysis,* **14,** 352-359.

[160] Hvattum, E. and Ekeberg, D. 2003. Study of the collision-induced radical cleavage of flavonoid glycosides using negative electrospray ionization tandem quadrupole mass spectrometry. *Journal of Mass Spectrometry,* **38,** 43-49.

[161] Parejo, I., Jáuregui, O., Viladomat, F., Bastida, J. and Codina, C. 2004. Characterization of acylated flavonoid-O-glycosides and methoxylated flavonoids from Tagetes maxima by liquid chromatography coupled to electrospray ionization tandem mass spectrometry. *Rapid Communications in Mass Spectrometry,* **18,** 2801-2810.

[162] Wolfender, J. L., Waridel, P., Ndjoko, K., Hobby, K. R., Major, H. J. and Hostettmann, K. 2000. Evaluation of Q-TOF-MS/MS and multiple stage IT-MSn for the dereplication of flavonoids and related compounds in crude plant extracts. *Analysis,* **28,** 895-906.

[163] De Rijke, E., Zappey, H., Ariese, F., Gooijer, C. and Brinkman, U. A. T. 2003. Liquid chromatography with atmospheric pressure chemical ionization and electrospray ionization mass spectrometry of flavonoids with triple-quadrupole and ion-trap instruments. *Journal of Chromatography A,* **984,** 45-58.

[164] Ma, Y.-L., Cuyckens, F., Heuvel, H. V. D. and Claeys, M. 2001. Mass spectrometric methods for the characterisation and differentiation of isomeric O-diglycosyl flavonoids. *Phytochemical Analysis,* **12,** 159-165.

[165] Cuyckens, F., Shahat, A. A., Pieters, L. and Claeys, M. 2002. Direct stereochemical assignment of hexose and pentose residues in flavonoid O-glycosides by fast atom bombardment and electrospray ionization mass spectrometry. *Journal of Mass Spectrometry,* **37,** 1272-1279.

[166] Fulcrand, H. L. N., Mane, C., Preys, S. B., Mazerolles, G. R., Bouchut, C., Mazauric, J.-P., Souquet, J.-M., Meudec, E., Li, Y., Cole, R. B. and Cheynier, V. R. 2008. Direct mass spectrometry approaches to characterize polyphenol composition of complex samples. *Phytochemistry,* **69,** 3131-3138.

for detecting low-molecular-weight polar and nonpolar compounds. MALDI, as a special soft ionization technique is applied mostly to the analysis of large organic molecules such as proteins, peptides, and polysaccharides. The two most popular interfaces are *electrospray ionization* (ESI) and *atmospheric pressure chemical ionization* (APCI).

ESI is one of the most versatile ionization sources that provide a simple, real-time means of analyzing a wide range of polar molecules (100-200,000 Dalton range). It becomes the preferred choice for detecting polar compounds separated by liquid chromatography and capillary electrophoresis **(Figure 18)**.

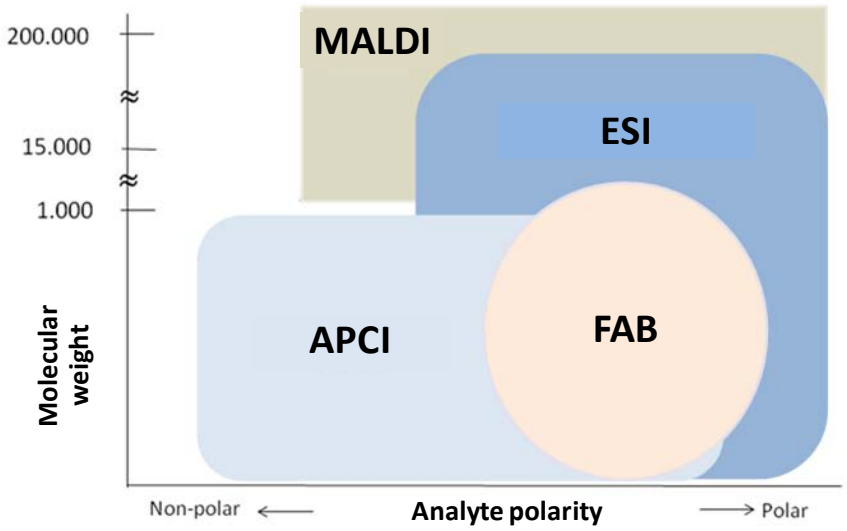


Figure 18. The application spectrum of different ion-source techniques.

In ESI, the sample is dissolved in a polar, volatile solvent and passed through the electrospray needle, which has a high potential difference (with respect to the counter electrode) applied to it. This high voltage between the needle and nozzle causes the fluid to form a Taylor cone, which is enriched with negative or positive ions at the tip. A spray of charged droplets is ejected from the Taylor cone by the electric field. The droplets shrink through evaporation, assisted by a warm flow of nitrogen gas passing across the front of the ionization source. Ions are formed at atmospheric pressure and pass through a small aperture into the high vacuum at the mass analyzer. The sample arrives to the spray chamber as a fine mist of droplets or spray. A drying gas, e.g. N_2, at a fairly high temperature causes the evaporation of any solvent from the droplets. A voltage gradient between the tip of the spray needle and the entrance to the transfer capillary, as well as a pressure difference from atmospheric pressure to vacuum encourages appropriately charged ions to move into the capillary and on

towards the skimmers. Pressure is reduced by pumping most of the vaporized sample and mobile phase to waste (no concentration takes place); only a tiny fraction of the sample is drawn into the MS itself.

This forces the spraying of charged droplets from the needle with a surface charge of the same polarity to the charge on the needle. The droplets are repelled from the needle towards the source-sampling cone on the counter electrode. As the droplets traverse the space between the needle tip and the cone, any solvent evaporates. This is circled on the **Figure 19** and **20**.

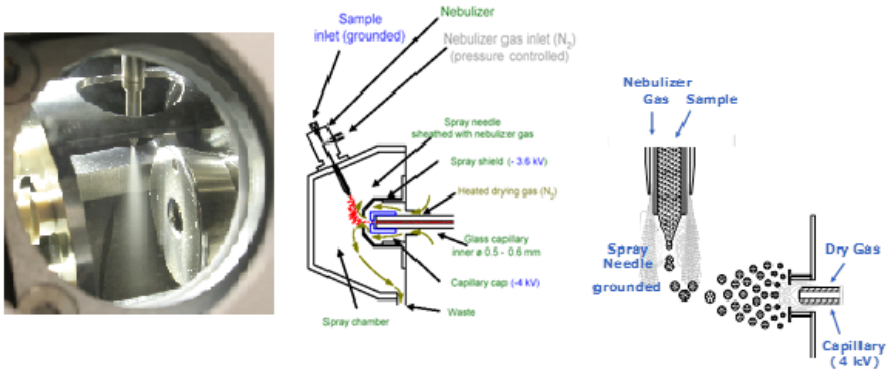


Figure 19. ESI ion source with its schematic diagram

As the solvent evaporates, the droplet shrinks until it reaches the point that the surface tension can no longer sustain the charge (the Rayleigh limit) at which point a "Coulombic explosion" occurs and the droplet is ripped apart. This produces smaller droplets that can repeat the process as well as naked charged analyte molecules.

These charged analyte molecules (which are not strictly ions) can be singly or multiply charged. This is a very soft method of ionization as very little residual energy is retained by the analyte upon ionization. It is the generation of multiply charged molecules that enables high-molecular-weight components such as proteins to be analyzed, since the mass range of the mass spectrometer is greatly increased, as it actually measures the *mass-to-charge ratio* rather than mass *per se*. The major disadvantage of the technique is that very little (usually no) fragmentation occurs,

although this may be overcome through the use of tandem mass spectrometric techniques such as MS/MS or MS^n[167,168].

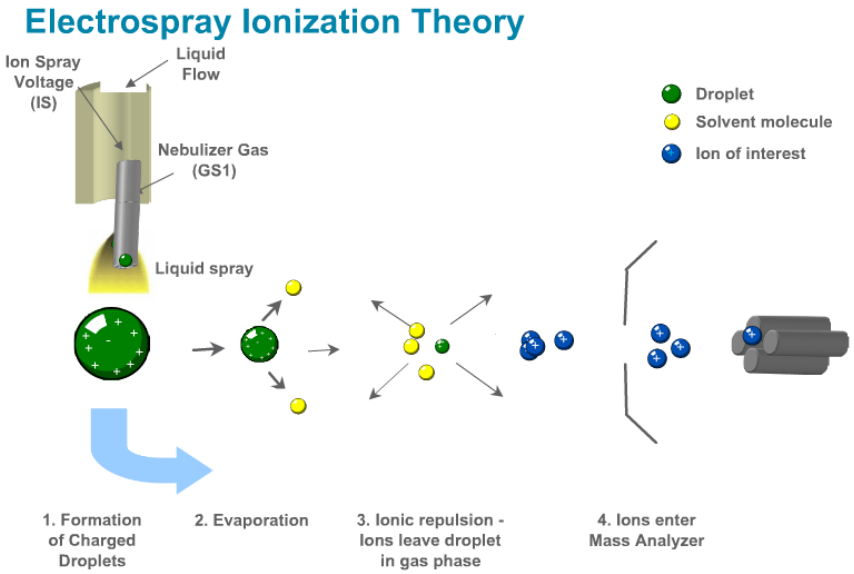


Figure 20. Electrospray Ionization theory scheme.

A trace of formic acid or acetic acid is often added to aid protonation of the analyte molecules in the negative and positive ionization mode. In negative ionization mode, ammonia solution or a volatile amine is added to aid the deprotonating of the analyte molecules.

7.2.2.2 Mass analyzer

This is a device that can separate and sort ions coming from the ion source according to their *m/z* values. Therefore, an ion source has to be coupled to mass analyzer as the "heart" of the mass spectrometer.

The analyzer uses electrical or magnetic fields, or a combination of the two, to move the ions from the region where they are produced, to a detector, where they produce a signal which is amplified. Since the motion and separation of ions is based on electrical or magnetic fields, the mass-to-charge ratio, and not only the mass, is

[167] Bristol Biogeochemistry Research Centre, University of Bristol (2005) High performance liquid chromatography mass spectrometry (HPLC/MS). URL http://www.bris.ac.uk/nerclsmsf/techniques/hplcms.html. Accessed 09 Jan. 2013.

[168] Anon. 2007. LC/MS applications guide. Dionex Corporation.

important. The analyzer is operated under high vacuum, so that the ions can travel to the detector with a sufficient yield[169].

Time-of-flight (TOF)

Ion separation in a TOF analyzer is based on Newton's third law. The underlying principle is that ions of different masses with equal kinetic energy have different velocities. Thus, ions are separated according to the time difference between a start signal and the pulse generated when an ion hits the detector. The faster or lighter the ion, the shorter the time of flight and the resulting time-of-flight spectrum can be converted into a mass spectrum. There is no need to know the exact potentials and distances of the spectrometer, as the time/mass conversion is made by calibration with ions of known masses. ESI can be connected to the TOF analyzers through orthogonal acceleration where the ions generated by the ion source enter the TPF analyzer perpendicular to its main axis.

In the time-of-flight (TOF) analyzer, ions formed in an ion source are extracted and accelerated to a high velocity by an electric field into an analyzer consisting of a long, straight "drift tube". The ions pass along the tube until they reach a detector. After the initial acceleration phase, the velocity reached by an ion is inversely proportional to its mass (strictly, inversely proportional to the square root of its *m/z* value). Since the distance from the ion origin to the detector is fixed, the time taken for an ion to traverse the analyzer in a straight line is inversely proportional to its velocity and hence proportional to its mass (strictly, proportional to the square root of its *m/z* value). Thus, each *m/z* value has its characteristic time of flight from the source to the detector[170]. TOF technology presents numerous advantages, such as high mass resolution, high mass accuracy, theoretically unlimited mass range, and relatively low cost. A representation of the TOF analyzer and its schematic diagram are represented in the following figure **(Figure 21)**.

Mass resolution is related to the length of the drift tube. Spatial limitations determine the length of the drift tube, so that one popular configuration uses an electrostatic mirror (reflectron) to increase the effective length of the drift tube, and thus improve mass resolution. With sufficient path length, the TOF can provide higher mass resolution than the quadrupole, so it is useful for structural work.

[169] De Hoffmann, E. and Stroobant, V. 2007. *Mass Spectrometry: Principles and Applications* Wiley.

[170] Watson, J. T. and Sparkman, O. D. 2008. *Introduction to Mass Spectrometry: Instrumentation, Applications, and Strategies for Data Interpretation* Wiley.

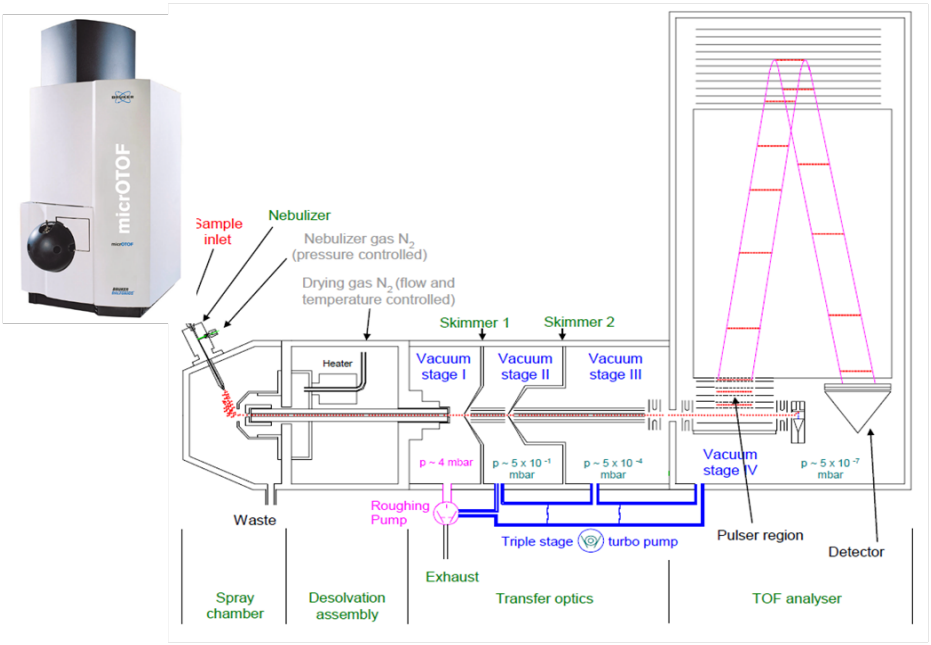


Figure 21. A schematic diagram of a time-of-flight mass spectrometer (TOF-MS).

Off-axis or orthogonal acceleration TOF-MS of ions from continuous ion beams has been known since the 1960s. However, it was only in the late 1980s and early 1990s that the current range of TOF-MS instruments with greatly improved resolving power and mass accuracy was developed. The key features enabling accurate mass measurement include high efficiency in gating ions from an external continuous source (e.g. ESI, APCI), simultaneous correction of velocity and spatial dispersion, and increased mass resolving power. TOF-MS has been supported more by the digital electronics revolution than by MS technologies with greater reliance on analog-signal processing. In this case, TOF-MS refers to orthogonal acceleration technology unless otherwise stated[171]. TOF-MS uses the differences in transit time through a flight/drift zone to separate ions of different masses. The principle is that smaller ions, being lighter, will reach the detector faster than will heavier ions. It operates in a pulsed mode so that ions must be produced or extracted in pulses. An electric field is used to accelerate all the ions into the field-free drift zone where lighter ions have a higher velocity than heavier ones and reach the detector sooner.

[171] Allwood, J. W. and Goodacre, R. An introduction to liquid chromatography–mass spectrometry instrumentation applied in plant metabolomic analyses. *Phytochemical Analysis,* **21,** 33-47.

Quadrupole-time-of-flight (QTOF)

The term "hybrid" applies to various mass-spectrometer designs that are composites of existing technologies such as double-focusing, magnetic sectors. One of the most innovative designs, the quadrupole time-of-flight (QTOF) mass spectrometer, couples a TOF instrument with a quadrupole instrument. This pairing results in the best combination of several performance characteristics: accurate mass measurement, the ability to carry out fragmentation experiments, and high-quality quantitation.

Quadrupoles use a set of four rods and a carefully controlled electric field to isolate selected ions from the sample. Ions of a selected mass-to-charge ratio (*m/z*) are then passed to an electron multiplier for detection, providing a selective response for the desired analyte.

The development of quadrupole–time-of-flight (TOF) mass spectrometers have rapidly been embraced by the analytical community as powerful and robust instruments with unique capabilities. In particular, they combine the high-performance of time-of-flight analysis in both the mass spectrometry (MS) mode as precursor and tandem MS (MS/MS) mode as product ion, with the well-accepted and widely used techniques of electrospray ionization (ESI) and atmospheric pressure chemical ionization (APCI). These are now applied to problems which range from nano-spray analysis of biological samples to liquid chromatography (LC)/MS/MS of pharmaceutical preparations at high flow rates.
Their rapid acceptance is due not only to the attractive combination of high sensitivity and high mass accuracy for both precursor and product ions, but also to the simplicity of operation for those already familiar with LC/MS analysis on quadrupole and triple quadrupole instruments[172].

The tandem transmission QTOF mass spectrometer has the advantage of mass accuracy to the nearest 0.1 millimass unit. Tandem Mass Spectrometry (MS/MS) is a technique where structural information on sample molecules is compiled by using multiple stages of mass selection and mass separation. Meanwhile, hybrid mass spectrometers are instruments equipped with two or more different types of mass analyzers coupled together.

[172] Chernushevich, I. V., Loboda, A. V. and Thomson, B. A. 2001. An introduction to quadrupole–time-of-flight mass spectrometry. *Journal of Mass Spectrometry,* **36,** 849-865.

The high-mass accuracy results from the orthogonal arrangement between the longitudinal axis of the ion beam coming from the quadrupole and the direction in which the ion beam is extracted into the TOF instrument. A schematic presentation of QTOF-MS is shown in **Figure 22.**

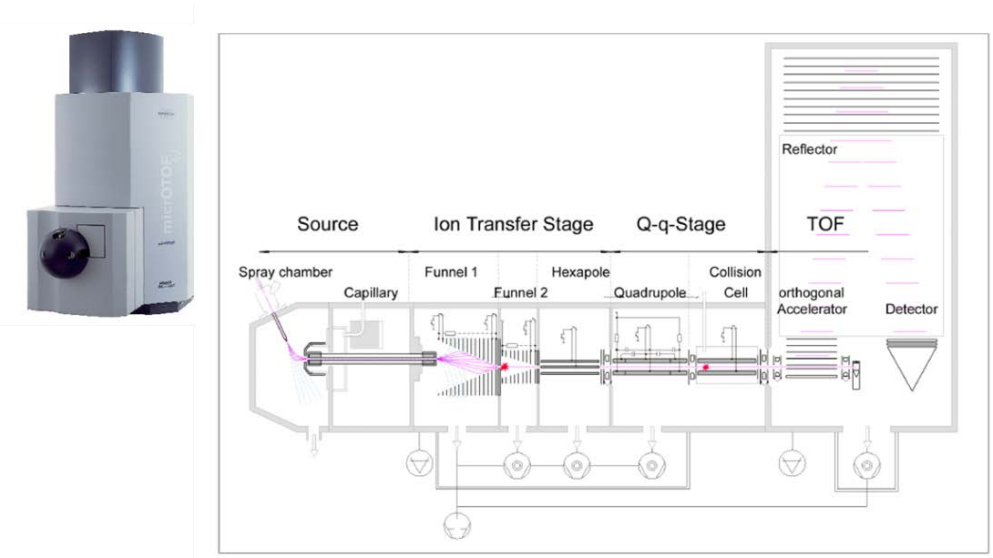


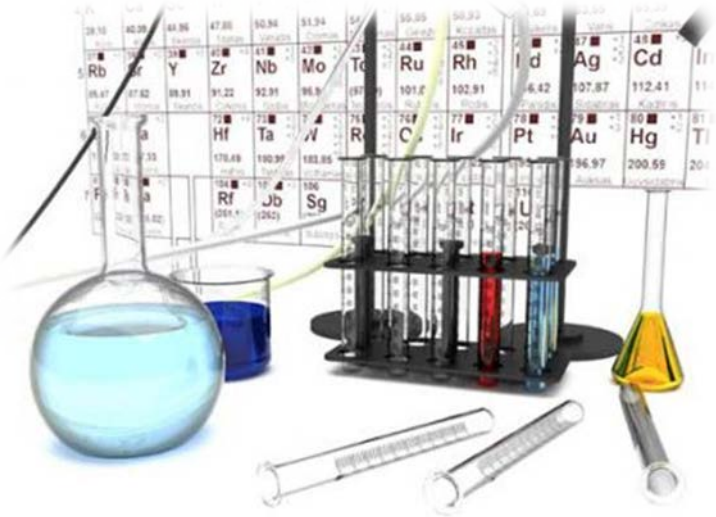
Figure 22. A schematic diagram of a quadrupole-time-of-flight mass spectrometer (QTOF-MS).

7.2.2.3. MS detector

The role of the detector is to convert the energy of incoming particles into a current signal that is registered by the electronic devices and transferred to the computer of the acquisition system of the mass spectrometer. The detector records the relative abundances of individual *m/z* values[173]. Since the mass analyzer and the detector require low pressure for operation, the instrument also needs a pumping system.

Relative ion current (signal) is plotted vs. *m/z*, producing mass spectrum. The peak at the highest *m/z* ratio is not always the molecular ion species ($[M-H]^-$ in the negative mode and $[M+H]^+$ in the positive mode), because adducts with solvent and/or acid molecules and also molecular complexes ($[2M-H]^-$ or $[2M+H]^+$) can be generated[174,175].

[173] Allwood, J. W. and Goodacre, R. An introduction to liquid chromatography-mass spectrometry instrumentation applied in plant metabolomic analyses. *Phytochemical Analysis,* **21,** 33-47.
[174] Barnes, S., Kirk, M. and Coward, L. 1994. Isoflavones and their conjugates in soy foods: extraction conditions and analysis by HPLC-mass spectrometry. *Journal of Agricultural and Food Chemistry,* **42,** 2466-2474.
[175] Tian, Q., Li, D. and Patil, B. S. 2002. Identification and determination of flavonoids in buckwheat (Fagopyrum esculentum Moench, Polygonaceae) by high-performance liquid chromatography with electrospray ionisation mass spectrometry and photodiode array ultraviolet detection. *Phytochemical Analysis,* **13,** 251-256.



Experimental Section

Chapter 1

Phytochemical Characterization of Green Beans (*Phaseolus vulgaris L.*) by Using High-performance Liquid Chromatography Coupled with Time-of-flight Mass Spectrometry

This work has been accepted as research article in:

Phytochemical Analysis

Research Article

Phytochemical Analysis

(wileyonlinelibrary.com) DOI 10.1002/pca.2385

Published online in Wiley Online Library

Phytochemical Characterisation of Green Beans (*Phaseolus vulgaris L.*) by Using High-performance Liquid Chromatography Coupled with Time-of-flight Mass Spectrometry

Ibrahim M. Abu-Reidah, David Arráez-Román, Jesús Lozano-Sánchez, Antonio Segura-Carretero and Alberto Fernández-Gutiérrez

Phytochemical characterization of green beans (*Phaseolus vulgaris L.*) by using high-performance liquid chromatography coupled with time-of-flight mass spectrometry

Abstract

Although green beans (*Phaseolus vulgaris L.*) constitute a major agro-industrial crop worldwide and involve many bean varieties, the phytochemical composition and how this fluctuates among varieties is not well known.

The purpose of this work was to conduct a comprehensive characterisation of the phytochemical compounds found in three varieties of green bean. Hydromethanol extracts from green beans were analysed by high-performance liquid chromatography (HPLC) coupled with electrospray time-of-flight mass spectrometry (ESI–TOF–MS) in order to characterise phytochemical compounds from different varieties of *P. vulgaris L.* The compounds were characterised based on interpreting their mass spectrum provided by the TOF–MS as well as by comparison with information from the literature (some compounds have been described previously in Fabaceae).

In this work, 72 phytochemical compounds were tentatively characterised by HPLC-ESI–TOF–MS. These compounds were classified as, 10 phenolic acids, 59 flavonoids, two lignans and an iridoid. Notably, out of the 72 compounds, 54 are new and their isomers have been characterised for the first time in green beans. The phytochemical composition of three different varieties of *P. vulgaris L.* have been characterised using HPLC-ESI-TOF-MS. A total of 72 phytochemical compounds were characterised, 54 being reported in green beans for the first time. Among these were the main flavonoids detected. These results highlight the influence that variety can exert on the quality of phytochemicals. Given that new phytochemical compounds have been characterised, this study offers a useful approach for improving and updating the food-composition tables.

Keywords: *HPLC-ESI-TOF-MS; green beans; phenolic compounds; phytochemical composition; Fabaceae (Leguminosae).*

Introduction

According to many epidemiological studies, regular consumption of an enriched diet in plant foods (e.g. fruits and vegetables) is known to protect against or reduce the risk of developing cardiovascular disease, cancer, and other chronic disorders. The link between vegetable intake and human health has stimulated multidisciplinary research to identify specific plant components that may offer health benefits[1]. The most well-known antioxidant constituents of fruits and vegetables, which may aid in oxidation prevention, protection and delay, are vitamins C and E, carotenoids, minerals (selenium, zinc and sulphur), as well as some peptides and phenolic compounds[2]. Among these, phenolic compounds, ubiquitous secondary plant metabolites, may be categorized into three groups, namely, simple phenols, phenolic acids, (Hydroxycinnamic and benzoic acids derivatives), and flavonoids (e.g. flavones, flavanones, flavanonols, flavanols, isoflavones, and lignans)[3]. Many factors can influence the content in phenolic compounds, such as variety, geographic area, among others.

These compounds not only have major effect on sensorial and nutritional quality of fruits, vegetables, and other food plants[1], but also may contribute to oxidative stability. Moreover, they are of considerable interest and have received increasing attention in recent years due to their bioactive functions, which could be related to antioxidant properties, although the molecular mechanism of these compounds in relation to many diseases could have different cellular targets. These phytochemicals have been reported to have a positive impact on health, e.g. curbing the cancer-tumour development, and providing anti-bacterial, anti-viral, antispasmodic, and anti-inflammatory properties[4,5].

Green bean (*Phaseolus vulgaris L.*, Fabaceae), an economically important food crop worldwide, has a host of common names used locally, such as: French beans, string beans (old name), and snap beans. Being the green (unripe) pods of any kind of edible bean, these pods are harvested for fresh consumption, freezing or canning[6]. Green beans are relatively rich in antioxidants and have been classified in the top 10 common vegetables in relation to antioxidant content and activity[7]. In some places of the world, green beans are deemed a major source of flavonoids[8].

Green beans considered as a functional food with health properties which have been discussed extensively in the literature. For example, the antiglycaemic, antioxidant, and hypolipidemic effects of green bean pods in diabetic rats have previously been described[9]. These positive effects have been linked to the antioxidant activity supported by the flavonoids present in green bean pod extracts[9,10]. The study of the specific consumption of green beans and its relation to reducing cancer risk has also been reported elsewhere[11]. Furthermore, isoflavones, a major subclass of flavonoids which are the most abundant in Fabaceae[12], have been proposed to play a role in learning and memory during aging, bone metabolism, thrombogenicity, and the prevention of climacteric syndrome as well as several types of cancer[13].

In green bean, pharmaceutical interest in phytochemical compounds due to their health properties has stimulated multidisciplinary research on their chemical composition. Since previous studies have focused only on flavonol identification[8,14-16], little is known about the other phytochemical compounds.

Over the last two decades, several studies using different analytical techniques have examined the phenolic composition of green beans, mainly flavonoids such as

flavonols [8,14,16-18]. The usual technique to analyse the phytochemical composition in food is high-performance liquid chromatography (HPLC). Recently, an improvement in chromatographic performance has been achieved by the introduction of narrow-bore columns packed with very small particles and high flow rate with delivery systems operating at high backpressures. The major advantages of this improvement over conventional HPLC are the improved resolution, shorter retention times, higher sensitivity, and the better performance. The coupling of HPLC with mass spectrometry (MS) detectors offers a potent analytical tool, which has been applied in recent publications characterizing food products. The inherent characteristics of time-of-flight mass spectrometry (TOF-MS) such as the accurate mass data, elevated sensitivity and a high mass resolving power[19] make it one of the most desirable detection methods.

In this context, the aim of the present work was to characterize the phenolic profile and other secondary plant metabolites present in three green bean varieties (whole) pods. This is the first study available in which this vegetable has been characterized using HPLC-ESI-TOF-MS technique. Results of this preliminary work could aid in the understanding of the health properties of green beans.

Experimental

Chemicals

All chemicals were of analytical reagent grade and used as received. HPLC-grade acetonitrile and methanol were purchased from Labscan (Dublin, Ireland). Acetic acid of analytical grade (assay>99.5 %) was purchased from Fluka (Buchs, Switzerland). Double-deionized water with a conductivity of <18.2 MΩ was obtained with a Milli-Q system (Millipore, Bedford, MA).

Plant materials

Commercial pods used in this study were from three different green bean (*Phaseolus vulgaris L.*) varieties cultivated in the zone of Spain as follow: Perona: San Agustin/Almeria, Helda: Las Norias de Daza/Almeria and Strike: San Agustin/Almeria. The green bean pods were produced between the period Octuber 2010 and February 2011. Samples in their maturation stage were purchased from local commercial market, which were washed with distilled water, frozen at -25°C and then lyophilized.

Preparation of green beans extracts

To extract the secondary plant metabolites from all varieties, the samples were treated according to the extraction process previously described in the literature[20]. Briefly, lyophilized samples (0.5 g) were dissolved with 16 mL of MeOH/H_2O (80:20, v/v) which were sonicated for 30 min in an ultrasonic bath. Then, the mixture was centrifuged at 4000 *g* for 15 min and the supernatant was collected in a round bottom flask. The solvent was evaporated under vacuum using rotary evaporation at 38°C. Finally, the dry residue was dissolved with 0.5 mL of MeOH/H_2O (80:20, v/v), passed through a 0.22 µm syringe filter, and stored at -20°C until analysis.

HPLC conditions

The phenolic compounds in pod extracts were separated on an Agilent 1200 series Rapid Resolution LC (Agilent Technologies, CA, USA) equipped with a vacuum degasser, auto-sampler, binary pump, and diode array detector (DAD). The chromatographic separation was performed in a Zorbax C18 analytical column (4.6

mm x 150 mm, 1.8µm particle size) purchased from Agilent Technologies (Palo Alto, CA, USA). The mobile phases were acidified water (0.5% acetic acid, v/v) and acetonitrile as eluent A and B, respectively. The chromatographic method consisted in the following multistep linear gradient: 0 min, 0% B; 10 min, 20% B; 15 min, 30% B; 20 min, 50% B; 25 min, 75% B; 30 min, 100% B; 32 min 0% B; and finally a 3-min post-run was used after each analysis. The injection volume was 10 µL and the column temperature was maintained at 25ºC. The flow rate was set at 0.80 mL/min throughout the gradient.

ESI-TOF-MS conditions

The HPLC system was coupled to a time-of-flight mass spectrometer (micrOTOF^{TM}, Bruker Daltonics GmbH, Bremen, Germany), equipped with a model G1607A ESI interface (Agilent Technologies) operating in negative ion mode. At this stage, the use of a T-type splitter (split=1:3) was required for coupling with the MS detector as the flow at the TOF detector had to be 0.2 mL/min to achieve reproducible results and stable spray. Thus, in this study the flow which arrived into the ESI-TOF-MS detector was 0.2 mL/min. The optimum values of source parameters were: capillary voltage of +4 kV; drying-gas temperature, 190°C; drying-gas flow, 9 L/min; nebulizing-gas pressure, 2 bar; and end-plate offset, -0.5 kV. The values of transfer parameters were: capillary exit, -120 V; skimmer 1, -40 V; hexapole 1, -23 V; RF hexapole, 50 Vpp; and skimmer 2, -22.5 V. The source and transfer parameters were optimized in ensure good sensitivity, to reach reasonable resolution within the mass range of the target compounds (50- 1100 *m/z*), and to improve the ionization performance.

External mass spectrometer calibration was performed passing a solution containing sodium acetate clusters (5 mM sodium hydroxide in water/2-propanol 1/1 (v/v), with 0.2% of acetic acid) in quadratic high-precision calibration (HPC) regression mode. With this method, an exact calibration curve was achieved based on numerous cluster masses, each differing by 82 Da ($C_2H_3NaO_2$). The calibration solution was injected at the beginning of the run using a Cole Palmer syringe pump (Vernon Hills, Illinois, USA) and all the spectra were calibrated prior to polyphenol characterization. The accurate mass data for the molecular ions were processed through Data Analysis 4.0 software (Bruker Daltonics, Bremen, Germany) which provided a list of possible elemental formulas by using the Generate Molecular Formula ™ editor. The latter uses a CHNO algorithm that provides standard functionalities such as minimum/maximum elemental range, and ring-plus double-bond equivalents, as well as a sophisticated comparison of the theoretical with the measured isotope pattern (mSigma value) for increasing the confidence in the suggested molecular formula. According to the literature, the widely accepted accuracy threshold for confirming elemental compositions was established at 5 ppm for most of the compounds[21]. However, the identification of the phytochemical compounds was based on the accurate mass measurements of the pseudomolecular [M-H]-, also on the previous related literature from the Fabaceae family members, since no commercial standards were available for all green beans detected compounds.

Results and Discussion

Characterization of phenolic compounds and their derivatives

The characterization process was conducted using the elution order, the interpretation of their mass spectrum provided by the TOF-MS and the information

previously reported in the literature (some compounds have been previously described in Fabaceae). **Figure 1** shows the resulting base peak chromatograms (BPC) of samples from the Strike (A), Helda (B), and Perona (C) varieties.

The main phytochemicals tentatively characterized are included in Table 1. A total of 72 compounds belonging to various metabolite classes and their derivatives were tentatively characterized in the different green bean varieties based on the information generated by the HPLC-ESI-TOF-MS analytical technique such as retention time, calculated *m/z*, molecular formula, error, mSigma values, and some fragments produced by the ionization conditions.

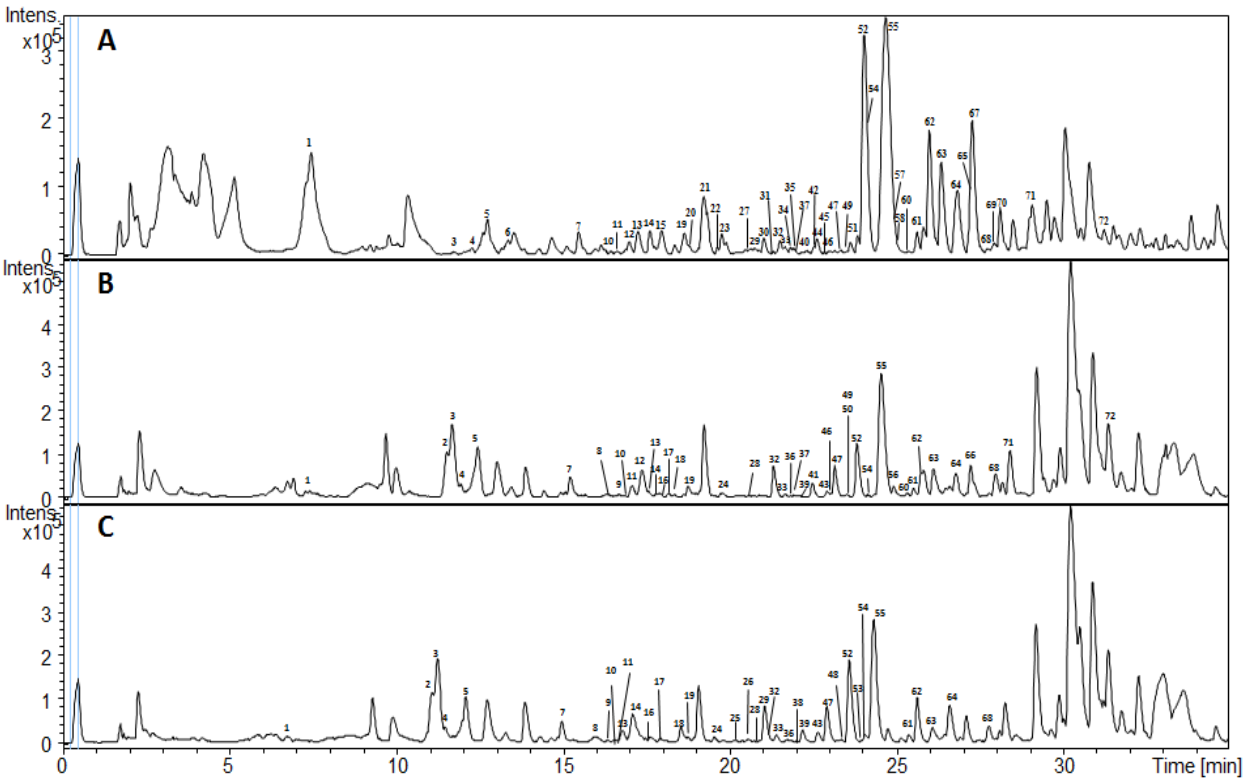


Figure 1. Base peak chromatograms (BPC) of green bean extracts obtained by HPLC-ESI-TOF-MS (Three varieties: A: Strike; B: Helda; and C: Perona).

The phytochemicals which were tentatively characterized in the green beans can be divided into different classes such as phenolic acids (Hydroxybenzoic and

Hydrxyocinnamic acids), flavonoids (flavonols, flavanols, flavones, dihydroflavonols, flavanones, isoflavones, and isoflavanones), lignans and other polar compounds. **Figure 2** shows the qualitative distribution of the main phytochemical compounds characterized in all varieties.

Phenolic acids

Hydroxybenzoic acid derivatives

Different compounds belonging to Hydroxybenzoic-acid group were characterized in the samples analysed. The spectra generated for peaks 3 and 4 gave deprotonated molecules at *m/z* 255 and 329, respectively, which could be attributed to Hydroxybenzoic acid derivatives reported in the Fabaceae (Leguminosae) family[22]. In this sense, according to the data generated by TOF-MS and the literature reported previously in this family[23], peak 3 was proposed as (+)-Piscidic acid. Peak 4, which was proposed as vanillic acid ß-glycoside, presented a fragment, produced by the ionization conditions, at *m/z* 167, which was tentatively attributed to the loss of glucose moiety.

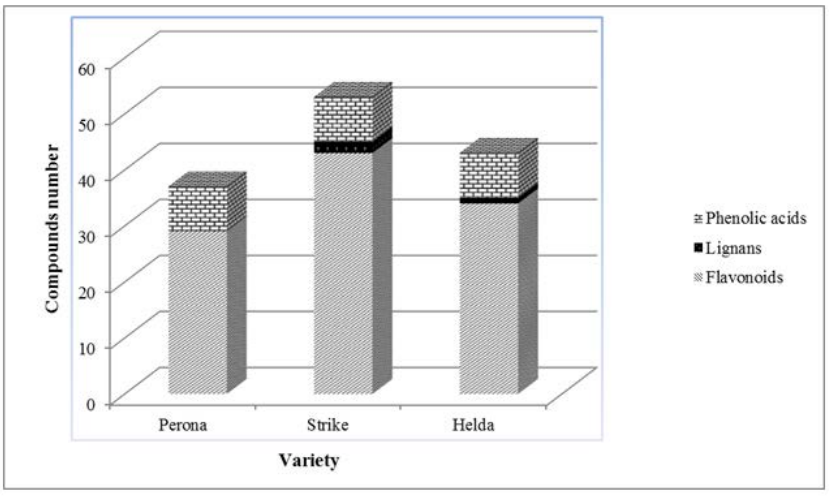


Figure 2. Phytochemical compounds characterized in three different varieties of green beans.

Peak 5, with experimental *m/z* 461 and molecular formula $C_{19}H_{25}O_{13}$ generated by TOF analyser; was tentatively proposed as Sibiricose A3. It elutes after (+)-Piscidic acid and vanillic acid ß-glycoside, both compounds being more polar than Sibiricose A3 (11.21, 11.44 and 12.06 min, respectively). These compounds have been detected in all green beans varieties.

On the other hand, peaks 2 and 6 have been detected with *m/z* at 431 with different retention times (Table 1). The fragment ion appears at *m/z* 137, which could refer to salicylic acid. Thus, peaks 2 and 6 were assigned as primeveroside salicylic acid and its isomer.

Hydroxycinnamic acid derivatives

For Hydroxycinnamic acid derivatives, five compounds belonging to this group were characterized. Peaks 7 and 16 had the same deprotonated molecule (*m/z* 355) and the molecular formula generated ($C_{19}H_{25}O_{13}$). According to the literature, these compounds were assigned to different isomers of ferulic acid ß-glycoside[24]. Both compounds presented a fragment at *m/z* 193, generated by the ionization conditions, corresponding to the loss of ferulic acid moiety. The isomer which elutes at 18.05 min was characterized only in the P and H varieties, while the other was detected in all varieties.

Peak 10 gave a deprotonated molecule at *m/z* 325 at 16.49 min being proposed as coumaroylglucose. This compound has been previously reported in the literature[24]. The product ion at *m/z* 163 generated by the experimental conditions is justified by the loss of the acidic moiety (coumaric acid).

Peak 12 was tentatively assigned to chlorogenic acid. Indeed, this compound has already been described in the literature[25]. The mass spectral data showed a fragment at *m/z* 191 (**Figure 3A**) corresponding to the loss of 162 Da (quinic acid). The ion found at *m/z* 341 (peak 13) at 16.85 min may correspond to caffeoylglucose[24], confirmed by the neutral loss of glucose moiety with the deprotonated ion at *m/z* 179 (caffeic acid). For all derived compounds from Hydroxycinnamic acids, the mass spectral data showed a fragment generated by the ionization conditions which can be justified by the release of the acidic moiety.

Flavonoids

Flavonol derivatives

The main components of the phytochemical fraction and its derivatives in green beans hydro-methanol extracts were flavonols and derivatives (**Figure 2**). Indeed, flavonols have been widely described in green beans[8,14-16]. A total of 41 flavonols were detected in the different varieties, the majority of these compounds being kaempferol and quercetin derivatives.

Thus, peaks 8, 18, and 43 gave the same deprotonated molecule at *m/z* 609 and the molecular formula ($C_{27}H_{29}O_{16}$).The MS spectrum has demonstrated the presence of the fragment ion at m/z *285* (corresponding to kaempferol aglycone), this could be interpreted by the loss of two hexoside moieties [M-H-292]-.Therefore, these compounds have been tentatively postulated as kaempferol 3,7-dihexoside or isomers[16].

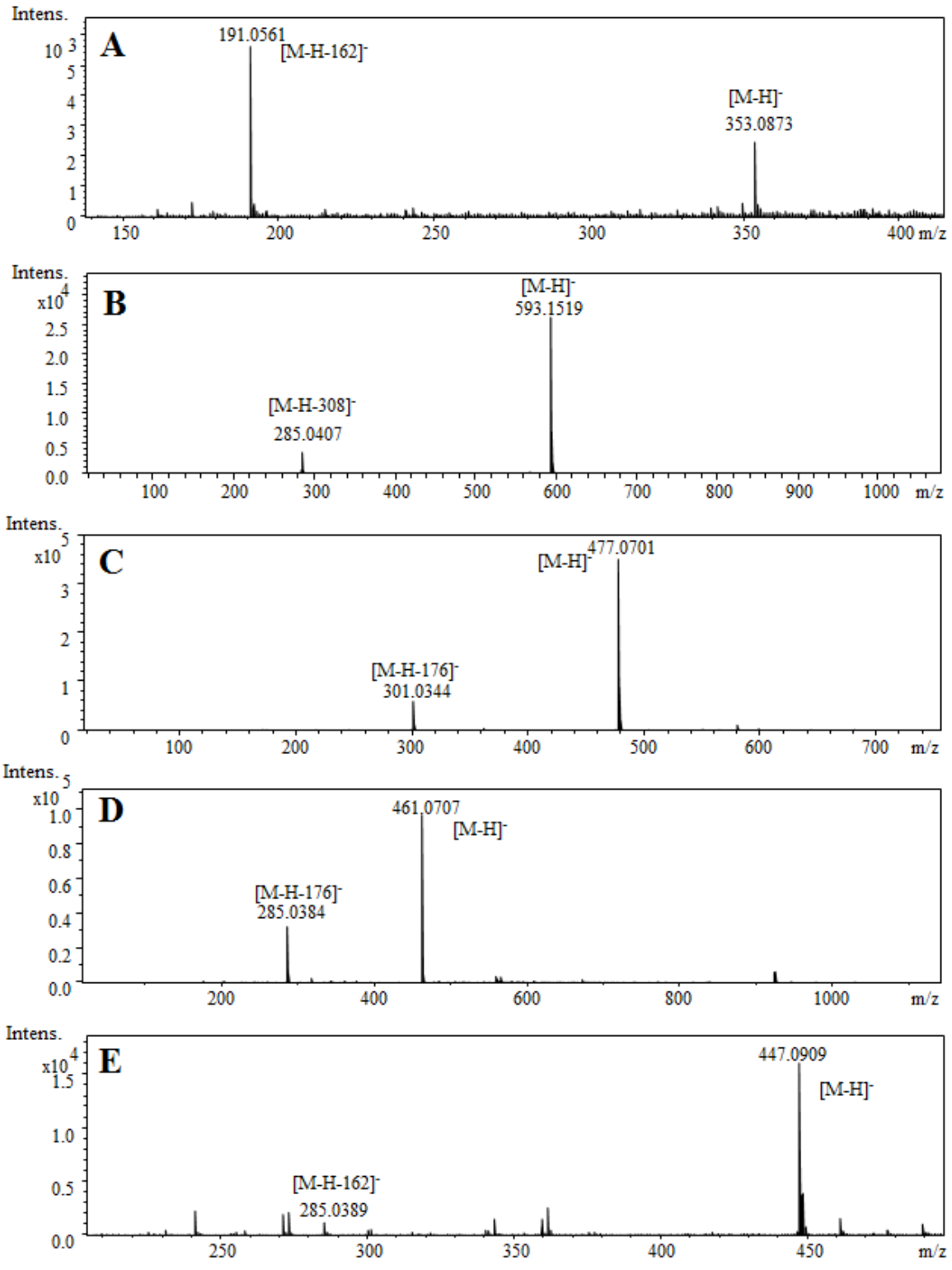


Figure 3. Mass spectra of the major compounds in green beans obtained by HPLC-ESI-TOF-MS.

On the other hand, peak 9 had a pseudomolecular ion at *m/z* 887. This compound was tentatively proposed as kaempferol -O-trihexosyl-pentoside (see **Figure 4a**). Identification of this compound was based on the acceptable MS spectrum data; thereafter, the yielded ion at m/z 447 was postulated as kaempferol hexoside obtained by the Smart Formula Editor possessing acceptable Error and mSigma values. Nevertheless, this compound has already been reported in the Leguminosae[26].

In the characterization of these compounds we relied on the MS data and previous reported bibliography either data reported about the tested matrix or from other matrices of the same botanical family. This on account of no commercial standards for all the characterized compounds could be found.

Peak 11 gave a deprotonated molecule at *m/z* 639 at 16.58 min; and it was tentatively characterized as quercetin 3-glucuronide-7-glucoside. Anyway, this compound was detected in all green beans extracts additionally it demonstrated the ion at *m/z* 301 which could refer to the aglycone (quercetin). The compound at *m/z* 771 at 16.93 min (peak 14) was tentatively proposed as glucopyranosylrutin. This compound had a fragment at *m/z* 301 which could be attributed to quercetin (**Figure 4b**).

Peaks 17, 39, 41 had the same deprotonated ions at *m/z* 595 and identical molecular formula ($C_{26}H_{27}O_{16}$). These compounds, which were detected at different retention times, were characterized as quercetin 3-vicianoside or isomer (**Figure 4c**). These compounds have been previously reported in *Vicia angustifolia* (Fabaceae)[27]. Furthermore, compounds 39 and 41 showed a fragment ion at *m/z* 301 corresponding to quercetin aglycone.

Peaks 19, 33 and 48, which generated the same deprotonated molecule, were proposed as kaempferol-3-O-rutinoside-7-O-glucose or isomers according to the mass spectra, which showed two fragments, at *m/z* 593 and *m/z* 285 (**Figure 4d**). Furthermore, these compounds have been previously reported in *Cladrastis kentukea* (Fabaceae)[28].

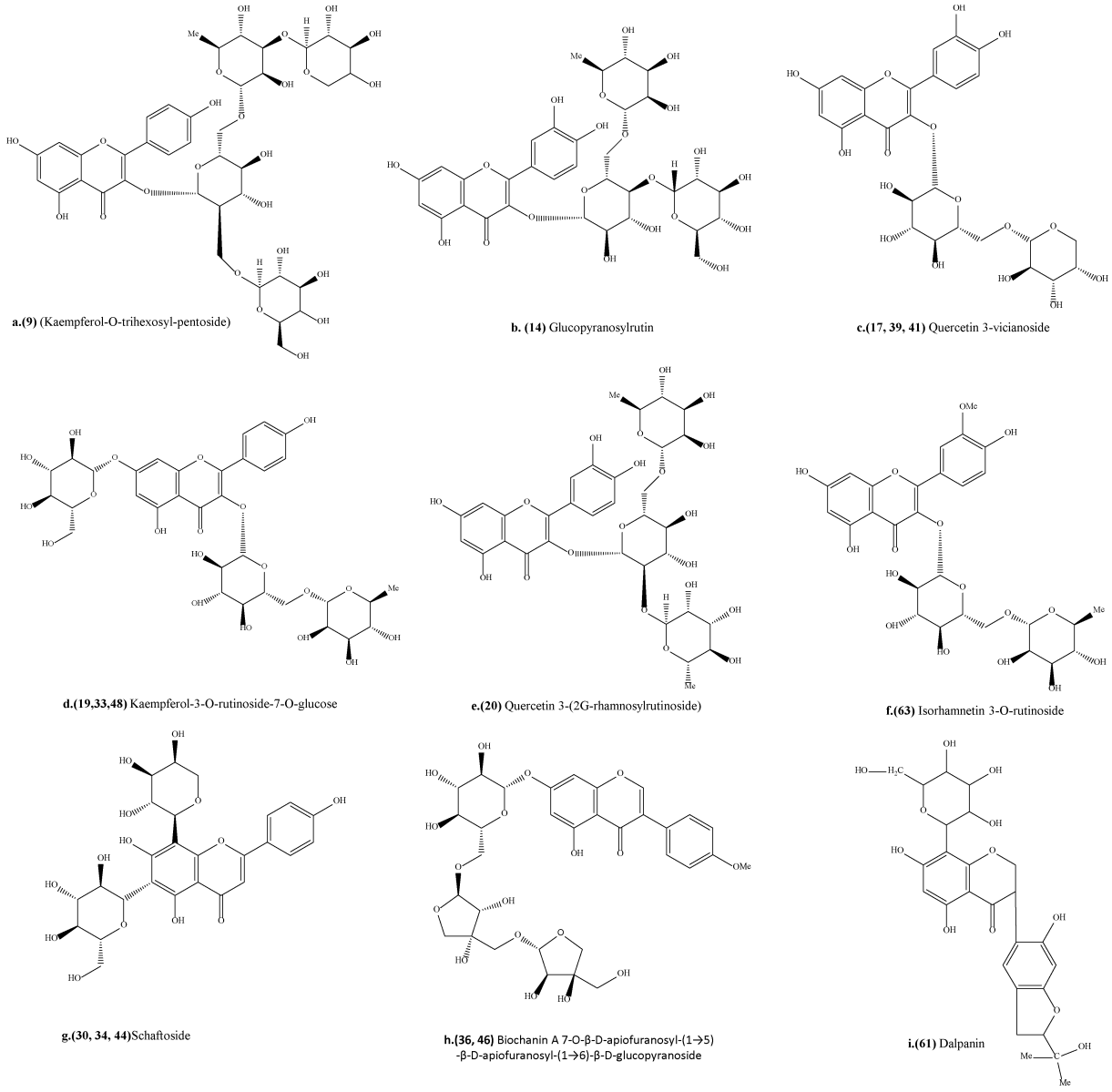


Figure 4. Structures of new phytochemical compounds characterized in green beans.

On the other hand, peak 20 was tentatively proposed as quercetin 3-(2G-rhamnosylrutinoside) (**Figure 4e**). The product ion at *m/z* 609 may justify the loss of rutin. This compound has been described in Fabaceae[22,28].

Peaks 15, 31, and 57, at the retention times 17.56, 21.09, and 24.37 min, respectively, had the same deprotonated molecule (*m/z* 755) and molecular formula ($C_{33}H_{39}O_{20}$) generated by TOF analyser. Although several authors have described these compounds as derivatives of either kaempferol or quercetin[22,28], with the use of the information generated by TOF analyser, these deprotonated molecules may be assigned to different quercetin 3-(2G-rhamnosylrutinoside) isomers as well as kaempferol-3-O-rutinoside-7-O-glucose isomers.

Peaks 21, 29, 32, and 40 had the same deprotonated molecule at *m/z* 741 and identical molecular formula $C_{32}H_{37}O_{20}$. These compounds have been suggested as quercetin 3-O-xylosylrutinoside or isomers, which have been previously reported in green beans[8,15]. For all peaks, the product ion at *m/z* 301 can be justified by the release of quercetin.

Peaks 22 and 27 gave the deprotonated molecule at *m/z* 579 with the following molecular formula $C_{26}H_{27}O_{15}$. Both compounds were tentatively suggested to be kaempferol 3-sambubioside or isomers[29]. The fragmented ion at 285 m/z arose from the neutral loss of 294 Da.

Peaks 23, 45, and 61 which eluted at 19.71, 22.78, and 25.59 min had the same deprotonated molecule at *m/z* 593. These compounds were proposed as kaempferol 3-rutinoside or isomers previously described in green beans[8,14,15,30]. In all cases, the ionization conditions of ESI-TOF analyser generated the product ion

at *m/z* 285. It can be explained by the release of kaempferol, which arose from the loss of 308 Da (rutinoside) (see **Figure 3B**).

By the use of The Smart Formula Editor; two compounds at (Rt 21.76 and 22.46 min) with the identical molecular formula $C_{21}H_{19}O_{11}$, were detected at *m/z* 447. Based on the data acquired from ESI-TOF-MS and the appearance of the aglycone at *m/z* 301 (quercetin), these compounds have been assigned as quercetin rhamnoside isomers[22].

Peaks 38 and 52 at (*Rt* 21.96 and 23.55 min) were tentatively proposed as quercetin rutinoside (rutin) or isomers. This compound has previously been described in green beans [8,15,25,30]. Moreover, the process showed the fragmented ion at *m/z* 301, which may indicate quercetin, this being explained by the neutral loss of rutinoside moiety $[M-H-308]^-$.

Peak 47 at *m/z* 725.1940 with molecular formula $C_{32}H_{37}O_{19}$ has been suggested to be kaempferol 3-(2G-xylosylrutinoside), based on the MS data and literature[8,15]. Two different fragment ions were identified in the spectral mass: fragment ion at *m/z* 593, this representing a loss of xylose moiety $[M-H-132]^-$, and the loss of glucose moiety resulting in the fragment ion at *m/z* 285 (kaempferol).

Peak 51 (Rt 23.48 min) was tentatively assigned as ombuin-3-ß-rutinoside, given that this compound has been found in *Lathyrus davidii* (Fabaceae)[31].

Peaks 53, 56, and 58 gave the same *m/z* 463 and the identical molecular formula $C_{21}H_{19}O_{12}$ yielded the fragmented ion at *m/z* 301 (quercetin) by losing a hexose moiety $[M-H-162]^-$. These compounds were proposed as quercetin 3-galactoside or quercetin 3-glucoside, which has been reported in beans and other Fabaceae[32,33].

The spectra generated by peak 55 gave deprotonated molecules at *m/z* 477 with the molecular formula $C_{21}H_{17}O_{13}$. This compound was proposed as quercetin 3-O-

glucuronide (see **Figure 3C**). Its mass spectra showed a fragment ion at *m/z* 301(quercetin) due to the neutral loss of glucuronic acid [M-H-176]$^-$. Indeed, its presence has been reported in green beans[8,14,15,30]. Nonetheless, this compound has been reported to be a potential lipid peroxidation inhibitor[15].

Peak 62, with experimental *m/z* 623 and molecular formula ($C_{28}H_{31}O_{16}$) generated by TOF analyser, was assigned as isorhamnetol 3-O-rutinoside[34] (**Figure 4f**). This compound showed the loss of rutinoside moiety (308 Da) and the appearance of the ion at *m/z* 315 in its MS spectrum.

Peak 63 at the retention time 26.56 min was tentatively assigned to be kaempferol 3-O-ß-glucuronide (**Figure 3D**). This compound was proposed on the basis of the acceptable MS data and fragmentation pattern by losing the glucuronic acid [M-H-176]$^-$ and giving rise to a fragmented ion at *m/z* 285 (kaempferol). Moreover, this compound has been previously discussed in green beans elsewhere[8,14,15].

Peak 67 gave a deprotonated molecule at *m/z* 447. The spectral mass showed a fragment ion at *m/z* 285 (kaempferol), and the neutral loss of glucose moiety [M-H-162]$^-$ (**Figure 3E**). This compound was proposed as kaempferol 3-O-glucoside[16,18].

In the same manner, peak 69 at *m/z* 491 was proposed as isorhamnetin 3-glucuronide, which has demonstrated a characteristic fragmentation involving cleavage of the glucuronic acid [M-H-176]$^-$ producing the ion at *m/z* 315, which indicates isorhamnetin[35].

Flavone derivatives

In recent years, scientific and public interest in flavones has grown enormously due to their presumed beneficial effects against several diseases such as atherosclerosis, osteoporosis, diabetes mellitus, and certain cancers[36]. This is the

first work available to report flavone glycosides in green beans. The MS spectra generated by the peaks 30, 34, and 44 displayed the identical $[M-H]^-$ deprotonated ion at *m/z* 563. The molecular formula suggested for these compounds was $C_{26}H_{27}O_{14}$. These compounds have demonstrated the same fragmentation pattern due to the loss of (hexose + pentose) moieties $[M-H-294]^-$ and the appearance of apigenin at *m/z* 269. Thus, these compounds have been suggested to be schaftoside isomers[37] (**Figure 4g**).

Peak 65 gave a deprotonated molecule at *m/z* 665 with the molecular formula $C_{29}H_{29}O_{18}$. This compound has been suggested as luteolin-O-6´-malonyl-apiosylglycoside.The presence of luteolin aglycone, has already been described in green beans[38,39]. Moreover, this proposed compound was previously discussed in cucumber[20].

Peak 70 was proposed as apigenin 7-O-glucoside. This compound generated a deprotonated molecule at *m/z* 431 and presented a fragment at *m/z* 269 corresponding to apigenin[40].

Peak 72 shows an *m/z* at 461 with the molecular formula $C_{22}H_{21}O_{11}$. Relying on the acceptable data obtained by TOF-MS and the literature on Fabaceae[41], this compound was tentatively proposed as chrysoeriol 7-glucoside.

Flavanone derivatives

Six compounds belonging to flavanone derivatives were characterized in different varieties. Among these, peaks 25 and 71 were assigned as hesperetin 7-glucoside and naringenin 7-glucoside, respectively, which have been previously reported[42].

Peaks 37 and 50 showed the same deprotonated ions at *m/z* 593 with the molecular formula $C_{28}H_{33}O_{14}$. Both mass spectra showed the same fragment ion at *m/z* 285. These compounds could be characterized as isosakuranetin 7-rutinoside or isosakuranetin 7-neohesperidoside or their isomers.

Peaks 64 and 66 had also shown the same deprotonated ions at *m/z* 609 with identical molecular formula $C_{28}H_{33}O_{15}$. Both isomers generated the same fragment ion at *m/z* 301 corresponding to hesperetin. The resulting loss of 308 Da may represent rutinoside or neohesperidoside. These compounds could be characterized as hesperetin 7-rutinoside or hesperetin 7-neohesperidoside or their isomers.

Flavanol derivatives

The spectra generated for peak 28 gave deprotonated molecules at m/z 711 with the molecular formula C39H35O13, which could be attributed to (+)-Catechin-3-O-ß-D-gluco-(2,6-bis-cinnamoyl)-pyranoside. This compound has been previously reported in the Fabaceae[43].

Isoflavonoid derivatives

Isoflavones are a class of secondary plant metabolites, related to flavonoids, which are most commonly produced by Fabaceae, and known for their antioxidant activity. HPLC-ESI-TOF-MS analysis has enabled the detection of eight isoflavones and one isoflavanone.

Thus, peaks 36 and 46 had the same deprotonated molecule at *m/z* 709 with identical molecular formula $C_{32}H_{37}O_{18.}$ These compounds were proposed as biochanin A 7-O-[ß-D-apiofuranosyl-(1→5)-ß-d-apiofuranosyl-(1→6)-ß-D-glucopyranoside] or isomer[44] (**Figure 4h**).

Compounds at retention times 7.6 and 27.27 min were characterized as dalbinol and formononetin 7-O-glucoside respectively. These compounds have been previously reported in Fabaceae[29,45].

Finally, dalpanin (**Figure 4i**) was tentatively suggested for peak 60, which had a deprotonated ion at *m/z* 533. This compound has also been reported in the literature[8].

Lignans derivatives

Regarding to lignans, two compounds which gave identical molecular formulae ($C_{26}H_{33}O_{11}$) were detected. These compounds were proposed as isolariciresinol 9-O-ß-D-glucopyranoside or isomers, due to the loss of glucose moiety [M-H-162]$^-$. Thus, the appearance of the fragment ion at *m/z* 359 could correspond to the isolariciresinol aglycone which could justify this characterization.

Other polar compounds

Kutkoside, in the iridoid class (peak 26), was tentatively characterized in the green bean hydro-methanol extracts according to the TOF-MS data and the literature.

In summary, the high separation capacity from the HPLC together with the sensitivity, mass accuracy, and true isotopic pattern provided by the ESI-TOF-MS, enabled the characterization of 72 phytochemical compounds present in three varieties of green beans (*Phaseolus vulgaris* L.) in less than 35 min. It is also important to highlight that, for 54 of these; our study is the first report available. According to the results, it bears underlining that all the varieties are qualitatively rich in flavonoids, especially the Strike variety. Furthermore, the data compiled could offer a useful approach for improving and updating food-composition tables.

Table 1. Proposed phytochemical compounds in three methanol extracts of green beans

Peak #	Rt	[M-H]- m/z	Molecular Formula	Error (ppm)	mSigma	ISCID (MS) fragments	Assignment	Variety	Reference
Phenolic acids									
Hydrobenzoic acid derivatives									
2	10.89	431.1209	$C_{18}H_{23}O_{12}$	3.1	34.6	137	Primeveroside salicylic acid **I**	P,H	NA
3	11.21	255.0509	$C_{11}H_{11}O_{7}$	0.7	9.2	-	(+)-Piscidic acid	P, H, S	23
4	11.44	329.0856	$C_{14}H_{17}O_{9}$	6.7	0.7	167	Vanillic acid ß-glucoside	P, H, S	22
5	12.06	461.1288	$C_{19}H_{25}O_{13}$	2.7	10.1	-	Sibiricose A3	P, H, S	NA
6	12.23	431.1194	$C_{18}H_{23}O_{12}$	0.2	3.7	137	Primeveroside salicylic acid **II**	S	NA
Hydrocinnamic acid derivatives									
7	15.36	355.1036	$C_{16}H_{19}O_{9}$	0.3	7.5	193	Ferulic acid ß-glucoside **I**	P, H, S	24
10	16.49	325.0937	$C_{15}H_{17}O$ 8	2.4	34.7	163	Coumaroylglucose	P, H, S	24
12	16.61	353.0871	$C_{16}H_{17}O_{9}$	2.1	44.5	191	Chlorogenic acid	H, S	25
13	16.85	341.0860	$C_{15}H_{17}O_{9}$	5.2	11.2	179	Caffeoylglucose	P, H, S	24
16	18.05	355.1024	$C_{16}H_{19}O_{9}$	2.9	18.5	193	Ferulic acid ß-glucoside **II**	P, H	24
Flavonoids									
Flavonol derivatives									
8	15.90	609.1441	$C_{27}H_{29}O_{16}$	3.3	23.7	285	Kaempferol 3,7-dihexoside **I**	P, H	16
9	16.34	887.2480	$C_{38}H_{47}O_{24}$	1.9	24.4	447	Kaempferol -O-trihexosyl-pentoside	P, H	26
11	16.58	639.1202	$C_{27}H_{27}O_{18}$	0.1	30.8	301	Quercetin 3-glucuronide-7-glucoside	P, H, S	NA
14	16.93	771.1984	$C_{33}H_{39}O_{21}$	0.7	17.7	301	Glucopyranosylrutin	P, H, S	NA
17	18.22	595.1669	$C_{27}H_{31}O_{15}$	0.1	31.1	-	Quercetin 3-vicianoside **I**	P, H	27
18	18.32	609.1454	$C_{27}H_{29}O_{16}$	1.1	33.5	-	Kaempferol 3,7-dihexoside **II**	P, H	16
15	17.56	755.2043	$C_{33}H_{39}O_{20}$	0.4	23.1	-	Quercetin 3-(2G-rhamnosylrutinoside) or Kaempferol-3-O-rutinoside-7-O-glucose **I**	S	22,28
19	18.50	755.2028	$C_{33}H_{39}O_{20}$	1.6	3.6	593	Kaempferol-3-O-rutinoside-7-O-glucose **I**	P, H, S	28
20	18.94	755.2042	$C_{33}H_{39}O_{20}$	0.3	37.8	609	Quercetin 3-(2G-rhamnosylrutinoside)	S	22
21	19.02	741.1857	$C_{32}H_{37}O_{20}$	3.6	21.8	301	Quercetin 3-O-xylosylrutinoside **I**	S	8,15
22	19.55	579.1334	$C_{26}H_{27}O_{15}$	3.8	19.3	285	Kaempferol 3-sambubioside **I**	S	29
23	19.71	593.1499	$C_{27}H_{29}O_{15}$	2.3	5.7	285	Kaempferol 3-rutinoside **I**	S	8,14,15,30
27	20.71	579.1329	$C_{26}H_{27}O_{15}$	4.6	12	-	Kaempferol 3-sambubioside **II**	S	29
28	20.86	711.2105	$C_{39}H_{35}O_{13}$	3	48.4	-	(+)-Catechin-3-O-ß-D-gluco(2,6-bis-cinnamoyl)-pyranoside	P, H	43
29	21.00	741.1886	$C_{32}H_{37}O_{20}$	0.3	7.8	301	Quercetin 3-O-xylosylrutinoside **II**	P, H, S	8,15
31	21.09	755.2037	$C_{33}H_{39}O_{20}$	0.4	61.2	-	Quercetin 3-(2G-rhamnosylrutinoside) or Kaempferol-3-O-rutinoside-7-O-glucose **II**	S	22,28
32	21.26	741.1886	$C_{32}H_{37}O_{20}$	0.4	7.6	301	Quercetin 3-O-xylosylrutinoside **III**	P, H, S	8,15
33	21.38	755.2024	$C_{33}H_{39}O_{20}$	2.1	11.8	593, 285	Kaempferol-3-O-rutinoside-7-O-glucose **II**	P, H, S	28
35	21.76	447.0936	$C_{21}H_{19}O_{11}$	0.7	11.9	301	Quercetin rhamnoside **I**	S	22

38	21.96	609.1076	$C_{26}H_{25}O_{17}$	3.5	10.6	301	Quercetin rutinoside (Rutin)	P	8,15,25 30
39	22.18	595.1279	$C_{26}H_{27}O_{16}$	4.3	9.2	301	Quercetin 3-vicianoside **II**	P, H	27
40	22.26	741.1863	$C_{32}H_{37}O_{20}$	2.8	35.4	-	Quercetin 3-O-xylosylrutinoside **IV**	S	8,15
41	22.43	595.1279	$C_{26}H_{27}O_{16}$	4.4	6.9	301	Quercetin 3-vicianoside **III**	H	27
42	22.46	447.0931	$C_{21}H_{19}O_{11}$	0.5	11.6	-	Quercetin rhamnoside **II**	S	22
43	22.52	609.1420	$C_{27}H_{29}O_{16}$	6.8	7	285	Kaempferol 3,7-dihexoside **III**	P, H	16
45	22.78	593.1521	$C_{27}H_{29}O_{15}$	1.5	29.4	285	Kaempferol 3-rutinoside **II**	S	8,14,15,30
47	22.88	725.1940	$C_{32}H_{37}O_{19}$	0.7	5.5	593, 285	Kaempferol 3-(2G-xylosylrutinoside)	P, H, S	8,15
48	22.93	755.2021	$C_{33}H_{39}O_{20}$	2.5	31	285	Kaempferol 3-O-rutinoside-7-O-glucoside **III**	P, H	28
51	23.48	637.1752	$C_{29}H_{33}O_{16}$	3.5	28.8	-	Ombuin-3ß-rutinoside	S	31
52	23.55	609.1456	$C_{27}H_{29}O_{16}$	0.9	2.7	301	Quercetin rutinoside (Rutin) **I**	P, H, S	8,15 25,30
53	24.04	463.0869	$C_{21}H_{19}O_{12}$	2.8	5.9	301	Quercetin 3-galactoside or Quercetin 3-glucoside **I**	P	18,33
54	24.21	579.1330	$C_{26}H_{27}O_{15}$	4.4	5.4	301	Quercetin 3-O-ß-xylopyranosyl-(1→2)-O-α-rhamnopyranoside	P, H, S	32
55	24.27	477.0655	$C_{21}H_{17}O_{13}$	0.4	6	301	Quercetin 3-O-glucuronide	P, H, S	8,14,15,30
56	24.29	463.0858	$C_{21}H_{19}O_{12}$	5.1	9.2	301	Quercetin 3-galactoside or Quercetin 3-glucoside **II**	H	18,33
57	24.37	755.2009	$C_{33}H_{39}O_{20}$	4.1	14.6	-	Quercetin 3-(2G-rhamnosylrutinoside) or Kaempferol-3-O- rutinoside-7-O-glucose **III**	S	22,28
58	24.96	463.0879	$C_{21}H_{19}O_{12}$	0.6	7.6	301	Quercetin 3-galactoside or Quercetin 3-glucoside **II**	S	18,33
61	25.59	593.1502	$C_{27}H_{29}O_{15}$	1.8	8.4	285	Kaempferol 3-rutinoside **III**	P, H, S	14,8,15,30
62	25.96	623.1600	$C_{28}H_{31}O_{16}$	2.8	26.8	315	Isorhamnetol 3-O-rutinoside	P, H, S	46
63	26.56	461.0719	$C_{21}H_{17}O_{12}$	1.4	3.4	285	Kaempferol 3-O-ß-glucuronide	P, H, S	14,15,30,8
67	27.21	447.0929	$C_{21}H_{19}O_{11}$	0.8	12.1	285	Kaempferol 3-O-glucoside	S	16,18
69	27.45	491.0808	$C_{22}H_{19}O_{13}$	4.7	23.5	315	Isorhamnetin 3-glucuronide	S	35
Flavones									
30	21.06	563.1391	$C_{26}H_{27}O_{14}$	2.8	23.6	269	Schaftoside **I**	S	37

34	21.61	563.1390	$C_{26}H_{27}O_{14}$	2.9	6.3	-	Schaftoside **II**	S	37
44	22.58	563.1393	$C_{26}H_{27}O_{14}$	2.4	7.5	269	Schaftoside **III**	S	37
65	27.08	665.1328	$C_{29}H_{29}O_{18}$	4.7	3.2	621, 447	Luteolin-O-6´-malonyl-apiosylglycoside	S	20
70	29.67	431.0986	$C_{21}H_{19}O_{10}$	0.6	11	269	Apigenin 7-O-glucoside	S	40
72	31.11	461.1092	$C_{22}H_{21}O_{11}$	2	3	-	Chrysoeriol 7-glucoside	H, S	41
Dihydroflavonols									
24	20.14	449.1086	$C_{21}H_{21}O_{11}$	0.8	14.1	303	Taxifolin 3-O-rhamnoside	P, H	47
Flavanones									
Isoflavones									
1	7.60	425.1232	$C_{23}H_{21}O_{8}$	2.3	41.2	-	Dalbinol	P, H, S	45
36	21.81	709.1975	$C_{32}H_{37}O_{18}$	1.4	28.5	-	Biochanin A 7-O-[ß-D-apiofuranosyl-(1→5)-ß-d-apiofuranosyl -(1→6)-ß-D-glucopyranoside]	P, H, S	44
46	22.85	709.1982	$C_{32}H_{37}O_{18}$	0.4	26.7	-	Biochanin A 7-O-[ß-D-apiofuranosyl-(1→5)-ß-d-apiofuranosyl -(1→6)-ß-D-glucopyranoside **I**	H, S	44
68	27.27	429.1221	$C_{22}H_{21}O_{9}$	7	36.1	-	Formononetin 7-O-glucoside (Ononin)	P, H, S	29
Isoflavanone									
60	25.51	533.1639	$C_{26}H_{29}O_{12}$	4.8	17.1	-	Dalpanin	H, S	**¡Error! Marcador no definido.**
Lignans									
49	23.27	521.2010	$C_{26}H_{33}O_{11}$	3.5	32.4	359	Isolariciresinol 9-O-ß-D-glucopyranoside **I**	H, S	NA
59	25.46	521.2015	$C_{26}H_{33}O_{11}$	2.6	24	359	Isolariciresinol 9-O-ß-D-glucopyranoside **II**	S	NA
Other polar compounds ***Iridoid***									
26	20.32	511.1460	$C_{23}H_{27}O_{13}$	0.5	43.3	193	Kutkoside	P	43

Notes: ISCID: stands for In-Source Collision Induced Dissociation. P, H, S; refer to green bean varieties analysed: Perona, Helda, and Strike, respectively. NA: Not available. *Rt:* Retention time. I, II: imply other isomers of the same m/z

References

1 Tomás-Barberan, F. A., Ferreres, F., Gil, M. I. and Atta ur R. 2000. Antioxidant phenolic metabolites from fruit and vegetables and changes during postharvest storage and processing. In Studies in Natural Products Chemistry pp. 739-795, Elsevier.

2 Tomás-Barberán, F. A. and Espín de Gea, J. C. 2001. Phenolic compounds and related enzymes as determinants of quality in fruits and vegetables. *Journal of Agricultural and Food Chemistry,* **81,** 853-876.

3 De La Rosa, L. A., Álvarez-Parrilla, E. and González-Aguilar, G. A. 2009. Fruit and Vegetable Phytochemicals: Chemistry, Nutritional Value and Stability Wiley.

4 Johnson, I. T. 2004. New approaches to the role of diet in the prevention of cancers of the alimentary tract. *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis,* **551,** 9-28.

5 Dicko, M. H., Gruppen, H., Traore, A. S., Voragen, A. G. J. and Van Berkel, W. J. H. 2006. Phenolic compounds and related enzymes as determinants of sorghum for food use. *Biotechnology and Molecular Biology Reviews,* **1,** 20-37.

6 Prohens, J. and Nuez, F. 2007. *Vegetables II: Fabaceae, Liliaceae, Solanaceae, and Umbelliferae* Springer.

7 Nunes, M. C. N. 2009. *Color Atlas of Postharvest Quality of Fruits and Vegetables* Wiley.

8 Price, K. R., Colquhoun, I. J., Barnes, K. A. and Rhodes, M. J. C. 1998. Composition and content of flavonol glycosides in green beans and their fate during processing. *Journal of Agricultural and Food Chemistry,* **46,** 4898-4903.

9 Pari, L. and Venkateswaran, S. 2003. Effect of an aqueous extract of Phaseolus vulgaris on the properties of tail tendon collagen of rats with streptozotocin-induced diabetes. *Brazilian Journal of Medical and Biological Research,* **36,** 861-870.

10 Roman-Ramos, R., Flores-Saenz, J. L. and Alarcon-Aguilar, F. J. 1995. Anti-hyperglycemic effect of some edible plants. *Journal of Ethnopharmacology,* **48,** 25-32.

11 Takezaki, T., Gao, C. M., Wu, J. Z., Ding, J. H., Liu, Y. T., Zhang, Y., Li, S. P., Su, P., Liu, T. K. and Tajima, K. 2001. Dietary protective and risk factors for esophageal and stomach cancers in a low-epidemic area for stomach cancer in Jiangsu Province, China: comparison with those in a high-epidemic area. *Japanese Journal of Cancer Research,* **92,** 1157-1165.

12 Boué, S. M., Wiese, T. E., Nehls, S., Burow, M. E., Elliott, S., Carter-Wientjes, C. H., Shih, McLachlan, J. A. and Cleveland, T. E. 2003. Evaluation of the Estrogenic Effects of Legume Extracts Containing Phytoestrogens. *Journal of Agricultural and Food Chemistry,* **51,** 2193-2199.

13 Marin, F. R., Perez-Alvarez, J. A. and Soler-Rivas, C. 2005. Isoflavones as functional food components. In *Studies in Natural Products Chemistry*, Vol 32. Rahman A-U (Ed.). Elsevier, 1177-1207.

14 Hempel, J. and Böhm, H. 1996. Quality and Quantity of Prevailing Flavonoid Glycosides of Yellow and Green French Beans (Phaseolus vulgaris L.). *Journal of Agricultural and Food Chemistry,* **44,** 2114-2116.

15 Plumb, G. W., Price, K. R. and Williamson, G. 1999. Antioxidant properties of flavonol glycosides from green beans. *Redox Report,* **4,** 123-127.

16 Bolling, B. W., Chen, E. and Parkin, K. L. 2007. Quinone reductase inducing and antioxidant activities of aqueous isolates of green bean (Phaseolus vulgaris L.). *Food Research International,* **40,** 182-190.

17 Ewald, C., Fjelkner-Modig, S., Johansson, K., Sjöholm, I. and Åkesson, Br. 1999. Effect of processing on major flavonoids in processed onions, green beans, and peas. *Food Chemistry,* **64,** 231-235.

18 Escarpa, A. and González M. 2000. Identification and quantitation of phenolics from green beans by high-performance liquid chromatography. *Chromatographia,* **52,** 33-38.

19 Benotti, M. J., Ferguson, P. L., Rieger, R. A., Iden, C. R., Heine, C. E. and Brownawell, B. J. 2003. HPLC/TOF-MS: An Alternative to LC/MS/MS for Sensitive and Selective Determination of Polar Organic Contaminants in the Aquatic Environment. In Liquid Chromatography/Mass Spectrometry, MS/MS and Time of Flight MS. *American Chemical Society*, 109-127.

20 Abu-Reidah, I. M., Arráez-Román, D., Quirantes-Piné, R., Fernández-Arroyo, S., Segura-Carretero, A. and Fernández-Gutiérrez, A. 2012. HPLC-ESI-Q-TOF-MS for a comprehensive characterization of bioactive phenolic compounds in Cucumber whole fruit extract. *Food Research International,* **46,** 108-117.

21 Gross, J. H. 2011. Mass Spectrometry (2nd Edn.). Springer-Verlag Berlin Heidelberg; Berlin, Heidelberg; 100-106.

22 Regos I, Urbanella A, Treutter D. 2009. Identification and Quantification of Phenolic Compounds from the Forage Legume Sainfoin (Onobrychis viciifolia). *Journal of Agricultural and Food Chemistry,* **57,** 5843-5852.

23 Ae, N., Otani, T. and Arihara, J. 1995. Effects of specific compounds exuded from roots on phosphorus nutrition, Genetic Manipulation of Crop Plants to Enhance Integrated Nutrient Management in Cropping Systems-1. Phosphorus. *FAO-ICRISAT Expert Consultancy Workshop*: India; 117-128.

24 Winter, M. and Herrmann, K. 1986. Esters and glucosides of hydroxycinnamic acids in vegetables. *Journal of Agricultural and Food Chemistry,* **34,** 616-620.

25 Escarpa, A. and González, M. C. 2001. Approach to the content of total extractable phenolic compounds from different food samples by comparison of chromatographic and spectrophotometric methods. *Analytica Chimica Acta,* **427,** 119–127.

26 Kite, G. C., Veitch, N. C., Boalch, M. E., Lewis, G. P., Leon, C. J. and Simmonds, M. S. J. 2009. Flavonol tetraglycosides from fruits of Styphnolobium japonicum (Leguminosae) and the authentication of Fructus Sophorae and Flos Sophorae. *Phytochemistry,* **70,** 785-794.

27 Takemura, M., Nishida, R., Mori, N. and Kuwahara, Y. 2002. Acylated flavonol glycosides as probing stimulants of a bean aphid, Megoura crassicauda, from Vicia angustifolia. *Phytochemistry,* **61,** 135-140.

28 Kite, G. C., Rowe, E. R., Lewis, G. P. and Veitch, N. C. 2011. Acylated flavonol tri- and tetraglycosides in the flavonoid metabolome of Cladrastis kentukea (Leguminosae). *Phytochemistry,* **72,** 372-384.

29 Cui, B., Nakamura, M., Kinjo, J. and Nohara, T. 1993. Constituents of leguminous plants. XXXV. Chemical constituents of Astragali Semen. *Chemical and Pharmaceutical Bulletin,* **41,** 178-182.

30 Berger, M., Küchler, T., Maaßen, A., Busch-Stockfisch, M. and Steinhart, H. 2007. Correlations of ingredients with sensory attributes in green beans and peas under different storage conditions. *Food Chemistry,* **103,** 875-884.

31 Park, S. Y., Kim, J. S., Lee, S. Y., Bae, K. H. and Kang, S. S. 2008. Chemical constituents of Lathyrusdavidii. *Journal of Natural Products,* **14,** 281-288.

32 Soicke, H., Görler, K. and Waring, H. 1990. Flavonol glycosides from Moghaniafaginea. *Planta Medicina,* **56,** 410-412.

33 Prati, S., Baravelli, V., Fabbri, D., Schwarzinger, C., Brandolini, V., Maietti, A., Tedeschi, P., Benvenuti, S., Macchia, M. and Marotti, I. 2007. Composition and content of seed flavonoids in forage and grain legume crops. *Journal of Separation Science,* **30,** 491-501.

34 Zallocchi, E. M. and Pomilio, A. B. 1994. Flavonoids from Argentinian species of Phaseolus. *Fitoterapia,* **65,** 467.

35 Sobolev, V. S., Sy, A. A. and Gloer, J. B. 2008. Spermidine and Flavonoid Conjugates from Peanut (Arachis hypogaea) Flowers. *Journal of Agricultural and Food Chemistry,* **56,** 2960-2969.

36 Cermak, R. 2008. Effect of dietary flavonoids on pathways involved in drug metabolism. *Expert Opinion on Drug Metabolism & Toxicology,* **4,** 17-35.

37 Kitagawa, I., Chen, W.-Z., Hori, K., Kobayashi, M. and Ren, J. 1999. Chemical Studies of Chinese Licorice-Roots. II. Five New Flavonoid Constituents from the Roots of Glycyrrhiza aspera PALL. Collected in Xinjiang. *Chemical and Pharmaceutical Bulletin,* **46,** 1511-1517.

38 Miean, K. H. and Mohamed, S. 2001. Flavonoid (Myricetin, Quercetin, Kaempferol, Luteolin, and Apigenin) Content of Edible Tropical Plants. *Journal of Agricultural and Food Chemistry,* **49,** 3106-3112.

39 Vasilopoulou, E., Georga, K., Joergensen, M. B., Naska, A. and Trichopoulou, A. 2005. The Antioxidant Properties of Greek Foods and the Flavonoid Content of the Mediterranean Menu. *Current Medicinal Chemistry - Immunology Endocrine & Metabolic Agents,* **5,** 33-45.

40 Kaneta, M., Hikichi, H., Endo, S., and Sugiyama, N. 1980. Identification of flavones in sixteen Leguminosae species. *Agricultural and Biological Chemistry,* **44,** 1407-1408.

41 Niassy, B., Lobstein, A., Um, B. H., Anton, R., and Kone, M. E. K. 2005. Flavonoids from Tephrosiadeflexa and T. albifoliolis. *Biochemical Systematics and Ecology,* **33,** 309-312.

42 De Rijke, E., Zappey, H., Ariese, F., Gooijer, C. and Brinkman, U. A. T. 2004. Flavonoids in Leguminosae: Analysis of extracts of T. pratense L., T. dubium L., T. repens L., and L. corniculatus L. leaves using liquid chromatography with UV, mass spectrometric and fluorescence detection. *Analytical and Bioanalytical Chemistry,* **378,** 995-1006.

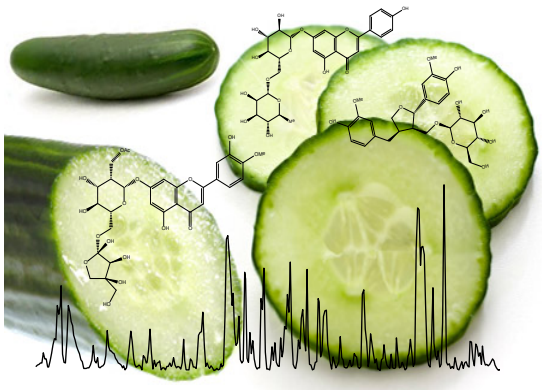
43 Lokvam, J., Coley, P. D., and Kursar, T. A. 2004. Cinnamoyl glucosides of catechin and dimeric procyanidins from young leaves of Inga umbellifera (Fabaceae). *Phytochemistry,* **65,** 351-358.

44 Silva, V. C. D., Carvalho, M. G. D., Borba, H. R. and Silva, S. L. D. C. 2008. Anthelmintic activity of flavonoids isolated from roots of Andiraanthelmia (Leguminosae). *Revista Brasileira de Farmacognosia,* **18,** 573-576.

45 Kim, Y. S., Ryu, Y. B., Curtis-Long, M. J., Yuk, H. J., Cho, J. K., Kim, J. Y., Kim, D. K., Lee, W. S., and Park, K. H. 2011. Flavanones and rotenoids from the roots of Amorpha fruticosa L. that inhibit bacterial neuraminidase. *Food Chemical Toxicology,* **49,** 1849-1856.

46 Zallocchi, E.M., and Pomilio, A. B. 1994. Flavonoids from Argentinian species of Phaseolus. *Fitoterapia,* **65,** 467.

47 Silva, V. C. D., Carvalho, M. G. D., and Silva, S. L. D. C. 2007. Chemical constituents from roots of Andiraanthelmia (Leguminosae). *Revista Latinoamericana de Quimica* **35,** 13-19.



Chapter 2

HPLC-ESI-Q-TOF-MS for a comprehensive characterization of bioactive phenolic compounds in cucumber whole fruit extract

This work has been published in:
Food Research International

Food Research International 46 (2012) 108–117

Contents lists available at SciVerse ScienceDirect

Food Research International

journal homepage: www.elsevier.com/locate/foodres

HPLC–ESI-Q-TOF-MS for a comprehensive characterization of bioactive phenolic compounds in cucumber whole fruit extract

Ibrahim M. Abu-Reidah, David Arráez-Román, Rosa Quirantes-Piné, Salvador Fernández-Arroyo, Antonio Segura-Carretero, Alberto Fernández-Gutiérrez

HPLC-ESI-Q-TOF-MS for a comprehensive characterization of bioactive phenolic compounds in Cucumber (*Cucumis Sativus L.*) whole fruit extract

Abstract

Cucumber (*Cucumis sativus L.*) belongs to Cucurbitaceae family and it is widely cultivated and consumed in many varieties for its edible fruit (Cucumber) either in fresh or processed (pickled) forms. It is an important part of the traditional Mediterranean diet since antiquity due to its nutritional and medicinal properties. *C. sativus* is a rich source of polyphenols and it is used not only as food, but also in folk medicine, health-care and cosmetology. Therefore, due to the important role that cucumber plays in those fields, we have developed a simple and rapid method to characterize the polyphenolic compounds in an extract of *Cucumis sativus L.* using Reversed-Phase High-Performance Liquid Chromatography coupled to Electrospray Ionization Quadropole Time-of-Flight Mass Spectrometry (RP-HPLC-ESI-Q-TOF-MS). The accuracy of mass data generated by Q-TOF-MS together with the fragmentation pattern of the full scan run of MS/MS analysis have been an useful tool to tentatively characterize 73 phenolic compounds in the extract of the studied matrix. Most of these compounds are being reported here for the first time in Cucumber, highlighting the importance of Cucurbitaceae family as a source of natural bioactive antioxidants.

Keywords: *Cucurbitaceae, Cucumis sativus L., Phenolic Compounds, High-Performance Liquid Chromatography, Quadropole time-of-flight mass spectrometry.*

Introduction

The increased consumption of fruits and vegetables is an effective strategy to increase antioxidant intake and to reduce oxidative stress and it may lead to lower risk of developing chronic diseases, such as cancer and cardiovascular disease[1]. Nowadays, there is a rising interest in antioxidants from natural sources present in the diet because *in vitro* evidences have suggested that they can affect numerous cellular processes like gene expression, apoptosis, platelet aggregation and intercellular signalling, which can have anti-carcinogenic and anti-atherogenic implications[2].

Phenolic compounds are secondary plant metabolites, which play a decisive role in the sensory quality of fruits, vegetables and other plants. These compounds, ubiquitous in the plant kingdom, are one of the most widely occurring groups of phytochemicals; moreover they are of considerable interest and have received a rising attention in recent years due to their bioactivity. Phenolic compounds comprise a wide variety of molecules with a polyphenol structure (i.e. several hydroxyl groups on aromatic rings), but also molecules with one phenol ring, such as phenolic acids and phenolic alcohols. Polyphenols are divided into several classes according to the number of phenol rings that they contain and to the structural elements that bind these rings. The main groups of phenolic compounds are: flavonoids, phenolic acids, tannins (hydrolysable and condensed), stilbenes and lignans.

Cucumber (*Cucumis sativus L.*) belongs to the Cucurbitaceae family, commonly known as Cucurbits or gourd family[3,4]. Among the 52 species reported in Cucumis sub-family, *Cucumis sativus L.* is one of the most

important species extensively cultivated in many varieties for its edible fruit, cucumber, which is a very important part of traditional Mediterranean diet since antiquity. Cucumber whole fruit is widely consumed mainly fresh in salads or fermented (pickles), moreover the consumption of the skin peel as well as the seeds maximises its nutritional benefits[5]. The leaves, stems, roots, seeds and peels of *Cucumis sativus L.* are reportedly used in folk medicine as anti-diarrheal, de-toxicant, anti-gonorrheal, anti-inflammatory, hypertension reducing agent, diabetes mellitus and serum lipids regulator, antioxidant, and analgesic[3-8].

Although cucumbers are a particularly rich source of phenolic compounds[9,10], it can be noted that the phenolic constituents of *Cucumis sativus* have not been extensively investigated[11]. Most studies concerning *Cucumis sativus L.* had reported the foliar, stems, roots, flowers and the general plant phenolics[7,12-18].

Additionally, it was noticed that several analytical techniques had been used such as RP-HPLC-UV and GC-MS to identify flavonoids and other phenolic compounds in cucumber fruit[19-22].

A comprehensive characterization of phenolic compounds using advanced and powerful techniques sounds to be crucial. For this reason, suitable methods need to be established for the characterization of phenolic compounds in vegetable matrices[13].

Thus, the aim of the present work was the development of a comprehensive and reliable method for the characterization of bioactive phenolic compounds in cucumber *(Cucumis sativus L.)* cultivated in Spain by using High-

Performance Liquid Chromatography coupled to Electrospray Ionization Quadropole-Time-of-Flight Mass Spectrometry (HPLC-ESI-Q-TOF-MS) as a powerful analytical technique. Furthermore, the obtained results could contribute to a detailed knowledge of the phenolic compounds consumed which may contribute to a better understanding of their influence on biological, nutritional and medicinal properties.

Materials and methods

Plant Material

Cucumber sample was collected from a commercial market. After purchasing, the samples were transported directly to the lab, washed with distilled water and stored at 4° C until use.

Sample treatment

Fresh cucumber samples were crushed and frozen at -25°C. Then, the samples are placed on the lyophilizer (Christ Alpha 1-2 LD Freeze dryer, Shropshire, UK) shelf, which was pre-cooled to -50 ºC for 1 hour at 1mbar. Afterward, 0.5 gram of lyophilized Cucumber was extracted using 16 mL of 80:20 (v/v) methanol/H_2O and sonicated for 30 min at room temperature (22°C). Then, the mixture was centrifuged for 15 min at 4000 *rpm* and the supernatant was collected in a rounded bottom flask. Afterwards, the solvent was evaporated by using a rotary evaporator under vacuum at 40ºC and the dry residue was resolved in 0.5 mL of 80:20 (v/v) methanol/H_2O. Finally, the extract was filtered through a 0.2 µm syringe filter and stored at −20°C until analysis[23].

Chemicals and Reagents

HPLC-grade acetonitrile and methanol were purchased from Labscan (Dublin, Ireland). Acetic acid of analytical grade (assay>99.5 %) was purchased from Fluka (Switzerland). Water was purified by using a Milli-Q system (Millipore, Bedford, USA). Other unmarked reagents were of analytical grade.

Separation by HPLC

Separation of phenolic compounds from cucumber extract was performed on an Agilent 1200 series Rapid Resolution LC (Agilent Technologies, CA, USA) consisting of a vacuum degasser, an auto-sampler and a binary pump. This instrument was equipped with an Agilent Zorbax C18 column (4.6 x 150 mm, 1.8µm) from Agilent Technologies. Acidified water (0.5% acetic acid, v/v) and acetonitrile were used as mobile phases A and B respectively. Gradient was programmed as the follows: 0 min, 0% B; 10 min, 20% B; 15 min, 30% B; 20 min, 50% B; 25 min, 75% B; 30 min, 100% B; 32 min 0% B, and finally, the initial conditions was held for 8 min as a re-equilibration step. The flow rate was set at 0.80 mL/min throughout the gradient. The effluent from the HPLC column was splitted using a T-type splitter before being introduced into the mass spectrometer (split=1:3). Thus in this study the flow which arrived into the ESI-Q-TOF-MS detector was 0.2 mL/min. The column temperature was maintained at 25° C and the injection volume was 10 µL.

ESI-Q-TOF-MS Analysis

The HPLC system was coupled to Quadrupole-Time-of-Flight (micrOTOF-Q^{TM}, Bruker Daltonik GmbH, Bremen, Germany), an orthogonal accelerated Q-TOF mass spectrometer, equipped with an electrospray ionization source (ESI).

Parameters for analysis were set using negative ion mode with spectra acquired over a mass range from m/z 50 to 1100. The optimum values of the ESI-MS parameters were: capillary voltage, +4.0 kV; drying gas temperature, 190° C; drying gas flow, 9.0 L/min; nubilizing gas pressure, 29 Psi; collision RF, 150 Vpp; transfer time 70 µs, and pre-pulse storage, 5 µs. Moreover, automatic MS/MS experiments were performed adjusting the collision energy values as follows: m/z 100, 20 eV; m/z 500, 30 eV; m/z 1000, 35 eV, and using nitrogen as collision gas.The MS data were processed through Data Analysis 4.0 software (Bruker Daltonics, Bermen, Germany) which provided a list of possible elemental formulas by using the Generate Molecular Formula ™ editor. The editor uses a CHNO algorithm, which provides standard functionalities such as minimum/maximum elemental range, electron configuration, and ring-plus double bonds equivalents, as well as a sophisticated comparison of the theoretical with the measured isotope pattern (mSigma value) for increasing the confidence in the suggested molecular formula. The widely accepted accuracy for confirmation of elemental compositions has been established to be 5 ppm[24].

During the HPLC method development, an external instrument calibration was performed using a Cole Palmer syringe pump (Vernon Hills, Illinois, USA) directly connected to the interface, passing a solution of sodium formate cluster containing 5 mM sodium hydroxide and 0.2 % formic acid in water/isopropanol 1:1 (v:v). Using this method, an exact calibration curve based on numerous cluster masses each differing by 68 Da (NaCHO2) was obtained. Due to the compensation of temperature drift in the Q-TOF, this

external calibration provided accurate mass values for a complete run without the need for a dual sprayer set up for internal mass calibration.

Results and Discussion

Identification process

An overview of all the characterized compounds in the cucumber extract by HPLC-ESI-Q-TOF-MS using the negative mode is given in **Table 1**. These compounds are summarized along with their retention time, *m/z* experimental and calculated, tolerance, error (ppm), mSigma value, molecular formula generated by the software for the detected deprotonated molecule, classification order in the list of possibilities (sorted with respect to sigma value), MS/MS fragments and the proposed assignment.

In the present work 73 phenolic compounds have been tentatively identified in *Cucumis sativus L.* by using the combination of MS and MS/MS data and the information previously reported in the literature. Thus, a comprehensive study of most of the characterized compounds is described below according their families.

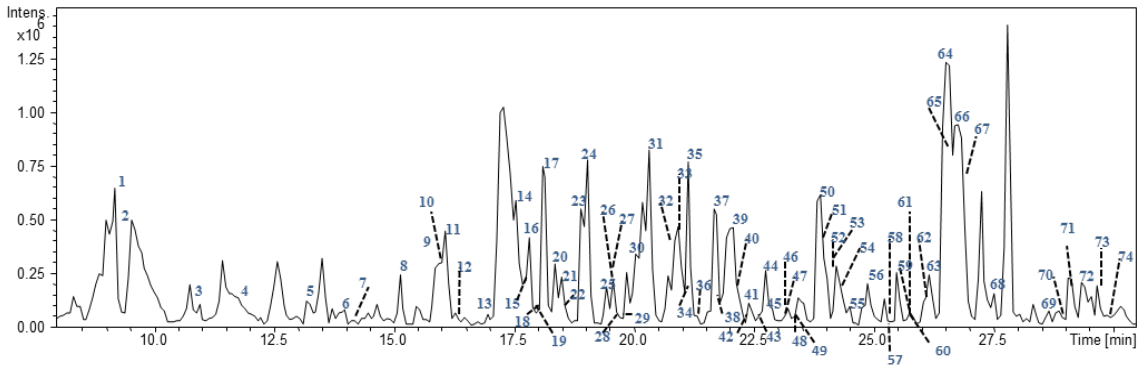


Figure 1. Base peak chromatogram of cucumber extract

In the first 5 min of the chromatogram analysis time, organic acids and sugars were detected; however, these compounds have not been discussed in this

work since our study was concentrated upon phenolic compounds. **Figure 1** shows the base peak chromatogram (BPC) of the cucumber extract in negative ionization mode. It is important to note that many of the proposed compounds in this work have been reported in cucumber for the first time. **Figure 2** shows the structures of several tentatively identified compounds in the studied extract.

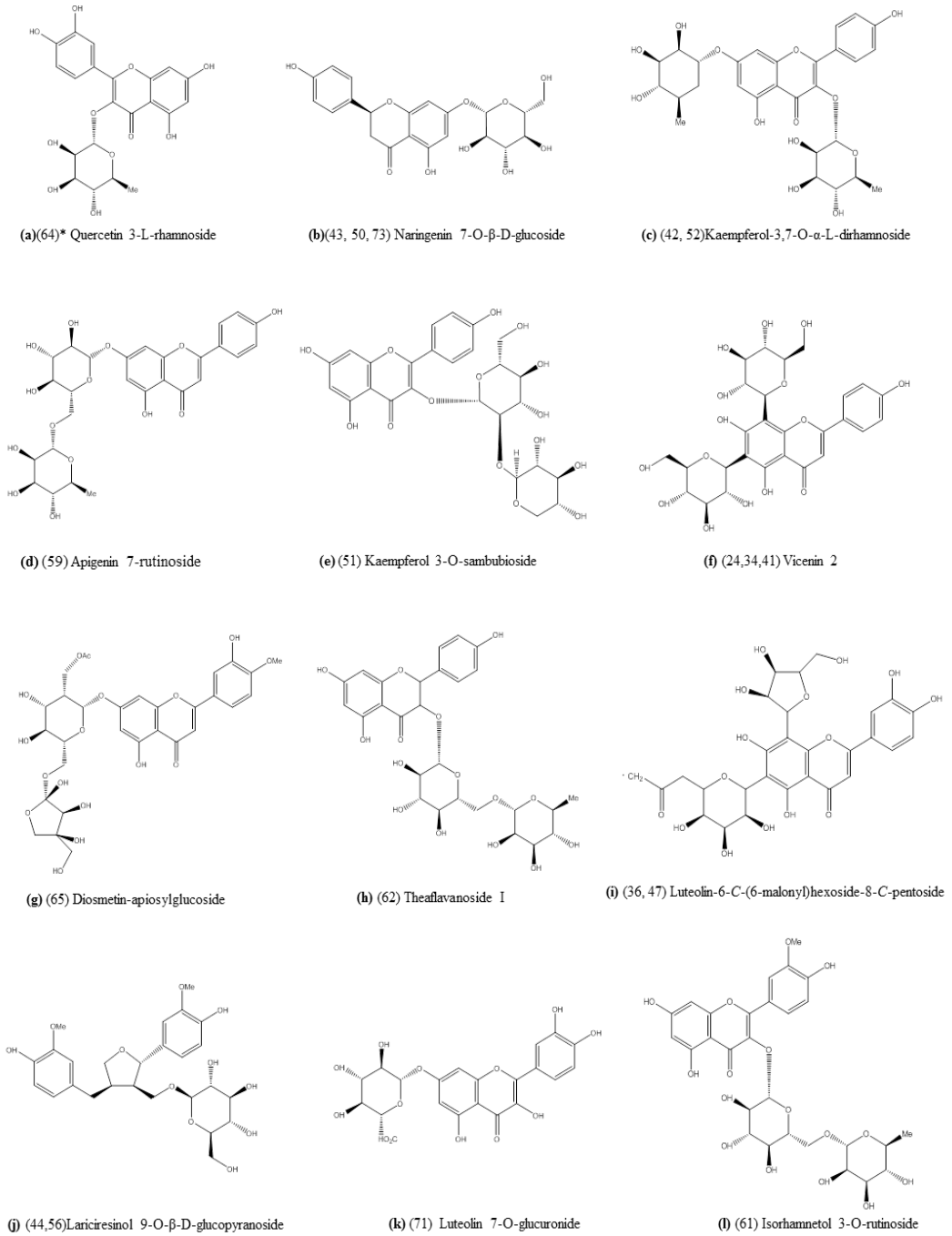


Figure 2. Chemical configurations of several proposed phenolic compounds in *Cucumis sativus L.*

Simple phenolic acid derivatives

Several conjugated and glycosylated forms of caffeic, *p*-coumaric and ferulic acids have been detected, in accordance with previous data cited in literature[14-17,25-27].

Thus, compounds **1, 2, 4** and **7** with m/z 299.0772 were characterized as salicylic acid O-glycoside according with the MS data and fragmentation pattern by the neutral loss of a glycoside moiety (162 Da), which has also been reported in cucumber leaves[26]. The fragment ion at m/z 137.0236 was already reported elsewhere[28-30].

The ion at m/z 329.0878, which appeared at the retention times 11.19 and 13.2 min, were identified as vanillic acid glycosides, due to the loss of a glucose moiety and from the yielded ion at m/z 167.035. Moreover, vanillic acid was already reported in cucumber by Herrmann et al.[31].

Compounds **6** and **10** yielded a fragment ion at 179.0345 corresponding to caffeic acid, which has been previously described in cucumber[16,25,31-35], by losing a glucose moiety (162 Da). These fragments belong to m/z 341.0878 precursor ions. Thus, these compounds have been proposed as caffeoyl glucose[15].

The detected ions at m/z 355.1034, which correspond to peaks **8, 15, 16, 29,** showed a loss of glucose moiety resulting in a fragment at m/z 193.0566, which corresponds to ferulic acid. Therefore, they were suggested as feruloyl glucose. This compound has been previously reported by Bellés et al. in *Cucumis melo* and *Cucumis sativus* leaves[14,15].

Chlorogenic acid has been identified in peaks **9** and **18**. It has been proposed according to its MS and MS/MS fragmentation pattern and appearance of a

fragment at m/z 191.0563 (quinic acid) and due to the loss of caffeic acid. This compound has been also previously mentioned by Maksimovic et al. in cucumber leaves and roots[14-16].

Compound **11** (R_t 16.10 min) exhibited a deprotonated molecule at m/z 325.0932 and a MS/MS fragment at m/z 163.0390 due to the neutral loss of a sugar moiety $[M-H-162]^-$. In keeping with the findings of Billett et al. [27] and the observed fragmentation pattern this compound has been suggested as *p*-coumaric acid-*O*-glycoside. *p*-coumaric acid and its isomer were proposed for peaks **38** and **45**, which has been reported previously[12,14-16,18,25,28,34-36].

Peaks **12, 17,** and **22** have been tentatively proposed as sinapic acid-hexosides that belong to hydroxycinnamic acid category. These compounds have been found in cucumber here for the first time according with the MS and MS/MS data obtained.

Flavonoid glycosides

Flavonoids have a basic structure consisting of two aromatic benzene rings separated by an oxygenated heterocyclic ring. Compounds belonging to various flavonoid classes, such as flavones, flavanones and flavonols and others have been detected in the cucumber sample analysed in this study. *Cucumis sativus L.* contains a good number of C-glycosides and O-glycosides flavonoids isomers as **Table 1** shows. The presence of C-glycosides flavonoids has been described in many Cucurbitaceae species[19] and they give a different fragmentation pattern than O-glucosyl flavones. A series of deprotonated ions resulted at m/z [M-H-18], [M-H-60], [M-H-90], [M-H-120], [M-H-180], and [M-

H-210] due to the cleavage of the C-hexosyl and C-pentosyl rings, have been reported as characteristic fragments of C-glycosides flavonoids[35].

Thus, peak **14** (R_t 17.57 min) has the precursor ion at *m/z* 609.1486. The ESI-MS spectra of this peak gave a MS/MS with a characteristic fragmentation pattern of 6,8-di-*C*-hexosyl flavones, in which it was shown the losses of ions at [M-H-18] (-H_2O), [M-H-60], [M-H-120], in addition to the fragment ions at *m/z* 369.0614 (aglycone +83) and 399.0715 (aglycone +113) which represent the aglycone plus the residues of the sugars that remained linked to it and therefore indicated the aglycone as a trihydroxyflavone (luteolin)[37]. On the basis of these MS and MS/MS spectrum data peak **14** was tentatively identified as luteolin-6, 8-di-C-hexoside.

The compound at m/z 755.2063 (R_t 16.96 min) with molecular formula $C_{33}H_{40}O_{20}$ showed a neutral loss of glucose moiety [M-H-162]$^-$ to form the fragment ion at m/z 593.1489 which corresponds to the characteristic fragment mass of O-glycoside. Relying on the glucose loss and the literature this compound was proposed as saponarin 4'-*O*-glucoside[12,38].

At m/z 447.0939 five peaks with identical molecular formula $C_{21}H_{20}O_{11}$ were detected (R_t 21.12, 21.65, 24.30, 26.93, 28.56 min). These peaks showed the same intense fragment at m/z 285.040 which can be justified by the elimination of a glucose residue (162 Da) from the precursor ion. This fragmentation pattern corresponds to the characteristic fragment mass of O-glycoside. The results are in accordance with the presence of kaempferol-*O*-glucoside, orientin or isoorientin. These compounds have been already mentioned in the bibliography for cucumber and muskmelon[15,18,19, 31,32,34,39].

It was not possible to accurately distinguish between them because of the lack of commercial standards. On the other hand, peak **64** with the same precursor ion at m/z 447.0934, it was suggested to be quercetin 3-L-rhamnoside (**Figure 2a**) on the basis of the presence of a fragment at m/z 301.0398 (quercetin) and the neutral loss of rhamnose moiety (146 Da).

Peaks **43, 50,** and **73** with identical molecular formula ($C_{21}H_{22}O_{10}$) were tentatively proposed as naringenin-*O*-glucosides (**Figure 2b**) and its isomers based on the presence of fragment ion at m/z 271.0617 (naringenin) and the loss of glucose moiety, resulting in an ion at $[M-H-162]^-$. To our knowledge this compound has been reported in Cucumber for the first time.

The ion at m/z 577.1583 exhibited two peaks (R_t 22.47 and 24.21 min) which were suggested to be kaempferol-*O*-dirhamnosides (**Figure 2c**). These compounds shown a loss of two rhamnoside moieties (292 Da) and the fragment ion 285.0416 was showed in the MS/MS spectra.

On the other hand peak **59** (**Figure 2d**) was assigned as apigenin-7-rutinoside relying on the noticeable fragment ion at m/z 269.0446 and the neutral loss of 308 Da which represents a rutinoside moiety.

Five signals have been detected at m/z 579.1368 (R_t 18.91, 19.43, 19.56, 20.05, 23.88). Four of them showed identical fragments (561.1239, 519.1161, 489.1019, 459.0935, 399.0715, 369.0601) indicating the loss of ions at $[M-H-H_2O]^-$, $[M-H-60]^-$, $[M-H-90]^-$, $[M-H-120]^-$, $[M-H-180]^-$ and $[M-H-210]^-$ respectively, characteristic of the cleavage in the sugar moiety. These compounds with the identical molecular formula $C_{26}H_{27}O_{15}$ were proposed to

be luteolin-C-hexosyl-C-pentosides. The fragmentation pattern is illustrated in **Figure 4A**.

Peak **51** with m/z 579.1368 has shown different fragmentation pattern (447.0942 and 285.0408) which resulted from the loss of ions at [M-H-132] and [M-H-132-162]. According to these data, this compound was suggested to be kaempferol-3-*O*-sambubioside (**Figure 2e** and **Figure 4B**). As far as we know, this compound has been reported in cucumber for the first time.

The Smart Formula Editor provided identical molecular formula for peaks **20, 46,** and **49** at m/z 609.1508. The loss of rutinoside $[M-H-308]^-$ and the appearance of the ion at m/z 301.0338 in the MS/MS spectra could corresponds to the aglycone quercetin. The acceptable data of MS with the produced fragment ions of MS/MS have lead to propose the precursor ion at m/z 609.1525 to be quercetin rutinoside (rutin) and its isomers since they showed the same fragmentation pattern. This compound was already mentioned in the literature[16,40].

With m/z 593.1512 ($C_{27}H_{29}O_{15}$), three peaks were detected at retention times 19.04, 21.07 and 22.40 min. These peaks showed the same fragmentation pattern (m/z 575.1331, 533.1260, 503.1202, 473.1078 and 383.2607) which indicated the losses of $[M-H-H_2O]^-$, $[M-H-60]^-$,$[M-H-90]^-$, $[M-H-120]^-$ and $[M-H-180]^-$ respectively.

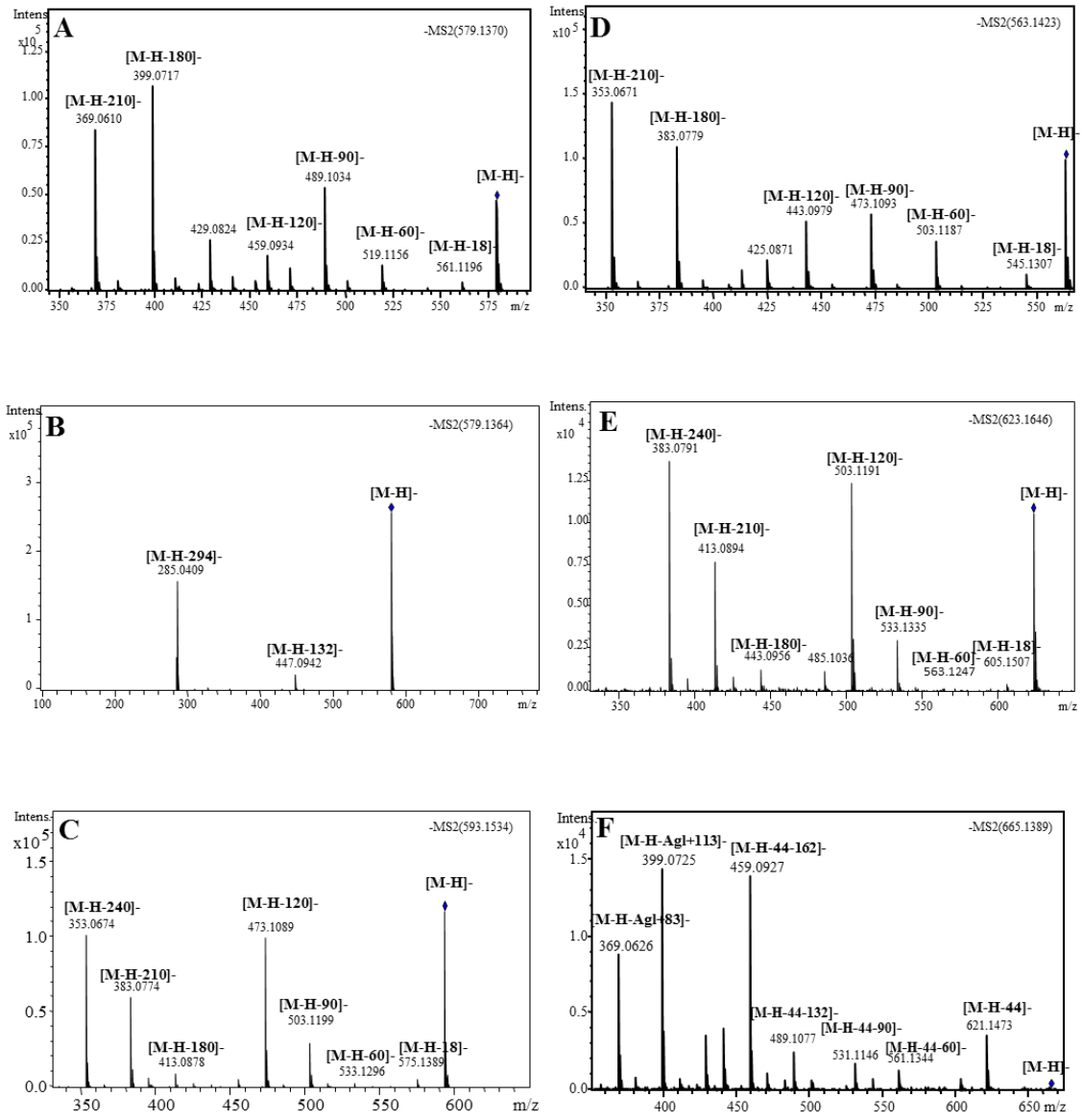


Figure 4.MS/MS spectra of phenolic compounds detected in *Cucumis sativus L.*

The precursor ion has already been mentioned in the bibliography in Cucurbitaceae family[12,19,31,32,41] as isosaponarin, saponarin, vicenin-2 or isovitexin-2"-O-ß-D-glucoside. The fragmentation pattern obtained for the above-mentioned isomers have revealed that they belong to C-flavonoids. Therefore, these compounds were assigned vicenin-2 isomers (**Figure 2f** and **Figure 4C**).

The compound **65** showed a fragment ion at m/z 299.0540; therefore, it was assigned as diosmetin-apiosylglucoside[42] (**Figure 2g**). It is worth noting that this compound has been reported in cucumber for the first time.

Compound **40** has been assigned to luteolin-6-C-rhaminoside-8-C-hexoside based on the MS/MS fragmentation pattern. This compound has shown the C-glycoside flavonoids fragmentation pattern ions[40].

Peak **62** showed a fragment ion at m/z 271.0617, which indicates naringenin, and the neutral loss of two moieties of rhamnose. Thus, this compound was tentatively proposed as theaflavanoside I (**Figure 2h**).

On the other hand, peak **57** (R_t 25.36), with m/z 593.1539, showed a fragment at *m/z* 285.0421 (kaempferol), which corresponds to the loss of rutinoside [M-H-308]$^-$. Thus it was suggested to be kaempferol-3-*O*-rutinoside. This compound has been already reported by Krauze-Baranowska et al. in cucumber flowers[18,19].

The Q-TOF-MS spectra of peaks **36, 47, 58, 60, 63,** and **66** displayed [M-H]$^-$ molecular ion at m/z 665.1359. The molecular formula suggested for these compounds were $C_{29}H_{29}O_{18}$. Compounds **36** and **47** showed a MS spectra characteristic of luteolin di-*C*-glycosides acylated with dicarboxylic acids, since an ion produced by decarboxylation [M-44] was observed in the MS/MS experiments (**Figure 4F**). Therefore, both peaks were proposed as luteolin-*C*-(6-malonyl) hexoside-*C*-pentoside (**Figure 2i**) relying on their MS and MS/MS data. This compound was already discussed in sweet pepper[43].

Compounds **58**, **60**, **63** and **66** were suggested as luteolin-*O*-6´-malonyl-apiosylglycosides, showed by the acceptable data of MS, fragmentation pattern of MS/MS and the data obtained from literature[40,44]. These suggested compounds have shown the same fragment ions at m/z 621.1476, 579.1310, 489.1019, 447.0913 and 285.0409 which correspond to the neutral losses of [M-H-44] (-CO_2), [M-H-44-42] (-malonyl group, $OCCH_2COOH$), [M-H-132] (-apiosyl), [M-H-162] (-glucosyl) and finally the appearance of 285.0408 which represents luteolin. This compound has been reported in cucumber for the first time and its proposed fragmentation pathway is shown in **Figure 3**.

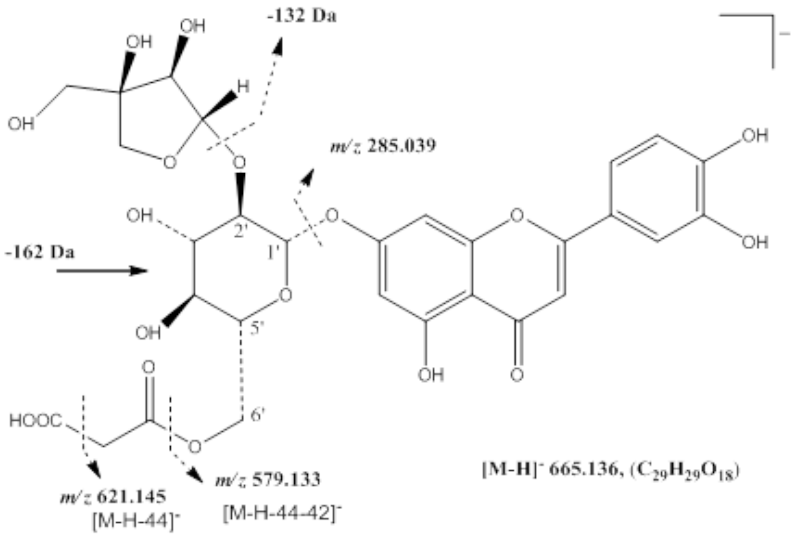


Figure 3. Proposed fragmentation pathway of Luteolin-O-6´-malonyl-apiosylglycoside (m/z 665).

Table 1. Proposed phenolic compounds detected in Cucumis Sativus L. extract obtained by HPLC-ESI-Q-TOF-MS.

Peak no.	Retention Time	Experimental m/z	Calculated m/z	Tolerance	Error	mSigma	Formula	Classification order	MS/MS Fragments (% Relative Abundance)	Proposed Compounds	Reference
1	9.20	299.0780	299.0772	7	2.5	39	$C_{13}H_{15}O_8$	2nd (2)	**137.0296(24)**	Salicylic acid O-glycoside (Isomer 1)	26
2	9.39	299.0772	299.0772	6	5	82.5	$C_{13}H_{15}O_8$	1st (2)	**137.0243(50)**	Salicylic acid O-glycoside (Isomer 2)	26
3	11.19	329.0888	329.0878	4	3.1	13.7	$C_{14}H_{17}O_9$	1st(2)	**167.0352(23)**	Vanillic acid glycoside (Isomer 1)	NA
4	11.92	299.0776	299.0772	5	1.1	18.7	$C_{13}H_{15}O_8$	1st(2)	**137.0257(12)**	Salicylic acid O-glycoside (Isomer 3)	26
5	13.20	329.0904	329.0878	10	7.8	7.5	$C_{14}H_{17}O_9$	1st(3)	**167.0353(100)**	Vanillic acid glycoside (Isomer 2)	NA
6	14.01	341.0893	341.0878	5	4.3	33.5	$C_{15}H_{17}O_9$	1st(2)	**179.03442(30)**	Caffeoyl glucose (Isomer 1)	15
7	14.56	299.0789	299.0772	5	5.6	21.8	$C_{13}H_{15}O_8$	1st(2)	**137.0240(100)**	Salicylic acid O-glycoside (Isomer 4)	26
8	15.18	355.1043	355.1035	4	2.3	101.7	$C_{16}H_{19}O_9$	1st(2)	**193.0566(6)**	Feruloyl glucose(Isomer 1)	15
9	15.89	353.0883	353.0878	4	1.4	42	$C_{16}H_{17}O_9$	1st(2)	**191.0563(100)**	Chlorogenic acid (Isomer 1)	14, 16
10	15.92	341.0884	341.0878	4	1.8	29.8	$C_{15}H_{17}O_9$	1st(2)	**179.0345(42)**	Caffeoyl glucose (Isomer 2)	15
11	16.10	325.0932	325.0929	4	0.8	18.5	$C_{15}H_{17}O_8$	1st(1)	**163.0390(20)**, 119.0500(100)	p-coumaric acid-*O*-glycoside	27
12	16.27	385.1141	385.1140	4	0.2	25.6	$C_{17}H_{21}O_{10}$	1st(2)	**223.0665(7)**	Sinapic acid-hexoside	NA
13	16.96	755.2063	755.2040	4	3	55.1	$C_{33}H_{39}O_{20}$	1st(3)	**593.1489(33)**, 455.0972(100), 413.0879(81)	Saponarin 4'-O-glycoside	12, 38

14	17.57	609.1486	609.1461	5	5	22.5	$C_{27}H_{29}O_{16}$	1st(3)	591.1321(5),549.1168(1), 489.1041(100),399.0715(79), 369.0614(72)	Luteolin-6,8-di-C-hexoside	NA
15	17.65	355.1039	355.1035	4	1.3	101.7	$C_{16}H_{19}O_9$	1st(2)	**193.0505(46)**	Feruloyl glucose (Isomer 2)	15
16	17.83	355.1037	355.1035	4	0.6	32	$C_{16}H_{19}O_9$	1st(2)	**193.0478(2)**	Feruloyl glucose (Isomer 3)	15
17	18.14	385.1148	385.1140	4	2.1	95.7	$C_{17}H_{21}O_{10}$	1st(2)	**223.0666(5)**	Sinapic acid-hexoside	NA
18	18.24	353.0880	353.0878	4	0.7	101.7	$C_{16}H_{17}O_9$	1st(2)	**191.0565(100)**	Chlorogenic acid (Isomer 2)	14, 16
19	18.25	623.1683	623.1676	4	1	60.2	$C_{28}H_{31}O_{16}$	1st(2)	605.1477(4),533.1295(31), 533.1240(31),503.1232(100), 413.9638(11), 383.0473(8)	Lucenin-2-methyl ether	NA
20	18.37	609.1476	609.1461	5	2.5	19.2	$C_{27}H_{29}O_{16}$	1st(3)	**447.0944(20)**, **301.0372(58)**	Quercetin 3-rutinoside (Rutin)	16, 40
21	18.52	741.1911	741.1884	4	3.7	19.1	$C_{32}H_{37}O_{20}$	1st(2)	**595.1305(100),463.0870(5) ,301.0380(21)**	Quercetin3-O-pentosyl-rutinoside	NA
22	18.57	385.1152	385.1140	4	3.2	18.7	$C_{17}H_{21}O_{10}$	1st(2)	**223.0613(36)**	Sinapic acid-hexoside	NA
23	18.91	579.1369	579.1355	4	2.3	18	$C_{26}H_{27}O_{15}$	1st(3)	561.1239(5),519.1102(5), 489.1019(31),459.0918(90), 399.0710(100), 369.0600 (66)	Luteolin-C-hexosyl-C-pentoside	NA
24	19.04	593.1534	593.1512	4	3.8	21.9	$C_{27}H_{29}O_{15}$	1st(3)	575.1398(5),533.1282(3), 503.1198(25), 473.1067(84),413.0857, 383.0760(51), 353.0660(87)	Vicenin-2 (Isomer 1)	12, 41
25	19.43	579.1361	579.1355	4	1	26.4	$C_{26}H_{27}O_{15}$	1st(2)	561.1238 (4), 519.1103 (6), 489.1018 (34),459.0918 (86), 399.0712 (93), 369.0603 (60)	Luteolin-C-hexosyl-C-pentoside	NA

26	19.56	579.1361	579.1355	4	0.9	19.9	$C_{26}H_{27}O_{15}$	1st(2)	561.1238(6),519.11555 (17), 489.1018(33), 459.0918 (82), 399.0712 (99), 369.0603 (55)	Luteolin-C-hexosyl-C-pentoside	NA
27	19.58	563.1374	563.1406	6	5.7	39.2	$C_{26}H_{27}O_{14}$	1st (5)	545.1267(14), 503.1163(13), 473.1059(39), 443.0977 (65), 383.0761(100),353.0629 (81)	Apigenin-6-C-pentosyl-8-C-hexoside	NA
28	19.65	623.1645	623.1618	5	4.4	10.3	$C_{28}H_{31}O_{16}$	1st(4)	605.1485(3), 563.1422(1), 533.1335(22), 503.1191(91), 413.0869(56), 383.0763(100)	Lucenin-2-methyl ether	NA
29	19.80	355.1037	355.1035	4	0.6	19.3	$C_{16}H_{19}O_9$	1st(2)	**193.05026(60)**	Feruloyl glucose (Isomer 4)	15
30	20.05	579.1370	579.1355	4	2.4	14.6	$C_{26}H_{27}O_{15}$	1st(3)	561.1234 (4), 519.1114 (12), 489.1018(50),459.0907(18), 399.0707(100),369.0601 (79)	Luteolin-C-hexosyl-C-pentoside	NA
31	20.32	563.1423	563.1406	4	3	51.8	$C_{26}H_{27}O_{14}$	1st(3)	545.1278(7), 503.1172(26), 473.1069(40), 443.0965(36), 383.0764(76),353.0658 (100)	Apigenin-6-C-pentosyl-8-C-hexoside	NA
32	20.93	563.1416	563.1406	4	1.8	42.5	$C_{26}H_{27}O_{14}$	1st(3)	545.1277 (5), 503.1173 (22), 473.1068 (43), 443.0964(32), 383.0765(71),353.0657 (100)	Apigenin-6-C-pentosyl-8-C-hexoside	NA
33	20.95	463.0908	463.0882	7	5.5	118.6	$C_{21}H_{19}O_{12}$	1st(2)	**301.0283(7)**	Quercetin-*O*-glucoside	19
34	21.07	593.1524	593.1512	4	2	17.9	$C_{27}H_{29}O_{15}$	1st(3)	575.1331(7),533.1255(9), 503.1157(42), 73.1107(100), 413.0866(26), 383.2607(1), 353.0608(2)	Vicenin-2 (Isomer 2)	12, 41
35	21.12	447.0937	447.0933	4	0.9	24.4	$C_{21}H_{19}O_{11}$	1st(2)	**285.04034(22)**	Kaempferol-O-glycoside or Orientin or Isoorientin	15,18,19, 31,32,34,
36	21.33	665.1389	665.1359	5	4.4	13.5	$C_{29}H_{29}O_{18}$	1st(4)	621.1472(25), 603.1375(5), 561.1344(9), 531.1084(12), 489.1010(17), 459.0909(97),	Luteolin-*C*-(6-malonyl) hexoside-*C*-pentoside	NA

									399.0709(100), 369.0603(61)		
37	21.65	447.0939	447.0933	4	1.4	24.2	$C_{21}H_{19}O_{11}$	1st(2)	**285.0405(20)**	Kaempferol-O-glycoside or Orientin or Isoorientin	15,18,19, 31,32,34,
38	21.83	163.0401	163.0401	4	0.3	56.5	$C_9H_7O_3$	1st(1)	119.0493(100)	p-Coumaric acid (Isomer 1)	12,14-16, 18, 25, 28, 34-36
39	21.93	563.1419	563.1406	4	2.3	25.6	$C_{26}H_{27}O_{14}$	1st(3)	545.1279(5), 503.1167 (5), 473.1070(58), 443.0966(84), 383.0757(43), 353.0656(64)	Apigenin-6-C-pentosyl-8-C-hexoside	NA
40	22.20	593.1526	593.1512	4	2.3	18.7	$C_{27}H_{29}O_{15}$	1st(3)	575.1340(4), 519.1195(9), 503.1192(17), 489.1025(40), 473.1094(25),399.0702(100), 369.0621(93)	Luteolin 6-C-rhaminoside-8-C-hexoside	NA
41	22.40	593.1539	593.1512	5	4.5	21.9	$C_{27}H_{29}O_{15}$	1st(4)	575.1393(5), 533.1260(4), 503.1170(56), 473.1078(63), 413.0864(41), 383.0756(81), 353.0608(3)	Vicenin-2 (Isomer 3)	12, 41
42	22.47	577.1583	577.1563	4	3.6	12.8	$C_{27}H_{29}O_{14}$	1st(3)	**285.0416 (19)**	Kaempferol-*O*-dirhamnoside	NA
43	22.67	433.1155	433.1140	4	3.5	25	$C_{21}H_{21}O_{10}$	1st(2)	**271.0617(100)**	Naringenin-*O*-glucoside	NA
44	22.74	521.2031	521.2028	4	0.5	31.6	$C_{26}H_{33}O_{11}$	1st(2)	**359.1501(43)**	Lariciresinol-glucopyranoside	NA
45	22.82	163.0408	163.0401	5	4.6	56.5	$C_9H_7O_3$	1st(1)	119.0488 (100)	p-Coumaric acid (Isomer 2)	12,14-16, 18, 25, 28, 34-36
46	23.01	609.1508	609.1461	8	7.6	39.5	$C_{27}H_{29}O_{16}$	1st(5)	**447.0934(100), 301.0357(89)**	Quercetin3-rutinoside (Rutin) (Isomer 1)	16, 40
47	23.14	665.1403	665.1359	7	6.5	44.6	$C_{29}H_{29}O_{18}$	1st(5)	621.1452 (28), 561.1302(14),	Luteolin-*C*-(6-malonyl) hexoside-	NA

									531.1069(20),489.1040(32), 459.0927(21), 399.0713(88), 369.0594(100)	*C*-pentoside	
48	23.29	695.1875	695.1829	7	6.7	24	$C_{31}H_{35}O_{18}$	1st(3)	548.1191(100), 447.0961(38), **301.03387(91)**	7-O-(Carbethoxy-methyl) rutin	NA
49	23.36	609.1500	609.1461	7	6.3	43.3	$C_{27}H_{29}O_{16}$	1st(4)	**301.0338(65)**	Quercetin3-rutinoside (Rutin) (Isomer 2)	16, 40
50	23.41	433.1144	433.1140	4	0.8	27	$C_{21}H_{21}O_{10}$	1st(2)	**271.0619(100)**	Naringenin-*O*-glucoside	NA
51	23.88	579.1364	579.1355	4	1.5	32.7	$C_{26}H_{27}O_{15}$	1st(2)	**447.09423(9)**, **285.0408(61)**	Kaempferol-3-*O*-sambubioside	NA
52	24.21	577.1564	577.1563	5	0.2	43.7	$C_{27}H_{29}O_{14}$	1st(3)	**285.0417(16)**	Kaempferol-*O*-dirhamnoside	NA
53	24.23	463.0883	463.0882	4	0.2	25.3	$C_{21}H_{19}O_{12}$	1st(2)	301.0343(33), **300.0282(100)**	Quercetin-*O*-glucoside	19
54	24.30	447.0945	447.0933	4	2.8	28	$C_{21}H_{19}O_{11}$	1st(2)	**285.0408(100)**	Kaempferol-O-glycoside or Orientin or Isoorientin	15,18, 19, 25, 28, 37, 38
55	24.77	461.1112	461.1089	4	4.9	25.5	$C_{22}H_{21}O_{11}$	1st(2)	**371.0745(11)**, 341.0677(100)	Isoorientin 3'-O-methyl ether (Isoscoparin)	43
56	24.87	521.2030	521.2028	4	0.3	10.4	$C_{26}H_{33}O_{11}$	1st(2)	**359.1512 (41)**	Lariciresinol glucopyranoside	NA
57	25.36	593.1552	593.1512	5	3.2	43	$C_{27}H_{29}O_{15}$	1st(3)	**285.0421(100)**	Kaempferol 3-O-rutinoside	19
58	25.42	665.1423	665.1359	10	9.5	47	$C_{29}H_{29}O_{18}$	1st(6)	621.1476(52), **579.1310(42)**, 489.1019(53), **447.0913(4), 285.0408(33)**	Luteolin-*O*-6´-malonyl-apiosylglycoside	NA
59	25.71	577.1538	577.1563	10	4.3	34.6	$C_{27}H_{29}O_{14}$	2nd(7)	**269.0446(100)**	Apigenin-7-rutinoside	NA
60	25.76	665.1362	665.1359	6	6	38.1	$C_{29}H_{29}O_{18}$	1st(6)	621.1474(100),**579.1337(85),**	Luteolin-*O*-6´-	NA

									489.1000(9),447.0902(4), 285.0384 (49)	malonyl-apiosylglycoside	
61	25.83	623.1652	623.1618	6	5.5	44	$C_{28}H_{31}O_{16}$	1st(4)	**315.0501(100)**	Isorhamnetol 3-O-rutinoside	NA
62	26.01	565.1572	565.1563	4	1.7	21.2	$C_{26}H_{29}O_{14}$	1st(3)	**271.0617(100)**	Theaflavanoside I	NA
63	26.14	665.1379	665.1359	5	3	8.5	$C_{29}H_{29}O_{18}$	1st(3)	621.1468(100), **579.1338(7), 489.1009(6), 447.0926(2), 285.0391(3)**	Luteolin-*O*-6´´-malonyl-apiosylglycoside	NA
64	26.48	447.0934	447.0933	4	0.4	24.5	$C_{21}H_{19}O_{11}$	1st(2)	**301.0398(44)**, 271.0250(9)	Quercitrin; Quercetin 3-L-rhamnoside	NA
65	26.63	593.14886	593.15119	5	3.9	40	$C_{27}H_{29}O_{15}$	1st(2)	**299.0540(100)**	Diosmetin-apiosylglucoside	NA
66	26.75	665.1386	665.1359	5	4.1	18.9	$C_{29}H_{29}O_{18}$	1st(5)	621.1482(100), 579.1351(3), 489.1016(6),447.0943(2), **285.0409(20)**	Luteolin-*O*-6´´-malonyl-apiosylglycoside	NA
67	26.93	447.0935	447.0933	4	0.5	27.5	$C_{21}H_{19}O_{11}$	1st(2)	**285.0403(24)**	Kaempferol-O-glycoside or Orientin or Isoorientin	15,18,19, 31,32,34,
68	27.35	489.1040	489.1039	4	0.4	17.1	$C_{23}H_{21}O_{12}$	1st(2)	**285.0413(100)**	Kaempferol 3-O-(6"-O-acetyl) glycoside	NA
69	28.56	447.0960	447.0933	6	6	24.2	$C_{21}H_{19}O_{11}$	1st(2)	**285.0475(14)**	Kaempferol-O-glycoside or Orientin or Isoorientin	15,18,19, 31,32,34,
70	29.05	431.0993	431.0984	4	2.2	25.9	$C_{21}H_{19}O_{10}$	1st(2)	285.0403(65), **284.0339(100)**	Kaempferol 3-O-α-L-rhamnoside	19
71	29.37	461.1099	461.1089	4	2	7.6	$C_{22}H_{21}O_{11}$	1st(2)	**285.0420(13)**	Luteolin 7-O-glucuronide	NA

72	29.64	433.1145	433.1140	4	1.1	23.6	$C_{21}H_{21}O_{10}$	1st(2)	**271.0617(90)**	Naringenin-*O*-glucoside	NA
73	29.84	621.1488	621.1461	5	4.3	10.7	$C_{28}H_{29}O_{16}$	1st(2)	**489.1031(11), 285.0411(59)**	Luteolin 7-O-(2-apiosyl-6-acetyl) glucoside	NA

NA: not available. * Bold entries referred to the main aglycones' fragments

Peaks **27, 31, 32** and **39** (R_t 19.58, 20.32, 20.92, 21.93 min) presented the molecular ion at 563.1419. The MS/MS fragmentation is typical of a flavone-C-diglycoside with $[M-H-18]^-$ (loss of H_2O), $[M-H-60]^-$,$[M-H-90]^-$, $[M-H-120]^-$, $[M-H-180]^-$, [M-H-210]. The higher intensity of the [M-H-90] ion relative to [M-H-120] is indicative of a 6-C-pentosyl-8-C-hexosyl-substitution[40] (see **Figure 4D**). Therefore, these compounds were proposed as apigenin-6-C-pentosyl-8-C-hexosides.

Peaks **33** and **53** were identified as quercetin-*O*-glucosides, showed by the neutral loss of glucose moiety $[M-H-162]^-$ and the appearance of fragment ions (300.0282) (100% relative intensity) which can be referred to [quercetin-H]-[19].

Compounds **44** and **56,** with the retention times 22.74 and 24.87, showed the same precursor ion and the fragment ion at m/z 521.2031 and 359.1501, respectively. The fragment ion refers to lariciresinol[24], and it was shown the neutral loss of glucose moiety (162 Da). Thus, these compounds were tentatively proposed as lariciresinol-glucopyranosides (**Figure 2j**).

The compound at the retention time 29.37 was tentatively suggested to be luteolin-7-*O*-glucuronide (**Figure 2k**), confirmed by the neutral loss of glucuronic acid (176 Da) and the appearance of the fragment ion 285.0420 (luteolin).

Peaks **19** and **28**, showing a molecular ion at 623.1645, yielded MS/MS fragments (see **Figure 4E**) which are characteristic ions of C-glycoside flavonoids. Thus, these compounds were tentatively characterized as lucenin-2-methyl ether.

Finally, compound **61** was tentatively assigned to isorhamantinol 3-O-rutinoside (**Figure 2I**), relying on the clear loss of rutinoside (308 Da) and the appearance of fragment ion at 315.0501 with 100% of relative abundance which represent isorhamnetin.

Conclusion

In the present work HPLC-ESI-Q-TOF-MS has confirmed to be a powerful analytical technique for the separation and detection of phenolic compounds in *Cucumis sativus L.* Thus, by using the proposed method, 73 compounds were tentatively identified accomplished with the accurate mass determination of the deprotonated ions which was obtained from the MS data and MS/MS fragmentation pattern. To our knowledge, this work is the first comprehensive study of the phenolic composition of cucumber whole fruit extract. In this context, the obtained data indicate qualitatively that cucumber is an abundant source of phenolic compounds.

These results could explain the past and current usage of *Cucumis sativus L.* as food, in folk medicine; also support the further uses of cucumber in health, nutrition and cosmetology. Nevertheless, these data also could serve to improve the data in the food composition tables. Moreover these results would be used in the quantification of the characterized compounds.

References

1 Song, W, Derito, C. M. , Liu, M. K, He, X., Dong, M., and Liu, R. H. 2010. Cellular Antioxidant Activity of Common Vegetables. *Journal of Agriculture and. Food Chemestry*. **58,** 6621–6629.

2 Leopoldini, M., Russo, N., and Toscan, M. 2011. The molecular basis of working mechanism of natural polyphenolic antioxidants, *Food Chemistry*, **125,** 288–306.

3 Gill, N. S., Sood, S., Muthuraman, A., Garg, M., Kumar, R., Bali, M., and Sharma, P. D. 2010. Antioxidant, Anti-inflammatory and Analgesic Potential of Cucumis sativus Seed Extract. *Latino America Journal Pharmacy,* **29,** 927-32.

4 Islam, S., Jalaluddin, M., and Hettiarachchy, N. S. 2011. Bio-active compounds of bitter melon genotypes (Momordica charantia L.) in relation to their physiological functions. *Functional Foods in Heals and Disease*, **2,** 61-74.

5 Hedges, L. J., and Lister, C. E. 2005. Nutritional attributes of salad vegetables, Crop and Food Research Confidential Report No. 1473, pp.2.

6 Dixit, Y., and Kar, A. 2010. Protective Role of Three Vegetable Peels in Alloxan Induced Diabetes Mellitus in Male Mice. *Plant Foods for Human Nutrition,* **65,**284–289.

7 Kai, H., Baba, M., and Okuyama, T. 2008. Inhibitory Effect of Cucumis sativus on Melanin Production in Melanoma B16 Cells by Down regulation of Tyrosinase Expression. *Planta Medicina*, **74,** 1785–1788.

8 Mukherjee, S., Das, J. K., and Roy, C. K. 2009. Screening of Antioxidant Activity and Phenolic Content in Cucumis Sativus L. Seeds. *Pharmacology online*, **2,** 711-714.

9 Ferro-Luzzi, A., Cialfa, E., Leclercql, C., and Toti, E. 1994. The Mediterranean diet revisited. Focus on fruit and vegetables. *International Journal of Food Sciences and Nutrition*, **45,** 291-300.

10 Murcia, M. A., Jiménez, A. M., and Martinez-Tomé, M. 2009. Vegetables antioxidant losses during industrial processing and refrigerated storage, *Food Research International*, **42,** 1046–1052.

11 Abou-Zaid, M. M., Lombardo, D. A., Kite, G. C., Grayer, R. J., and Veitch, N. C. 2001. Acylated Flavones C-glycosides from Cucumis sativus. *Phytochemistry* , **58,** 167–172.

12 Gómez-Romero, M., Zurek, G., Schneider, B., Baessmann, C., Segura-Carretero, A., and Fernández-Gutiérrez, A. 2011. Automated Identification of Phenolics in Plant-Derived Foods by using Library Search Approach. *Food Chemistry*, **124,** 379–386.

13 Ibrahim, T. A., El-Hefnawy, H. M., and El-Hela, A. A. 2010. Antioxidant potential and phenolic acid content of certain cucurbitaceous plants cultivated in Egypt. *Natural Product Research*, **16,** 1537–1545.

14 Bellés, J. M., López-Gresa, M. P., Fayos, J., Pallás, V., Rodrigo, I., and Conejero, V. 2008. Induction of cinnamate 4-hydroxylase and phenylpropanoids in virus-infected cucumber and melon plants. *Plant Science*, **174,** 524–533.

15 Maksimovic, J. D., Bogdanovic, J., Maksimovic, V., and Nikolic, M. 2007. Silicon modulates the metabolism and utilization of phenolic compounds in cucumber (Cucumis sativus L.) grown at excess manganese. *Journal of Plant Nutrition and Soil Science*, **170,**739–744.

16 Daayf, F., Bel-Rhild, R., and Belanger, R. R. 1997. Methyl Ester of p-Coumaric Acid: A Phytoalexin-Like Compound From Long English Cucumber Leaves. *Journal of Chemical Ecology*, **6,** 1517-1526.

17 McNally, D. J., Wurms, K. V., Labbé, C., Quideau, S., and Bélanger, R. R. 2003. Complex C-Glycosyl Flavonoid Phytoalexins from Cucumis sativus. *Journal of Natural Products*, **66,** 1280-1283.

18 Krauze-Baranowsk, M., and Cisowski, W. 2001. Flavonoids from some species of the genus Cucumis. *Biochemical Systematics and Ecology,* **29,** 321-325.

19 Hertog, M. G. L., Hollman, P. C. H., and Katan, M. B. 1992. Content of potentially anticarcinogenic flavonoids of 28 vegetables and 9 fruits commonly consumed in the Netherland. *Journal of Agriculture and Food Chemistry,* **40,** 2379-2383.

20 Lugasi, A., and Hovari, J. 2000. Flavonoid aglycons in foods of plant origin I. Vegetables. *Acta Alimentaria,* **29,**345-352.

21 Chu, Y. H., Chao-Lin, C., and Hsia-Fen, H. 2000. Flavonoid content of several vegetables and their antioxidant activity. *Journal of the Science of Food and Agriculture*, **80,** 561-566.

22 Peñalvo, J. L., Haajanen, K. M., Botting, N., and Adlercreutz, H. 2005. Quantification of lignans in food using isotope dilution gas chromatography/mass spectrometry, *Journal of Agriculture and Food Chemistry,* **53,** 9342-9347.

23 Garcia-Salas, P., Morales-Soto, A., Segura-Carretero, A., and Fernández-Gutiérrez, A. 2010. Phenolic-Compound-Extraction Systems for Fruit and Vegetable Samples. *Molecules,* **15,** 8813-8826.

24 Ferrer, I., García-Reyes, J. F., Mezcua, M., Thurman, E. M., and Fernández-Alba, A. R. 2005. Multi-residue pesticide analysis in fruits and vegetables by liquid chromatography–time-of-flight mass spectrometry. *Journal of Chromatography A*, **1082,** 81–89.

25 Monties, B., Bouillan, M. L., and Chopin, J., 1976. C-Diholosylflavones Dans Les Feuilles Du Melon (Cucumis Melo), *Phytockemistrj*, **15,** 1053-1056.

26 Daayf, F., Ongena, M., Boulanger, R., El-Hadrami, I., Bélanger, R. R. 2000. Induction of Phenolic Compounds in Two Cultivors of Cucumber by Teatment of Healty and Powdery Mildew-Infected Plants With Extracts of Reynoutria sachalinensis. Journal of Chemical Ecology, **7,** 1579-1593.

27 Billett, E. E., Grayer-Barkmeijerc, R. J., Johanson, B., and Harborne, J. B. 1981. The Effect of Blue Light on Free and Esterified Phenolic Acids in Etiolated Gherkin Tissues. *Phytochemistry*, **6,** 1259-1263.

28 Zhang, P. Y., Wang, J. C., Liu, S. H., and Chen, K. S. 2009. A novel burdock fructooligosaccharide induces changes in the production of salicylates, activates defence enzymes and induces systematic acquired resistance to Colletotrichum orbiculare in cucumber seedlings. *Journal of Phytopathology*, **4,** 201-207.

29 Fofana, B., Mcnally, D. J., Labbé, C., Boulanger, R., Benhamou, N., Séguin, A., and Bélanger, R. R. 2002. Milsana-induced resistance in powdery mildew-infected cucumber plants correlates with the induction of chalcone synthase and chalcone isomerise. *Physical and Molecular Plant Pathololog*y, **61,** 121-132.

30 Segarra, G., Jáuregui, O., Casanova, E., and Trillas, I. 2006. Simultaneous quantitative LC–ESI-MS/MS analyses of salicylic acid and jasmonic acid in crude extracts of Cucumis sativus under biotic stress. *Phytochemistry,* **67,** 395–401.

31 Herrmann, K. 1977. Review on nonessential constituents of vegetables I. Cucumbers, melons, squashes and pumpkins, bell peppers, eggplants, peas, beans, and broad beans. *Z. Lebensm. Unters.-Forsch,* **165,** 87-98.

32 Re, R., Pellegrini, N., Proteggente, A., Pannala, A, Yang, M., and Rice-Evans, C. 1999. Antioxidant activity applying an improved ABTS radical cation decolorization assay. *Free Radical Biology & Medicine,* **26,** 1231–1237.

33 Laporta, O., Pérez-Fons, L., Mallavia, R., Caturla, N., and Micol, V. 2007. Isolation, characterization and antioxidant capacity assessment of the bioactive compounds derived from Hypoxis rooperi corm extract (African potato), *Food Chemistry*, ***101,*** 1425–1437.

34 Al-Duais, M., Müller, L., Böhm, V., and Jetschke, G. 2009. Antioxidant capacity and total phenolics of Cyphostemma digitatum before and after processing: use of different assays. *European Food Research and Technology,* **228,** 813–821.

35 Wang, Y., Yang, L., He, Y. Q., Wang, C. H., Welbeck, E. W., Bligh, S. W. A., and Wang, Z. T. 2008. Characterization of fifty-one flavonoids in a Chinese herbal prescription Longdan Xiegan Decoction by high-performance liquid chromatography coupled to electrospray ionization tandem mass spectrometry and photodiode array detection. *Rapid Communications in Mass Spectrometry,* **22,** 1767–1778.

36 Laporta, O., Pérez-Fons, L., Mallavia, R., Caturla, N., and Micol, V. 2007. Isolation, characterization and antioxidant capacity assessment of the bioactive compounds derived from Hypoxis rooperi corm extract (African potato), *Food Chemistry,* **101,** 1425–1437.

37 Breiter, T., Laue, C., Kressel, G., Gröll, S., Engelhardt, U., H., and Hahn, A. 2011. Bioavailability and antioxidant potential of rooibos flavonoids in humans following the consumption of different rooibos formulations, *Food Chemistry,* **128,** 338–347.

38 Krauze-Baranowska, M., and Cisowski, W. 1994. High-performance liquid chromatographic determination of flavone C-glycosides in some species of the Cucurbitaceae family, *Journal Chromatography A.,* **675,** 240-243.

39 Aboellil, A. H. 2007. Trilogy, a Product of Neem (Azadirachta indica) Induces Resistancein Cucumber against Podosphaera xanthii. *Journal of Microbiology Research.* **5,** 402-414.

40 Kastner, J. 2009. Chinese Nutrition Therapy: Dietetics in Traditional Chinese Medicine (TCM) Georg Thieme Verlag.

41 Delazar, A., Gibbons, S., Kosari, A. R., Nazemiyeh, H., Modarresi, M., Nahar, L., and Sarker, S. D. 2006. Flavone C-Glycosides and Cucurbitacin Glycosides from Citrullus Colocynthis, *DARU,* **3,**109-114.

42 Justesen, U. 2000. Negative atmospheric pressure chemical ionisation low-energy collision activation mass spectrometry for the characterisation of flavonoids in extracts of fresh herbs, *Journal of Chromatography A*, **902,** 369–379.

43 Marín, A., Ferreres, F., Tomás-Barberán, F. A., and Gil, M. I. 2004. Characterization and Quantitation of Antioxidant Constituents of Sweet Pepper (Capsicum annuum L.). *Journal of Agriculture and Food Chemistry,* **52,** 3861-3869.

44 Lin, L. Z., Lu, S., and Harnly, J. M. 2007. Detection and Quantification of Glycosylated Flavonoid Malonates in Celery, Chinese Celery, and Celery Seed by LC-DAD-ESI/MS, *Journal of Agriculture and Food Chemistry,* **55,** 1321-1326.

Chapter 3

Profiling of phenolic and other polar constituents from hydro-methanolic extract of watermelon (Citrullus lanatus) by means of accurate-mass spectrometry (HPLC–ESI–QTOF–MS)

This work has been published in:

Food Research International

Food Research International 51 (2013) 354–362

Contents lists available at SciVerse ScienceDirect

Food Research International

journal homepage: www.elsevier.com/locate/foodres

Profiling of phenolic and other polar constituents from hydro-methanolic extract of watermelon (*Citrullus lanatus*) by means of accurate-mass spectrometry (HPLC–ESI–QTOF–MS)

Ibrahim M. Abu-Reidah, David Arráez-Román, Antonio Segura Carretero, Alberto Fernandez-Gutierrez

Profiling of phenolic and other polar constituents from hydro-methanolic extract of watermelon (*Citrullus lanatus*) by means of accurate-mass spectrometry (HPLC-ESI-QTOF-MS)

Abstract

Watermelon, *Citrullus lanatus* (formerly C*itrullus vulgaris*), is a natural and rich source of the phytochemical compounds. In this regard, the use of high-performance liquid chromatography coupled with electrospray-quadropole-time-of-flight mass spectrometry (HPLC-ESI-QTOF-MS) has shown to be a powerful technique for the characterization of phenolic and other polar compounds from a hydro-methanolic extract of watermelon. Thus, in the present work, 71 polar compounds such as phenolic acids, flavonoids, iridoids, coumarins, lignan, and other phenolic derivatives have been detected and characterized by using MS and MS/MS data provided by the QTOF-MS, in addition to using the relevant literature on the same botanical family. Watermelon flesh was found to contain an array of diverse phytochemical components. These results indicate that watermelon offers a good source of natural phyto-components.

Keywords: *Watermelon (Citrullus lanatus), Cucurbitaceae, Phenolic compounds, phytochemicals, High-Performance Liquid Chromatography, Electrospray-Quadropole-Time-of-Flight Mass Spectrometry.*

Introduction

A wide array of phytochemicals produced by plants have antioxidant properties and are broadly studied for their beneficial effects on the health of animals and humans, based on scavenging of free radicals generated by environmental and metabolic factors[1]. Phytochemicals, such as phenolic compounds, are considered beneficial for human health and well-being, lowering the risk of degenerative diseases such as cardiovascular diseases and certain cancers, by several mechanisms: free-radical neutralization; protection and regeneration of other dietary antioxidants (i.e. vitamin E); and the chelating of pro-oxidant metal ions[2]. Interest in the role of phenolic antioxidants in human health has prompted research into the separation and characterization of functional phenolic components in various plant-derived foods[3].

Cucurbitaceae family has a considerable economic importance, comprise approximately 122 genera and 900 species[4], many of them used for food and/or as folk medicines. Watermelon (*Citrullus lanatus*), native of southern Africa, represents a large portion of the Mediterranean diet, dating to some 3000 years ago[5].

Watermelon fruit is consumed fresh (eaten directly or in fruit salad) or is processed as juice, jam, pickled rind, etc., and the seeds are also eaten (seed cultivars). This highly consumed fruit, once being only a summer food is currently becoming an everyday fruit. Moreover, *Citrullus lanatus* embodies one of the most widely cultivated crops in the world, occupying the largest production of all Cucurbits. Recently, the world production of watermelon was estimated that to be around 99 million metric tonnes[6] (FAOSTAT, 2010). Medicinally, watermelons are mildly diuretic, being effective in the treating dropsy and renal stones, reducing hypertension, preventing erectile dysfunction, acting as an antioxidant, and treating enlarged liver, jaundice and giardiasis[7-9].

It bears mentioning that previous chemical research has revealed that cucurbits could be referred to the following active groups: amino acids, essential oils, sterols, triterpene saponins, cucurbiticins aglycones or glycosides, phenyl amines, and phenolic compounds (mainly flavonoids)[10,11].

Recently, it has been reported that phenolic compounds may be the principal hydrophilic compounds contributing to the hydrophilic antioxidant activity in watermelon[12]. The effectiveness of watermelon extracts has been recently described in promoting a laxative effect, this property being due partly to phenolic compounds[13].

Phenolic compounds with different structures or levels are likely to have different functional properties. Therefore, it is crucial to analyse the composition of phenolic compounds in plant foods before their health-promoting properties can be adequately studied. Although most studies on *Cucumis* species have focused on the total content of phenols and antioxidant activity[14-16], limited data are available on the phenolic and phyto-components of these plant foods.

Indeed, it is indispensable to have high-quality and comprehensive data on bioactive compounds in foods. For this, adequate analytical methodologies are necessary for their reliable and extensive characterization.

Nowadays, phytochemicals' identification in complex plant matrices is a difficult task due to the complexity of their structures[17] and the limited standards commercially available. High performance liquid chromatography (HPLC) is one of the most commonly used separation techniques used to determine these kinds of bioactive compounds in plant matrix. Lately, the capability of HPLC to separate polyphenols has been well known[18], being the most commonly used separation technique for determining these compounds. Currently, mass spectrometry (MS) is the detection system mainly used due to its high sensitivity and its great potential for identifying compounds[19]. Moreover, the use of MS/MS aid enabling the detection of hundreds of compounds within a single extracts supplying very useful structural information.

On the other hand, it is important to mention that no report has been previously described the phenolic composition of watermelon. Thus, in the present work, a comprehensive and robust method to characterize phenolic components and other polar compounds present in methanolic extract of watermelon flesh (edible part) has been established by high-performance liquid chromatography coupled to electrospray ionization-quadrupole time-of-flight mass spectrometry (HPLC-ESI-

QTOF-MS) as a powerful analytical tool providing accurate mass measurements and complementary structural information based on the MS and MS/MS spectra.

However, it is important to indicate that measuring the exact mass of an unknown substance, by using of ESI-TOF-MS it can directly lead to the atomic composition and sum formula. A list of possible isobaric substances is found when searching in chemical abstracts or other databases like *Scifinder, Kegg Ligand, Massbank,* among others. Additional independent information besides exact mass is necessary to base the decision on which of these isobaric is the correct match. Therefore, by combining the retention time, exact mass and the MS/MS fragmentation pattern, it will result in a generally clear identification of the unknown compounds. Information from already reported literature on the botanical family may also be used in the characterization process.

Materials and methods

Chemicals

HPLC-grade acetonitrile was purchased from Labscan (Dublin, Ireland). Acetic acid was of an analytical grade (assay >99.5%) and purchased from Fluka (Switzerland). Water was purified by using a Milli-Q system (Millipore, Bedford, USA). Other reagents unmarked were of an analytical grade.

Sample preparation

Plant material

Four samples of watermelon fruits (cv. Fashion) were purchased from local commercial market, cultivated in the period between June and July (2012) in the region of Almeria/Spain. After purchasing samples were washed with tap water, dried, then peeled, and subsequently peeled inner flesh was frozen at -25°C and then lyophilized.

Phenolic compound extraction

Freeze-dried watermelon sample (0.5 g) was extracted with 16 mL of methanol:water 80:20 (v:v) at room temperature using the ultrasonic path for 30 min. The supernatant was collected in a round-bottom flask and evaporated to dryness under reduced pressure by using rotary evaporator at 38 °C to remove the extraction solvent. Then, the dried extract was recovered with 80% aqueous

methanol (0.5 mL) and filtered through a 0.22-µm disposable filter disk. Finally, the sample was kept at -20 °C until analysis.

Liquid chromatography

Analysis of the phytochemical fraction of watermelon-fruit-pulp extract was performed by HPLC-ESI-QTOF-MS following our previously used protocol with some adjustments[20]. Thus, separation has been performed using a Rapid Resolution liquid chromatography (RRLC) 1200 Series system (Agilent Technologies, Palo Alto, CA, USA), equipped with a vacuum degasser, an auto-sampler, and a binary pump. Watermelon Hydro-methanolic extract was separated using a reversed-phase C18 analytical column (4.6x150 mm, 1.8 µm particle size; ZORBAX Eclipse Plus) protected by a guard cartridge of the same packing, and maintained at 25°C. The mobile phase, consisting acetic acid (0.5%) in Milli-Q water deionised water (A) and acetonitrile (B), was pumped at 0.8 mL/min into the HPLC system and with injection volume of 5 µL with the following gradient elution program: 0–2 min, 1–3% B; 2–15 min, 3–9% B; 15–32 min, 9–35% B; 32–39 min, 35–100% B. Subsequently, the B content was decreased to the initial conditions and for 6 min the column was re-equilibrated (total run time 35 min).

Mass spectrometry

The HPLC system was online coupled to a quadrupole time-of-flight (QTOF; micrOQTOF™) mass-spectrometer (Bruker Daltonik GmbH, Bremen, Germany). The effluent from the HPLC system was split using a ‘‘T” before being introduced into the mass spectrometer (split ratio 1:3); thus, the flow rate entering into ESI source was 0.2 mL/min. The MS/MS acquisition was performed in negative ionization.

The QTOF detector operated in the following parameters in negative-ion mode: nebulizer 2 bars; dry gas flow 8 L/min and temperature of 210°C; endplate offset -500 V; capillary voltage +4 kV; mass range 50–1100 m/z. External instrument and post-run internal-mass scale calibrations were made using a 74900-00-05 Cole Palmer syringe pump (Vernon Hills, Illinois, USA) with sodium acetate clusters. The elemental composition of each phytochemical compound was selected according to the accurate masses and the isotopic pattern through (Bruker Compass DataAnalysis™, consisting of Smart Formula tool developed by Bruker Daltonik

GmbH); which provides a list of possible molecular formulae by combining accurate mass and isotopic distribution, reflected in their error and mSigma values, respectively. Besides the observed MS and MS/MS spectra and obtained by QTOF-MS analysis, the characterization of compounds was based on the comparison with those found in the relevant literature which have been the main tool for the identification of (poly) phenols compounds.

Results and discussion

In the present work, a total of 71 phytochemical compounds have been tentatively characterized, mainly phenolic compounds: phenolic acids (hydroxybenzoic and hydroxycinnamic acids), flavonoids derivatives, lignans, iridoids, coumarins, stilbenoids, and others. The tentative assignment of these compounds together with the retention time, observed m/z, molecular formula, mSigma, error, and MS/MS fragments are listed in **Table 1**.

Also, **Figure 1** shows the base peak chromatogram (BPC) resulting from the optimal gradient elution program and the optimal MS conditions in negative ionization mode for the watermelon hydro-methanolic extract.

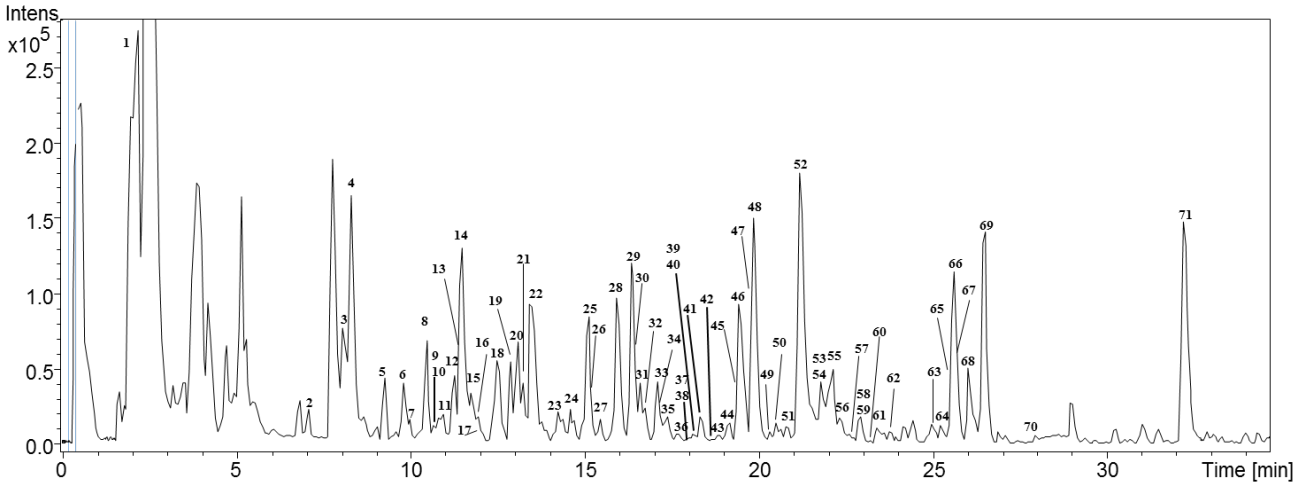


Figure 1 Base peak chromatogram (BPC) of watermelon extract by HPLC-ESI-QTOF-MS

Phenolic acid derivatives

Hydroxybenzoic acid derivatives

The proposed method has been useful for the characterization of 11 hydroxybenzoic acid derivatives.

Thus, the ions detected at 8.23 and 8.62 min produced an MS/MS spectrum corresponding to the characteristic ions at [M-2H-153.0195]− and [M-2H-109.0277]−, which are consistent with the loss of glucose moiety and CO_2, respectively. Thus, these compounds have been proposed as protocatechuic acid glucoside isomers (**Figure 2a**)[21].

On the other hand, compounds **9**, **13** and **70** with m/z at 315.1072 have demonstrated identical molecular formula $C_{14}H_{19}O_8$. Information from the MS/MS spectra showed fragmented ions at m/z 151.0400 and 109.0304, which are consistent with the appearance of vanillin and the loss of CO_2. It is to say that vanillin was previously described in bottle gourd (Cucurbitaceae)[22].Accordingly; peaks **9**, **13,** and **70** have been tentatively assigned as isomers of vanillin hexoside.

Parent ions observed at m/z 301.0571 was tentatively suggested to be either pyrogallol *O*-glucuronide or phloroglucinol *O*-glucuronic acid. In the ESI-QTOF spectra, the fragment ion detected at m/z 125.0235, may be produced by the neutral loss resulting from a loss of the glucuronic acid [M-H-176]−.

Similarly, the molecular ion at m/z 329.0880, with formula $C_{14}H_{17}O_9$, has been assigned as vanillic acid hexoside shown by presence of a fragment ion at m/z 167.0349, which indicates vanillic acid. It bears noting that this compound was previously described in cucumber with the same fragmentation pattern (**Figure 3A**)[23]. Based on the accepted data from MS and MS/MS, peaks **8** and **19** showing m/z at 299.0770 have been assigned to be isomers of salicylic acid *O*-hexoside (**Figure 3B**). The MS/MS analysis revealed a fragment ion at 137.0236 which may refer to the loss of hexose [M-H-162]−. Likewise, this compound has been already described elsewhere in Cucurbitaceae[23,24].

Hydroxycinnamic acid derivatives

Three compounds (**5, 16, and 25**), with the identical molecular formula $C_{15}H_{17}O_8$, have been detected at m/z 325.0916. In MS/MS spectra the loss of the hexose moiety [M-H-162]− has been displayed together with other loss of H_2O [M-H-162-18]−. Therefore, they were tentatively characterized as *p*-coumaric acid glucoside

isomers (**Figure 2b**). Moreover, these data agree with other reports in the literature[23,25].

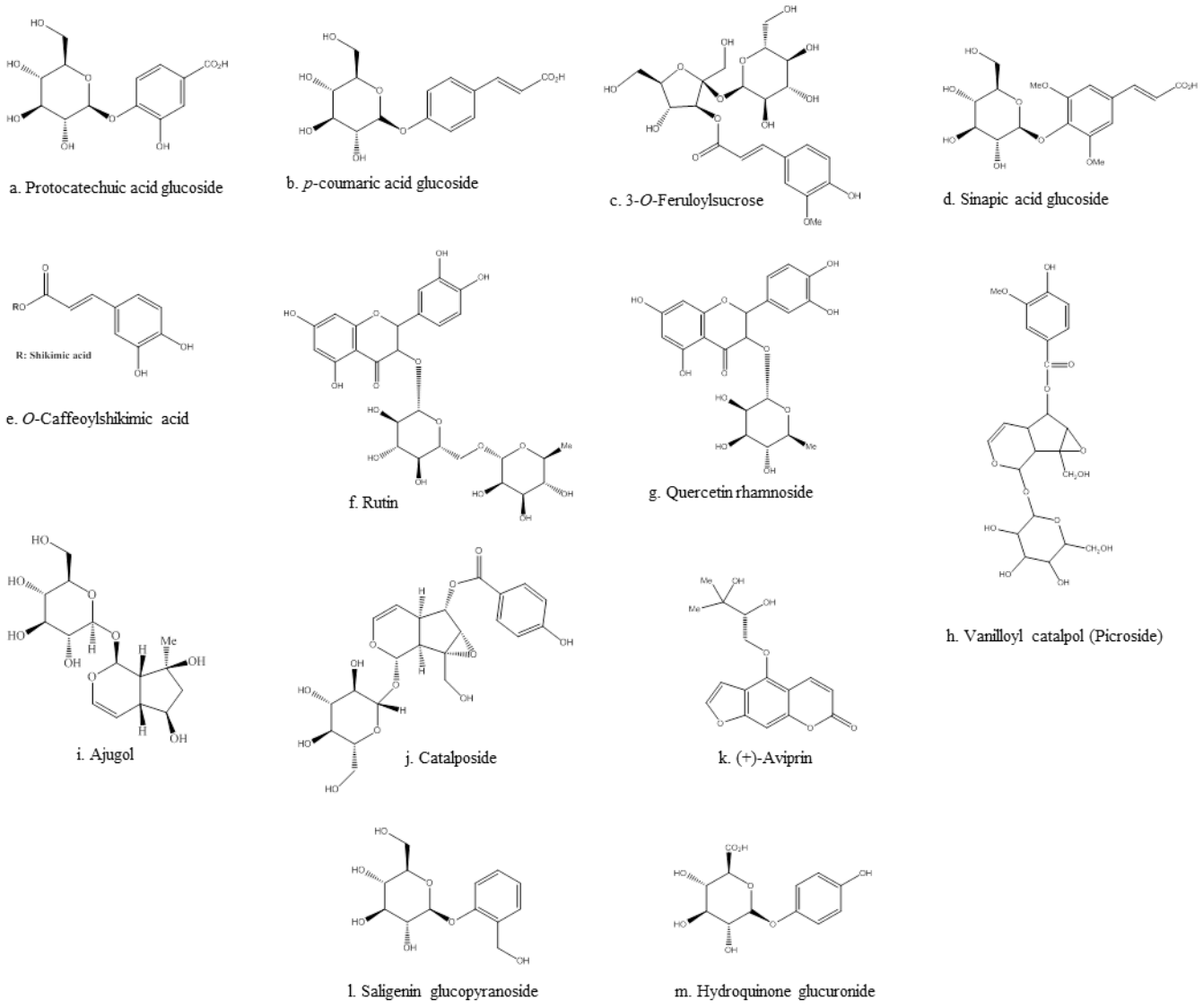


Figure 2. Chemical structures of some characterized compounds in the polar extract of watermelon *Citrullus lanatus.*

The ion at m/z 659.1826 was proposed as tri-*O*-caffeoylshikimic acid, relying on the appropriate information from MS data and MS/MS spectra, which demonstrated a successive loss of two caffeoyl moieties [M-2H-162-162]$^-$[26].

Furthermore, with the identical molecular formula $C_{16}H_{19}O_9$ obtained by the Data Analysis 4.0 program, four peaks at retention times 10.82, 13.60, 16.32, and 17.11 min have been detected. These isomers were proposed as ferulic acid hexoside[23,27]. The assignment of these compounds was based on the MS data and

MS/MS fragments in which the main precursor ion at 355.1028 showed the fragment ions at m/z 193.0512 and 175.0396, which correspond to the successive loss of a hexose and H_2O.

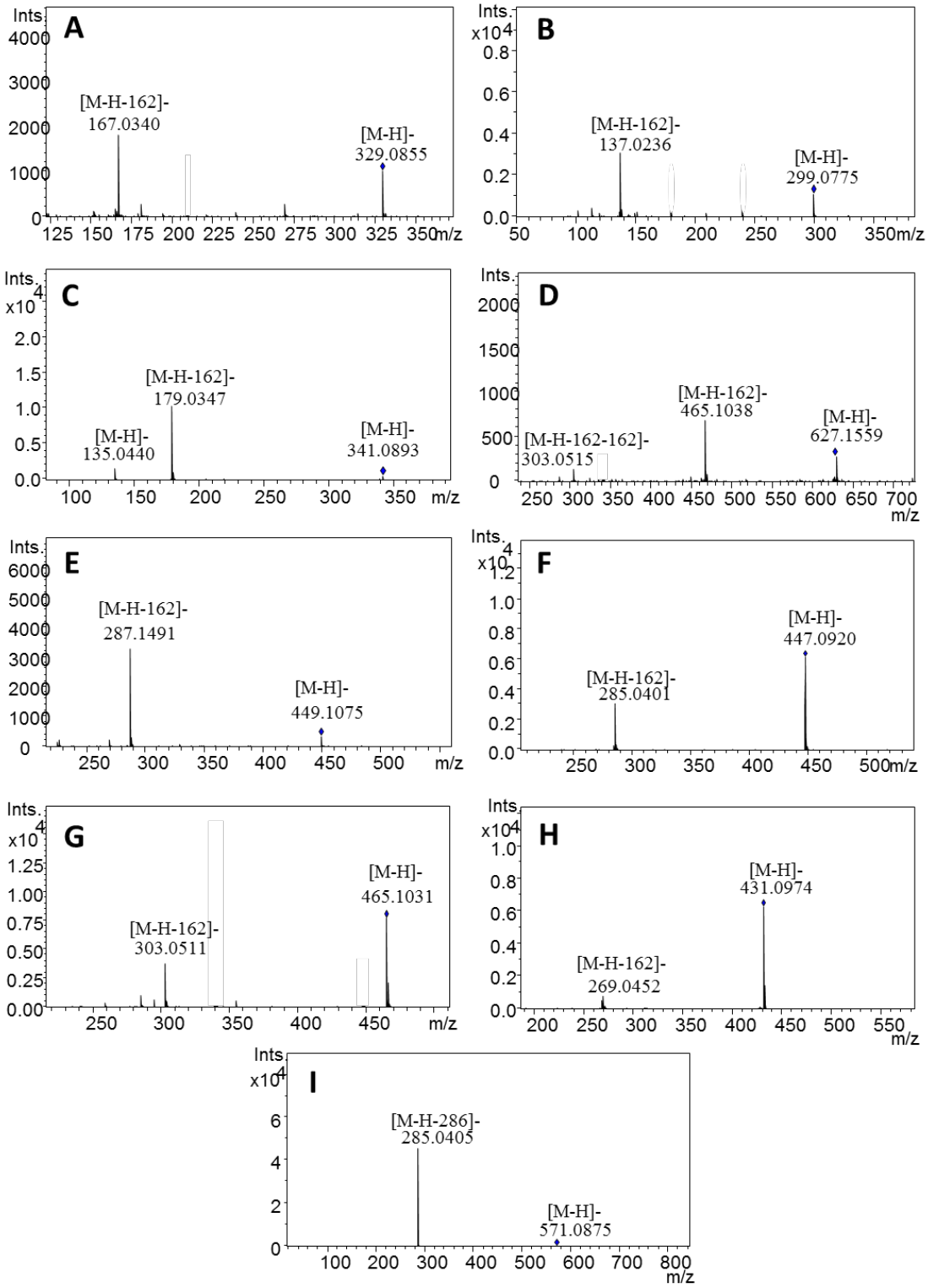


Figure 3. MS/MS spectra of some characterized phenolic glycosides in watermelon. Figures: **A;** Vanillic acid hexoside, **B;** Salicylic acid *O*-hexoside, **C;** Caffeoylhexose, **D;** Taxifolin dihexoside (Calodendroside A), **E;** Eriodictyol 7-glucoside or Dihydrokaempferol 7-glucoside, **F;** Luteolin-*O*-hexoside, **G;** Taxifolin-*O*-hexoside, **H;** Apigenin-*O*-hexoside, and **I;** Dihydrophilonotisflavone.

The MS/MS fragmentation of compound **11** (m/z at 517.1556) originated a fragment ion at m/z 193.0517. This compound, based on MS data and the MS/MS fragmentation pattern, was suggested to be 3-*O*-feruloylsucrose. Moreover, this compound has been detected to be abundant in highly reddish beet roots[28].
Similarly, at retention times 12.56 and 14.24 min, two isomers were detected at m/z 341.0877 with identical formula $C_{15}H_{17}O_9$. In accordance with the literature[23,27] and based on data from QTOF-MS analysis, these compounds were identified as caffeoylhexoses (**Figure 3C**).

Four dicaffeoylshikimic acids (**21, 28, 34,** and **45**) at the retention times 13.57, 15.92, 17.29, and 19.09 min were tentatively characterized in the QTOF-MS using negative ionisation mode. The MS/MS spectrum of the precursor ion at m/z 497.1285 revealed the product ions at m/z 335.0768 and 179.0348, which correspond to the loss of caffeoyl moiety and CO_2, in accordance with MS/MS data previously reported in *Helichrysum obconicum*[26].

In agreement with our previous report[23], peak **32** was proposed as sinapic acid glucoside (**Figure 2d**). The data obtained from QTOF-MS analysis showed a molecular ion at m/z 385.1134 and daughter ions at m/z 223.0619 and 205.0506, which correspond to a successive loss of hexose moiety and H_2O, respectively.

The spectrum in the negative ionization mode also showed two deprotonated molecular ions at m/z 471.1510 (compounds **37** and **49**) and a MS/MS fragment ion at m/z 193.0497, this latter ion may be originated from the neutral loss of arabinose and xylose moieties (-278 Da). Depending on the data obtained from QTOF-MS, together with the accepted error, mSigma values, and data from the literature[29], compounds **37** and **49** were tentatively identified as *O*-(trans)feruloyl-arabinofuranosyl-xylopyranose.

On the other hand, compounds **40** and **50** ($C_{16}H_{15}O_8$) detected at m/z 335.0749 were tentatively suggested to be *O*-caffeoylshikimic acid (**Figure 2e**). Additionally, MS/MS fragments of this latter compound were consistent with previous reports on Cucurbitaceae (cucumber and melon plants)[27].

Finally, *O*-feruloylquinide was tentatively proposed for peak **59**, which possesses the molecular ion at m/z 349.0931 —that is, this compound has been described previously in Cucurbitaceae[27].

Flavonoid derivatives

In relation to the family of flavonoid derivatives, 23 compounds with their isomers have been characterized using the proposed method in the watermelon extract analysed.

Thus, peaks **31** and **56** showed identical molecular formula $C_{27}H_{29}O_{15}$. The MS data were acceptable, but because of the low peak intensity it was difficult to find fragments in the MS/MS analysis except peak **31,** which displayed the fragment ion at m/z 431.0987. Additionally, this compound with same exact mass and formula has been detected in Cucurbitaceae[23]. Accordingly, both compounds were identified as kaempferol hexoside-rhamnoside.

In addition, peak **35** ($C_{28}H_{31}O_{16}$) with m/z at 623.1596, was suggested to be either isorhamnetin 3-*O*-rutinoside or lucenin-2-methyl ether, noting that both compounds have been previously reported in gourds[23]. In fact, a very low-intensity spectrum was detected in the MS/MS spectrum analysis for this compound, with which it was not possible to distinguish between the two isomers.

With the molecular ion at m/z 609.1488 and the formula $C_{27}H_{29}O_{16}$ (peak **42**), one signal was detected in the QTOF-MS analysis. In the MS/MS spectrum a fragment ion was revealed at m/z 447.0911 corresponding to quercetin hexoside. In this sense, according to the literature, this compound was tentatively considered to be quercetin 3-rutinoside (rutin) (**Figure 2f**)[23,30].

On the other hand, taxifolin dihexoside (Calodendroside A) was proposed according to its precursor ion at m/z 627.1557. The two fragment ions at m/z 465.1038 and 303.0515 could be interpreted as a successive loss of two hexose moieties [M-H-162-162]−, resulting in the appearance of taxifolin aglycone with a m/z at 303.0515 (**Figure** 3D).

At retention times 18.44 and 23.44 min, the precursor ion at m/z 449.1097 was detected. The MS/MS spectrum has shown a product ion at m/z 287.0786, which corresponds to a neutral loss of glucose moiety [M-H-162]− (**Figure 3E**). Therefore, these compounds were proposed to be either eriodictyol 7-glucoside or dihydrokaempferol 7-glucoside. It was not possible to distinguish between these compounds, since no commercial standards were available for all the compounds detected in the present study. Similarly, the appearance of the fragment ion at m/z 271.0722 (naringenin) in the peak **48**, was associated with the loss of disaccharides [M-H-162-146]−. Therefore, this compound has been tentatively considered to be either naringenin 7-neohesperidoside or naringenin 7-rutinoside.

The information provided from the MS/MS spectra made it possible to differentiate between isomers. In this context, six signals (**47**, **51**, **57**, **63**, **65**, and **69**) were detected, showing identical m/z (447.0919) and the molecular formula ($C_{21}H_{19}O_{11}$).

Therefore, luteolin 6-C-glucoside (isoorientin) was postulated for peak **47**, depending on the MS data and the fragmentation pattern in the MS/MS spectrum which exhibited the product ions at m/z 429.0783, 357.0591, and 327.0519, indicating *C*-glycoside fragmentation pattern by the loss of [M-H-18]−, [M-H-90]−, and [M-H-120]−[31]. The fragmentation pattern found for each ion was useful in confirming the proposed structure (**Figure 4**). Additionally, this compound has been previously reported in Cucurbitaceae[32]. On the other hand, the MS/MS spectrum of peaks **51**, **57** and **69** showed a fragment ion at m/z 285.0413, which corresponds to a loss of hexose moiety [M-H-162]−. By reference to the literature[33], these isomers were assigned as luteolin-*O*-hexoside (**Figure 3F**).

Peak **63** ($C_{21}H_{19}O_{11}$), which was detected at 25.14 min, showed the daughter ion at m/z 301.0309, which corresponds to the presence of quercetin in the structure; achieved by the loss of the rhamnose moiety [M-H-146]−[23]. Thus, this compound was characterized as quercetin rhamnoside (quercitrin) (**Figure 2g**).

The product ion detected at m/z 303.0511, (peak **52**) corresponds to taxifolin, which resulted from the neutral loss of a hexose moiety [M-H-162]− (**Figure 3G**).

Therefore, this compound with the main ion at m/z 465.1031 was tentatively proposed to be taxifolin-*O*-hexoside.

It bears mentioning that peak **54** ($C_{28}H_{35}O_{13}$) was tentatively considered to be 2″-*O*-pentosyl-6-C-hexosyl-luteolin. This fragmentation pattern for *C*-glycosyl and *O*-glycosylated flavones was previously reported elsewhere in the literature[34].

As for compounds **55**, **64**, and **66** at 21.96, 25.44, and 25.71 min, all gave the [M-H]− ions at m/z 431.0973, having the identical molecular formula ($C_{21}H_{19}O_{10}$).

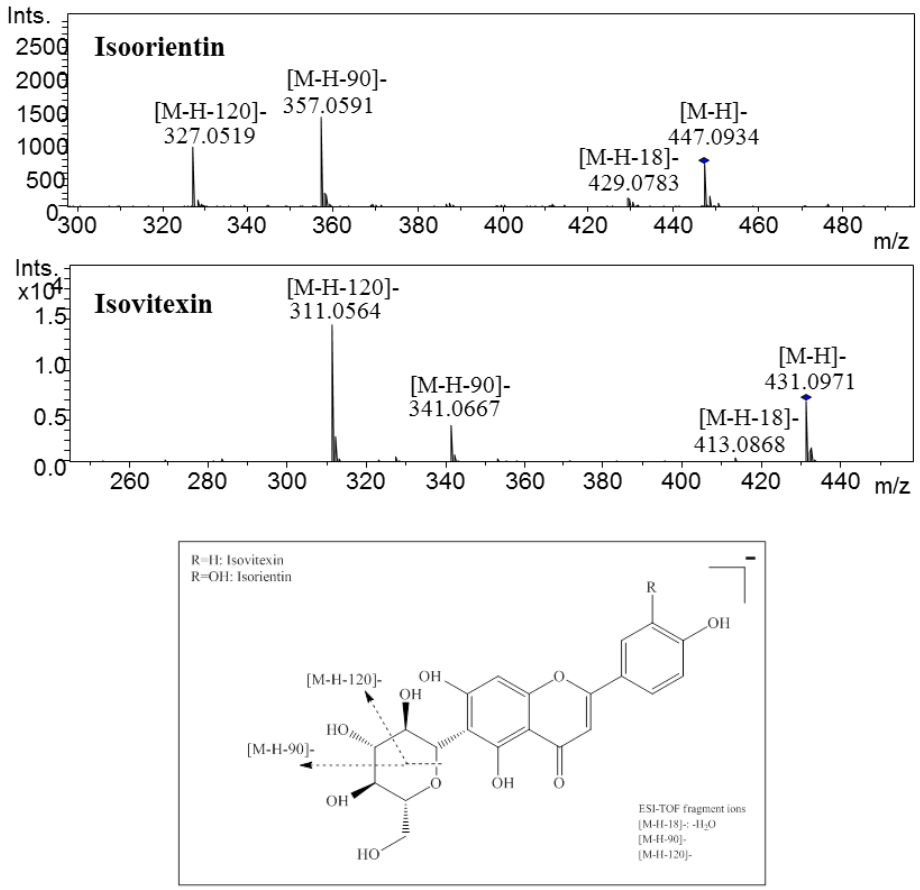


Figure 4. Proposed fragmentation pattern of Isovitexin and Isoorientin (6-C-glycosides) together with their MS/MS spectra.

In this sense, in the MS/MS spectrum, compound **55** has displayed the fragment ions at m/z 413.0868, 341.0667 and 311.0564 in accord with the loss of [M-H-18]−, [M-H-90]−, and [M-H-120]−. A comparison of these MS/MS ions with the literature made it possible to differentiate vitexin from its iso-group[35]. Moreover, this

compound has previously been reported in Cucurbitaceae[32]. Thus, accordingly, this compound was suggested to be isovitexin[35](**Figure 4**).

Table 1. Tentatively characterized compounds in the polar fraction of watermelon

Peak No.	*Rt* (min)	Observed *m/z*	Calculated *m/z*	Empirical formula	Error (ppm)	mSigma	Major MS/MS fragments	Proposed Compound	Reference
Amino acid									
1	2.22	174.0885	174.0884	$C_6H_{12}N_3O_3$	-0.3	1.7	131.0828	Citrulline	40
Hydroxybenzoic acid derivatives									
3	8.23	315.0715	315.0722	$C_{13}H_{15}O_9$	2.2	38.8	153.0195, 109.0277	Protocatechuic acid glucoside I	21
4	8.62	315.0730	315.0722	$C_{13}H_{15}O_9$	-2.8	12.7	153.0177, 152.0115, 109.0261	Protocatechuic acid glucoside II	21
6	9.84	301.0571	301.0565	$C_{12}H_{13}O_9$	-2.0	41.7	125.0235	Phloroglucinol glucuronide	-
8	10.48	299.0770	299.0772	$C_{13}H_{15}O_8$	0.7	1.3	137.0236	Salicylic acid-*O*-hexoside I	24
9	10.68	315.1072	315.1085	$C_{14}H_{19}O_8$	4.3	5.0	151.0397	Vanillin hexoside I	-
12	11.30	659.1826	659.1829	$C_{28}H_{35}O_{18}$	0.5	15.7	497.1281, 335.0764	Tri-*O*-caffeoylshikimic acid I	26
13	11.50	315.1070	315.1085	$C_{14}H_{19}O_8$	4.9	10.2	151.0400, 109.0304	Vanillin hexoside II	-
15	11.64	329.0880	329.0878	$C_{14}H_{17}O_9$	-0.6	9.2	167.0349	Vanillic acid hexoside	23
19	13.23	299.0767	299.0772	$C_{13}H_{15}O_8$	1.7	2.9	137.0235	Salicylic acid-*O*-hexoside II	23,24
39	18.20	389.1249	389.1242	$C_{20}H_{21}O_8$	-1.7	35.3	Nd	Salicin 2-benzoate	-
70	28.58	315.1087	315.1085	$C_{14}H_{19}O_8$	-0.4	16.1	151.0410	Vanillin hexoside III	-
Hydroxycinnamic acid derivatives									
5	9.67	325.0927	325.0929	$C_{15}H_{17}O_8$	0.4	21.1	163.0416, 145.0296	*p*-coumaric acid glucoside I	23,25
10	10.82	355.1028	355.1035	$C_{16}H_{19}O_9$	1.8	5.5	193.0512, 175.0396	Ferulic acid hexoside I	23,27
11	10.85	517.1556	517.1563	$C_{22}H_{29}O_{14}$	1.3	11.6	193.0517	3-*O*-feruloylsucrose	28
16	11.91	325.0916	325.0929	$C_{15}H_{17}O_8$	3.9	33.5	163.0423	*p*-coumaric acid glucoside II	23,25
18	12.56	341.0877	341.0878	$C_{15}H_{17}O_9$	0.3	1.7	179.0351, 135.0421	Caffeoylhexose I	23,27
20	13.43	461.1646	461.1664	$C_{20}H_{29}O_{12}$	4.1	7.4	133.0502	Decaffeoylacetoside or Descaffeoylverbascoside	-
21	13.57	497.1278	497.1301	$C_{22}H_{25}O_{13}$	4.5	2.0	335.0767, 179.0344	Dicaffeoylshikimic acid I	26
22	13.60	355.1021	355.1035	$C_{16}H_{19}O_9$	3.8	8.8	193.0505, 175.0379	Ferulic acid hexoside II	23,27
23	14.24	341.0893	341.0878	$C_{15}H_{17}O_9$	-4.3	18.0	179.0347, 135.0440	Caffeoylhexose II	23,27
25	15.10	325.0926	325.0929	$C_{15}H_{17}O_8$	0.8	13.2	163.0392, 145.0292	*p*-coumaric acid glucoside III	23,25
28	15.92	497.1289	497.1301	$C_{22}H_{25}O_{13}$	2.3	7.4	335.0768, 179.0348	Dicaffeoylshikimic acid II	26
30	16.32	355.0330	355.1035	$C_{16}H_{19}O_9$	0.5	3.0	193.0502, 175.0398	Ferulic acid hexoside III	23,27
32	16.59	385.1134	385.1140	$C_{17}H_{21}O_{10}$	1.5	3.4	223.0619, 205.0506	Sinapic acid glucoside	23

33	17.11	355.1027	355.1035	$C_{16}H_{19}O_9$	2.3	5.1	193.0509, 175.0402	Ferulic acid hexoside IV	23,27
34	17.29	497.1285	497.1301	$C_{22}H_{25}O_{13}$	3.1	7.8	335.0786, 179.034	Dicaffeoylshikimic acid III	26
37	18.15	471.1510	471.1508	$C_{21}H_{27}O_{12}$	-0.5	21.3	193.0497	*O*-(trans)feruloyl-arabinofuranosyl-xylopyranose I	29
40	18.30	335.0749	335.0772	$C_{16}H_{15}O_8$	-1.0	6.5	179.0351, 161.0272, 135.0451	*O*-Caffeoylshikimic acid I	36,27
45	19.09	497.1292	497.1301	$C_{22}H_{25}O_{13}$	1.8	10.6	335.0769, 179.0361	Dicaffeoylshikimic acid IV	26
49	20.48	471.1497	471.1508	$C_{21}H_{27}O_{12}$	2.3	7.8	193.0489	*O*-(trans)feruloyl-arabinofuranosyl-xylopyranose II	29
50	20.52	335.0771	335.0772	$C_{16}H_{15}O_8$	0.3	36.0	179.0330, 161.0272, 135.0473	*O*-Caffeoylshikimic acid II	36,27
59	22.94	349.0931	349.0929	$C_{17}H_{17}O_8$	-0.5	23.5	193.0502	*O*-feruloylquinide	27
Flavonoids derivatives									
31	16.56	593.1527	593.1512	$C_{27}H_{29}O_{15}$	-2.6	23.3	431.0987	Kaempferol rhamnoside-hexoside I	23
35	18.11	623.1596	623.1618	$C_{28}H_{31}O_{16}$	3.4	53.8	Nd	Isorhamnetin 3-*O*-rutinoside or Lucenin-2-methyl ether	23
41	18.44	449.1097	449.1089	$C_{21}H_{21}O_{11}$	-1.7	12.3	287.0786	Eriodictyol 7-glucoside or Dihydrokaempferol 7-glucoside I	-
42	18.46	609.1488	609.1461	$C_{27}H_{29}O_{16}$	-4.4	40.7	447.0911	Quercetin 3-rutinoside (rutin)	23,30
44	18.54	627.1557	627.1567	$C_{27}H_{31}O_{17}$	1.6	16.3	465.1038, 303.0515	Taxifolin dihexoside (Calodendroside A)	-
47	19.63	447.0934	447.0933	$C_{21}H_{19}O_{11}$	6.4	12.7	357.0591, 327.0519	Luteolin 6-*C*-glucoside (Isoorientin)	31,32
48	19.85	579.1736	579.1719	$C_{27}H_{31}O_{14}$	-2.9	14.7	271.0722	Naringenin 7-neohesperidoside or Naringenin 7-rutinoside I	-
51	20.84	447.0943	447.0933	$C_{21}H_{19}O_{11}$	-2.2	31.2	285.0437	Luteolin-*O*-hexoside I	33
52	21.16	465.1031	465.1038	$C_{21}H_{21}O_{12}$	1.6	2.0	303.0511	Taxifolin-*O*-hexoside I	-
54	21.76	579.2090	579.2083	$C_{28}H_{35}O_{13}$	-1.20	10.90	357.1264, 339.1236, 327.1161	2″-*O*-pentosyl-6-*C*-hexosyl-luteolin	34
55	21.96	431.0973	431.0984	$C_{21}H_{19}O_{10}$	2.5	3.3	413.0868, 341.0667,3 23.0572, 311.0564, 269.0455	Isovitexin	32,35
56	22.70	593.1507	593.1512	$C_{27}H_{29}O_{15}$	0.8	19.6	Nd	Kaempferol rhamnoside-hexoside II	23
57	22.84	447.0919	447.0933	$C_{21}H_{19}O_{11}$	3	19.5	285.0413	Luteolin-*O*-hexoside II	33
60	23.20	513.1785	513.1766	$C_{27}H_{29}O_{10}$	-3.7	59.2	Nd	Icariside II	-
62	23.44	449.1094	449.1089	$C_{21}H_{21}O_{11}$	-1.1	23.2	287.0574	Eriodictyol 7-glucoside or Dihydrokaempferol 7-glucoside II	-
63	25.14	447.0932	447.0933	$C_{21}H_{19}O_{11}$	0.1	73.3	301.0309	Quercitrin; Quercetin rhamnoside I	23
64	25.44	431.0971	431.0984	$C_{21}H_{19}O_{10}$	2.9	21.6	269.0452	Apigenin-*O*-hexoside I	-
65	25.57	447.0920	447.0933	$C_{21}H_{19}O_{11}$	2.9	1.6	285.0404	Luteolin-*O*-hexoside III	33

66	25.71	431.0975	431.0984	$C_{21}H_{19}O_{10}$	2	15.9	269.0466	Apigenin-*O*-hexoside II	-
67	26.01	461.1076	461.1089	$C_{22}H_{21}O_{11}$	2.8	39.6	299.0577	Chrysoeriol-*O*-hexoside I	-
68	26.51	461.1074	461.1089	$C_{22}H_{21}O_{11}$	3.4	6.6	299.0547	Chrysoeriol-*O*-hexoside II	-
69	26.83	447.0925	447.0933	$C_{21}H_{19}O_{11}$	1.8	17.1	285.0413	Luteolin-*O*-hexoside IV	33
71	32.17	571.0875	571.0882	$C_{30}H_{19}O_{12}$	1.3	49.1	285.0405	Dihydrophilonotisflavone	-
Lignan									
53	21.71	521.2001	521.2028	$C_{26}H_{33}O_{11}$	5.2	14.7	359.1493	Isolariciresinol 9'-Β-D-glucopyranoside I	23
Iridoid									
17	12.02	347.1341	347.1348	$C_{15}H_{23}O_{9}$	1.9	4.3	285.0983, 183.0658, 139.0694	Ajugol	-
24	14.98	511.1435	511.1457	$C_{23}H_{27}O_{13}$	4.3	13.0	349.0916	Vanilloyl catalpol (Picroside) I	37
29	16.27	481.1910	481.1927	$C_{20}H_{33}O_{13}$	3.5	14.9	319.0810, 163.0401, 155.0344, 137.0241	Catalposide	-
43	18.47	511.1450	511.1457	$C_{23}H_{27}O_{13}$	1.3	27.4	349.0937	Vanilloyl catalpol (Picroside) II	37
Coumarin									
26	15.13	145.0294	145.0295	$C_{9}H_{5}O_{2}$	0.7	6.4	117.0326	Coumarin	38
27	15.45	303.0871	303.0874	$C_{16}H_{15}O_{6}$	1.1	8.4	255.0652	(+)-Aviprin I	-
35	17.31	303.0872	303.0874	$C_{16}H_{15}O_{6}$	0.8	7.9	255.0652	(+)-Aviprin II	-
58	22.92	455.1576	455.1559	$C_{21}H_{27}O_{11}$	-3.8	15.2	231.1017, 309.1211	Obtusoside	-
Other phenolic compounds									
2	7.12	285.0990	285.0980	$C_{13}H_{17}O_{7}$	-3.5	4.8	123.0459	Saligenin glucopyranoside	39
7	9.86	305.1032	305.1031	$C_{16}H_{17}O_{6}$	-0.3	2.8	Nd	Cimifugin	-
14	11.62	285.0617	285.0616	$C_{12}H_{13}O_{8}$	-0.3	15.0	152.0117, 108.0217, 109.0292	Hydroquinone glucuronide	-
38	18.18	551.1759	551.1770	$C_{26}H_{31}O_{13}$	2.1	39.9	Nd	Glehlinoside C	-
46	19.23	471.1476	471.1449	$C_{28}H_{23}O_{7}$	-5.7	76.0	355.0999, 193.0495	Leachianol G	-
61	23.31	287.0921	287.0925	$C_{16}H_{15}O_{5}$	1.4	19.0	257.0819, 135.0454	Shikonine	-

On the other hand, compounds **64** and **66** with m/z 431.0971 have been assigned as apigenin-*O*-hexoside isomers. Their characterization was based on MS and MS/MS data demonstrating the fragment ion at m/z 269.0452 (apigenin), which resulted from the loss of hexose moiety (-162 Da) (**Figure 3H**).

Two isomers at 26.01 and 26.51 min, displayed the same molecular formula ($C_{22}H_{21}O_{11}$). The MS/MS spectra of these compounds yielded the product ions at m/z 299.0577 indicating chrysoeriol, which corresponds to the loss of the hexose moiety [M-H-162]−. Based on MS and MS/MS data, both compounds have been tentatively characterized as chrysoeriol *O*-hexoside.

Finally, peak **71** ($C_{30}H_{19}O_{12}$) was tentatively proposed to be dihydrophilonotisflavone (**Figure 3I**), given the acceptable data from MS and MS/MS spectra.

Lignan derivatives

One lignan glycoside (compound **53** at 21.71 min) was detected in the QTOF-MS, which showed an MS/MS spectrum product ion at m/z 359.1493 (indicating isolariciresinol), this being achieved after a neutral loss of glucose moiety (-162 Da). Therefore, the precursor ion was suggested to be isolariciresinol 9'-ß-D-glucopyranoside[23].

Iridoid derivatives

Concerning the Iridoids family, four compounds were detected and characterized in watermelon.

Thus, based on the appropriate data obtained from the QTOF-MS analysis, compounds **24** and **43** ($C_{23}H_{27}O_{13}$) were tentatively characterized as vanilloyl catalpol (Picroside) (**Figure 2h**), according with its MS, MS/MS spectra, and the literature in which a similar fragmentation pattern has been reported elsewhere[37].

Peaks **17** and **29** were tentatively characterized as ajugol (**Figure 2i**) and catalposide (**Figure 2j**), respectively.

Coumarin derivatives

Peak **25** (m/z 145.0294) has been assigned as Coumarin. It should be mentioned that this compound has been previously identified in melon (Cucurbitaceae)[38].

In the same manner, two compounds (**27** and **35**) gave the [M-H]− ions at m/z 303.0871 and the identical molecular formula $C_{16}H_{15}O_6$. Consequently, these two isomers were characterized as (+)-Aviprin (**Figure 2k**). Last but not least, compound **58** (Rt 22.92 min) was tentatively assigned as obtusoside.

Other phenolic and polar compounds

Finally, saligenin glucopyranoside (**Figure 2l**) was proposed for peak **2** at 7.12 min. This compound displayed the [M-H]− at m/z 285.0990 and the product ion at m/z 123.0459 (saligenin) resulted from a loss of glucose moiety [M-H-162]−, this compound has previously been reported in Cucurbitaceae[39].

On the other hand, peak **14** at 11.62 min was suggested to be hydroquinone glucuronide (**Figure 2m**). Its characterization was based on the accepted MS data and the appearance of the fragment ion at m/z 109.0292, which correspond to hydroquinone by losing a glucuronic acid [M-H-176]−.

Finally, in the polar aqueous-methanolic extract a non-essential amino acid (Citrulline) was detected and characterized (Peak **1** at 2.22 min). It has been previously reported that watermelon is a rich source of this amino acid[40], which is suggested to have an efficient hydroxyl radical scavenger and a strong antioxidant capacity[41]. Moreover, this compound has been described in Cucurbitaceae fruits, including cucumber, bitter melon, pumpkin, muskmelon, bottle gourd, wax gourd, and dishrag gourd[40].

Other phenolic compounds not discussed in the text, are shown in (**Table 1**) with their fragmentation pattern and their tentative assignment.

Conclusion

In the present study, the coupling of HPLC with ESI-QTOF-MS has been demonstrated to be a useful analytical technique for separating and characterizing phytochemical compounds (mainly phenolic compounds) in *Citrullus lanatus*. Notably, 71 phenolic and polar compounds were tentatively identified by the method proposed using the accurate mass determination of the deprotonated ions (MS data and MS/MS fragmentation pattern). The present report represents a preliminary study for the phytochemicals in watermelon and also marks the first characterization of these compounds in watermelon pulp. In this context, the data compiled indicate that watermelon is qualitatively a rich source of phenolic compounds such as flavonoids. Moreover, the data may also be useful for improving food-composition tables.

References

1 Meskin, M. S., Bidlack, W. R., Davies, A. J. and Omaye, S. T. 2002. *Phytochemicals in Nutrition and Health* Taylor and Francis.

2 Watson, R. R., and Preedy, V. R. 2010. *Fruit and Vegetables in Health Promotion: Fruits and Vegetables.* Academic Press.

3 Shahidi, F. and Naczk, M. 1995. *Food Phenolics: Sources, Chemistry, Effects, Applications* Technomic Publishing Company.

4 SIMPSON, M. G. *Plant Systematics* Elsevier Science.

5 Vaughan, J. G. and Geissler, C. A. 2009. *The New Oxford Book of Food Plants*. Oxford University Press.

6 FAOSTAT. 2010. Retrieved on, December 01, 2012, from; http://faostat3.fao.org/home/index.html#SEARCH_DATA.

7 Erukainure, O. L., Oke, O. V., Daramola, A. O., Adenekan, S. O., and Umanhonlen, E. E. 2010. Improvement of the Biochemical Properties of Watermelon Rinds Subjected to Saccharomyces cerevisae Solid Media Fermentation. *Pakistan Journal of Nutrition,* **9,** 806-809.

8 Hassan, L. E. A., Koko, W. S., Osman, E-B. E., Dahab, M. M., and Sirat, H. M. 2011. In vitro antigiardial activity of Citrullus lanatus Var. citroides extracts and cucurbitacins isolated Compounds. *Journal of Medicinal Plants Research,* **5,** 3338-3346.

9 Figueroa, A., Sanchez-Gonzalez, M. A., Wong, A., and Arjmandi, B. H. 2012. Watermelon Extract Supplementation Reduces Ankle Blood Pressure and Carotid Augmentation Index in Obese Adults With Prehypertension or Hypertension. *American Journal of Hypertension,* **25,** 640-643.

10 Jeffrey, C. 1990. *Systematics of the Cucurbitaceae: An overview*. In D. M. Bates, R.W. Robinson, and C. Jefferey (Eds.), Biology and utilization of the Cucurbitaceae (pp. 3–28). Cornell University Press.

11 Tarazona-Díaz, M. P., Viegas, J., Moldao-Martins, M., and Aguayo, E. 2011. Bioactive compounds from flesh and by-product of fresh-cut watermelon cultivars. *Journal of the Science of Food and Agriculture*, **91,** 805-812.

12 Tlili, I., Hdider, C., Lenucci, M. S., Ilahy, R., Jebari, H., and Dalessandro, G. 2011. Bioactive compounds and antioxidant activities during fruit ripening of watermelon cultivars. *Journal of Food Composition and Analysis*, **24,** 923-928.

13 Sharma, S., paliwal, S., Dwivedi, J., and Tilak, A. 2011. First report on laxative activity of Citrullus lanatus. *Pharmacologyonline,* **2,** 790-797.

14 Ismail, H. I., Chan, K. W., Mariod, A. A., and Ismail, M. 2010. Phenolic content and antioxidant activity of cantaloupe (Cucumis melo) methanolic extracts. *Food Chemistry,* **119,** 643-647.

15 Tarazona-Díaz, M. P., Viegas, J., Moldao-Martins, M., and Aguayo, E. 2011. Bioactive compounds from flesh and by-product of fresh-cut watermelon cultivars. *Journal of the Science of Food and Agriculture*, **91,** 805-812.

16 Tlili, I., Hdider, C., Lenucci, M. S., Ilahy, R., Jebari, H., and Dalessandro, G. 2011. Bioactive compounds and antioxidant activities during fruit ripening of watermelon cultivars. *Journal of Food Composition and Analysis*, **24,** 923-928.

17 Crozier, A., Jaganath, I. B., and Clifford, M. N. 2009. Dietary phenolics: Chemistry, bioavailability and effects on health. *Natural Product Reports,* **26,** 1001-1043.

18 Almela, L., Sánchez-Muñoz, B., Fernández-López, J. A., Roca, M. J., and Rabe, V. 2006 Liquid chromatograpic-mass spectrometric analysis of phenolics and free radical scavenging activity of rosemary extract from different raw material. *Journal of Chromatography A,* **1120,** 221-229.

19 Borrás Linares, I., Arráez-Román, D., Herrero, M., Ibáñez, E., Segura-Carretero, A. and Fernández-Gutiérrez, A. 2011. Comparison of different extraction procedures for the comprehensive characterization of bioactive phenolic compounds in *Rosmarinus officinalis* by reversed-phase high-performance liquid chromatography with diode array detection

coupled to electrospray time-of-flight mass spectrometry. *Journal of Chromatography A,* **1218,** 7682-7690.

20 Abu-Reidah, I. M., Arráez-Román, D., Lozano-Sánchez, J., Segura-Carretero, A., and Fernández-Gutiérrez, A. 2012. Phytochemical Characterization of Green Beans (Phaseolus vulgaris L.) by Using High-performance Liquid Chromatography Coupled with Time-of-flight Mass Spectrometry. *Phytochemical Analysis*, **24,** *105–116.*

21 Rodríguez-Medina, I. C., Segura-Carretero, A., and Fernández-Gutiérrez, A. 2009. Use of high-performance liquid chromatography with diode array detection coupled to electrospray-Qq-time-of-flight mass spectrometry for the direct characterization of the phenolic fraction in organic commercial juices. *Journal of Chromatography A,* **1216,** 4736-4744.

22 Tachi, I., Hayashi, A., and Sato, A. 1957. Biosynthesis of lignin. I. Changes in the contents of lignin and other chemical constituents during growth of bottle gourds. *Bulletin of the Agricultural Chemical Society of Japan,* **21,** 279-283.

23 Abu-Reidah, I. M., Arráez-Román, D., Quirantes-Piné, R., Fernández-Arroyo, S., Segura-Carretero, A., and Fernández-Gutiérrez, A. 2012 HPLC-ESI-QTOF-MS for a comprehensive characterization of bioactive phenolic compounds in cucumber whole fruit extract. *Food Research International,* **46,** 108-117.

24 Zhang, P. Y., Wang, J. C., Liu, S. H., and Chen, K. S. 2009. A Novel Burdock Fructo-oligosaccharide Induces Changes in the Production of Salicylates, Activates Defence Enzymes and Induces Systemic Acquired Resistance to Colletotrichum orbiculare in Cucumber Seedlings. *Journal of Phytopathology,* **157,** 201-207.

25 Billett, E. E., Grayer-Barkmeijer, R. J., Johnson, C. B., and Harborne, J. B. 1981. The effect of blue light on free and esterified phenolic acids in etiolated gherkin tissues. *Phytochemistry,* **20,** 1259-1263.

26 Gouveia, S., and Castilho, P. C. 2011 Characterisation of phenolic acid derivatives and flavonoids from different morphological parts of Helichrysum obconicum by a RP-HPLC-DAD-(-)-ESI-MSn method. *Food Chemistry,* **129,** 333-344.

27 Bellés, J. M. a., López -Gresa, M. a. P., Fayos, J. n., Pallás, V., Rodrigo, I., and Conejero, V. 2008. Induction of cinnamate 4-hydroxylase and phenylpropanoids in virus-infected cucumber and melon plants. *Plant Science,* **174,** 524-533.

28 Kujala, T., Vienola, M., Klika, K., Loponen, J., and Pihlaja, K. 2002. Betalain and phenolic compositions of four beetroot (Beta vulgaris) cultivars. *European Food Research and Technology,* **214,** 505-510.

29 Hosny, M., and Rosazza, J. P. N. 1997. Structures of Ferulic Acid Glycoside Esters in Corn Hulls, *Journal of Natural Products*. **60,** 219-222.

30 Cai, J., and Wang, W. 2005. Study on content of total flavonoids in leaves of cucumber. Shipin Kexue (Beijing, China). *Food Science,* **8,**194-197.

31 Gattuso, G., Caristi, C., Gargiulli, C., Bellocco, E., Toscano, G., and Leuzzi, U. 2006. Flavonoid Glycosides in Bergamot Juice (Citrus bergamia Risso). *Journal of Agricultural and Food Chemistry*, **54,** 3929-3935.

32 Krauze-Baranowska, M., and Cisowski, W. 1994. High-performance liquid chromatographic determination of flavone C-glycosides in some species of the Cucurbitaceae family. *Journal of Chromatography A,* **675,** 240-243.

33 Siciliano, T., De Tommasi, N., Morelli, I., and Braca, A. 2004. Study of Flavonoids of Sechium edule (Jacq) Swartz (Cucurbitaceae) Different Edible Organs by Liquid Chromatography Photodiode Array Mass Spectrometry. *Journal of Agricultural and Food Chemistry,* **52,** 6510-6515.

34 Ferreres, F., Gil-Izquierdo, A., Andrade, P. B., Valentão, P., and Tomás-Barberán, F. A. 2007. Characterization of C-glycosyl flavones O-glycosylated by liquid chromatography-tandem mass spectrometry. *Journal of Chromatography A,* **1161,** 214-223.

35 Pikulski, M., and Brodbelt, J. S. 2003. Differentiation of flavonoid glycoside isomers by using metal complexation and electrospray ionization mass spectrometry. *Journal of the American Society for Mass Spectrometry,* **14,** 1437-1453.

36 Bastos, D., Saldanha, L., Catharino, R., Sawaya, A., Cunha, I., Carvalho, P., and Eberlin, M. 2007. Phenolic Antioxidants Identified by ESI-MS from Yerba Maté (Ilex paraguariensis) and Green Tea (Camelia sinensis) Extracts. *Molecules*, **12,** 423-432.

37 Hong, J.-L., Qin, X.-Y., Shu, P., Wu, G., Wang, Q., and Qin, M.-J. 2010. Analysis of catalpol derivatives by characteristic neutral losses using liquid chromatography combined with electrospray ionization multistage and time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, **24,** 2680-2686.

38 Arora, R., Kaur, M., and Gill, N. S. 2011. Antioxidant activity and pharmacological evaluation of Cucumis melo var. agrestis methanolic seed extract. *Research Journal of Phytochemistry,* **5,** 146-155.

39 Moing, A., Aharoni, A., Biais, B., Rogachev, I., Meir, S., Brodsky, L., et al. 2011.Extensive metabolic cross-talk in melon fruit revealed by spatial and developmental combinatorial metabolomics. *New Phytologist,* **190,** 683-696.

40 Rimando, A. M., and Perkins-Veazie, P. M. 2005. Determination of citrulline in watermelon rind. *Journal of Chromatography A,* **1078,** 196-200.

41 Fang, Y.-Z., Yang, S., and Wu, G. 2002. Free radicals, antioxidants, and nutrition. *Nutrition*, **18,** 872-879.

Chapter 4

Extensive characterization of bioactive phenolic constituents from Globe Artichoke (*Cynara scolymus* L.) by HPLC-DAD-ESI-QTOF-MS

Ibrahim M. Abu-Reidah, David Arráez-Román, Antonio Segura-Carretero, Alberto Fernández-Gutiérrez

This work has been submitted to:
Food Chemistry

Extensive characterization of bioactive phenolic constituents from Globe Artichoke (*Cynara scolymus* L.) by HPLC-DAD-ESI-QTOF-MS

Abstract

The aim of this work was to characterize the phenolic compounds in artichoke (edible flower heads) by using HPLC coupled to DAD-ESI-QTOF-MS, which proved useful in characterizing 61 phenolic and other polar compounds. Notably, of the 61 compounds characterized, 33 new phenolic compounds with their isomers have been tentatively characterized in artichoke for the first time, namely: 3 hydroxybenzoic acids, 17 hydroxycinnamic acids, 4 lignans, 7 flavones, 2 flavonols, and one phenol derivative. Moreover, a total of 28 isomers of previously described phenolics have also been detected. The data compiled from the qualitative polyphenol characterization indicate that the artichoke extract analyzed (Blanca de Tudela variety) could be regarded as a bioactive functional food and also as a promising source of antioxidant phenolic compounds.

Keywords: *Artichoke (Cynara scolymus L.), Asteraceae, phenolic compounds, flavonoids, Mediterranean Diet, HPLC-DAD-ESI-QTOF-MS.*

Introduction

The incidence of many chronic disorders, such as cardiovascular diseases and certain cancers, could be reduced by improved nutrition, particularly through increased consumption of diets rich in fruit and vegetables. Such health-promoting properties of edible plants are related mainly to the presence of secondary metabolites, known as phytochemicals[1]. Phytochemicals in fruits and vegetables are extremely diverse, including ascorbic acid, carotenoids, and phenolic compounds, among others[2]. Phenolic compounds are plant secondary metabolites, which defend the plants against stress, whether biotic (i.e. bacteria, fungi, viruses) or abiotic (i.e. temperature, ultraviolet radiation), and also participate in plant growth and reproduction. In general, phenolic compounds are abundant in the plant kingdom and are divided into several classes, i.e. hydroxybenzoic acids, hydrocinnamic acids, phenolic acids, flavonoids (flavonols, flavones, flavanols, and flavanones), and lignans[2].

The domesticated or globe artichoke (*Cynara scolymus* L.), belonging to the family Asteraceae, is a traditionally consumed vegetable in the Mediterranean area[3], where its commercial production contributes substantially to the agro-economy. The edible fractions of artichoke plants include the inner bracts and receptacle ("heart") of the immature flowers (heads, buds or capitula), which are consumed fresh, canned or frozen worldwide. The tender inner parts constitute nearly 35-55% of the fresh weight of the head, depending on the variety and the harvesting time[4].

In comparison to other vegetables, artichoke contains high levels of total polyphenols[4-6]. Nevertheless, globe artichoke is also regarded a promising source of biopharmaceuticals, such as luteolin and mono-/di-caffeoylquinic acids that are responsible for its therapeutic effects[7]. Also, the antioxidant properties of artichoke are thought to be related to its abundant phenolic composition[8]. For all that the above, artichoke is considered a functional food.

Many studies have demonstrated that artichoke has major medicinal properties, including antioxidative, anticarcinogenic, antigenotoxic, cholesterol-lowering,

hepatoprotective, bile-expelling, diuretic, and anti-inflammatory, as well as antifungal, anti-HIV, and antibacterial[8- 15].

In addition, the biological activities of artichoke have been reported in various studies, mainly the strong antioxidative effects, which are attributed to caffeoylquinic acid derivatives, and flavonoids such as luteolin glycosides. Caffeoylquinic acids are present in artichoke as mono- and dicaffeoyl esters, its isomers being described as the most abundant molecules[7,16]. Regarding bioavailability data, *in vivo* studies have demonstrated that the caffeoyl esters are adsorbed, metabolized, and excreted as methylated phenolic acids[7,8,16].

While the action mechanisms of the plant and its active principles are not fully known, caffeoylquinic acids and flavonoids seem to play major roles in its pharmacological properties[17]. Although the chemical components of artichoke leaves have been studied extensively[10,13,16,18], research on the edible portion of the artichoke head remains incomplete.

Several analytical methods have been reported for separating and determining phenolic compounds in artichoke[6,17,19-21]. Nevertheless, methodological improvement is still needed for an extensive characterization of the phenolic compounds in the edible part of artichoke flower heads. Thus, in the present work, a short, quick, and robust High Performance Liquid Chromatography coupled to Diode Array Detector-Electrospray Ionization Quadrupole Time-of-Flight Mass Spectrometry (HPLC-DAD-ESI-Q-TOF-MS) method has been developed in an attempt to fully characterize the phenolic compounds present in artichoke from Blanca de Tudela variety, to provide some evidence for the study of the functional aspects of this important traditional plant food, as well as to gain a better understanding of their potential use in human health. Moreover, this study could be of use for improving the food composition data and tables.

Materials and Methods

Chemicals and reagents

Acetonitrile and methanol of analytical or HPLC grade were purchased from Labscan (Dublin, Ireland). Acetic acid of analytical grade (assay>99.5%) was purchased from Fluka (Switzerland). Water was purified by using a Milli-Q system (Millipore, Bedford, USA). Other unmarked reagents were of analytical grade.

Plant material

Artichoke (cv. Blanca de Tudela) samples, purchased from a commercial market, were from a crop cultivated in Granada (Spain) in the year 2011. After being purchased, the samples were transported directly to the lab, washed with distilled water and stored at 4°C until used.

Sample treatment

Fresh artichoke samples were crushed and frozen at -25°C and afterwards were lyophilized using (Christ Alpha 1–2 LD Freeze dryer, Shropshire, UK) shelf, which was pre-cooled to -50°C for 1 h at 1 mbar.

Extraction of phenolic compounds

The extraction procedure was the same as that used by Abu-Reidah et al.[22] with some modifications. Thus, 0.5 g of lyophilized artichoke was extracted using 16 mL of 80:20 (v/v) methanol/H_2O and sonicated for 30 min at room temperature (22°C). Then, the mixture was centrifuged for 15 min at 3800 x *g* and the supernatant was collected in a round-bottom flask. Subsequently, the solvent was evaporated using a rotary evaporator under vacuum at 40°C and the dry residue was resolved in 0.5 mL of 80:20 (v/v) methanol/H_2O. Finally, the extract was centrifuged again and the supernatant was filtered through a 0.2 µm syringe filter and stored at -20°C until analyzed.

HPLC-DAD-ESI-QTOF-MS analysis

Analyses were performed using an Agilent 1200 series Rapid Resolution LC (Agilent Technologies, CA, USA) consisting of a vacuum degasser, an auto-sampler and a binary pump and diode array detection (DAD) system. This instrument was equipped with an Agilent Zorbax C18 column (4.6×150 mm, 1.8 µm) from Agilent Technologies. Acidified water (0.5% acetic acid, v/v) and acetonitrile were used as mobile phases A and B, respectively. The gradient was programmed as follows: 0 min, 0% B; 10 min, 20% B; 15 min, 30% B; 20 min, 50% B; 25 min, 75% B; 30 min, 100% B; 37 min 0% B; and finally, the initial conditions were maintained for 3 min as a re-equilibration step. The flow rate was set at 0.80 mL/min throughout the gradient. The effluent from the HPLC column was split using a T-type splitter before being placed in the mass spectrometer (split=1:3). Thus the flow which arrived to the ESI-Q-TOF-MS detector was 0.2 mL/min. The column temperature was maintained at 25°C and the injection volume was 5 µL.

The HPLC system was coupled to Quadrupole-Time-of-Flight (micrOTOF-Q™, Bruker Daltonik GmbH, Bremen, Germany), an orthogonal accelerated Q-TOF mass spectrometer, equipped with an electrospray ionization source (ESI). Analysis parameters were set using a negative ion mode with spectra acquired over a mass range from m/z 50 to 1100. The optimum values of the ESI-MS parameters were: capillary voltage, +4.0 kV; drying gas temperature, 190 °C; drying gas flow, 9.0 L/min; nubilizing gas pressure, 29 psi; collision RF, 150 Vpp; transfer time 70 µs, and pre-pulse storage, 5 µs. Moreover, automatic MS/MS experiments were performed adjusting the collision energy values as follows, using nitrogen as collision gas: m/z 100, 20 eV; m/z 500, 30 eV; m/z 1000, 35 eV

The MS and MS/MS data were processed through Data Analysis 4.0 software (Bruker Daltonics, Bremen, Germany) which provided a list of possible elemental formulas by using the Generate Molecular Formula™ editor. The editor uses a CHNO algorithm, which provides standard functionalities such as minimum/maximum elemental range, electron configuration, and ring-plus double-bond equivalents, as well as a sophisticated comparison of the theoretical with the measured isotope

pattern (mSigma value) for increasing the confidence in the suggested molecular formula. The widely accepted accuracy for confirmation of elemental compositions has been established to be 5 ppm[23]. During the development of the HPLC method, an external instrument calibration was performed using a Cole Palmer syringe pump (Vernon Hills, Illinois, USA) directly connected to the interface, passing a solution of sodium acetate cluster. With this method, an exact calibration curve based on numerous cluster masses was drawn. Due to the compensation of temperature drift in the Q-TOF, this external calibration provided accurate mass values for a complete run without the need for a dual sprayer set up for internal mass calibration.

Results and discussion

In the present work, a qualitative analysis of the phenolic composition in an extract from globe artichoke flower heads was performed using HPLC-DAD-ESI-QTOF-MS operated in negative ionization mode. The method used was helpful to detect and characterize 61 phenolic compounds, of which 34 were tentatively characterized for the first time in artichoke heads (**Table 1**). The base peak chromatogram (BPC) corresponding to the aqueous-methanol extract from globe artichoke heads (edible part) is illustrated in **Figure 1**. That is, all the data on the metabolites shown here are presented according to the recommended IUPAC numbering system for caffeoylquinic acids[24].

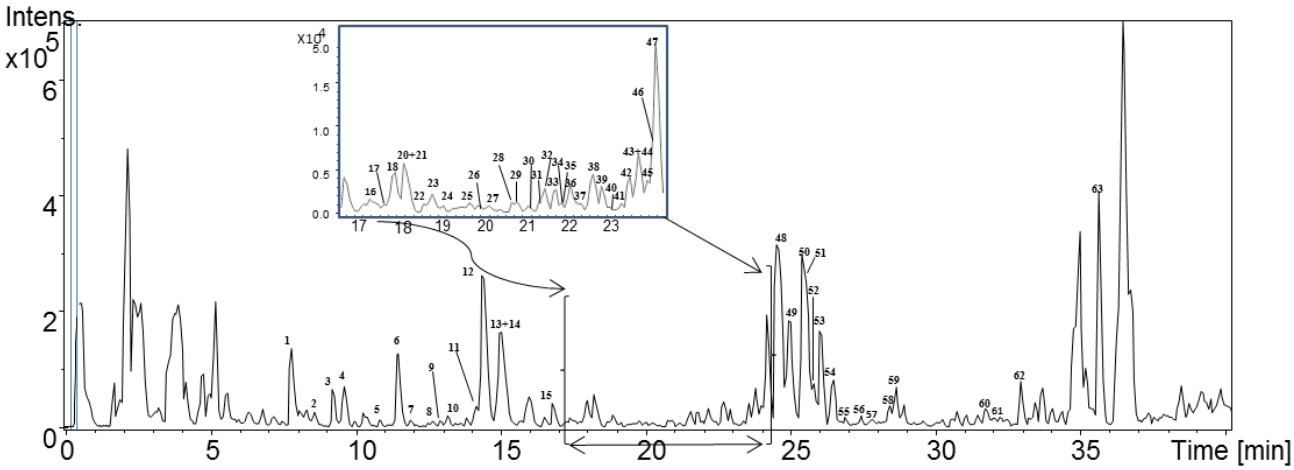


Figure 1. BPC (Base Peak Chromatogram) of the artichoke obtained by HPLC-ESI-QTOF-MS

All the compounds detected (**Table 1**) were tentatively characterized by means of their detectable UV spectrum, MS data, together with the interpretation of the observed MS/MS spectra in comparison with those found in the literature. The already identified phenolic compounds from the same botanical family were also utilized in the characterization process when applicable. The characterization process was based on the latter mentioned process, since no commercial standards were available for the all artichoke phenolic compounds detected in this work.

Identification of the well-known phenolic compounds in artichoke

The hydroxycinnamic derivatives detected in this work belong to mono- and di-caffeoylquinic acid compounds. These data agree with the results in the literature cited[7,16].

In this context, five isomers were detected at different retention times (9.61, 11.50, 14.34, 15.03, and 16.86 min) in the ESI-QTOF in negative ion mode. All compounds gave the same $[M-H]^-$ ion at m/z 353.0878 in accordance with the molecular formula ($C_{16}H_{17}O_9$). Their molecular ions showed a fragment ion at m/z 191.0562, which represents a quinic acid, resulting from the neutral loss of caffeic acid $[M-2H-163]^-$. Therefore, these compounds have been assigned to mono-caffeoylquinic acid and its isomers[19,25].

On the other hand, compounds 18, 42, 45, 48, 49, and 53, with a precursor ion at m/z 515.1188 and identical molecular formula ($C_{25}H_{23}O_{12}$), have been identified as isomers of dicaffeoylquinic acid. The MS/MS data gathered demonstrated different fragment ions at m/z 353.0877 $[M-H-162]^-$, 191.0567 $[M-H-162]^-$ and 179.0346 $[M-H-12]^-$, which showed the typical fragmentation pattern of di-caffeoylquinic acids[19,25].

Compounds 7 and 24 (11.97 and 19.31 min), with the identical molecular formula $C_{15}H_{17}O_8$, showed a fragment ion at m/z 163.0393 by the loss of glucose moiety [M-H-162]. These compounds were suggested to be p-coumarylglucose isomers[19].

With regard to flavonol derivatives, two compounds were detected at 18.01 and 21.95 min, showing the same $[M-H]^-$ at m/z 609.1449 and a product ion at m/z 301.0263. Both compounds were suggested to be rutin isomers, also in agreement with the literature[16,18,19,26].

As for flavone derivatives present in the extract, several compounds have been characterized. Thus, compound 35 (21.91 min) showed a precursor ion at m/z 593.1511, yielding a fragment ion at m/z 285.0427 by the neutral loss of the rutinoside moiety $[M-H-308]^-$, which corresponds to luteolin. Accordingly, this compound was proposed as luteolin 7-rutinoside (sclymoside)[21].

On the other hand, the compounds detected at 22.89, 25.64, and 26.85 min, with identical molecular formula $C_{21}H_{19}O_{11}$, were proposed as isomers of luteolin-7-O-glucoside (cynaroside), based on MS and MS/MS data and other works[8,18,26,27]. In the same manner, peaks 50 and 51 were assigned as Apigenin 7-glucoside (cosmoside) and apigenin 7-glucuronide, respectively[6,28].

Compound 44, with precursor ion at m/z 579.1718, has been identified as naringenin 7-rutinoside (narirutin) which belongs to the flavanones family[27,28].

Other well-known compounds belonging to other different phenolic families were also detected (**Table 1**).

Table1. MS and MS/MS data together with the tentative assignment of phenolics characterized from HPLC-QTOF/MS analysis

Peak	*Rt*	MW	[M-H]-	Formula	Error	mSigma	Order*	UV λ (nm)	MS/MS fragmentation**	Tentative Assignment	Reference
Hydroxybenzoic derivatives											
1	7.79	299.0770	299.0774	C 13 H 15 O 8	-0.6	9.6	1st(2)	250	137.0243(100)	Salicylic acid O-hexoside	-
2	8.60	315.0735	315.0730	C 13 H 15 O 9	0.3	3.4	1st(2)	255	152(100), 108.0224(39)	Protocatechuic acid 4-O-β-hexoside	-
5	10.88	359.0993	359.0982	C 15 H 19 O 10	0.4	5	1st(2)	280	197.0457(100), 153.056 (30)	Syringic acid O-hexoside	-
Hydroxycinnamic derivatives											
3	9.17	515.1418	515.1421	C 22 H 27 O 14	-2.8	39.9	1st(2)	243, 299sh, 328	191.0547(43)	Chlorogenic acid glycoside I	(31
4	9.61	353.0899	353.0876	C 16 H 17 O 9	0.5	1.7	1st(2)	242, 299sh, 329	191.0555(77)	Monocaffeoylquinic acid I	19,25
6	11.50	53.0899	353.0878	C 16 H 17 O 9	0.1	16.1	1st(2)	243, 298, 328	191.0562(100)	Monocaffeoylquinic acid II	19,25
7	11.97	325.0917	325.0935	C 15 H 17 O 8	-1.9	8.7	1st(2)	-	163.0393(30)	p-Coumarylglucose I	19
8	12.71	705.1674	705.1647	C 32 H 33 O 18	3.6	43.6	1st(3)	330	513.1040(100), 339.0529(36)	Caffeoylquinic acid dimer I	32
9	12.90	515.1418	515.1394	C 22 H 27 O 14	2.5	7.8	1st(2)	242, 299sh, 327	353.0859(3), 191.0563(87)	Chlorogenic acid glycoside II	31
10	13.42	327.1099	327.1089	C 15 H 19 O 8	-1	41.5	1st(2)	-	165.0551(100)	Dihydroxypropiophend-hexoside	-
11	14.24	341.1031	341.0880	C 15 H 17 O 9	-0.5	54.6	1st(2)	-	179.0349(100)	Caffeoyl-hexoside	-
12	14.34	353.0887	353.0875	C 16 H 17 O 9	0.9	5.2	1st(2)	241, 300sh, 329	191.0559(100)	Monocaffeoylquinic acid III	19,25
13	15.03	353.0883	353.0870	C 16 H 17 O 9	2.2	2.7	1st(2)	243, 300sh, 328	191.0556(100)	Monocaffeoylquinic acid IV	19,25
14	15.05	163.0403	163.0400	C 9 H 7 O 3	0.5	1.1	1st(1)	282	119.0498(100)	p-Coumaric acid	4
15	16.86	353.0875	353.0875	C 16 H 17 O 9	0.9	5	1st(2)	242, 299sh, 329	191.0563(100)	Monocaffeoylquinic acid V	19,25
16	17.36	337.0945	337.0933	C 16 H 17 O 8	-1.3	18.7	1st(2)	312	191.0558(100)	3-p-Coumarylquinic acid	33
17	17.71	607.1314	705.1677	C 32 H 33 O 18	-0.6	9.2	1st(2)	256, 334	513.1030(100), 339.0524(30)	Caffeoylquinic acid dimer II	32
18	17.90	515.1207	515.1188	C 25 H 23 O 12	1.4	41.7	1st(2)	247, 326	353.0869(100), 191.057(55), 179.0353(65)	Dicaffeoylquinic acid I	19,25
20	18.13	705.1674	705.1682	C 32 H 33 O 18	-1.3	8.2	1st(3)	258, 326	513.1016(100), 339.0528(39)	Caffeoylquinic acid dimer III	32
23	18.64	367.0330	367.1018	C 17 H 19 O 9	4.5	10.7	1st(1)	325	191.056 (100),173.0444 (24)	3-O-Feruloylquinic acid I	33
24	19.31	325.0940	325.0939	C 15 H 17 O 8	-3.1	9.1	1st(2)	312	163.0442 (30)	p-Coumarylglucose II	19
25	19.58	705.1644	705.1657	C 32 H 33 O 18	2.2	28.6	1st(2)	300	513.1018(100), 339.0530(32)	Caffeoylquinic acid dimer IV	32
26	20.07	705.1649	705.1668	C 32 H 33 O 18	0.6	21	1st(2)	299	513.1023(100), 339.0539(26)	Caffeoylquinic acid dimer V	32
29	21.03	677.1716	677.1733	C 31 H 33 O 17	-1.4	18.7	1st(3)	-	191.0511(70), 179.0351(22)	Dicaffeoylquinic acid glucoside I	35,36
30	21.16	531.1156	531.1147	C 25 H 23 O 13	-0.6	6.4	1st(2)	266,296,337	191.0555(100)	Mono-hydroxylated dicaffeoylquinic acid	34
32	21.51	677.1716	677.1739	C 31 H 33 O 17	-1.8	19.7	1st(3)	-	191.0521(72), 179.0349(24)	Dicaffeoylquinic acid glucoside II	35,36
34	21.82	677.1719	677.1725	C 31 H 33 O 17	-0.3	28.4	1st(2)	-	191.0551(56)	Dicaffeoylquinic acid glucoside III	35,36
37	22.25	705.1652	705.1676	C 32 H 33 O 18	-0.5	11.6	1st(2)	329	513.1035(100), 339.0533(19)	Caffeoylquinic acid dimer VI	32
40	23.36	367.1022	367.1027	C 17 H 19 O 9	2	14	1st(2)	-	191.0514(100)	3-O-Feruloylquinic acid II	33
42	23.74	515.1195	515.1186	C 25 H 23 O 12	1.7	3.8	1st(2)	246, 299sh, 326	353.0877(100), 191.0567(8),179.0346(29)	Dicaffeoylquinic acid II	19,25
45	24.04	515.1195	515.1180	C 25 H 23 O 12	1.6	6.6	1st(2)	240, 334	353.0886(100), 191.0560(16), 179.0359(32)	Dicaffeoylquinic acid III	19,25
48	24.48	515.1195	515.1187	C 25 H 23 O 12	1.6	6.6	1st(2)	242, 299sh, 328	353.0876(31), 191.0563(100), 179.0348(5)	Dicaffeoylquinic acid IV	19,25
49	24.88	515.1195	515.1188	C 25 H 23 O 12	1.3	4	1st(2)	241, 299sh, 328	353.0876(100), 191.0565(46), 179.0353(27)	Dicaffeoylquinic acid V	19,25
53	25.94	515.1195	515.1184	C 25 H 23 O 12	2.2	7.6	1st(2)	245, 300sh, 327	353.088(100), 191.0562(63), 179.0358(22)	Dicaffeoylquinic acid VI	19,25

56	27.37	499.1248	499.1238	C 25 H 23 O 11	1.3	6.6	1st(2)	249, 297sh, 328	337.0950(16),191.0561(100)	3-p-Coumaroyl-4-caffeoylquinic acid I	33
Flavonol derivatives											
19	18.01	609.1435	609.1449	C 27 H 29 O 16	1.9	25.6	1st(2)	260, 358	301.0263(21)	Rutin I	19,16,18,26
36	21.95	609.1435	609.1462	C 27 H 29 O 16	-0.1	16.4	1st(2)	-	301.0255(19)	Rutin II	19,16,18,26
Flavone derivatives											
21	18.17	623.1258	623.1250	C 27 H 27 O 17	0.6	3.6	1st(2)	267, 335	447.0917(40), 461.0728(50), 285.0395(26)	Luteolin -O-glucuronide-glucoside or Luteolin -O-glucoside-glucuronide I	-
22	18.24	607.1314	607.1282	C 27 H 27 O 16	3.6	1.7	1st(3)	268, 335	431.0975(74), 269.0456(16)	Apigenin 4'-O-ß-D-glucoside 7-O-ß-D-glucuronide	37
27	20.43	623.1262	623.1250	C 27 H 27 O 17	0.6	16.9	1st(2)	327	461.0698(38), 447.1006(29),285.0390(20)	Luteolin 7-O-glucurono(1→6)-glucoside or Luteolin -O-glucoside-glucuronide II	-
33	21.81	693.1318	693.1310	C 30 H 29 O 19	-0.2	13.2	1st(2)	334	649.1408(100), 607.1218(4), 445.0761(30), 269.0457(15)	Apigenin 7-O-glucuronide-4'-O-(6-O-malonylglucoside	38
35	21.91	593.1538	593.1511	C 27 H 29 O 15	0.2	10.4	1st(2)	267	285.0427(8)	Luteolin 7-rutinoside (Scolymoside)	21
38	22.67	461.0737	461.0713	C 21 H 17 O 12	2.8	2.1	1st(2)	254, 268, 346	285.0401(100)	Luteolin 7-glucuronide	18
39	22.89	447.0933	447.0919	C 21 H 19 O 11	3.1	10.4	1st(2)	266, 346	285.0407(57)	Luteolin-7-O-glucoside (cynaroside) I	8,18,27,26
46	24.11	563.1416	563.1403	C 26 H 27 O 14	0.7	49.4	1st(2)	-	269.0467(100)	Apiin	29
47	24.14	577.1571	577.1554	C 27 H 29 O 14	1.5	9.6	1st(2)	267, 334	269.0462(100)	Isorhoifolin (apigenin-7-O-rutinoside)	30
50	25.44	431.0992	431.0984	C 21 H 19 O 10	-0.1	5.1	1st(2)	278	269.0459(15)	Apigenin 7-glucoside (Cosmoside) I	28,6
51	25.56	445.0798	445.0774	C 21 H 17 O 11	0.4	1.5	1st(2)	277	269.0462(100)	Apigenin 7-glucuronide I	28,26
52	25.64	447.0935	447.0907	C 21 H 19 O 11	5.7	10.4	1st(2)	-	285.0411(30)	Luteolin-7-O-glucoside (cynaroside) II	8,18,27,26
55	26.85	447.0939	447.0940	C 21 H 19 O 11	-1.5	8	1st(2)	-	285.0409(100)	Luteolin-7-O-glucoside (cynaroside) III	8,18,27,26
58	28.21	269.0471	269.0464	C 15 H 9 O 5	-3	14.6	1st(1)	267, 336	225.0579 (5)	Apigenin I	8,27
59	28.61	473.1090	473.1081	C 23 H 21 O 11	1.8	150	1st(2)	-	269.0455(100)	Apigenin-7-O-(6'acetyl) glucoside I	39
60	31.90	473.1092	473.1091	C 23 H 21 O 11	-0.4	29.2	1st(2)	267, 334	269.0471(14)	Apigenin-7-O-(6'acetyl) glucoside II	39
61	32.20	285.0406	285.0405	C 15 H 9 O 6	0.0	3.6	1st(1)	334	175.0412(3), 151.0039(4), 133.0335(4)	Luteolin	8,27
63	35.56	269.0451	269.0458	C 15 H 9 O 5	-0.9	8.6	1st(2)	267, 334	225.0560(4),151.0030(9)	Apigenin II	8,27
Lignan derivatives											
28	20.85	535.1813	535.1802	C 26 H 31 O 12	3.6	22.7	1st(2)	275	357.1349(55)	1-Hydroxypinoresinol 1-O-ß-D-glucoside	40
43	23.84	519.1859	519.1847	C 26 H 31 O 11	4.8	7.7	1st(2)	-	357.1341(100)	Pinoresinol 4-O-ß-D-glucoside	41
54	26.83	561.1987	561.1974	C 28 H 33 O 12	0.5	15.7	1st(2)	253, 276	357.1320(100)	Pinoresinol-acetylhexoside	-
62	32.87	357.1336	357.1341	C 20 H 21 O 6	0.7	3.1	1st(1)	280	342.1103(13), 151.0400(100), 136.0166 (21)	(+)-Pinoresinol	44
Flavanone derivatives											
44	23.94	579.1729	579.1718	C 27 H 31 O 14	0.3	13.3	1st(2)	335	271.0615(100)	Naringenin 7-rutinoside (Narirutin)	28,27
Other polar compounds											
31	21.39	441.1755	441.1749	C 21 H 29 O 10	1.2	26.8	1st(2)	277	133.0605(5)	Lusitanicoside (chavicol ß-rutinoside)	44
41	23.56	485.1456	485.1476	C 25 H 25 O 10	-4.6	26.1	1st(2)	-	265.0905(100),173.0426(86)	unknown	-
57	27.40	539.0998	539.1758	C 25 H 31 O 13	2.3	18.1	1st(2)	-	377.1227(59), 307.0824(87), 223.0599(30)	unknown	-

* Order implies: classification order (number of possibilities). ** In parenthesis (relative Intensity %).

- sh: shoulder.
- I, II, III... stand for isomers.

New compounds detected and characterized in artichoke

Notably, new phenolic compounds were detected and characterized in the hydro-methanolic extract of artichoke analyzed. The characterization of these compounds was confirmed by the exact mass data (MS), and the MS/MS fragmentation pattern, absorption UV spectrum as well as data bases and the literature focusing on some species of the Asteraceae family.

Some phenolic compounds detected according to their phenolic family classification are discussed.

Hydroxybenzoic derivatives

Compound 1 (7.79 min) exhibited a deprotonated molecule at m/z 299.0774 and a MS/MS fragment at m/z 137.0243 corresponding to salicylic acid, due to the neutral loss of a hexose moiety $[M-H-162]^-$. Thus, this compound has been suggested to be salicylic acid hexoside (**Figure 2a**). Otherwise, peak 2 at the precursor ion m/z 315.0730 displayed a molecular formula $C_{13}H_{15}O_9$ given by the formula editor. In the MS/MS spectra, ions at 152.0113 $[M-2H-162]^-$ (loss of glucose moiety) and 108.0224 $[M-H-44]^-$ (loss of CO_2) were observed. Thus, it was tentatively characterized as protocatechuic acid 4-O-ß-hexoside. Compound 5, at 10.88 min, showed pseudo-molecular ion at m/z 359.0982, and released a MS/MS fragment ion at m/z 197.0457 (syringic acid), which was resulted from the neutral loss of a hexose moiety (**Figure 2b**). Based on the correct MS and MS/MS data, this compound was designated as syringic acid hexoside.

Hydroxycinnamic acid derivatives

The main newly characterized compounds in the artichoke extract were 33 phenolics, of which17 belonged to caffeoylquinic acid derivatives (hydroxycinnamic acid derivatives). In this context, peaks 3 and 9, with identical molecular formula ($C_{22}H_{27}O_{14}$), were detected at 9.17 and 12.90 min. These peaks demonstrated the same UV absorption bands and a fragment ion at m/z 191.0547 (see **Table 1**), which was ascribed to quinic acid. Therefore, isomers 3 and 9 were assigned to

chlorogenic acid glucoside, which was previously reported in *Erigeron breviscapus* (Asteraceae)[31].

With m/z at 705.16 ($C_{32}H_{33}O_{18}$), six peaks were detected at the following retention times: 12.71, 17.71, 18.13, 19.58, 20.07, and 22.25 min. The ESI-QTOF data of these compounds showed the same fragmentation pattern of the main ions at m/z 513.1040 and 339.0529. These compounds have been proposed to be isomers of caffeoylquinic acid dimer[32].

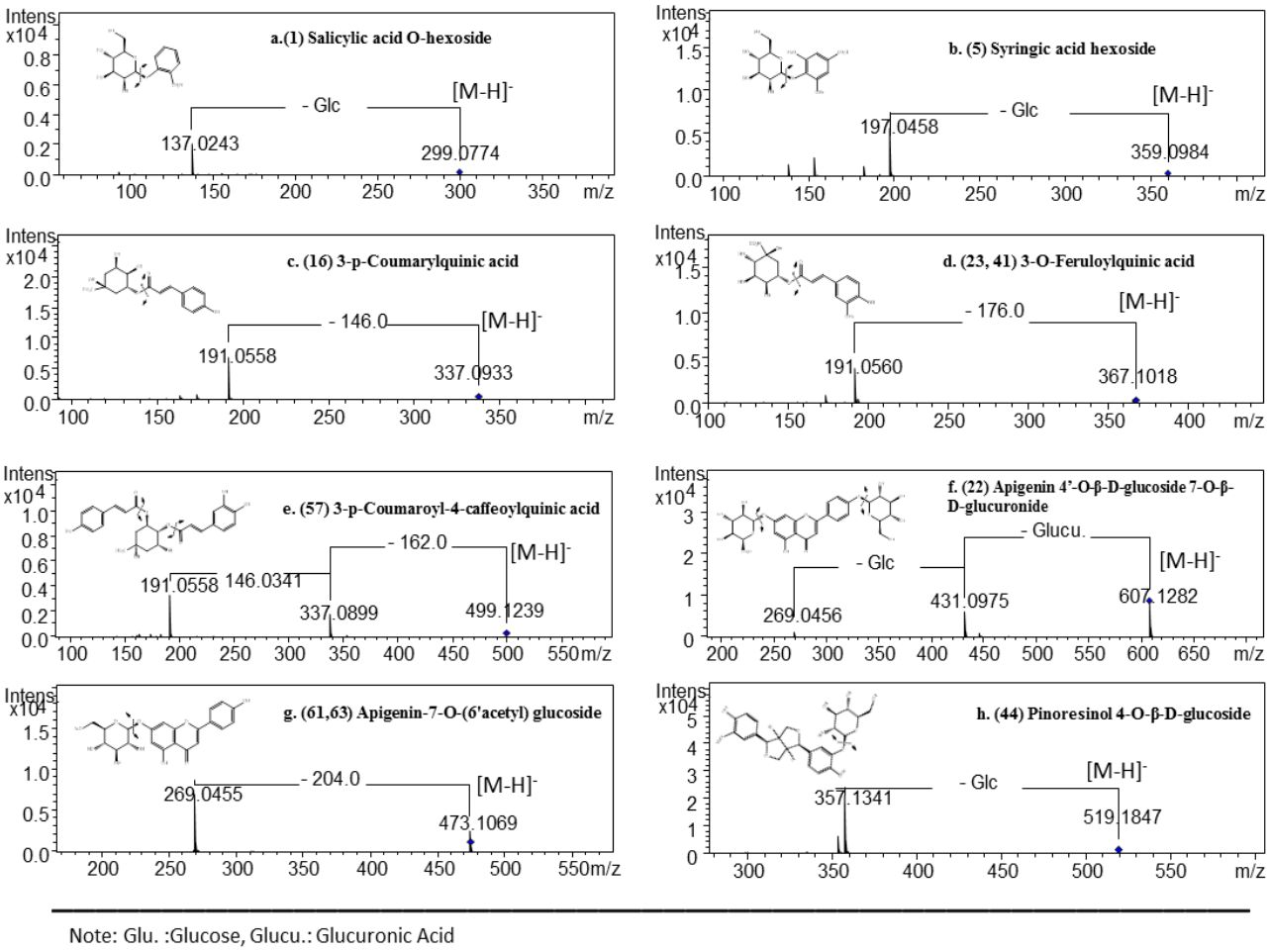


Figure 2. MS/MS spectra and postulated fragmentation pattern of newly detected phenolic compounds in artichoke by means of ESI-QTOF-MS in negative ionization mode.

Peak 10, with the pseudo-molecular ion at m/z 327.1089, was assigned to dihydroxypropiophend-hexoside, according to the correct MS data, and the MS/MS fragment ion at m/z 165.0551 arose from the loss of a hexose moiety $[M\text{-}H\text{-}162]^{-}$.

Compound 11 (14.24 min) showed a fragment ion at m/z 179.0349 $[M-H-162]^-$ by the neutral loss of hexose moiety; therefore, it was designated as caffeoyl-hexoside. It bears mentioning that this compound has not previously been reported in artichoke.

The peak at m/z 337.0933 (17.36 min) was suggested to be 3-p-coumarylquinic acid (**Figure 2c**). This compound was characterized based on the MS information and MS/MS data that showed a fragment ion at m/z 191.0558 $[M-H-162]^-$, representing the loss of quinic acid. Furthermore, this compound has been discussed in aster (Asteraceae)[33]. Similarly, the fragment ion at m/z 191.0560 was observed in the MS/MS analysis of compounds 23 and 40, demonstrating the molecular ion at m/z 367.1018 (**Figure 2d**). Thus, these compounds were tentatively identified as 3-O-feruloylquinic acid isomers[33].

Compound 30, with the molecular formula $C_{25}H_{23}O_{13}$, yielded $[M-H]^-$ at m/z 531.1147 showing a MS/MS fragment ion at m/z 191.0555, which corresponds to quinic acid. According with the data compiled, this compound was tentatively characterized as mono-hydroxylated dicaffeoylquinic acid[34].

Another three peaks detected at 21.03, 21.51 and 21.82 min were suggested to be isomers of dicaffeoylquinic acid glucoside. The characterization of these compounds was based on the correct data of MS and MS/MS fragmentation pattern that showed a daughter ion at m/z 191.0551 (indicates the presence of quinic acid in the structure). Furthermore, this compound has been reported in some members plants of Asteraceae[35,36].

Finally, compound 56 displayed a $[M-H]^-$ ion at m/z 499.1239 and its MS/MS spectrum yielded fragment ions at m/z 337.0899 [M-H-caffeic acid-H]$^-$ and 191.0558 referring to a loss of quinic acid (**Figure 2e**). Therefore, this compound was tentatively identified as 3-p-Coumaroyl-4-caffeoylquinic acid[33].

Flavone derivatives

Compounds 21 and 27 at 18.17 and 20.43 min gave an identical fragmentation pattern by the MS/MS spectra, showing a fragment ion at 447.0917 $[M-H-176]^-$ and at m/z 461.0728 $[M-H-162]^-$ which corresponded to the loss of glucuronic acid and the loss of glucose moiety, respectively. The fragment ion at m/z 285.0395 $[M-H-338]^-$ could be explained by the loss of glucuronic acid and the glucose moiety. Thus, compounds 21 and 27 were proposed to be isomers of either luteolin -O-glucuronide-glucoside or luteolin-*O*-glucoside-glucuronide.

Peak 22 gave a $[M-H]^-$ ion at m/z 607.1282 with the molecular formula $C_{27}H_{27}O_{16}$. The MS/MS spectra yielded fragment ions at m/z 431.0975 $[M-H-176]^-$ and at m/z 269.0456 $[M-H-162]^-$, indicating the presence of apigenin in the structure (**Figure 2f**). Thus, compound 22 was tentatively assigned to apigenin-4'-O-ß-D-glucoside-7-O-ß-D-glucuronide[37]. To our knowledge, this compound is reported in *Cynara scolymus* L. for the first time.

Compound 33, which displayed a molecular ion at m/z 693.1310, was characterized as apigenin 4(-O-(6-O-malonyl-ß-D-glucoside)-7-O-6-D-glucuronide. The characterization of this compound was based on the acceptable MS data along with the MS/MS fragmentation pattern, which showed the following spectral ions: 649.1408, 607.1218, 445.0761, and 269.0457, which agree with the neutral losses of [M-H-44] (-CO2), $[M-H-44-28]^-$ (-malonyl group, OCCH2CO), $[M-H-162]^-$ (-glucose), and finally the appearance of 269.0457, which represents apigenin (see **Figure 3A**). Additionally, this compound was reported in flowers of *Centaurea cyanus* (Asteraceae)[38].

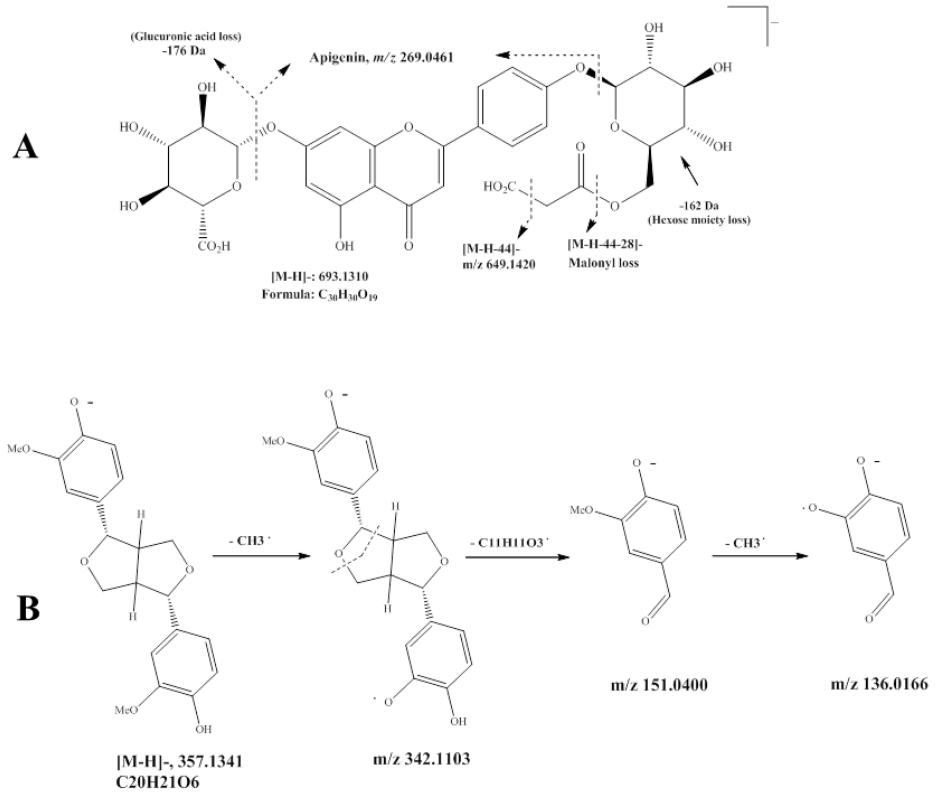


Figure 3. Hypothesized fragmentation pattern of deprotonated (A) –(+) pinoresinol and (B) Apigenin 7-O-glucuronide-4'-O-(6-O-malonylglucoside).

Last but not least, two peaks detected at 28.61 and 31.90 min (59 and 60) displayed the same molecular ion at m/z 473.1081 and a product ion at m/z 269.0455 (referred to apigenin). This fragment ion resulted from the successive loss of acetyl group (-42 Da) and a glucose moiety $[M\text{-}H\text{-}acytl\text{-}162]^-$ (**Figure 2g**). Therefore, these compounds were characterized as isomers of apigenin-7-O-(6'-acetyl) glucoside[39].

Lignan derivatives

Concerning the lignan compound derivatives, four compounds were detected in the artichoke extract analyzed.

Compound 28 gave $[M\text{-}H]^-$ ion at m/z 535.1802, and its MS/MS spectrum yielded an ion at m/z 357.1349 $[M\text{-}H\text{-}178]^-$, which represents the presence of pinoresinol. Based on the acceptable MS data and the MS/MS fragmentation pattern and the data reported in literature[40], compound 28 was tentatively suggested to be 1-

Hydroxypinoresinol 1-O-ß-D-glucoside. This is the first available report of this compound being characterized in artichoke.

On the other hand, the compound detected at 23.84 min (m/z 519.1847), with the molecular formula $C_{26}H_{31}O_{11}$, showed a neutral loss of glucose moiety [M-H-162]$^-$ to form the fragment ion at m/z 357.1341, which corresponds to presence of pinoresinol (**Figure 2h**). This compound was tentatively proposed to be pinoresinol 4-O-ß-D-glucoside[41]. It bears pointing out that the molecular mass of this compound has previously been reported in artichoke as unknown[28].

The molecular ion at m/z 561.1974 (26.83 min) showed the molecular formula $C_{28}H_{33}O_{12}$. This compound was suggested as pinoresinol-acetylhexoside, based on data from UV absorption and MS/MS fragment ion at 357.1320 demonstrated by the neutral loss (-204 Da), which can be interpreted to be due to the loss of acetyl-hexoside[42,43]. This is the first available characterization of pinoresinol-acetylhexoside in nature.

Finally, peak 62 detected at 32.87 min has been proposed to be (+)-Pinoresinol, relying on the acceptable MS and MS/MS data obtained by ESI-QTOF-MS analysis[44]. The proposed fragmentation pattern for this compound can be seen in **Figure 2B**.

Other polar compounds detected

Lastly, in the hydro-methanolic extract, three polar compounds were also detected, two unknown and one compound which could be tentatively characterized for the first time in artichoke as chavicol ß-rutinoside (lusitanicoside) (compound 31), based on MS data and the appearance of the fragment ion at m/z 133.0605 (chavicol), which corresponds to the neutral loss of rutinoside moiety (-308 Da)[44].

Conclusion

Despite the many previous studies that have been published dealing with the phenolic composition of the artichoke, the present work characterizes new and

novel phenolic compounds. The analysis of artichoke composition by HPLC-DAD-ESI-QTOF-MS has proved to be a powerful tool for selectively screening artichoke hydro-methanolic extract for the occurrence of phenolics. Thus, 61 phenolic compounds were characterized in artichoke. Of these, 33 compounds (isomers) have never been reported before in artichoke material, and are described here for the first time. Moreover, another 28 phenolic compounds already reported in artichoke have also been characterized in this work. The method used could provide more chemical information on artichoke composition, which may be useful for further research into understanding its effects on humans. Moreover, the information presented should help consumers and food technologists to know the benefits of using this traditional plant in contemporary diets as potential sources of antioxidants.

Nevertheless, the qualitative data in this study may support the ancient and current use of this plant in pharmaceutical, medicinal, and dietary contexts. Moreover, these findings may also offer a better understanding of the bioactive components in artichoke that contribute to human health and well-being.

References

1 Segasothy, M. and Phillips, P. A. 1999. Vegetarian diet: panacea for modern lifestyle diseases? *QJM,* **92,** 531-544.

2 De La Rosa, L. A., Alvarez-Parrilla, E. and Gonzalez-Aguilar, G. A. 2009. *Fruit and Vegetable Phytochemicals: Chemistry, Nutritional Value and Stability* Wiley.

3 FAO Statistical Database. 2010. http://faostat.fao.org/site/339/default.aspx, accessed 11.05.2012.

4 Lattanzio, V. and Van Sumere, C. F. 1987. Changes in phenolic compounds during the development and cold storage of artichoke (Cynara scolymus L.) heads. *Food Chemistry,* **24,** 37-50.

5 Brat, P., Georgé, S. p., Bellamy, A., Chaffaut, L. D., Scalbert, A., Mennen, L., Arnault, N., and Amiot, M. J. 2006. Daily Polyphenol Intake in France from Fruit and Vegetables. *The Journal of Nutrition,* **136,** 2368-2373.

6 Lombardo, S., Pandino, G., Mauromicale, G., Knödler, M., Carle, R., and Schieber, A. 2010. Influence of genotype, harvest time and plant part on polyphenolic composition of globe artichoke [Cynara cardunculus L. var. scolymus (L.) Fiori]. *Food Chemistry,* **119,** 1175-1181.

7 Jun, N. J. , Jang, K. C., Kim, S.C., Moon, D. Y., Seong, K. C., Kang, K. H., Tandang, L., Kim, P. H., Cho, S. K. and Park, K. H. 2007. Radical Scavenging Activity and Content of Cynarin (1,3-dicaffeoylquinic acid) in Artichoke (Cynara scolymus L.). *Journal of Applied Biological Chemistry,* **50,** 244-248.

8 Mulinacci, N., Prucher, D., Peruzzi, M., Romani, A., Pinelli, P., Giaccherini, C., and Vincieri, F. F. 2004. Commercial and laboratory extracts from artichoke leaves: estimation of caffeoyl esters and flavonoidic compounds content. *Journal of Pharmaceutical and Biomedical Analysis,* **34,** 349-357.

9 Agarwal, R., and Mukhtar, H. 1996. Cancer chemoprevention by polyphenols in green tea and artichoke. *Advances in experimental medicine and biology,* **401,** 35-50.

10 Gebhardt, R. 1997. Antioxidative and Protective Properties of Extracts from Leaves of the Artichoke (Cynara scolymus L.) against Hydroperoxide-Induced Oxidative Stress in Cultured Rat Hepatocytes. *Toxicology and Applied Pharmacology,* **144,** 279-286.

11 Brown, J. E., and Rice-Evans, C. A. 1998. Luteolin-rich artichoke extract protects low density lipoprotein from oxidation In vitro. *Free Radical Research,* **29,** 247-255.

12 McDougall, B., King, P. J., Wu, B. W., Hostomsky, Z., Manfred, G., and Robinson, W. E. Jr. 1998. Dicaffeoylquinic acid and dicaffeoyltartaric acid are selective inhibitors of human immunodeficiency virus type 1 integrase. *Antimicrobal Agents Chemotherapy,* **42,** 140–146.

13 Shimoda, H., Ninomiya, K., Nishida, N., Yoshino, T., Morikawa, T., Matsuda, and Yoshikawa, M. 2003. Anti-Hyperlipidemic sesquiterpenes and new sesquiterpene glycosides

from the leaves of artichoke (Cynara scolymus L.): structure requirement and mode of action. *Bioorganic & amp; Medicinal Chemistry Letters,* **13,** 223-228.

14 Zhu, X. F., Zhang, H. X., and Lo, R. 2005. Antifungal activity of Cynara scolymus L. extracts. *Fitoterapia,* **76,** 108-111.

15 Miadokova, E., Nadova, S., Vlckova, V., Duhova, V., Kopaskova, M., Cipak, L., Rauko, P., Mucaji, P. and Grancai, D. (2008). Antigenotoxic effect of extract from Cynara cardunculus L. *Phytotherapy Research,* **22,** 77-81.

16 Pinelli, P., Agostini, F., Comino, C., Lanteri, S., Portis, E., and Romani, A. 2007. Simultaneous quantification of caffeoyl esters and flavonoids in wild and cultivated cardoon leaves. *Food Chemistry,* **105,** 1695-1701.

17 Häusler, M., Ganzera, M., Abel, G., Popp, M., and Stuppner, H. 2002. Determination of caffeoylquinic acids and flavonoids in *Cynara scolymus L.* by high performance liquid chromatography. *Chromatographia,* **56,** 407-411.

18 Coinu, R., Carta, S., Urgeghe, P. P., Mulinacci, N., Pinelli, P., Franconi, F., and Romani, A. 2007. Dose-effect study on the antioxidant properties of leaves and outer bracts of extracts obtained from Violetto di Toscana artichoke. *Food Chemistry,* **101,** 524-531.

19 Sánchez-Rabaneda, F., Jáuregui, O., Lamuela-Raventós, R. M., Bastida, J., Viladomat, F., and Codina, C. 2003. Identification of phenolic compounds in artichoke waste by high-performance liquid chromatography–tandem mass spectrometry. *Journal of Chromatography A,* **1008,** 57-72.

20 Wang, M., Simon, J. E., Aviles, I. F., He, K., Zheng, Q.-Y., and Tadmor, Y. 2003. Analysis of Antioxidative Phenolic Compounds in Artichoke (Cynara scolymus L.). *Journal of Agricultural and Food Chemistry,* **51,** 601-608.

21 Schütz, K., Muks, E., Carle, R., & Schieber, A. 2006. Quantitative Determination of Phenolic Compounds in Artichoke-Based Dietary Supplements and Pharmaceuticals by

High-Performance Liquid Chromatography. *Journal of Agricultural and Food Chemistry,* **54,** 8812-8817.

22 Abu-Reidah, I. M., Arráez-Román, D., Lozano-Sánchez, J., Segura-Carretero, A., and Fernández-Gutiérrez, A. 2012. Phytochemical characterization of green beans (Phaseolus vulgaris L.) by using high performance liquid chromatography coupled with time-of-flight mass spectrometry. *Phytochemical Analysis*, DOI 10.1002/pca.2385.

23 Abu-Reidah, I. M., Arráez-Román, D., Quirantes-Piné, R., Fernández-Arroyo, S., Segura-Carretero, A., and Fernández-Gutiérrez, A., 2012. HPLC-ESI-Q-TOF-MS for a comprehensive characterization of bioactive phenolic compounds in cucumber whole fruit extract. *Food Research International,* **46,** 108-117.

24 IUPAC Commission on the Nomenclature of Organic Chemistry (CNOC) and IUPAC-IUB Commission on Biochemical Nomenclature (CBN). Nomenclature of cyclitols. Recommendations, 1973. 1976. *Biochemical Journal,* **153,** 23–31.

25 Moglia, A., Lanteri, S., Comino, C., Acquadro, A., de Vos, R., and Beekwilder, J. 2008. Stress-Induced Biosynthesis of Dicaffeoylquinic Acids in Globe Artichoke. *Journal of Agricultural and Food Chemistry,* **56,** 8641-8649.

26 Ferracane, R., Pellegrini, N., Visconti, A., Graziani, G., Chiavaro, E., Miglio, C., and Fogliano, V. 2008. Effects of Different Cooking Methods on Antioxidant Profile, Antioxidant Capacity, and Physical Characteristics of Artichoke. *Journal of Agricultural and Food Chemistry,* **56,** 8601-8608.

27 Azzini, E., Bugianesi, R., Romano, F., Di Venere, D., Miccadei, S., Durazzo, A., Foddai, M. S., Catasta, G., Linsalata, V. and Maiani, G. 2007. Absorption and metabolism of bioactive molecules after oral consumption of cooked edible heads of Cynara scolymus L. (cultivar Violetto di Provenza) in human subjects: A pilot study. *British Journal of Nutrition,* **97,** 963-969.

28 Schütz, K., Kammerer, D., Carle, R., and Schieber, A. 2004. Identification and Quantification of Caffeoylquinic Acids and Flavonoids from Artichoke (Cynara scolymus L.) Heads, Juice, and Pomace by HPLC-DAD-ESI/MSn. *Journal of Agricultural and Food Chemistry,* **52,** 4090-4096.

29 Wagener, H., and Kirmayer, W. 1957. Occurrence of apigeninglycosides in composites. *Naturwissen-schaften,* **44,** 307.

30 Pandino, G., Courts, F. L., Lombardo, S., Mauromicale, G., and Williamson, G. 2009. Caffeoylquinic Acids and Flavonoids in the Immature Inflorescence of Globe Artichoke, Wild Cardoon, and Cultivated Cardoon. *Journal of Agricultural and Food Chemistry,* **58,** 1026-1031.

31 Liao, S.-G., Zhang, L.-J., Li, C.-B., Lan, Y.-Y., Wang, A.-M., Huang, Y., Zhen, L., Fu, X. Z., Zhou, W., Qi, X. L., Guan, Z. Z., Wang, Y. L. 2010. Rapid screening and identification of caffeic acid and its esters in Erigeron breviscapus by ultra-performance liquid chromatography/tandem mass spectrometry. *Rapid Communications in Mass Spectrometry,* **24,** 2533-2541.

32 Schilling, S., Sigolotto, C.-I., Carle, R., and Schieber, A. 2008. Characterization of covalent addition products of chlorogenic acid quinone with amino acid derivatives in model systems and apple juice by high-performance liquid chromatography/electrospray ionization tandem mass spectrometry. *Rapid Communications in Mass Spectrometry,* **22,** 441-448.

33 Clifford, M. N., Zheng, W., and Kuhnert, N. 2006. Profiling the chlorogenic acids of aster by HPLC–MSn. *Phytochemical Analysis,* **17,** 384-393.

34 Santos, M. D., Lopes, N. P., and Iamamoto, Y. 2008. HPLC-ESI-MS/MS analysis of oxidized di-caffeoylquinic acids generated by metalloporphyrin-catalyzed reactions. *Quimica Nova,* **31,** 767-770.

35 Zhang, Y., Zhao, Q., Ma, J., Wu, B., and Zeng, X. 2010. Chemical Characterization of Phenolic Compounds in Erigeron Injection by Rapid-Resolution LC Coupled with Multi-Stage and Quadrupole-TOF-MS. *Chromatographia,* **72,** 651-658.

36 Jaiswal, R., Kiprotich, J., and Kuhnert, N. 2011 Determination of the hydroxycinnamate profile of 12 members of the Asteraceae family. *Phytochemistry,* **72,** 781-790.

37 Asen, S., and Horowitz, R. M. 1974. Apigenin 4′-O-ß-d-glucoside 7-O-ß-d)- glucuronide: The copigment in the blue pigment of Centaurea cyanus. *Phytochemistry,* **13,** 1219-1223.

38 Takeda, K., Osakabe, A., Saito, S., Furuyama, D., Tomita, A., Kojima, Y., Yamadera M, and Sakuta M. 2005. Components of protocyanin, a blue pigment from the blue flowers of Centaurea cyanus. *Phytochemistry,* **66,** 1607-1613.

39 Repčák, M., and Krausová, T. 2009. Phenolic glucosides in the course of ligulate flower development in diploid and tetraploid Matricaria chamomilla. *Food Chemistry,* **116,** 19-22.

40 Choi, S., Yang, M., Lee, K., Kim, K., and Lee, K. 2007. Lignan and terpene constituents from the aerial parts of saussurea pulchella. *Archives of Pharmacal Research,* **30,** 1067-1074.

41 Shoeb, M., MacManus, S. M., Jaspars, M., Trevidu, J., Nahar, L., Kong-Thoo-Lin, P., and Sarker, S. D. 2006. Montamine, a unique dimeric indole alkaloid, from the seeds of Centaurea montana (Asteraceae), and it's in vitro cytotoxic activity against the CaCo2 colon cancer cells. *Tetrahedron,* **62,** 11172-11177.

42 Willför, S. M., Smeds, A. I., & Holmbom, B. R. 2006. Chromatographic analysis of lignans. Journal of *Chromatography A,* **1112,** 64-77.

43 Kajdžanoska, M., Gjamovski, V., & Stefova, M. 2010. HPLC-DAD-ESI-MSn, identification of phenolic compounds in cultivated strawberries from Macedonia. *Macedonian Journal of Chemistry and Chemical Engineering,* **29,** 181–194.

44 Pan, L., Sinden, M. R., Kennedy, A. H., Chai, H., Watson, L. E., Graham, T. L. and Kinghorn, A. D. 2009. Bioactive constituents of Helianthus tuberosus (Jerusalem artichoke). *Phytochemistry Letters,* **2,** 15-18.

Chapter 5

Reversed-phase ultra-performance liquid chromatography coupled to electrospray ionization-quadrupole-time-of-flight mass spectrometry as a powerful tool for metabolic profiling of vegetables: *Lactuca sativa* as an example of its application

Ibrahim M. Abu-Reidah, María del Mar Contreras, David Arráez-Román, Antonio Segura-Carretero, Alberto Fernández-Gutiérrez

This work has been submitted to ***Chromatography A***.

Reversed-phase ultra-performance liquid chromatography coupled to electrospray ionization–quadrupole–time-of-flight mass spectrometry as a powerful tool for metabolic profiling of vegetables: *Lactuca sativa* as an example of its application

Abstract

Lettuce (*Lactuca sativa*), a leafy vegetal widely consumed worldwide, fresh cut or minimally processed, constitutes a major dietary source of natural antioxidants and bioactive compounds. In this study, reversed-phase ultra-performance liquid chromatography (RP-UPLC) coupled to electrospray ionization-quadrupole-time-of-flight mass spectrometry (ESI-QTOF-MS) was applied for the comprehensive profiling of polar and semi-polar metabolites from three lettuce cultivars (baby, romaine, and iceberg). The UPLC system allowed the use of a small-particle-size C18 column (1.8 □m) with very fine resolution for the separation of up to seven isomers, and the QTOF mass analyzer enabled sensitive detection with high mass resolution and accuracy in full scan. Thus, a total of 171 compounds were tentatively identified by matching their accurate mass signals and suggested molecular formula with those previously reported in family Asteraceae. Afterwards, their structures were also corroborated by the MS/MS data provided by the QTOF analyzer. Well-known amino acids, organic acids, sesquiterpene lactones, phenolic acids and flavonoids were characterized, e.g. lactucin, lactucopicrin, caftaric acid, chlorogenic acid, caffeoylmalic acid, chicoric acid, isochlorogenic acid A, luteolin, and quercetin glycosides. For this plant species, this is the first available report of several isomeric forms of the latter polyphenols and other types of components such as nucleosides, peptides, and tryptophan-derived alkaloids. Remarkably, 10 novel structures formed by the conjugation of known amino acids and sesquiterpene lactones were also proposed. Thus, the methodology applied is a useful option to develop an exhaustive metabolic profiling of plants that helps to explain their potential biological activities and folk uses.

Keywords: *Lettuce (Lactuca sativa), Asteraceae (Compositae), phenolic compounds, metabolic profiling, secondary metabolites, UPLC-ESI-QTOF-MS*

Introduction

Diets containing high proportions of fruit and vegetables have been advocated as one of the best practices to reduce the incidence of chronic disease in the modern world. These beneficial effects have been partially attributed to the increased consumption of phenolic compounds and other bioactive phytochemicals[1]. Although a great number of studies have focused on discovering and characterizing these metabolites, a large percentage still remain unknown and need to be identified before their relation to health can be fully understood[2]. For this, different pretreatment procedures and analysis techniques have been developed and recently reviewed[3]. Among the analytical methods, liquid chromatography (LC) coupled with mass spectrometry (MS) has become a useful tool in the metabolic profiling of plants extracts. For improved speed, resolution, and sensitivity, HPLC methods can be successfully transferred to UPLC, RRLC, and UHPLC, which operate with smaller particle sizes (<2 μm) of the stationary phase and at higher pressures (up to 600 bar)[4]. In addition, MS is a powerful structural-characterization technique, for which good prior separation is crucial because of the complexity of plant extracts and the presence of large numbers of isomers[5], which could be undifferentiated by MS.

The hybrid mass analyzer quadrupole-time-of-flight (QTOF) provides excellent mass accuracy over a wide dynamic range and measurements of the true isotope pattern that elucidates the molecular formula of unknown metabolites with a high degree of reliability, and taking into account the seven golden heuristic and chemical rules for selecting elemental compositions [6]. This instrument also performs tandem MS, which is useful as a structural confirmation tool when standard compounds are not available. Furthermore, QTOF provides high selectivity by the extracted-ion chromatogram (EIC) mode when there are overlapping peaks, where spectrophotometric detection could be limited. Thus, LC coupled to highly sensitive and high-resolution MS, such as QTOF, enables the separation and detection of minor compounds that could co-elute and be underestimated vs. major ions, not being identified or even detected by older methodologies. LC coupled to QTOF-MS has been applied in targeted analysis but is suitable also for the extensive profiling of hundreds of plant metabolites[7]. Among the ionization sources, electrospray ionization (ESI), particularly in the negative mode, is a good choice for studying small molecules[4,5]. In our previous studies, it has been demonstrated that RP-RRLC-ESI-QTOF-MS is well suited to the untargeted characterization of edible vegetables, and enabled a wide overview of the composition

(amino acids, nucleosides, organic acids, and phenolics) of zucchini[8]. Following a similar strategy, 73 phenolic compounds have been identified in cucumber[9].

Lettuce (*Lactuca sativa*), a leafy vegetable of the family Asteraceae (Compositae), and one of the most widely used food crops for the so-called " fourth-range" vegetables (cleaned, possibly chopped, mixed, ready to be seasoned and eaten fresh)[10]. Thus, minimally processed, consumed in salads, but also in soups, sandwiches and wraps, it retains many naturally occurring components, such as labile vitamins. The world production of lettuce is estimated to be >24 million of tons (FAOSTAT, 2010), being one of the major vegetable crops in the world[11]. Lettuce is a good source of vitamins, fiber, amino acids, minerals, such as potassium, and as well as phytochemicals[12,13]. Among these, dietary phenolics are of the particular interest due to their potential as antioxidants, not only as natural food preservatives, but also for their role on human heath[14], as commented above. In this respect, antioxidant activity of lettuce has been reported both *in vitro*[1] and *in vivo*[15]. Furthermore, lettuce contains other important phytochemicals such as sesquiterpene lactones, characteristic and diverse class from Asteraceae. These types of terpenoids, such as phenolic compounds, have different bioactivities and contribute to sensory properties of plants, but also, in some cases, present allergenic potency[16,17]. Several techniques have been used to identify phenolic compounds and other phytochemicals, e.g. LC coupled to UV and diode array detection (DAD), MS and MS/MS[18-23], and GC-MS[24]. However, only one NMR-based approach has been applied for more comprehensive characterization of lettuce[25]. Sensitivity is perhaps the most important requirement for metabolomic profiling, and NMR is several orders of magnitude less sensitive than MS[26]. Thus, the main objective of this study is to evaluate the polar and semi-polar fraction, including primary and secondary metabolites, of three highly consumed lettuce cultivars (baby, romaine, and iceberg) by RP-UPLC-ESI-QTOF-MS and -MS/MS. In addition, this methodology was used as starting point for structure elucidation of new molecules, based on the suggested molecular formula, structural information provided by MS/MS fragmentation, and the literature.

Experimental

Chemicals

Acetic acid, methanol, and acetonitrile were purchased from Fluka (Sigma-Aldrich, Steinheim, Germany) and Fisher chemical (Acros Organics, Geel, Belgium), respectively. Water was purified by a Milli-Q system from Millipore (Bedford, MA, USA).

All chemicals were of HPLC-MS grade, except methanol (analytical grade), and used as received.

Plant material

Fresh lettuce (cv. baby, romaine, and iceberg), purchased from a local market, were from crops cultivated in Almeria (Spain) in the year 2011 with no prior knowledge of their growing conditions. After being purchased, the samples were transported directly to the lab and washed with distilled water. Thereafter, each sample was freeze-dried and ground to a fine powder to facilitate its later extraction.

Extraction of lettuce metabolites

The polar and semi-polar fraction of lyophilized material (leaves) was extracted for the characterization of its phytochemical compounds as described by Abu Reidah et al. (2012)[9], with some modifications. In brief, an amount of lettuce (0.5 g freeze dried) of each sample was extracted with 16 ml of a mixture of methanol/water (80:20) (v/v) and sonicated for 30 min at room temperature. Then, the mixture was centrifuged for 15 min at 3,750 *rpm* and the supernatant collected in a round-bottom flask. Subsequently, the solvent was evaporated using a rotary evaporator under vacuum at 38°C and the dry residue was resolved in 0.5 mL of 80:20 (v/v) methanol/water. Finally, the extract was centrifuged again and the supernatant was filtered through a 0.2 µm syringe filter and stored at -20°C prior to analysis.

Analysis by RP-UPLC-ESI-QTOF-MS and -MS/MS

Analyses were performed using an ACQUITY UPLC (Waters, Millford, MA, USA), consisting of a vacuum degasser, an auto-sampler, and a binary pump system. The analytical conditions were as described previously[9], with some modifications. The column temperature was 25°C and the injection volume 5 µL. The UPLC system was coupled to a micrOTOF-Q™ (Bruker Daltonik GmbH, Bremen, Germany), equipped with an ESI interface. The spectra were acquired in negative-ion mode over a mass-to-charge (*m/z*) range from 50 to 1100. The MS and MS/MS parameters were set as described previously[9], and a sodium acetate cluster solution was used for external instrument calibration. The MS/MS analyses were made by automatic fragmentation, where the three most intense mass peaks where fragmented.

The MS and MS/MS data were processed through Data Analysis 4.0 software (Bruker Daltonics) which supplies tools to calculate molecular formulae (SmartFormula), simulate isotope patterns as well as theoretical monoisotopic, nominal, and average masses (IsotopePattern), and build common neutral losses (BuildingBlockEditor).

Therefore, the elemental composition of each compound was determined by the SmartFormula tool that generates a list of possible molecular formulae based on a CHNO algorithm. It provides standard functionalities such as minimum/maximum elemental range, electron configuration, and ring-plus double-bond equivalents, the deviation between the measured mass and theoretical mass (error) (Da and ppm), and a sophisticated comparison of the theoretical with the measured isotope patterns (msigma value) for increasing the confidence in the suggested molecular formula. The broadly accepted error and msigma (mσ value for confirmation of elemental compositions was set at 5 ppm and at 50, respectively. The use of isotopic-abundance pattern removes more than 95% of false candidates[4]. The other elements, which usually occur in smaller numbers, are dealt with using the straightforward trial-and-error method. The molecular formula generated for each compound, its MS/MS spectra and its comparison with spectra found in the literature was the main tool for putative identification of lettuce metabolites. For this, the following online public databases were consulted: PubChem (http://pubchem.ncbi.nlm.nih.gov), ChemSpider (http://www.chemspider.com), Kegg Ligand Database (http://www.genome.jp/kegg/ligand.html), SciFinder Scholar (https://scifinder.cas.org), Phenol-Explorer (www.phenol-explorer.eu), KNApSAcK Core System (http://kanaya.naist.jp/knapsack_jsp/top.html), Metlin (http://metlin.scripps.edu), and MassBank (http://www.massbank.jp), which is related to KNApSAcK database. The two latter databases, Metlin and MassBank, also include a repository of tandem mass spectrometry data.

Results and Discussion

RP-UPLC-ESI-QTOF-MS as a powerful tool for metabolic profiling

In this work, a simple and non-selective extraction procedure using a methanol/water solution after freeze-drying of lettuce samples was selected for its metabolic profiling by RP-UHPLC-ESI-QTOF-MS. The base peak chromatograms (BPC) of the extracts from baby, iceberg and romaine lettuces and the compounds identified are shown in **Fig. 1.** It revealed a complex mixture of polar and semi-polar metabolites, including some regions of the chromatograms that at first sight appeared to be flat (for an example, see zoomed regions). In agreement with literature[5,8], a certain tendency in the elution order of the compounds related to their chemical structure class was observed, appearing in the following order of increasing retention time (RT), and thus hydrophobicity: amino acids, organic acids, nucleosides, phenolic acids, peptides,

flavonoids, tryptophan-derived alkaloids, lignans, sesquiterpene lactones (main terpenoid components), and hydrolysable tannins (**Fig. 2a**). An overview of the qualitative distribution of the characterized compounds in all three cultivars is shown in **Fig. 2b**. In general, larger differences were found in the variety baby with respect to iceberg and romaine, which presented similar qualitative metabolic patterns, in agreement with Llorach et al.[21].

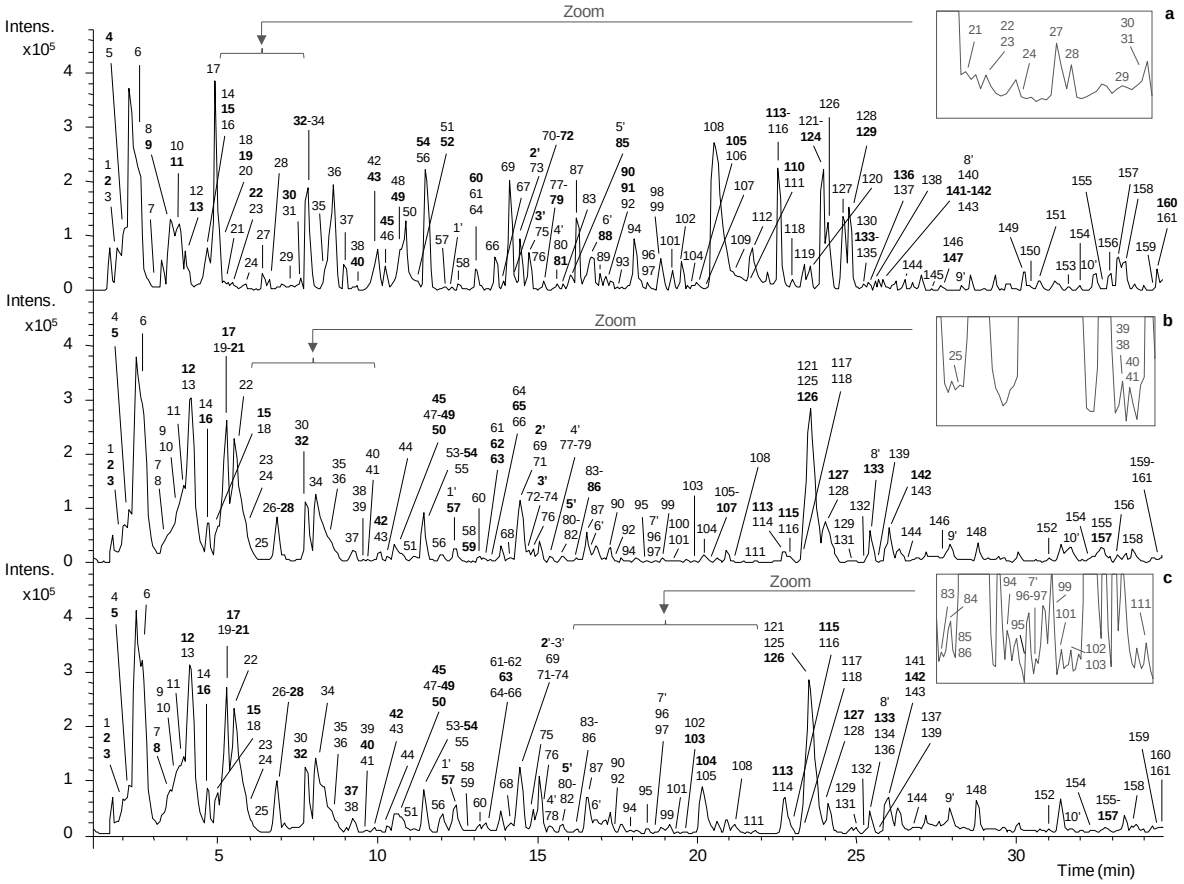


Fig. 1. Base peak chromatograms of the lettuce extracts obtained by RP-UPLC-ESI-QTOF-MS in negative ion mode: (a) baby, (b) iceberg, and (c) romaine. Compounds are numbered according to Tables 1, 2, and 3, and those most abundant in each peak are given in bold.

A total of 171 compounds were tentatively identified and, due to this high diversity, some were co-eluted with the solvent gradient employed. However, the chromatographic method was sufficiently selective to separate up to seven isomer forms of several compounds. For instance, chromatographic separation of isomers of hydrolysable tannin tri-4-hydroxyphenylacetyl glucoside and the hydroxycinnamate caffeoyl-hexose are presented in **Fig. 1a** and **1b**, respectively. The nature of the sugar substitution and/or the position affects the polarity of the isomers and therefore their retention time[27]. Furthermore, MS detection showed adequate specificity for the analysis of such complex extracts, enabling the separation of compounds with similar

retention times one by one by extracting the EIC, such as sugar esters of caffeic and sinapic acids represented in **Fig. 1b** and **1c**, respectively. By contrast, their spectrophotometric detection could be compromised, since compounds belonging to the same family present similar UV absorption bands[5,28]. Previous studies have characterized from four to 17 phenolic compounds in different green lettuce cultivars, using LC coupled with UV and DAD and/or ion trap (IT), TOF, and triple quadrupole (QqQ),[18-23,29] and GC coupled to single quadrupole (Q)[24], whereas an NMR study[25] characterized 34 water-soluble compounds. In this sense, our results demonstrate that the combination of the high resolving power and wide selective retention of the C18 (with particle size 1.8 µm) stationary phase using an UPLC system and the sensitive detection of the QTOF mass analyzer constitute a valuable technique for the metabolic profiling of plants.

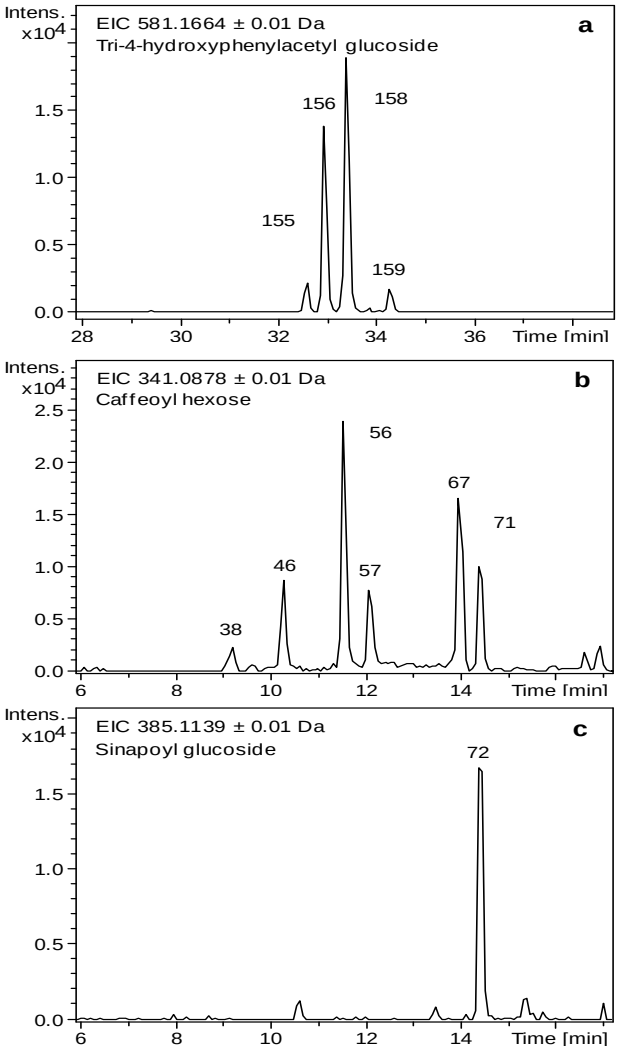


Figure1. Representative extracted ion chromatograms (EICs) of (a) try-3-hydroxyphenylacetyl glucoside, (b) caffeoyl-hexose, and (c) sinapoyl glucoside. Isomers are numbered according to Table 2.

Identification strategy with QTOF detection and interpretation of tandem MS results

Accurate mass measurement and isotope pattern provided by the QTOF was used as the main step towards the characterization of the ions detected by generating their molecular formula with high confidence (see Subsection 2.4).

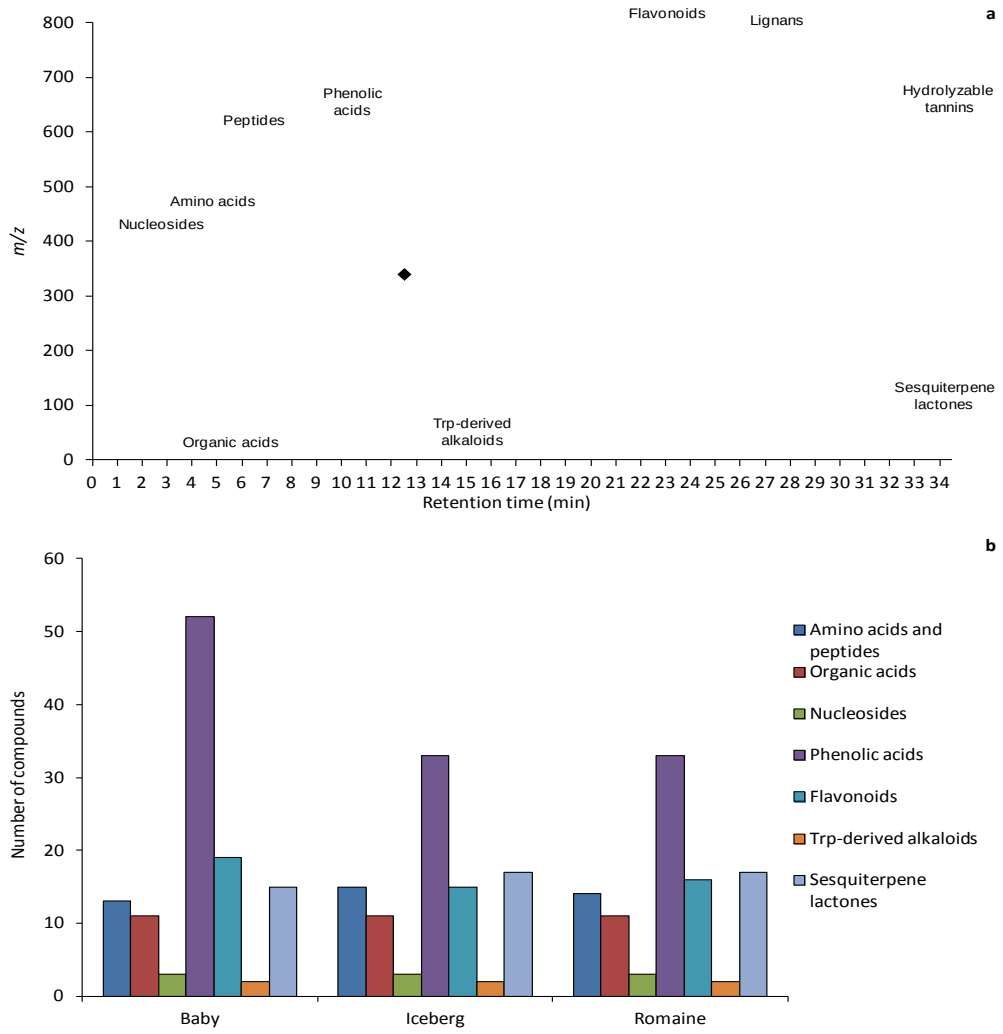


Fig. 2. (a) Survey view and classification of the metabolites characterized on the basis of their *m/z* and retention time, and (b) qualitative distribution depicted by the number of compounds characterized in each of the lettuce cultivars studied. Initial Maillard products are not taken into account.

After a thorough survey of the literature on Asteraceae, the putative identification of the compounds was corroborated by their MS/MS fragmentation. It made it possible to establish neutral losses related to functional groups (decarboxylation from carboxyl group, for instance), and substitutions (sugars, organic acids, etc.)[5,8,9,18], as well as to know the true molecular formula of the daughter ions for the unambiguous assignment

of the precursor ion. Thus, **Table 1** (non-phenolic metabolites) and **Table 2** (phenolic metabolites) list the molecular ions studied ($[M-H]^-$) with their retention time, observed and calculated *m/z*, molecular formula, mass error (ppm), m values, main fragment ions detected in MS/MS, and the proposed tentative identification. In some cases, where the MS/MS spectrum was unclear, possible identifications or general names are shown.

Non-phenolic metabolites

Amino acid and derivatives

Seven free amino acids, Asn (**1**) (*m/z* 131.0463, $C_4H_7N_2O_3$), Gln (**2**) (*m/z* 145.0618, $C_5H_9N_2O_3$), Leu and Ile (**16** and **18**) (*m/z* 130.087, $C_6H_{12}NO_2$), Tyr (**19**) (*m/z* 180.0663, $C_9H_{10}NO_3$), Phe (**32**) (*m/z* 164.0718, $C_9H_{10}NO_2$), and Trp (**54**) (*m/z* 203.0829, $C_{11}H_{11}N_2O_2$), were detected in the lettuce samples analyzed from RT 1.8 to 11.4 min (**Fig. 2a**). The amino acid composition of lettuce has previously been described[25,30]. In general, neutral losses of NH_3 (17.0265 u) and CO_2 (43.9898 u) from amino- and carboxy-terminal, respectively, and H_2O (18.0106 u) were observed independently of their chain substitution, except Leu/Ile isomers, which were not fragmented by our MS/MS conditions. In the case of Trp, its characteristic indole group (*m/z* 116.0500, C_8H_6N) was the most abundant fragment generated after sequential neutral losses of NH_3, CO_2 and C_2H_2 (26.0157 u). It is noteworthy that several dipeptides and tripeptides were also detected in negative ion mode and identified as Glu-Leu/Ile 1 (**23**) (*m/z* 259.1299, $C_{11}H_{19}N_2O_5$), Glu-Leu/Ile 2 (**24**) (*m/z* 259.1298, $C_{11}H_{19}N_2O_5$), Glp-Leu/Ile 1 (**83**) (*m/z* 241.1187, $C_{11}H_{17}N_2O_4$), Hyp-Gly-Leu/Ile (**86**) (*m/z* 300.1558, $C_{13}H_{22}N_3O_5$), Glp-Leu/Ile 2 (**92**) (*m/z* 241.1187, $C_{11}H_{17}N_2O_4$), isomers of Asp-Leu/Ile-Leu/Ile (**96**,**101** and **107**) (*m/z* 358.1969, 358.1974 and 358.1976, $C_{16}H_{28}N_3O_6$), and Asp-Leu/Ile-Phe (**117**) (*m/z* 392.1816, $C_{19}H_{26}N_3O_6$). The fragmentation of the peptides occurred at the peptidic bond (amide bond), but also losses of H_2O, CO_2 and NH_3 were detected in some cases. Thus, these results suggest that the negative ion mode could be used complementarily with respect to the positive ion mode to perform peptide-characterization studies. In the case of plant peptides, they have not received much attention compared to other food sources and other plant metabolites, probably because they are not abundant, making their detection difficult[31]. In this sense, □-Glu dipeptides have been described as characteristic metabolites in *Iridaceae* family more than 30 years ago, but their biological significance remains unknown[32].

Furthermore, initial Maillard products derived from the condensation of the latter amino acids and Glp (probably, derived from cyclization of Gln) (12 compounds), and several dipeptides (14 compounds) with reducing sugars present in lettuce, such fructose and glucose[25], were also tentatively identified (**Table 1**). In this regard, cv. baby presented a lower number of such compounds. As an example of the fragmentation pattern of the peptides identified in negative-ionization mode, **Fig. 3a** represents the MS/MS spectrum of the dipeptide Glu-Leu/Ile conjugated to hexose (**25**) (*m/z* 421.1815, $C_{17}H_{29}N_2O_{10}$). It gave rise to ions at *m/z* 331.1502 and *m/z* 259.1286, which represented the losses of three aldehyde groups (3×(HCHO), 90.0316 u) from the sugar and the sugar moiety ($C_6H_{10}O_5$, 162.0528 u), as has been previously reported for other glycosylated compounds[8]. Subsequently, the fragmentation of the dipeptide core occurred at the peptidic bond, with ions appearing that had *m/z* values of 128.0350 ($C_5H_6NO_3$) (negatively charged b_1 ion) and 130.0826 ($C_6H_{12}NO_2$) (negatively charged y_1 ion, which correspond with leucine or isoleucine) from the N- and C-terminal regions, respectively. In addition, the presence of a fragment ion at *m/z* 218.0699 ($C_8H_{12}NO_6$), which could be b_1 ion plus 3×(HCHO), indicates that the N-terminus was involved in the Maillard reaction, in agreement with previous findings[33]. Several initial Maillard reaction products and thermal cyclization of Gln have been previously found in other vegetables after drying[5,8,34]. This stage does not cause serious physical and flavor changes, in contrast to advanced stages of Maillard reaction browning[35], but could help to improve antioxidant properties of the compounds[33].

Table 1. Amino acids, peptides, organic acids, alkaloids, and terpenoids and derivatives detected in aqueous-methanolic extracts from green leaf lettuce cultivars baby (B), iceberg (I) and romaine (R) by RP-UPLC-ESI-QTOF-MS.

Nº	RT (min)	Observed *m/z* ($[M-H]^-$)	Calculated *m/z* ($[M-H]^-$)	Molecular formula	Error (ppm)	mσ	Major fragments *m/z* ($[M-H]^-$)(%)	Proposed Compound	Class	B	I	R	Ref.
1	1.8	131.0463	131.0462	$C_4H_7N_2O_3$	-0.8	32	114.0138 (100), 113.0313 (96)	Asn	Amino acid	+	+	+	25,30
2	1.8	145.0618	145.0619	$C_5H_9N_2O_3$	0.3	13	127.0487 (100), 102.0539 (18)	Gln	Amino acid	+	+	+	25,30
3	2.1	293.0991	293.099	$C_{10}H_{17}N_2O_8$	-0.3	22	203.0674 (6), 173.0561 (5), 131.0418 (100)	Asn-hexose	Amino acid	+	+	+	
4	2.1	191.0562	191.0561	$C_7H_{11}O_6$	-0.5	6	n.d.	Quinic acid	Organic acid	+	+	+	36
5	2.2	307.1147	307.1147	$C_{11}H_{19}N_2O_8$	-0.1	29	217.0847 (4), 187.073 (5), 145.0614 (100), 127.051 (9)	Gln-hexose	Amino acid	+	+	+	
6	2.4	133.0142	133.0142	$C_4H_5O_5$	0.6	2	115.0037 (100)	Malic acid	Organic acid	+	+	+	25, 36
7	3.5	278.125	278.1245	$C_{11}H_{20}NO_7$	-1.8	2	158.0808 (3), 116.0713 (100)	Val-hexose	Amino acid	+	+	+	
8	3.4	290.0891	290.0881	$C_{11}H_{16}NO_8$	-3.2	7	200.0560 (40), 170.0453 (7), 128.0334 (100)	Glp-hexose 1	Amino acid	+	+	+	
9	3.5	191.0202	191.0197	$C_6H_7O_7$	-2.3	3	111.0087 (100)	Citric acid	Organic acids	+	+	+	36
10	3.7	290.0887	290.0881	$C_{11}H_{16}NO_8$			n.d.	Glp-hexose 2	Amino acid	+	+	+	
11	3.9	191.0200	191.0197	$C_6H_7O_7$	-1.3	3	111.0085 (100)	Isocitric acid	Organic acid	+	+	+	
12	4.0	290.0889	290.0881	$C_{11}H_{16}NO_8$	-2.8	9	200.0551 (48), 170.0419 (9), 128.0348 (100)	Glp-hexose 3	Amino acid	+	+	+	
13	4.0	115.0037	115.0037	$C_4H_3O_4$	-0.1	5	n.d.	Fumaric acid	Organic acid	+	+	+	25, 36
14	4.6	147.0296	147.0299	$C_5H_7O_5$	1.8	4	103.0392 (43)	Citramalic acid	Organic acid	+	+	+	
15	4.6	117.0185	117.0193	$C_4H_5O_4$	6.9	4	n.d.	Succinic acid	Organic acid	+	+	+	25, 36
16	4.7	130.0868	130.0874	$C_6H_{12}NO_2$	4.5	2	n.d.	Leu/Ile	Amino acid	+	+	+	25,30
17	4.9	243.0622	243.0623	$C_9H_{11}N_2O_6$	0.3	5	200.0542 (9), 152.035 (21), 140.0372 (9), 110.0247 (100)	Uridine[a]	Nucleosides	+	+	+	38
18	5.0	130.0871	130.0874	$C_6H_{12}NO_2$	1.6	4	n.d.	Leu/Ile	Amino acid	+	+	+	25,30
19	5.0	180.0663	180.0666	$C_9H_{10}NO_3$	1.5	59	163.0403 (59), 119.0498 (100)	Tyr	Amino acid	+	+	+	30
20	5.2	342.1197	342.1194	$C_{15}H_{20}NO_8$	-0.8	7	180.0673 (100)	Tyr-hexose	Amino acid	+	+	+	
21	5.2	292.1406	292.1402	$C_{12}H_{22}NO_7$	-1.5	3	202.1080 (1), 172.0988 (2), 130.0871 (100)	Leu/Ile-hexose 1	Amino acid	+	+	+	
22	5.5	292.1406	292.1402	$C_{12}H_{22}NO_7$	-1.3	2	202.1076 (2), 172.0975 (1), 130.0871 (100)	Leu/Ile-hexose 2	Amino acid	+	+	+	
23	5.6	259.1299	259.1299	$C_{11}H_{19}N_2O_5$	0.0	15	215.1352 (15), 171.1493 (19), 130.0894 (38)	Glu-Leu/Ile 1	Peptide	+	+	+	
24	6.1	259.1298	259.1299	$C_{11}H_{19}N_2O_5$	0.4	10	215.1390 (12), 171.1514 (8), 130.0849 (46), 128.0309 (10)	Glu-Leu/Ile 2	Peptide	+	+	+	
25	6.5	421.1816	421.1828	$C_{17}H_{29}N_2O_{10}$	2.9	7	331.1502 (100), 259.1287 (50), 241.1110 (7), 218.0722 (21), 130.0826 (13), 128.0309 (23)	Glu-Leu/Ile hexose 1	Peptide	-	+	+	

26	6.7	421.1814	421.1828	$C_{17}H_{29}N_2O_{10}$	3.2	11	331.1500 (100), 259.1299 (72), 241.1110 (48), 130.0826 (6), 128.0365 (16)	Glu-Leu/Ile hexose 2	Peptide	-	+	+	
27	6.4	266.0888	266.0895	$C_{10}H_{12}N_5O_4$	2.7	30	134.0467 (100)	Adenosine[a]	Nucleoside	+	+	+	38
28	6.8	282.0846	282.0844	$C_{10}H_{12}N_5O_5$	-0.6	5	150.0419 (100), 133.0149 (10)	Guanosine	Nucleoside	+	+	+	
32	7.8	164.0718	164.0717	$C_9H_{10}NO_2$	-0.5	2	147.0451 (100), 103.0547 (18)	Phe	Amino acid	+	+	+	25,30
34	8.1	326.1247	326.1245	$C_{15}H_{20}NO_7$	-0.6	3	236.0938 (1), 206.0815 (3), 164.0717 (100), 147.0447 (7)	Phe-hexose	Amino acid	+	+	+	
37	8.9	218.1040	218.1034	$C_9H_{16}NO_5$	-3	6	146.0826 (100)	Pantothenic acid (vitamine B5)	Organic acid	+	+	+	13
39	9.2	421.1832	421.1828	$C_{17}H_{29}N_2O_{10}$	-1	19	331.1515 (100), 259.1286 (68), 241.1183 (48), 130.0873 (26), 128.0381 (8)	Glu-Leu/Ile-hexose 3	Peptide	-	+	+	
41	9.7	391.2077	391.2086	$C_{17}H_{31}N_2O_8$	2.2	61	301.1768 (100), 229.1569 (87), 130.0826 (4)	Val-Leu/Ile-hexose 1	Peptide	-	+	+	
44	10.2	391.208	391.2086	$C_{17}H_{31}N_2O_8$	1.8	23	301.1776 (100), 229.1559 (86), 130.0823 (10)	Val-Leu/Ile-hexose 2	Peptide	-	+	+	
51	11.2	365.1353	365.1354	$C_{17}H_{21}N_2O_7$	0.4	11	203.0830 (100), 159.0877 (1)	Trp-hexose 1	Amino acid	+	+	+	
53	11.4	365.1358	365.1354	$C_{17}H_{21}N_2O_7$	-0.9	9	203.0833 (100), 159.0935 (2)	Trp-hexose 2	Amino acid	-	+	+	
54	11.4	203.0829	203.0826	$C_{11}H_{11}N_2O_2$	-1.5	29	142.0662 (36), 116.0500 (100)	Trp	Amino acid	+	+	+	30
55	11.5	391.2076	391.2086	$C_{17}H_{31}N_2O_8$	2.8	8	301.1770 (100), 229.1559 (78), 227.1397 (34), 130.0863 (6)	Val-Leu/Ile-hexose 3	Peptide	-	+	+	
59	12.9	405.2202	405.2242	$C_{18}H_{33}N_2O_8$	10	34	315.1898 (100), 243.1719 (71), 130.0826 (7)	Leu/Ile-Leu/Ile-hexose 1	Peptide	-	+	+	
60	13.2	175.0613	175.0612	$C_7H_{11}O_5$	-0.9	11	157.0489 (7), 131.0709 (9), 115.0394 (100), 113.0616 (42)	Isopropylmalic acid	Organic acid	+	+	+	
62	13.3	403.1719	403.1722	$C_{17}H_{27}N_2O_9$	0.9	14	313.1411 (60), 241.1194 (100), 197.1264 (8), 130.0878 (15)	Glp-Leu/Ile-hexose	Peptide	-	+	+	
65	13.5	425.1909	425.1929	$C_{20}H_{29}N_2O_8$	4.1	31	335.1612 (100), 263.1416 (69), 164.0715 (5)	Val-Phe-hexose	Peptide	-	+	+	
68	14.0	405.2198	405.2242	$C_{18}H_{33}N_2O_8$	10.9	22	315.1941 (100), 243.1707 (72), 130.0899 (4)	Leu/Ile-Leu/Ile-hexose 2	Peptide	-	+	+	
74	14.6	405.2239	405.2242	$C_{18}H_{33}N_2O_8$	0.8	28	315.1922 (100), 243.1719 (86), 130.0875 (3)	Leu/Ile-Leu/Ile-hexose 3	Peptide	-	+	+	
76	14.8	215.0842	215.0826	$C_{12}H_{11}N_2O_2$	-7.6	11	171.0885 (100), 142.0661 (34), 116.0503 (60)	1,2,3,4-tetrahydro-β-carboline-3-carboxylic acid (Lycoperodine 1)[a]	Trp-derived alkaloids	+	+	+	40
78	15.4	259.0730	259.0724	$C_{13}H_{11}N_2O_4$	-2.2	15	215.0841 (48), 171.0921 (100), 142.0653 (26), 116.0511 (55)	1,2,3,4-tetrahydro-β-carboline-1,3-dicarboxylic acid	Trp-derived alkaloids	+	+	+	
80	15.6	387.1306	387.1297	$C_{17}H_{23}O_{10}$	-2.4	20	225.0779 (100)	Geniposide	Iridoid	+	+	+	
82	15.7	405.2242	405.2242	$C_{18}H_{33}N_2O_8$	0.2	21	315.1921 (100), 243.1732 (61), 130.0870 (11)	Leu/Ile-Leu/Ile-hexose 4	Peptide	-	+	+	
83	16.0	241.1187	241.1194	$C_{11}H_{17}N_2O_4$	2.7	6	197.1304 (100), 130.0851 (17)	Glp-Leu/Ile 1	Peptide	-	+	+	
86	16.1	300.1558	300.1565	$C_{13}H_{22}N_3O_5$	-2	88	256.1653 (13), 214.1553 (100), 187.1097 (4), 169.0567 (5), 130.0872 (24)	Hyp-Gly-Leu/Ile	Peptide	+	+	+	
92	17.2	241.1187	241.1194	$C_{11}H_{17}N_2O_4$	2.7	6	197.1295 (100), 130.0886 (16)	Glp-Leu/Ile 2	Peptide	+	+	+	
95	18.4	439.2073	439.2086	$C_{21}H_{31}N_2O_8$	3	17	349.1778 (100), 277.1555 (72), 164.0685 (10)	Leu/Ile-Phe-hexose	Peptide	-	+	+	
96	18.6	358.1969	358.1984	$C_{16}H_{28}N_3O_6$	4.1	33	340.1866 (21), 323.1601 (100), 243.1668 (26)	Asp-Leu/Ile-Leu/Ile 1	Peptide	+	+	+	

97	18.7	275.0926	275.0925	$C_{15}H_{15}O_5$	-0.4	18	213.0903 (100), 185.0970 (50)	Lactucin	Sesquiterpene lactone	+	+	+	22, 44
99	19.1	186.1138	186.1136	$C_9H_{16}NO_3$	-1.3	13	142.1204 (9), 125.0980 (100)	Amino oxononanoic acid	Organic acid	+	+	+	
100	19.2	439.2051	439.2086	$C_{21}H_{31}N_2O_8$	3.5	8	349.1757 (100), 277.1535 (68), 130.0899 (13)	Phe-Ile/Leu-hexose	Peptide	-	+	+	
101	19.3	358.1974	358.1984	$C_{16}H_{28}N_3O_6$	2.7	47	340.1843 (19), 323.1607 (100), 279.1738 (15), 243.1692 (15), 130.0866 (27)	Asp-Leu/Ile-Leu/Ile 2	Peptide	+	+	+	
107	20.4	358.1976	358.1984	$C_{16}H_{28}N_3O_6$	2.1	38	340.1864 (23), 323.1615 (100), 279.1684 (21), 243.1668 (9)	Asp-Leu/Ile-Leu/Ile 3	Peptide	+	+	-	
117	23.2	392.1816	392.1827	$C_{19}H_{26}N_3O_6$	2.9	22	374.1759 (31), 357.1476 (100), 313.1568 (16), 277.1542 (11), 164.0727 (32)	Asp-Leu/Ile-Phe	Peptide	-	+	+	
126	24.1	341.0698	341.0700	$C_{15}H_{17}O_7S$	0.8	11	96.9594 (100)	8-deacetylmatricarin-8-sulfate[a]	Sesquiterpene lactone	+	+	+	48
127	24.5	339.0533	339.0544	$C_{15}H_{15}O_7S$	4	9	96.9593 (100)	15-deoxylactucin-8-sulfate	Sesquiterpene lactone	+	+	+	22
139	25.7	561.2340	561.2341	$C_{29}H_{37}O_{11}$	0.3	63	427.1963 (100), 409.1913 (36), 151.0435 (53)	1ß-(4-hydroxyphenylacetyl)-15-*O*-ß-D-glucopyranosyl-5α, 6ßH-eudesma-3-en-12, 6α-olide[a]	Sesquiterpene lactone	-	+	+	46
142	25.8	187.0978	187.0976	$C_9H_{15}O_4$	-1.1	2	169.0887 (4), 143.1092 (3), 125.0971 (100)	Azelaic acid[a]	Organic acid	+	+	+	37
157	33.1	481.1122	481.114	$C_{25}H_{21}O_{10}$	3.8	71	257.0819 (69), 213.0924 (100), 151.0399 (34), 107.0469 (4)	Lactucopicrin-15-oxalate	Sesquiterpene lactone	+	+	+	22
160	34.4	411.1437	411.1449	$C_{23}H_{23}O_7$	3.1	6	277.1087 (15), 259.0934 (10), 215.1079 (100), 151.0409 (41), 107.0487 (9)	11ß,13-dihydrolactucopicrin	Sesquiterpene lactone	+	+	+	45
161	34.4	409.1284	409.1293	$C_{23}H_{21}O_7$	2	68	257.0768 (30), 213.0930 (100), 151.0350 (23), 107.0516 (4)	Lactucopicrin	Sesquiterpene lactone	+	+	+	22, 44

- Amino acids are denoted by the three letter code: Asparagine, Asn, aspartic acid, Asp, glutamine, Gln, glutamic acid, Glu, glycine, Gly, hydroxyproline, Hyp, isoleucine, Ile, leucine, Leu, phenylalanine, Phe, pyroglutamic acid, Glp, tryptophan, Trp, tyrosine, Tyr, valine, Val.

[a] These compounds have been previously identified in several Asteraceae spp., but, to our knowledge, they have been reported for first time in *L. sativa*

Organic acids derivatives

A total of 11 known organic acids were tentatively identified in all the lettuce cultivars (**Table 1**), concretely, quinic (**4**) (*m/z* 191.0562, $C_7H_{11}O_6$)[36] and malic acid (**6**) (*m/z* 133.0142, $C_4H_5O_5$)[25,36], isomers citric (**9**) (*m/z* 191.0202, $C_6H_7O_7$)[36] and isocitric acid (**11**) (*m/z* 191.0200, $C_6H_7O_7$), fumaric (**13**) (*m/z* 115.0037, $C_4H_3O_4$)[25,36], citramalic (**14**)

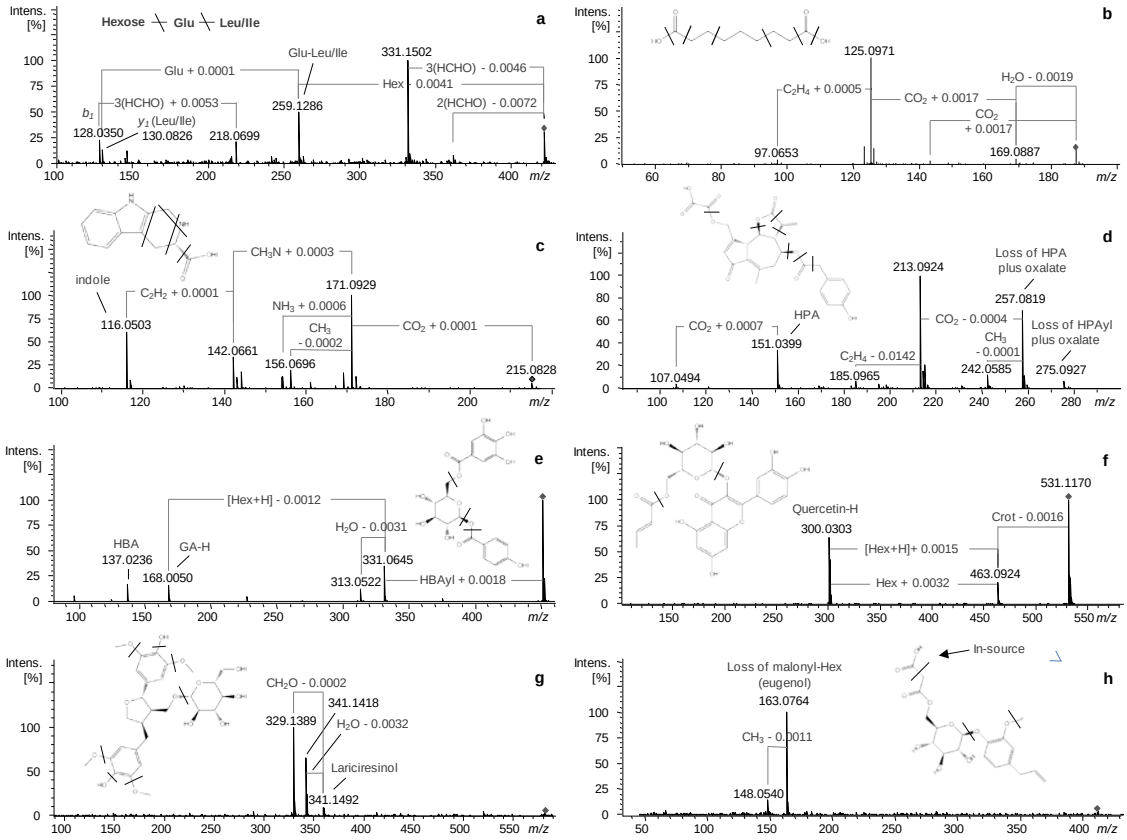


Fig. 3. Negative ESI-MS/MS spectra and structure of (a) dipeptide Glu-Leu/Ile conjugated to hexose (*m/z* 421.1816), (b) azelaic acid (*m/z* 187.0970), (c) lycoperodine 1 (*m/z* 215.0829), (d) lactucopicrin 15-oxalate (*m/z* 481.7841), (e) 2-*O*-*p*-hydroxybenzoyl-6-*O*-galloyl glucoside (*m/z* 451.0879), (f) quercetin-3-*O*-(6''-*O*-crotonyl)-ß-glucoside (*m/z* 531.1131), (g) (+)-5,5'-dimethoxy-9-*O*-ß-D-glucopyrnosyl lariciresinol, and (h) citrusin C malonate. In the case of the dipeptide, the fragment ions are labeled according to the nomenclature proposed by Roepstorff and Fohlman[85], and are 2 amu below the protonated counterpart. Crot, crotonyl; Glu, glutamic acid; Hex, hexose; HPA, hydroxyphenylacetic acid; HPAyl, hydroxyphenylacetyl; HBA, hydroxybenzoic acid; HBAyl, hydroxybenzoyl; Ile, isoleucine; Leu, leucine. The symbols and indicate the molecular ion $[M-H]^-$ and the possible sites of fragmentation, respectively.

(*m/z* 147.0296, $C_5H_7O_5$), succinic (**15**) (*m/z* 117.0185, $C_4H_5O_4$) [25,36], isopropylmalic (**60**) (*m/z* 175.0613, $C_7H_{11}O_5$) and azelaic acid (**142**) (*m/z* 187.0978, $C_9H_{15}O_4$)[37], and the amino derivatives, pantothenic acid (vitamine B5) (**37**) (*m/z* 218.1040, $C_9H_{16}NO_5$)[13] and amino oxononanoic acid (**99**) (*m/z* 186.1138, $C_9H_{16}NO_3$). Leafy vegetables showed higher concentrations of malic acid (on the order of 6 mg/g, approximately) than other

ones[36], explaining the high intensity of the corresponding BPC peaks in all lettuces (**Fig. 1**). In general terms, these compounds are small semi-polar organic molecules with RT between 2.1 and 13.2 min (**Fig. 2a**), except compounds **99** and **142**, which eluted later due to their increased alkyl chain and thus higher hydrophobicity.

Fragmentation patterns of compounds **6**, **9**, **11**, and **60**, were similar to those described by previous studies performed by IT[5] and QTOF[8], generating neutral losses of CO_2 from carboxylic group and/or H_2O. This fragmentation pattern was also found for molecular ion $[M-H]^-$ at *m/z* 187.0978 , which yielded fragment ions *m/z* 169.0887 $[M-H-H_2O]^-$, *m/z* 143.1092 $[M-H-CO_2]^-$, and *m/z* 125.0971 $[M-H-H_2O-CO_2]^-$ (**Fig. 3b**). Therefore, this compound was identified as azelaic acid, previously described in chicory (Asteraceae)[37]. Fragmentation of pantothenic acid, compound 37 at *m/z* 218.1040, produced a ion with *m/z* value of 146.0826, which corresponds to the loss of CO_2 and 2×(CH_2), whereas amino oxononanoic acid (*m/z* 186.1138) led to successive losses of CO_2 (*m/z* 142.1204) and their amino group, NH_3 (*m/z* 125.0980). This is the first study available to report the presence of isocitric, citramalic, azelaic and amino oxononanoic acids in lettuce. Organic acids strongly influence the organoleptic properties of vegetables, being responsible of sourness or acidity but also several of them, such as isopropylmalic and pantothenic acid, have a physiological role[5,13,36].

Nucleosides

Three nucleosides have been tentatively characterized in lettuce samples, uridine (**17**) (*m/z* 243.0622, $C_9H_{11}N_2O_6$), and adenosine (**27**) (*m/z* 266.0888, $C_{10}H_{12}N_5O_4$), previously reported in *Atractylodes lancea* (Asteraceae)[38], and guanosine (**28**) (*m/z* 282.0846, $C_{10}H_{12}N_5O_5$). In the case of adenosine and guanosine, purine nucleosides, the corresponding most prominent fragment ions at *m/z* 134.0467 and *m/z* 150.0419 were generated by the breakage of the glycosidic bond and losing the ribose moiety ($C_5H_8O_4$) (132.0423 u). In contrast, the pyrimidine nucleoside, uridine, released a main fragment ion with *m/z* value of 110.0247 ($C_5H_4NO_2$) after the loss of CHNO (43.0058 u) by the retro-Diels-Alder reaction and ring contraction and the subsequent loss of 3×(HCHO) from ribose, in agreement with MS^n experiments performed with IT[39].

Alkaloid derivatives

Alkaloids are a group of naturally occurring secondary metabolites that contain mostly basic nitrogen atoms. Two Trp-derived alkaloids with tetrahydro-ß-carboline structure, 1,2,3,4-tetrahydro-ß-carboline-3-carboxylic acid (lycoperodine 1) (**76**) (*m/z* 215.0842,

$C_{12}H_{11}N_2O_2$) and their carboxylic derivative (**78**) (*m/z* 259.0730, $C_{13}H_{11}N_2O_4$) were tentatively characterized in lettuce for the first time. Among these, lycoperodine 1 was previously identified in *Taraxacum formosanum* (Asteraceae)[40]. Regarding the MS/MS spectra, successive losses of CO_2 and CH_3N (29.0265 u), via the Retro-Diels-Alder reaction[41], and C_2H_2 were noted in both compounds, leading to the detection of the fragment ion at *m/z* 116.051 (indole group) and confirming the proposed structures. **Fig. 3c** shows the MS/MS spectrum of lycoperodine 1 as example. Antioxidant activity of these compounds has been previously reported[42].

Terpenoid derivatives: iridoid and sesquiterpene lactones

Iridoids

One iridoid (**80**) was detected at RT 15.6 min, before sesquiterpenoids. Geniposide was proposed for this compound since it gave the molecular ion $[M\text{-}H]^-$ at *m/z* 387.1306 corresponding to the molecular formula $C_{17}H_{23}O_{10}$. In MS/MS, aglycone form (genipin) at *m/z* 225.0779 ($C_{11}H_{13}O_5$) was detected, as well as a minor odd-electron ion at *m/z* value of 210.0519 ($C_{10}H_{10}O_5$), which corresponded to the loss of CH_3 from the methoxy group of genipin, in accordance with previous reported methoxylated compounds[5,8]. It bears noting that our detection system enabled us to identify such a minor bioactive compound, which possesses anti-diabetic, anti-inflammatory, detoxifying, anti-oxidative, and hypoglycemic properties[43], being reported for the first time in a member of the family Asteraceae.

Sesquiterpene lactones

Eight known sesquiterpene lactones eluted between 18.7 and 34.4 min were characterized, these being the most hydrophobic compounds of our study together with flavonoids and hydrolysable tannins (**Fig. 2a**). Lactucin-type guaianolides, lactucin (**97**) (*m/z* 275.0926, $C_{15}H_{15}O_5$)[22,44], 11ß,13-dihydrolactucopicrin (**160**) (*m/z* 411.1437, $C_{23}H_{23}O_7$)[45] and lactucopicrin (*m/z* 409.1284, $C_{23}H_{21}O_7$) (**161**)[22,44], are widespread in the genus *Lactuca*, including edible lettuces. In this sense, these compounds were found in all lettuce cultivars analyzed. The fragmentation pattern of lactucin leads to a loss of CO_2 from the lactone ring plus H_2O (*m/z* 213.0903) and C_2H_2 (*m/z* 185.0970), whereas lactucopicrin and its derivative, 11ß,13-dihydrolactucopicrin, gave respective daughter ions at *m/z* 257.0808 and *m/z* 259.0934, after loss of 4-hydroxyphenylacetic acid (152.0473 u), as described elsewhere[22]. In addition, the most prominent daughter ions were found at *m/z* 213.0848 and *m/z* 215.1079, by the loss of

CO_2, and minor fragments at *m/z* 151.0403 and *m/z* 151.0409 (4-hydroxyphenylacetic acid), respectively. A similar fragmentation pattern was also observed for the molecular ion $[M-H]^-$ at *m/z* 561.2340 ($C_{29}H_{37}O_{11}$) (**139**), which was detected only in iceberg and romaine lettuces. Based on the latter information and previously cited literature on the family Asteraceae, this compound was tentatively assigned to 1ß-(4-hydroxyphenylacetyl)-15-O-ß-D-glucopyranosyl-5α, 6ßH-eudesma-3-en-12, 6α-olide[46].

Molecular ions $[M-H]^-$ at *m/z* 341.0698 (**126**) and *m/z* 339.0541 (**127**) gave a prominent fragment ion at *m/z* 96.959, which corresponds to the sulfate group (HSO_4)[47], and thus their molecular formulas were generated considering the element "S" with the SmartFormula tool. It allowed the tentative assignment of compound **126** ($C_{15}H_{17}O_7S$) to 8-deacetylmatricarin-8-sulfate and **127** ($C_{15}H_{15}O_7S$) to 15-deoxylactucin-8-sulfate, previously described in Asteraceae[22,48]. The sesquiterpenoid lactucopicrin-15-oxalate (**157**) (*m/z* 481.1122, $C_{25}H_{21}O_{10}$) shows a similar fragmentation pattern to the one described above for lactucopicrin, but with the loss of oxalate (71.9847 u) plus 4-hydroxyphenylacetic acid to give a daughter ion at *m/z* 257.0808 (**Fig. 3d**)[22]. These three conjugates were identified in all lettuces cultivars.

These secondary metabolites are complementary to phenolics in determining both health-promoting and sensory properties of the Asteraceae vegetables, such as lettuce lactucin and lactucopicrin, which cause a bitter taste and bear sedative and analgesic properties as well as antimalarial activity[49,50], whereas sulfate conjugates might play a vital role in antimicrobial defense[46].

Phenolic metabolites
Phenolic acids and derivatives

More than 50 phenolic acid derivatives were found to be distributed among the three cultivars of lettuce analyzed, eluting between 7.3 and 31.6 min (**Fig. 2a**). The compounds were classified as hydroxybenzoic, hydroxycinnamic, and hydroxyphenylacetic acids, including their esters and related compounds (**Table 2**). Qualitative differences were found between the lettuce cultivars, baby being richer than the others (**Fig. 2b**). In general, losses of H_2O and CO_2 were regularly observed in the MS/MS spectra of phenolic acids, which have also been described by other authors using IT, QqQ, and QTOF[5,8,51].

Hydroxybenzoic derivatives

Among the hydroxybenzoic acids, simple forms were detected, such as two isomers of dihydroxybenzoic acid at *m/z* 153.0200 (**47**) and 153.0193 (**93**) with molecular formula $C_7H_5O_4$, and hydroxybenzoic acid at *m/z* 137.0246 (**64**) and 137.0238 (**138**) with molecular formula $C_7H_5O_3$, previously reported in Asteraceae[52-56]. In the case of glycoside esters, hydroxybenzoic acid hexose (**30**) (*m/z* 299.0782, $C_{13}H_{15}O_8$) and tachioside (**31**) (*m/z* 301.0935, $C_{13}H_{17}O_8$)[57], characterization was made according to the MS data and fragmentation pattern by the neutral loss of the glycosidic moiety. Aside from this loss, vanillic acid glucoside (**40**) (*m/z* 329.0877, $C_{14}H_{17}O_9$) showed the losses of CH_3 (*m/z* 152.0120) from the methoxy group of the aglycone (*m/z* 167.0351), CO_2 (*m/z* 123.0437) and CH_3 plus CO_2 (59.0133 u) (*m/z* 108.0220), in agreement with the bibliography[58]. A similar pattern was found for syringic acid hexose (**48**) (*m/z* $C_{15}H_{19}O_{10}$). The aglycone forms, vanillic acid and syringic acid, have previously been described in *L. sativa* cv. crispa after alkaline and acid hydrolysis[59] and in various Asteraceae species[52,54,55]. Notably, compounds **29** and **33** with a *m/z* of 331.0672 and 331.0677 ($C_{13}H_{15}O_{10}$) gave an odd electron daughter ion at *m/z* 168.010, corresponding to the loss of hexose plus H ($C_6H_{11}O_5$) (163.0606 u), and an even electron ion at *m/z* 169.014, loss of hexose. The compounds were identified as galloyl-hexose isomers and they were detected only in the cv. baby. The release of such unusual losses were also observed for two isomers of dihydroxybenzoyl-hexose (**36** and **43**) (*m/z* 315.0721 and 315.0723, $C_{13}H_{15}O_9$), found in the three lettuce cultivars. In addition, 2-*O*-*p*-hydroxybenzoyl-6-*O*-galloyl glucose was proposed as compound **150** (*m/z* 451.0879, $C_{20}H_{19}O_{12}$). In MS/MS, it yielded fragment ions at *m/z* 331.0645 (galloyl-glucose group), which suffer the subsequent loss of H_2O (*m/z* 313.0522) and hexose plus H (*m/z* 168.0050) (gallic acid-H), and *m/z* 137.0236 (hydroxybenzoic acid) (**Fig. 3e**). The presence of this compound in an antibacterial extract of *Melaleuca ericifolia*, a member of Myrtaceae family, and its structure were recently reported by Hussein et al.[60]. Likewise, two isomers constituted by hydroxybenzoyl-dihydroxybenzoyl-hexose (**151** and **153**) (*m/z* 435.0927 and 435.0923, $C_{20}H_{19}O_{11}$) were tentatively characterized based on their molecular formula (and thus mass weight) and MS/MS spectra, since no information related to them was found in the literature. These gave rise to the simultaneous presence of hydroxybenzoic and dihydroxybenzoic acid product ions. In addition, losses of the hydroxybenzoyl moiety (*m/z* 315.0762) and hydroxybenzoic acid (*m/z* 297.0649) were detected from the parent ion, and remaining part formed by dihydroxybenzoyl-hexose and dihydroxybenzoyl-hexosyl, respectively (**Table 2**). These novel hydroxybenzoic conjugates were found only in the cv. baby.

Hydroxycinnamic derivatives

Previous studies have shown that hydroxycinnamic acids contribute to the antioxidant properties of extracts of different cultivars of *L. sativa* and several Asteraceae species [21,61]. It bears mentioning that this phenolic group makes up the largest class of secondary metabolites, with 36 compounds tentatively identified in all three of the lettuce cultivars studied (baby, iceberg, and romaine) and with a different distribution in each (**Table 2**). In this sense, according to the MS and MS/MS data, caffeic acid (3,4-dihydroxycinnamic acid) (**87**) (*m/z* 179.0353, $C_9H_7O_4$) was detected in all lettuce cultivars, in accordance with the literature[28,63], and the characteristic loss of CO_2 (*m/z* 135.0449) in MS/MS. Two isomers of ferulic acid methyl ester were tentatively proposed for molecular ions $[M-H]^-$ at *m/z* 207.0668 ($C_{11}H_{11}O_4$) (**90** and **135**). In agreement with Gomez-Romero et al.[58], they showed demethylated fragment ions at *m/z* 192.0457 ($[M-CH_3-H]^-$) and *m/z* 177.0187 ($[M-2\times(CH_3)-H]^-$), which is characteristic of methoxylated cinnamic acids. One isomer was previously found in *Taraxacum formosanum*[62].

Table 2. Phenolic compounds detected in aqueous-methanolic extracts from green leaf lettuce cultivars baby (B), iceberg (I) and romaine (R) by RP-UPLC-ESI-QTOF-MS.

Nº	RT (min)	Observed *m/z* ($[M-H]^-$)	Calculated *m/z* ($[M-H]^-$)	Molecular formula	Error (ppm)	mσ	Major fragments *m/z* ($[M-H]^-$) (%)	Proposed Compound	Class	B	I	R	Ref.
29	7.3	331.0672	331.0671	$C_{13}H_{15}O_{10}$	-0.5	7	313.0562 (22), 169.0136 (8), 168.0065 (53), 125.0233 (29)	Galloyl-hexose 1	Hydroxybenzoic	+	-	-	
30	7.5	299.0782	299.0772	$C_{13}H_{15}O_8$	-3.1	19	137.0253 (100)	Hydroxybenzoic acid hexose	Hydroxybenzoic	+	+	+	
31	7.6	301.0935	301.0929	$C_{13}H_{17}O_8$	-2	7	139.0392 (100)	3-Methoxy-4-hydroxyphenyl-1-*O*-ß-D-glucoside (tachioside)[a]	Hydroxybenzoic	+	-	-	57
33	7.8	331.0677	331.0671	$C_{13}H_{15}O_{10}$	-1.8	2	313.0561 (19), 169.0140 (9), 168.0095 (50)	Galloyl-hexose 2	Hydroxybenzoic	+	-	-	
35	8.4	343.1034	343.1035	$C_{15}H_{19}O_9$	0.3	14	181.0508 (100), 163.0401 (16), 135.0447 (3)	Dihydrocaffeic acid hexose	Hydroxycinnamic	+	+	+	
36	8.6	315.0721	315.0722	$C_{13}H_{15}O_9$	0.1	2	153.0115 (100), 108.0208 (41)	Dihydroxybenzoic acid hexose 1	Hydroxybenzoic	+	+	+	
38	9.2	341.0892	341.0878	$C_{15}H_{17}O_9$	-4	20	179.0347 (37), 161.0271 (49), 135.0431 (15)	Caffeoyl-hexose 1[b]	Hydroxycinnamic	+	+	+	64
40	9.4	329.0877	329.0878	$C_{14}H_{17}O_9$	0.3	4	167.0351 (100), 152.0120 (20), 123.0437 (21), 108.0220 (26)	Vanillic acid glucoside	Hydroxybenzoic	+	+	+	
42	9.9	261.0073	261.0074	$C_9H_9O_7S$	0.5	5	181.0506 (100), 163.0407 (27), 135.0462 (16), 119.0502 (19)	Dihydrocaffeic acid sulfate	Hydroxycinnamic	+	+	+	
43	10.0	315.0723	315.0722	$C_{13}H_{15}O_9$	0.1	2	153.0189 (100), 109.0293 (32)	Dihydroxybenzoic acid hexose 2	Hydroxybenzoic	+	+	+	
45	10.2	313.0929	313.0929	$C_{14}H_{17}O_8$	-0.1	30	151.0397 (94), 107.0493 (13)	4-hydroxyphenylacetyl glucoside	Hydroxyphenylacetic	+	+	+	
46	10.3	341.0892	341.0878	$C_{15}H_{17}O_9$	-4	20	n.d.	Caffeoyl-hexose 2[b]	Hydroxycinnamic	+	-	-	64
47	10.3	153.02	153.0193	$C_7H_5O_4$	-4.4	17	109.0295 (100)	Dihydroxybenzoic acid 1[b]	Hydroxybenzoic	-	+	+	53,54,56
48	10.6	359.0982	359.0984	$C_{15}H_{19}O_{10}$	0.3	4	197.0458 (100), 182.0227 (27), 153.0561 (41), 138.0323 (28), 123.0097 (3)	Syringic acid hexose	Hydroxybenzoic	+	+	+	
49	10.7	311.0413	311.0409	$C_{13}H_{11}O_9$	-1.4	3	179.0356 (100), 149.0084 (66), 135.0454 (53)	Caffeoyltartaric acid (caftaric acid) 1[c]	Hydroxycinnamic	+	+	+	18,21,23, 28,29
50	10.9	311.0411	311.0409	$C_{13}H_{11}O_9$	-0.7	4	179.0354 (100), 149.0093 (44), 135.0441 (50)	Caffeoyltartaric acid (caftaric acid) 2[c]	Hydroxycinnamic	+	+	+	18,21,23, 28,29
52	11.2	353.0874	353.0878	$C_{16}H_{17}O_9$	1.2	8	191.0564 (100), 179.0357(78), 135.0447 (20)	Caffeoylquinic acid 1[d]	Hydroxycinnamic	+	-	-	18,21,23, 28,29,64
56	11.5	341.0884	341.0878	$C_{15}H_{17}O_9$	-1.7	12	179.0351 (100), 135.0453 (18)	Caffeoyl-hexose 3[b]	Hydroxycinnamic	+	+	+	64
57	12.1	341.0879	341.0878	$C_{15}H_{17}O_9$	-0.3	23	251.0580 (28), 179.0357 (81), 161.0247 (100), 135.0465 (23)	Caffeoyl-hexose 4[b]	Hydroxycinnamic	+	+	+	64
58	12.5	339.0714	339.0722	$C_{15}H_{15}O_9$	2.4	14	177.0195 (100)	Esculetin 6-O-glucoside (aesculin, vitamin C2)/7-O-glucoside (cichoriin)[a]	Coumarin	+	+	+	78,55
61	13.3	355.1040	355.1035	$C_{16}H_{19}O_9$	-1.4	20	193.0507 (100), 149.0622 (27)	Ferulic acid glucoside[a]	Hydroxycinnamic	+	+	+	65
63	13.4	341.0867	341.0878	$C_{15}H_{17}O_9$	3.4	11	179.0344 (100), 135.0444 (29)	Caffeoyl-hexose 5[b]	Hydroxycinnamic	-	+	+	
64	13.4	137.0246	137.0244	$C_7H_5O_3$	-1	11	n.d.	Hydroxybenzoic acid 1[b]	Hydroxybenzoic	+	+	+	52-55,62
66	13.7	295.0460	295.0459	$C_{13}H_{11}O_8$	-0.1	10	163.0397 (100), 119.0505 (27)	Coumaroyltartaric acid	Hydroxycinnamic	+	+	+	18,29
67	13.9	341.0872	341.0878	$C_{15}H_{17}O_9$	1.9	47	179.0351 (100), 135.0437 (15)	Caffeoyl-hexose 6[b]	Hydroxycinnamic	+	-	-	64
69	14.1	353.0878	353.0878	$C_{16}H_{17}O_9$	0	4	191.0565 (100), 179.0362 (1)	Caffeoylquinic acid (chlorogenic) 2[d]	Hydroxycinnamic	+	+	+	18,21,23, 28,29,64
70	14.2	639.1203	639.1203	$C_{27}H_{27}O_{18}$	0	29	463.0855 (100)	Quercetin hexose glucuronide 1	Flavonol	+	-	-	
71	14.4	341.0872	341.0878	$C_{15}H_{17}O_9$	1.9	47	179.0354 (100), 135.0456 (22)	Caffeoyl-hexose 7[b]	Hydroxycinnamic	+	+	+	64

72	14.4	385.1139	385.114	$C_{17}H_{21}O_{10}$	0.4	11	223.0623 (100), 208.0392 (20), 179.0719 (74), 164.0487 (28)	Sinapoyl glucoside	Hydroxycinnamic	+	+	+	
73	14.5	325.0931	325.0929	$C_{15}H_{17}O_8$	-0.7	8	163.0406 (100), 119.0507 (38)	*p*-Coumaroyl glucoside	Hydroxycinnamic	+	+	-	
75	14.7	639.1242	639.1203	$C_{27}H_{27}O_{18}$	0.3	12	463.0888 (100), 301.0388 (20)	Quercetin hexose glucuronide 2	Flavonol	+	-	-	
77	15.3	349.0931	349.0929	$C_{17}H_{17}O_8$	-0.5	25	215.1074 (100), 191.0615 (15)	Unknown 1, with quinic acid		+	+	-	
79	15.5	151.0400	151.0401	$C_8H_7O_3$	1.3	10	123.0436 (55), 107.0496 (100)	4-hydroxyphenylacetic acid[a]	Hydroxyphenylacetic	+	+	+	54,55
81	15.6	725.1196	725.1207	$C_{30}H_{29}O_{21}$	1.4	6	681.1307 (84), 505.0956 (100)	Quercetin 3-O-(6″-O-malonyl)-glucoside 7-O-glucuronide	Flavonol	+	+	+	21
84	16.1	177.0192	177.0193	$C_9H_5O_4$	0.7	25	149.0244 (20), 133.0293 (100), 105.0351 (23)	6,7-dihydroxycoumarin (esculetin)[a]	Coumarin	+	+	+	55
85	16.1	711.1389	711.1355	$C_{30}H_{31}O_{20}$	3.5	25	667.1544 (100), 505.0907 (4), 463.0906 (3)	Quercetin 3-O-(6″-O-malonyl)-glucoside 7-O-glucoside	Flavonol	+	+	+	21
87	16.2	179.0353	179.035	$C_9H_7O_4$	-1.6	0	135.0449 (100)	3,4-dihydroxycinnamic acid (caffeic acid)	Hydroxycinnamic	+	+	+	28,63
88	16.7	353.0872	353.0878	$C_{16}H_{17}O_9$	1.7	8	191.0564 (100), 179.0337 (1)	Caffeoylquinic acid 3[d]	Hydroxycinnamic	+	-	-	18,21,23, 28,29,64
89	17.0	579.2079	579.2083	$C_{28}H_{35}O_{13}$	0.8	9	417.1545 (100), 399.1428 (73), 384.1074 (7)	Syringaresinol ß-D-glucoside[a]	Lignan	+	-	-	76
90	17.1	207.0668	207.0663	$C_{11}H_{11}O_4$	-2.3	3	192.0457 (29), 177.0187 (100)	Ferulic acid methyl ester 1[a]	Hydroxycinnamic	+	+	+	62
91	17.2	337.0928	337.0929	$C_{16}H_{17}O_8$	0.3	16	191.0563 (100), 173.0448 (13), 163.0392 (14)	*p*-coumaroylquinic acid 1[e]	Hydroxycinnamic	+	-	-	18,29,64
93	17.5	153.0193	153.0193	$C_7H_5O_4$	0.4	6	135.0113 (32), 109.0297 (100)	Dihydroxybenzoic acid 2[b]	Hydroxybenzoic	+	-	-	53,54,56
94	18.0	295.0461	295.0459	$C_{13}H_{11}O_8$	-0.7	3	133.0141 (100), 179.0342 (33), 115.0040 (28)	Caffeoylmalic acid	Hydroxycinnamic	+	+	+	18,21,28, 65
98	18.9	621.2190	621.2189	$C_{30}H_{37}O_{14}$	-0.2	32	417.1593 (89), 399.1435 (100), 307.1405 (45), 189.0768 (79)	Unknown 2, with acetylhexose		+	-	-	
102	19.5	337.0931	337.0929	$C_{16}H_{17}O_8$	-0.7	18	191.0560 (100), 173.0442 (9), 163.0387 (2)	*p*-coumaroylquinic acid 2[e]	Hydroxycinnamic	+	-	+	18,29,64
103	19.7	275.1031	275.1037	$C_{14}H_{15}N_2O_4$	2.1	20	147.0446 (16), 127.0510 (100)	Cinnamic acid-dihydrothymine	Hydroxycinnamic	-	+	+	
104	19.9	473.0742	473.0725	$C_{22}H_{17}O_{12}$	-3.6	11	311.0418 (28), 293.0320 (31), 179.0359 (100), 149.0082(100), 135.0457 (9)	Dicaffeoyltartaric acid 1[e]	Hydroxycinnamic	+	+	+	18,21,23, 28,65
105	20.0	609.1447	609.1461	$C_{27}H_{29}O_{16}$	2.2	9	285.0397 (9)	Luteolin diglucoside[a]	Flavone	+	+	+	71
106	20.1	463.0875	463.0882	$C_{21}H_{19}O_{12}$	1.6	7	300.0292 (12), 301.0367 (59)	Quercetin hexose 1[f]	Flavonol	+	-	-	21,28,63, 67
108	20.5	473.0714	473.0725	$C_{22}H_{17}O_{12}$	2.3	3	311.0411 (29), 293.0308 (35), 179.0353 (100), 149.0090 (100), 135.0447 (7)	Dicaffeoyltartaric acid 2[e]	Hydroxycinnamic	+	+	+	18,21,23, 28,65
109	21.1	531.1134	531.1144	$C_{25}H_{23}O_{13}$	1.9	44	339.0499 (100), 295.0563(22), 191.0538 (17), 179.0349 (24)	Unknown 3, derived from caffeic and quinic acids	Other phenolic	+	-	-	
110	21.6	467.1914	467.1923	$C_{23}H_{31}O_{10}$	1.9	7	293.0297 (24), 219.0326 (14), 179.0354 (90), 149.0089 (100)	Unknown 4, derived from caffeic and tartaric acids	Other phenolic	+	-	-	
111	21.6	581.2235	581.2240	$C_{28}H_{37}O_{13}$	0.9	25	359.1492 (10), 341.1418 (62), 329.1389 (100)	(+)-5,5'-Dimethoxy-9-O-beta-D-glucopyranosyl lariciresinol (alangilignoside C)	Lignan	+	+	+	
112	21.7	473.0717	473.0725	$C_{22}H_{17}O_{12}$	1.8	4	311.0413 (12), 293.0305 (18), 179.0354 (63), 149.0092 (100), 135.0453 (7)	Dicaffeoyltartaric acid 3[e]	Hydroxycinnamic	+	-	-	18,21,23, 28,65
113	22.5	477.0665	477.0675	$C_{21}H_{17}O_{13}$	1.9	3	301.0353 (100)	Quercetin 3-glucuronide	Flavonol	+	+	+	18,21,28, 67
114	22.6	461.0722	461.0725	$C_{21}H_{17}O_{12}$	0.8	55	285.0409 (100)	Luteolin 7-glucuronide	Flavone	+	+	+	21,28

115	22.6	463.0875	463.0882	$C_{21}H_{19}O_{12}$	1.6	8	300.0278 (100), 301.0328 (73)	Quercetin hexose 2[f]	Flavonol	+	+	+	21,28,63, 67
116	22.6	447.0925	447.0933	$C_{21}H_{19}O_{11}$	1.6	8	285.0403 (100)	Luteolin 7-glucoside	Flavone	+	+	+	21,28
118	23.2	447.1283	447.1297	$C_{22}H_{23}O_{10}$	3.1	18	193.0512 (43), 175.0402 (71), 151.0392 (100)	Di(4-hydroxyphenylacetyl)-hexose 1	Hydroxyphenylacetic	+	+	+	
119	23.3	341.1235	341.1242	$C_{16}H_{21}O_8$	2	12	179.0717 (100)	Coniferyl alcohol glucoside (coniferoside)[a]	Monolignol	+	-	-	79
120	23.5	433.0785	433.0776	$C_{20}H_{17}O_{11}$	-2	67	300.0265 (100), 301.0297 (70)	Quercetin 3-arabinoside	Flavonol	+	-	-	68
121	23.7	447.1301	447.1297	$C_{22}H_{23}O_{10}$	-1.1	26	193.0517 (57), 175.0396 (100), 151.0392 (89)	Di(4-hydroxyphenylacetyl)-hexose 2	Hydroxyphenylacetic	+	+	+	
122	23.8	457.0773	457.0776	$C_{22}H_{17}O_{11}$	0.8	17	311.0422(40), 295.0453 (39), 277.0352 (38), 179.0363 (30), 163.0396 (100), 149.0132(6)	Caffeoyltartaric-*p*-coumaroyl acid 1[c]	Hydroxycinnamic	+	-	-	18,28
123	23.9	515.1189	515.1195	$C_{25}H_{23}O_{12}$	1.1	12	353.0873 (100), 191.0565 (44), 179.0351 (29), 135.0454 (4)	Dicaffeoylquinic acid 1[g]	Hydroxycinnamic	+	-	-	18,21,23, 28,29,61
124	23.9	549.0874	549.0886	$C_{24}H_{21}O_{15}$	2.2	9	505.0975 (100), 463.0867 (4), 301.0321 (28)	Quercetin malonylglucoside 1[c]	Flavonol	+	-	-	18,21,28, 67
125	24.0	579.1726	579.1719	$C_{27}H_{31}O_{14}$	-1.2	9	271.0603 (83)	Naringenin 7-neohesperidoside (naringin o narirutin)[a]	Flavanone	-	+	+	74
128	24.7	549.0882	549.0886	$C_{24}H_{21}O_{15}$	2.9	8	505.0975 (100), 463.0859 (4), 301.0339 (27), 300.0279 (32)	Quercetin malonylglucoside 2[c]	Flavonol	+	+	+	18,21,28, 67
129	24.7	515.1189	515.1195	$C_{25}H_{23}O_{12}$	1.1	12	353.0872 (100), 191.0561 (43), 179.0352 (27), 135.0465 (3)	Dicaffeoylquinic acid (isochlorogenic acid A) 2[g]	Hydroxycinnamic	+	+	+	18,21,23, 28,29,61
130	25.1	457.0756	457.0776	$C_{22}H_{17}O_{11}$	4.4	12	311.0411(70), 295.0451 (36), 277.0407 (33), 179.0363 (30), 163.0396 (100), 149.0090 (49)	Caffeoyltartaric-*p*-coumaroyl acid 2[c]	Hydroxycinnamic	+	-	-	18, 28
131	25.1	447.0922	447.0933	$C_{21}H_{19}O_{11}$	2.5	26	301.0354 (79), 300.0405 (100)	Quercetin 3-*O*-rhamnoside (quercitrin)	Flavonol	-	+	+	67,68
132	25.2	607.1694	607.1668	$C_{28}H_{31}O_{14}$	-4.1	47	299.0574 (100)	4'-methoxyapigenin rutinoside (acacetin 7-*O*-rutinoside, linarin)[a]	Flavone	-	+	+	75
133	25.2	609.1807	609.1825	$C_{28}H_{33}O_{15}$	2.9	12	301.0729 (100)	Hesperetin 7-O-rutinoside (hesperidin)[a]	Flavanone	+	+	+	52,74
134	25.2	431.0970	431.0984	$C_{21}H_{19}O_{10}$	3.2	9	269.0485 (15)	Apigenin glucoside[a]	Flavone	+	-	+	76
135	25.3	207.0668	207.0663	$C_{11}H_{11}O_4$	-2.4	33	177.0178 (100)	Ferulic acid methyl ester 2[a]	Hydroxycinnamic	+	-	-	40
136	25.3	515.1182	515.1195	$C_{25}H_{23}O_{12}$	2.6	8	353.0875 (100), 191.0561 (53), 179.0363 (30), 135.0489 (3)	Dicaffeoylquinic acid 3[g]	Hydroxycinnamic	+	-	+	
137	25.4	445.0741	445.0718	$C_{21}H_{17}O_{11}$	8	63	269.0451 (100)	Apigenin 7-*O*-glucuronide[a]	Flavone	+	-	-	
138	25.6	137.0238	137.0244	$C_7H_5O_3$	4.5	12	93.0325 (100)	Hydroxybenzoic acid 2[b]	Hydroxybenzoic	+	-	-	40,52-54
140	25.8	533.0943	533.0937	$C_{24}H_{21}O_{14}$	-1.2	29	489.1016 (100), 285.0361 (15)	Luteolin malonylhexose 1[a]	Flavonol	+	-	-	72
141	25.8	665.2090	665.2087	$C_{31}H_{37}O_{16}$	-0.4	40	417.1543 (100)	Syringaresinol malonylhexose	Lignan	+	-	+	
143	25.9	515.1176	515.1195	$C_{25}H_{23}O_{12}$	3.7	12	353.0888 (100), 191.0572 (15), 179.0374 (18)	Dicaffeoylquinic acid 4[g]	Hydroxycinnamic	+	+	-	18,21,23, 28,29,61
144	26.5	533.0926	533.0937	$C_{24}H_{21}O_{14}$	2.1	19	489.1023 (100), 285.0414 (98)	Luteolin malonylhexose 2[a]	Flavonol	+	+	+	72
145	27.4	337.0943	337.0929	$C_{16}H_{17}O_8$	-4	20	191.0550 (100)	*p*-coumaroylquinic acid 3[e]	Hydroxycinnamic	+	-	-	18,29,64
146	27.5	405.1178	405.1191	$C_{20}H_{21}O_9$	3.2	29	243.0634 (24)	2,3,5,4'-Tetrahydroxystilbene 2-*O*-ß-D-glucoside	Stilbene	+	+	-	
147	27.6	499.1241	499.1246	$C_{25}H_{23}O_{11}$	1.1	9	353.0913 (78), 191.0557 (100), 179.0354 (46)	*p*-Coumaroyl-caffeoylquinic acid[a]	Hydroxycinnamic	+	-	-	29
148	28.8	531.1131	531.1144	$C_{25}H_{23}O_{13}$	2.6	24	463.0924 (21), 301.0364 (43), 300.0303 (64)	Quercetin-3-*O*-(6"-*O*-crotonyl)-ß-glucoside[a]	Flavonol	-	+	+	69

149	30.2	411.1286	411.1297	$C_{19}H_{23}O_{10}$	2.7	6	163.0763 (100), 148.0540 (13)	Eugenol malonylglucoside (citrusin C malonate)[a]	Phenylpropene	+	-	-	80
150	30.4	451.0879	451.0882	$C_{20}H_{19}O_{12}$	0.6	24	331.0645 (36), 313.0522 (13), 167.9985 (16), 137.0236 (17)	2-*O*-*p*-hydroxybenzoyl-6-*O*-galloyl glucoside[a]	Hydrolybenzoic	+	-	-	60
151	30.7	435.0927	435.0933	$C_{20}H_{19}O_{11}$	-2.4	25	315.0762 (33), 297.0649 (21), 152.0111 (24), 153.0201 (27), 137.0223 (48)	Hydroxybenzoyl dihydroxybenzoyl-hexose1	Hydroxybenzoic	+	-	-	
152	30.9	593.1897	593.1876	$C_{28}H_{33}O_{14}$	-3.6	12	285.0765 (100)	Scorzocreticin 8-*O*-rhamnoside-1"-*O*-glucoside (scorzocreticoside II)[a]	Dihydroisocoumarin	-	+	+	81
153	31.6	435.0923	435.0933	$C_{20}H_{19}O_{11}$	-2.4	25	315.0717 (52), 297.0649 (11), 153.0176 (67), 137.0239 (78)	Hydroxybenzoyl dihydroxybenzoyl-hexose2	Hydroxybenzoic	+	-	-	
154	31.9	285.0410	285.0405	$C_{15}H_{9}O_{6}$	-1.9	2	217.0486 (2), 199.0373 (3), 175.0443 (2), 151.0091 (2), 133.0309 (3)	Luteolin	Flavone	+	+	+	19
155	32.6	581.1662	581.1664	$C_{30}H_{29}O_{12}$	0.5	18	175.0416 (44), 151.0396 (100), 143.0354 (36)	Tri-4-hydroxyphenylacetyl glucoside 1[c]	Hydrolyzable tannins	+	+	+	22
156	32.9	581.1659	581.1664	$C_{30}H_{29}O_{12}$	1	5	151.0396 (100), 143.0316 (45)	Tri-4-hydroxyphenylacetyl glucoside 2[c]	Hydrolyzable tannins	+	+	+	22
158	33.4	581.1662	581.1664	$C_{30}H_{29}O_{12}$	0.4	5	295.0843 (10), 175.0405 (17), 151.0401 (100), 143.0347 (43)	Tri-4-hydroxyphenylacetyl glucoside 3[c]	Hydrolyzable tannins	+	+	+	22
159	34.2	581.1656	581.1664	$C_{30}H_{29}O_{12}$	1.5	11	295.0825 (44), 175.0403 (100), 151.0406 (59)	Tri-4-hydroxyphenylacetyl glucoside 4[c]	Hydrolyzable tannins	+	+	+	22

[a]These compounds have been previously identified in several Asteraceae spp., but, to our knowledge, they have been reported for first time in L. sativa

[b]Six isomers of caffeoyl-hexose, including caffeic acid-4-ß-D-glucose, isomers 2,5- (gentisic acid) and 3,5- (□-resorcilic acid) and 3,4- (protocatechuic acid) dihydroxybenzoic acid, and isomers 2- (salicylic acid), 3- , and 4-hydroxybenzoic acid have been previously identified in Asteraceae spp. In the latter case, 3-hydroxibenzoic acid is the least frequent (see KNApSAcK core system database).

[c]Only one isomer have been previously described in L. sativa, in concrete caffeoyltartaric-p-coumaroyl acid, quercetin 3-O-(6-O-malonyl-ß-D-glucoside), and 2,3,4-tri-O-4-hydroxyphenylacetyl-ß-glucopyranose.

[d]Compounds could correspond to four possible caffeoylquinic acid isomers, 3-, 4-, 5-, and cis-5-caffeoylquinic acid, previously identified in green lettuces cultivars

[e]Isomers di- (chicoric acid) and meso-di-O-caffeoyltartaric acid, and 5-p-coumaroylquinic acid have been previously identified in L. sativa, whereas 3-, 4-, and cis-5-p-coumaroylquinic acid have been also described in other Asteraceae spp.

fQuercetin 3-glucoside (isoquercitrin), 7-glucoside, and 3-galactoside have been previously identified in green L. sativa, being isoquercitrin the most common.

[g]Isomers 3,5- (isochlorogenic acid A), cis-3,5-, and 4,5- (isochlorogenic acid B) dicaffeoylquinic acid have been previously reported in L. sativa, whereas other eight isomers more were also found in several Asteraceae spp.

Moreover, several glycosylated hydroxycinnamates were found and their fragmentation pattern was characterized by their losing a hexose residue, and thus the aglycone product ion could be detected in MS/MS[9]. In this way, dihydrocaffeic acid hexose (**35**) (*m/z* 343.1034, $C_{15}H_{19}O_9$) was tentatively identified. In addition to the ion corresponding to the aglycone (*m/z* 181.0508) derived from the precursor ion, the subsequent losses of H_2O (*m/z* 163.0401) and CO (27.9949 u) (*m/z* 135.0447) were also detected. Compounds **38**, **46**, **56**, **57**, **63**, **67,** and **71** showed *m/z* values of 341.0867-341.0892 ($C_{15}H_{17}O_9$), which were tentatively identified as isomeric forms of caffeoyl-hexose, in agreement with Clifford et al.[64]. The fragmentation pattern was quite similar for all of them, mainly the caffeic acid ion and as well as losses of H_2O (*m/z* 161.0243) and CO_2 (*m/z* 135.0444) were observed. Thus, it was impossible to distinguish the sugar ester, the position of linking, or their stereochemistry. According to this fragmentation pattern, ferulic acid glucoside (**61**) (*m/z* 355.1040, $C_{16}H_{19}O_9$) was also characterized[65], as was sinapoyl glucoside (**72**) (*m/z* 385.1139, $C_{17}H_{21}O_{10}$) and *p*-coumaroyl glucoside (**73**) (*m/z* 325.0931, $C_{15}H_{17}O_8$), showing the major daughter ions at *m/z* 193.0507 (ferulic acid ion), *m/z* 223.0623 (sinapic acid ion), and *m/z* 163.0406 (*p*-coumaric acid ion), respectively, with the subsequent losses of CO_2. Sinapoyl glucoside also led to the loss of CH_3 (*m/z* 208.0392), and CH_3 plus CO_2 (*m/z* 164.0487) from sinapic acid moiety, as shown above for methoxylated derivatives. On the other hand, compound **42** produced a molecular ion $[M\text{-}H]^-$ at *m/z* 261.0073 and a prominent MS/MS ion at *m/z* 181.0506, which is in keeping with the loss of sulfate (SO_3) (79.9568 u) and the presence of dihydrocaffeic acid[66]. The fragmentation pattern was similar to that found for compound **35**, fragments being subsequently detected at *m/z* 163.0407, *m/z* 135.0462, and *m/z* 119.0502, which corresponded to the loss of H_2O, CO and CO_2, respectively, in agreement with previous findings[58]. Thus, this compound was tentatively identified as dihydrocaffeic acid sulfate. In general, these compounds were detected in most of the three lettuce types.

Following our strategy, 18 compounds were characterized as esters formed between certain cinnamic acids and several organic acids, in agreement with literature [18,21,23,25,28,29,63]. These types of metabolites led to the detection of caffeic (*m/z* 179.0350) and *p*-coumaric acid (*m/z* 163.0401) and as well malic (*m/z* 133.0142), tartaric (*m/z* 149.0092), and quinic (*m/z* 191.0561) acid ions in MS/MS, depending on the compound. In this manner, two isomers of caffeoyltartaric acid (caftaric acid) (**49**

and **50**) (*m/z* 311.0411-311.0413, $C_{13}H_{11}O_9$) were found, three isomers of caffeoylquinic acid (**52**, **69**, and **88**) (*m/z* 353.0872-353.0878, $C_{16}H_{17}O_9$), coumaroyltartaric acid (**66**) (*m/z* 295.0460, $C_{13}H_{11}O_8$), three isomers of *p*-coumaroylquinic acid (**91**, **102** and **145**) (*m/z* 337.0928-337.0943, $C_{16}H_{17}O_8$), and caffeoylmalic acid (**94**) (*m/z* 295.0461, $C_{13}H_{11}O_8$). According to the literature cited above, at the start of our study only one isomer of caffeoyltartaric acid had been previously reported. In addition, caffeoylquinic isomers **52** and **88** could be 1-, 3-, *cis*-3-, 4-, *cis*-4- or *cis*-5-caffeoylquinic acid previously found different Asteraceae species [29,64], whereas compound **69** was present in all three lettuce cultivars, which could be identified as 5-caffeoylquinic acid (chlorogenic acid). This is one of the major phenolic components of green lettuce cultivars, including iceberg and romaine [18,21,25,65]. Related to these compounds, three molecular ions ($[M\text{-}H]^-$) were found, with *m/z* values ranging from 473.0714 to 473.0742 (**104**, **108** and **112**, $C_{22}H_{17}O_{12}$), which yielded MS/MS fragment ions at *m/z* 311.041 and *m/z* 293.031, and indicates the successive loss of the caffeoyl moiety and caffeic acid from the precursor ion, leaving the part formed by the caffeoyltartaric acid. Besides, the appearance of caffeic and tartaric acid ions in their MS/MS spectra confirmed the presence of a second caffeoyl and a tartaric moiety, in agreement with the literature[18]. Therefore, these compounds were identified as isomeric forms of dicaffeoyltartaric acid. In this respect, isomers di-*O*-caffeoyltartaric (chicoric acid), mainly, and *meso*-di-*O*-caffeoyltartaric acid have been detected in lettuce elsewhere[18,21,28]. Similarly, two isomers of caffeoyltartaric-*p*-coumaroyl acid (**122** and **130**) (*m/z* 457.0773 and 457.0756, $C_{22}H_{17}O_{11}$) were also identified, which led to MS/MS fragment ions at *m/z* 311.041 ($[\text{caffeoyltartaric-H}]^-$), *m/z* 295.045 ($[p\text{-coumaroyltartaric acid-H}]^-$), *m/z* 277.040 ($[p\text{-coumaroyltartaric-}H_2O\text{-H}]^-$), and as well as caffeic acid, *p*-coumaric acid and tartaric acid ions. One isomer has previously been found in green lettuce cultivars, also exhibiting a similar fragmentation pattern[18,28]. Finally, four isomers of dicaffeoylquinic acid (**123**, **129**, **136**, and **143**) (*m/z* 515.1176-515.1189, $C_{25}H_{23}O_{12}$) and a *p*-coumaroyl-caffeoylquinic acids (**147**) (*m/z* 499.1241, $C_{25}H_{23}O_{11}$) were also successfully characterized in lettuce extracts, according to the MS data and literature[18,29]. In this sense, only isomers 3,5 (isochlorogenic acid A)-, *cis*-3,5-, and 4,5 (isochlorogenic acid B)-dicaffeoylquinic acid have previously been reported in *L. sativa*[18,21,23,28]. Among these, isochlorogenic acid A was reported to be the most abundant and, thus, it could be assigned as compound **129** (**Fig. 1**). On the other hand, isomers 1-*p*-coumaroyl-5-caffeoylquinic and 1-*p*-coumaroyl-3-caffeoylquinic have been recently discovered in Asteraceae species and

reported in nature for first time[29], but the MS/MS data did not enable the correct assignment of compound **147** to one of them.

Hydroxyphenylacetic derivatives

According to the MS data, the aforementioned fragmentation pattern, and the above-cited literature, 4-hydroxyphenylacetic acid (**79**) (*m/z* 151.0400, $C_8H_7O_3$) was characterized, showing the typical decarboxylation of phenolic acids in MS/MS[58]. In relation to this compound, 4-hydroxyphenylacetyl-hexose (**45**) (*m/z* 313.0929, $C_{14}H_{17}O_8$) and isomers di-4-hydroxyphenylacetyl-hexose (**118** and **121**) (*m/z* 447.1283-447.1301) ($C_{22}H_{23}O_{10}$) were tentatively identified. The first derivative gave a major fragment corresponding to the loss of the sugar and the second derivatives led to the loss of 4-hydroxyphenylacetyl plus 4×(HCHO) (from the attached sugar) (*m/z* 193.051), and 4-hydroxyphenylacetic acid plus 4×(HCHO) (*m/z* 175.040), which resulted in the appearance of a daughter ion at *m/z* 151.0392 (4-hydroxyphenylacetic acid) in the spectra. The sugar would probably be glucose, according to the results of Sessa et al. [22]. All of these compounds were detected in the three lettuce cultivars studied. Among these, 4-hydroxyphenylacetic acid has been previously described in the family Asteraceae[54,55], while the presence of its glycosylated derivatives is reported for first time in this family.

Flavonoids

There is increasing interest in flavonoids due to the wide range of health benefits reported in recent decades. With the analytical technique applied, our study has revealed that this group is more widely spread in *L. sativa* than reported in previous studies[18,21,28,67]. It was the second group qualitatively richest in lettuce (**Fig. 2b**), and eluted between 14.2 and 31.9 min (**Fig. 2a**). In agreement with these studies, quercetin (flavonol) and luteolin (flavone) were the most representative and appeared to be linked to sugars and organic acids. In this sense, the conjugated group could be established by the mass difference between the conjugated parent ion and the non-conjugated product ion found in MS/MS: 132.0423 u (pentose, $C_5H_8O_4$), 146.0579 u (deoxyhexose, $C_6H_{10}O_4$), hexose, 308.1107 u (deoxyhexose-hexose, $C_{12}H_{20}O_9$), 324.1056 u (di-hexose, $C_{12}H_{20}O_{10}$), 68.0262 u (crotonyl, C_4H_4O), 86.0004 u (malonyl, $C_3H_2O_3$), 176.0321 u (glucuronide, $C_6H_8O_6$), 248.0532 u (malonyl plus hexose, $C_9H_{12}O_8$), and 338.0849 u (glucuronide plus hexose, $C_{12}H_{18}O_{11}$). Therefore, daughter ions corresponding to quercetin, *m/z* 300.0276 (odd electron ion) and/or 301.0354 (even

electron ion)[9], and luteolin, *m/z* 285.0405, were observed elsewhere in MS/MS. In this way, the following were tentatively identified: two isomers of quercetin hexose-glucuronide (**70** and **75**) (*m/z* 639.1203 and639.1242, $C_{27}H_{27}O_{18}$), quercetin 3-*O*-(6″-*O*-malonyl)-glucoside 7-*O*-glucuronide (**81**) (*m/z* 725.1196, $C_{30}H_{29}O_{21}$)[21], quercetin 3-*O*-(6″-*O*-malonyl)-glucoside 7-*O*-glucoside (**85**) (*m/z* 711.1389, $C_{30}H_{31}O_{20}$)[21], two isomers of quercetin hexose (**106** and **115**) (*m/z* 463.0875, $C_{21}H_{19}O_{12}$)[21,28,63,67], quercetin 3-glucuronide (**113**) (*m/z* 477.0665, $C_{21}H_{17}O_{13}$)[18,21,28,67], quercetin 3-arabinoside (**120**) (*m/z* 433.0785, $C_{20}H_{17}O_{11}$)[68], two isomers of quercetin malonylglucoside (**124** and **128**) (*m/z* 549.0874-549.0882, $C_{24}H_{21}O_{15}$)[18,21,28,67], quercetin 3-*O*-rhamnoside (quercitrin) (**131**) (447.0922, $C_{21}H_{19}O_{11}$)[67,68], and quercetin-3-*O*-(6"-*O*-crotonyl)-ß-glucoside (**148**) (*m/z* 531.1131, $C_{25}H_{23}O_{13}$)[69]. In addition, under our MS and MS/MS conditions, the neutral loss of CO_2 was a common feature of compounds presenting the malonyl group, such as several of the latter compounds. Be aware of this in-source fragmentation can affect the correct identification of the molecular ion $[M\text{-}H]^-$ of interest, which could appear in lower abundance than the product ion $[M\text{-}H\text{-}CO_2]^-$ in MS. This particularly labile group could be partially lost during ion transfer from a higher-pressure region of the source to a lower-pressure region, e.g. while passing through the nozzle, skimmers, or ion funnels, as suggested previously[70]. It is worth noting that this study found two isomers (**124** and **128**) of quercetin malonylglucoside, whereas only the isomer quercetin 3-*O*-(6-*O*-malonyl-ß-D-glucoside had previously been described in lettuce[18,21,28,67]. Moreover, compounds **70**, **75** and **148** are reported for first time in this species. For example, the fragmentation pattern of the latter compound, quercetin-3-*O*-(6"-*O*-crotonyl)-ß-glucoside, is shown in **Fig. 3f**. It was characterized by the losses of crotonyl (68.0262 u), hexose and hexose plus H moieties to give fragment ions at *m/z* 463.0924, *m/z* 300.0303 and *m/z* 301.0364, respectively.

Regarding luteolin conjugates, luteolin diglucoside (**105**) (*m/z* 609.1447, $C_{27}H_{29}O_{16}$)[71], luteolin 7-glucuronide (**114**) (*m/z* 461.0722, $C_{21}H_{17}O_{12}$)[21,28], luteolin 7-glucoside (**116**) (*m/z* 447.0925, $C_{21}H_{19}O_{11}$)[21,28], and two isomers of luteolin malonylhexose (**140** and **144**) (*m/z* 533.0943-533.0926, $C_{24}H_{21}O_{14}$)[72] were found. On the other hand, molecular ion ($[M\text{-}H]^-$) at *m/z* 285.0410 was tentatively identified as luteolin (**154**), based on the suggested molecular formula ($C_{15}H_9O_6$) and its MS/MS. In this sense, several minor daughter ions at *m/z* 217.0486 ($C_{12}H_9O_4$), 199.0373 ($C_{12}H_7O_3$), 175.0443 ($C_{10}H_7O_3$), *m/z* 151.0091 ($C_7H_3O_4$) ($^{1,3}A^-$ fragment) and *m/z* 133.0309 ($C_8H_5O_2$) ($^{1,3}B^-$) were detected, which helps to distinguish it from its isomers kaempferol and scutellarein. This

fragmentation pattern agrees with the results found by ESI(−)-MS/MS using IT as the mass analyzer[58,73] and with the databases consulted. As commented above, luteolin is characteristic of green lettuce cultivars, whereas its isomer kaempferol has been found in red lettuce cultivars and other Asteraceae species[19,21].

Finally, two flavanones, naringenin 7-neohesperidoside (naringin, narirutin) (**125**) (*m/z* 579.1726, $C_{27}H_{31}O_{14}$)[74] were also successfully assigned, as were hesperetin 7-*O*-rutinoside (hesperidin) (**133**) (*m/z* 609.1807, $C_{28}H_{33}O_{15}$)[52,74], and three flavones, 4'-methoxyapigenin rutinoside (acacetin 7-*O*-rutinoside, linarin) (**132**) (*m/z* 607.1694, $C_{28}H_{31}O_{14}$)[75], apigenin glucoside (**134**) (*m/z* 431.0970, $C_{21}H_{19}O_{10}$)[76], and apigenin 7-*O*-glucuronide (**137**) (*m/z* 445.0741, $C_{21}H_{17}O_{11}$)[77]. As commented above, neutral losses of the conjugated groups were observed, and daughter ions were detected at *m/z* 271.0603, *m/z* 301.0729, *m/z* 299.0574, and *m/z* 269.0585, indicating the presence of naringenin, hesperetin, methoxyapigenin, and apigenin in their structures, respectively. These compounds were distributed differently throughout the three lettuce cultivars studied (**Table 2**).

Lignan derivatives

Regarding lignans, compound **89** (*m/z* 579.2079) was proposed as syringaresinol ß-D-glucoside based on its generated molecular formula ($C_{28}H_{35}O_{13}$) and MS/MS spectra. This ion yielded fragment ions at *m/z* 417.1545, 399.1428 and 384.1074, which corresponded to the respective loss of the glucose moiety (syringaresinol), and then H_2O and CH_3. This was found only in cv. baby though previously cited in Asteraceae[76]. Similarly, syringaresinol malonylhexose (*m/z* 665.2090, $C_{31}H_{37}O_{16}$) (**141**) was characterized, this giving the same main daughter ion at *m/z* 417.1543 (syringaresinol) and being detected in baby and romaine. Based on compound **111**, with an *m/z* value of 581.2235 ($C_{28}H_{37}O_{13}$), it was tentatively identified as (+)-5,5'-dimethoxy-9-*O*-ß-D-glucopyranosyl lariciresinol (alangilignoside C) and was present in all of the lettuce cultivars. In the MS/MS spectra, the latter compound showed the loss of the glucose moiety plus 2×(CH_2O) (two methoxy groups) (222.0739 u), releasing lariciresinol (*m/z* 359.1492), and then H_2O (*m/z* 341.1379) and CH_2O (30.0105 u) (another methoxy group of lariciresinol) (*m/z* 329.1400) (**Fig. 3g**).

Hydrolysable tannins

Tannins can be divided into two groups, hydrolysable tannins and condensed tannins. The first are generally formed by three galloyl (tri-hydroxybenzoyl) groups linked to a monosaccharide nucleus, but also a 4-hydroxyphenylacetyl ester of glucose was detected in milky latex by Sessa et al.[22]. Notably, our methodology detected four isomers of this compound (**155**, **156**, **158**, and **159**) (*m/z* 581.1656-581.1662, $C_{30}H_{29}O_{12}$). The fragmentation pattern of these compounds was consistent with reports by the latter authors, observing the loss of 4-hydroxyphenylacetic acid plus 4-hydroxyphenylacetyl (286.0841 u) (*m/z* 295.0843), two 4-hydroxyphenylacetic acids plus 4-hydroxyphenylacetyl (438.1315 u) (*m/z* 143.0347) (sugar residue ion), and the ion corresponding to 4-hydroxyphenylacetic acid at *m/z* 151.0401.

Other phenolic compounds

In the three lettuce extracts, two coumarins were tentatively identified: 6,7-dihydroxycoumarin (esculetin) (**84**) (*m/z* 177.0192, $C_9H_5O_4$) and its glucosidic derivative (**58**) (*m/z* 339.0714, $C_{15}H_{15}O_9$). Esculetin produced the major fragment ion at *m/z* 133.0293, corresponding to the loss of CO_2, and subsequently C_2H_4 (*m/z* 105.0351). Since coumarins are structurally lactones derived from hydroxycinnamic acids by cyclization, the MS/MS of both types of compounds were also similar. In the case of the glycosylated derivative, the loss of the sugar moiety was observed, thereby detecting the aglycone ion (esculetin). Both coumarins have been previously described in Asteraceae[55,78]. Moreover, a monolignol, coniferyl alcohol glucoside (coniferoside) (**119**) (*m*/z 341.1235, $C_{16}H_{21}O_8$)[79], a stilbene, 2,3,5,4'-Tetrahydroxystilbene 2-*O*-β-D-glucoside (**146**) (*m/z* 405.1178, $C_{20}H_{21}O_9$), and a phenylpropene derivative, eugenol malonylglucoside (citrusin C malonate) (**149**) (*m/z* 411.1286, $C_{19}H_{23}O_{10}$)[80], and a dihydroisocoumarin, scorzocreticoside II (**152**) (*m/z* 593.1897, $C_{28}H_{33}O_{14}$)[81], were also tentatively proposed, and found mainly in cv. baby. For example, the MS/MS spectrum of the latter compound is shown in **Fig. 3h**. In this case, the loss of CO_2 (*m/z* 367.1411, $C_{18}H_{23}O_8$) from the malonyl moiety by in-source fragmentation was detected, as commented for malonic conjugates in Section 3.2.2.2. In MS/MS, the parent ion gave fragments at *m/z* 163.0763 ($C_{10}H_{11}O_2$) and *m/z* 148.0540 ($C_9H_8O_2$), which were the subsequent losses of glucose plus malonic moiety (eugenol) and CH_3 from the methoxy group attached to the benzene ring of eugenol, confirming its assignation. Likewise, these fragments were also detected in the MS analysis. This compound has recently

been discovered in *Leucosceptrum* japonicum (Lamiaceae)[80]. In addition, for all of these compounds, this is the first report available in lettuce.

Prediction of novel sulfate and amino acid conjugates of known sesquiterpene lactones

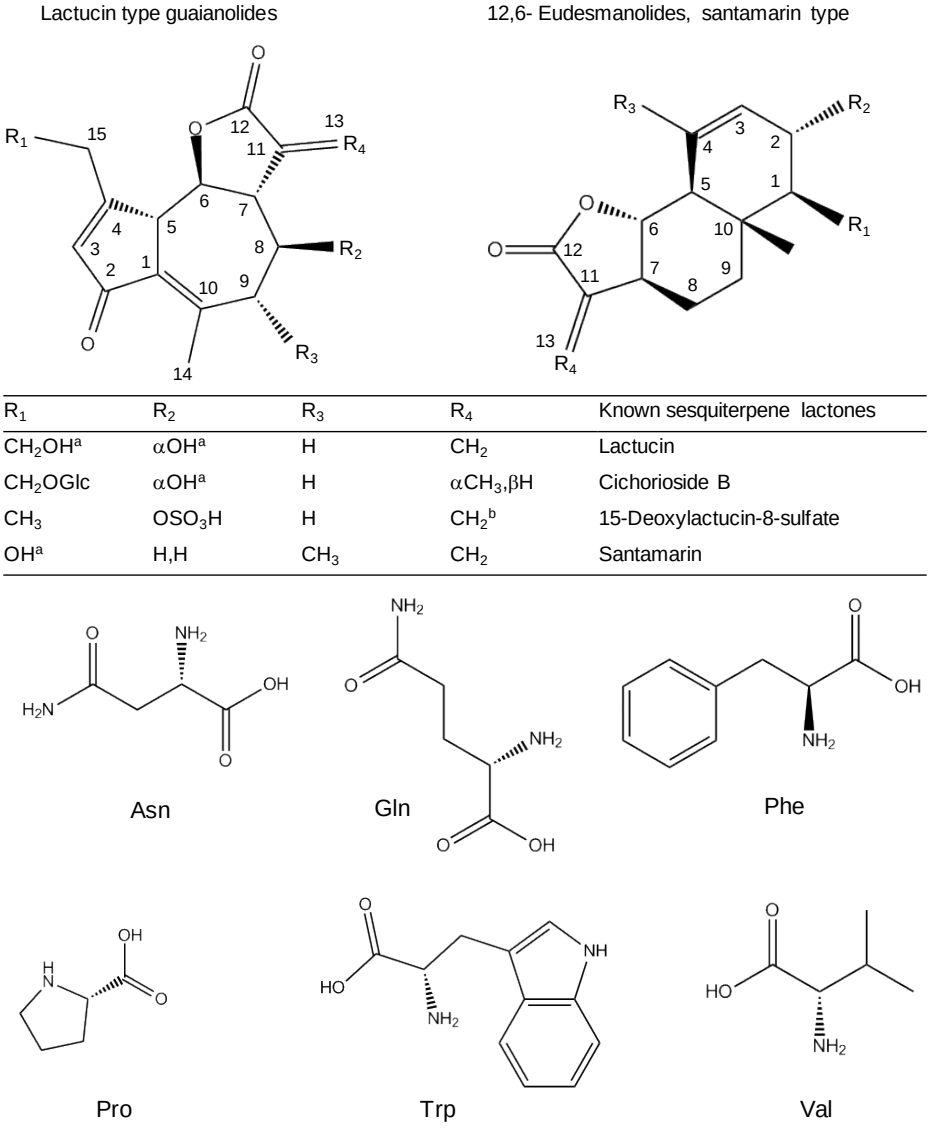
Along the consciously profiling of the extracts, 10 compounds with common features in their MS/MS spectra were found (**Table 3**). Notably, the sulfate ion was detected in all of them and, when "S" was considered to obtain their elemental composition, six compounds also contained "N". The amino acid component of the latter compounds was detected in MS/MS and it was also established by mass difference between the molecular parent ion and the non-conjugated fragment ion. Thus, based on the latter findings and the literature on Asteraceae, the compounds were tentatively identified as conjugated sulfate and amino acid forms of widely distributed sesquiterpene lactones found in the genus *Lactuca*[16]: **1'** (*m/z* 519.1185, $C_{21}H_{27}O_{13}S$), cichorioside B-sulfate, **2'** (*m/z* 357.0652, $C_{15}H_{17}O_8S$), dihydrolactucin-sulfate, **3'** (*m/z* 355.0495, $C_{15}H_{15}O_8S$), lactucin-sulfate, **4'** (*m/z* 471.1075, $C_{19}H_{23}N_2O_{10}S$), 15-deoxylactucin-8-sulfate-asparagine, **5'** (*m/z* 485.1232, $C_{20}H_{25}N_2O_{10}S$), 15-deoxylactucin-8-sulfate-glutamine, **6'** (*m/z* 454.1175, $C_{20}H_{24}NO_9S$), 15-deoxylactucin-8-sulfate-proline, **7'** (*m/z* 456.1321, $C_{20}H_{26}NO_9S$), 15-deoxylactucin-8-sulfate-valine, **8'** (*m/z* 504.1326, $C_{24}H_{26}NO_9S$), 15-deoxylactucin-8-sulfate-phenylalanine, **9'** (*m/z* 543.1445, $C_{26}H_{26}N_2O_9S$), 15-deoxylactucin-8-sulfate-tryptophan, and **10'** (*m/z* 327.0907, $C_{15}H_{19}O_6S$), santamarin-sulfate. Upon our MS/MS conditions, it was not possible to distinguish where the sulfate and amino acid moieties could be linked with the sesquiterpene lactone backbone. In this sense, the sulfate group should be located where free OH groups are found, concretely at the C-8 and C-15 positions of lactucin and at the C-8 and C-1 of cichorioside B and santamarin, respectively. Amino acids could probably be linked with C-13, as has been reported for similar compounds found in Asteraceae[82,83]. These types of compounds might be produced biogenetically via a Michael type attack of the endogenous amino acids to the sesquiterponoid lactones, as suggested by Cha et al. [84]. The suggested structure of these compounds is presented in **Figure 2**. This study is the first available to report the occurrence of amino acid conjugated to sesquiterpene lactones in *Lactuca* genus, highlighting the power of the methodology used for discovering unexplored novel sulfate metabolites in plants. Further studies are needed in order to confirm the structure of these novel compounds and elucidate their stereochemistry.

Table 3. Novel sulfate and amino acid conjugates of known sequiterpene lactones proposed by RP-UPLC-ESI-QTOF-MS and -MS/MS and detected in green leaf lettuce cultivars baby (B), iceberg (I) and romaine (R).

							Neutral loss										
N°	RT (min)	Observed *m/z* ([M-H]$^-$)	Calculated *m/z* ([M-H]$^-$)	Molecular formula	Error (ppm)	m □	$-NH_3$	$-CO_2$	$-SO_3$	$-(Aa-NH_3)$	$-Aa$	$[Aa-H]^-$	HSO_4^-	Proposed Compound	B	I	R
1'	12.4	519.1185	519.1178	$C_{21}H_{27}O_{13}S$	-1.3	7.8							96.9592	Cichorioside B-sulfate	+	+	+
2'	14.4	357.0652	357.0650	$C_{15}H_{17}O_8S$	-0.7	5.9			277.1076				96.9589	Dihydrolactucin-sulfate	+	+	+
3'	14.7	355.0495	355.0493	$C_{15}H_{15}O_8S$	-0.5	9.3							96.9589	Lactucin-sulfate	+	+	+
4'	15.4	471.1075	471.1079	$C_{19}H_{23}N_2O_{10}S$	0.9	13.1	454.0834	427.1097	391.1486	356.0811	339.0475	131.0417	96.9588	15-deoxylactucin-8-sulfate-Asn	+	+	+
5'	15.7	485.1232	485.1235	$C_{20}H_{25}N_2O_{10}S$	0.7	8.6			405.1672	356.0804	339.0550	145.0618	96.9591	15-deoxylactucin-8-sulfate-Gln	+	+	+
6'	16.7	454.1175	454.1177	$C_{20}H_{24}NO_9S$	0.4	29.6					339.0555	114.0570	96.9593	15-deoxylactucin-8-sulfate-Pro	+	+	+
7'	18.6	456.1321	456.1334	$C_{20}H_{26}NO_9S$	2.8	53.3				356.0764	339.0555	116.0701	96.9606	15-deoxylactucin-8-sulfate-Val	-	+	+
8'	25.5	504.1326	504.1334	$C_{24}H_{26}NO_9S$	1.5	16.3			259.0998[a]	356.0776	339.0564	164.0713	96.9593	15-deoxylactucin-8-sulfate-Phe	+	+	+
9'	27.8	543.1445	543.1443	$C_{26}H_{26}N_2O_9S$	-0.5	15.7				356.0837	339.0535		96.9594	15-deoxylactucin-8-sulfate-Trp	+	+	+
10'	31.6	327.0907	327.0908	$C_{15}H_{19}O_6S$	0.2	6.7		203.1461[b]	247.1337				96.9592	Santamarin-sulfate	+	+	+

[a]From fragment ion at *m/z* 339.0564

[b]After neutral loss of SO_3 from parent ion at *m/z* 327.0907



R_1	R_2	R_3	R_4	Known sesquiterpene lactones
CH_2OH[a]	αOH[a]	H	CH_2	Lactucin
CH_2OGlc	αOH[a]	H	$\alpha CH_3,\beta H$	Cichorioside B
CH_3	OSO_3H	H	CH_2[b]	15-Deoxylactucin-8-sulfate
OH[a]	H,H	CH_3	CH_2	Santamarin

Figure 2. Tentative proposed structures for novel sulfate and amino acid conjugates of known sesquiterpene lactones. [a]Possible sites of sulfation; [b]possible sites of conjugation to amino acids. Asn, asparagine; Glc, glucose; Gln, glutamine; Phe, phenylalanine; Pro, proline; Trp, tryptophan; Val, valine.

Conclusion

This study demonstrates the power of the RP-UPLC-ESI-QTOF-MS for "non-targeted" metabolic profiling of vegetables in one run without tedious sample pretreatments. It allowed the tentative characterization of a total of 171 compounds belonging to various structural classes: amino acids and peptides (42), organic acids (11), nucleosides (3), alkaloids (2), phenolic compounds (92), sesquiterpene lactones (17) and one iridioid. Our results indicate that genotype, at least in part, was an important

factor affecting the metabolic composition of lettuce. Furthermore, all the cultivars studied proved qualitatively rich in phenolic compounds (hydroxybenzoic, hydroxyphenylacetic and hydroxycinnamic acids, flavonols and flavones, mainly), especially the baby cultivar, underscoring that lettuce is a good source of natural antioxidants, besides its nutritional value. In addition, two antioxidant tryptophan-derived alkaloids were also characterized. Most of these compounds have not been previously reported in lettuce. Among them, 10 novel structures formed by the conjugation of known amino acids and sesquiterpene lactones were proposed. Thus, the methodology applied is useful to identify known phytochemicals, and as well as to discover new metabolites in an initial attempt prior to NMR and spectroscopy studies. These results could be useful in further chemosystematics, epidemiological studies, and as well for nutritional recommendations and programming, allowing the selection of salad ingredients with abundant bioactive phytochemicals.

References

1 Cano, A. and Arnao, M. B. 2005. Hydrophilic and lipophilic antioxidant activity in different leaves of three lettuce varieties. *International journal of food properties.,* **8,** 521-528.

2 Liu, R. H. 2003. Health benefits of fruit and vegetables are from additive and synergistic combinations of phytochemicals. *The American Journal of Clinical Nutrition,* **78,** 517S-520S.

3 Dai, J. and Mumper, R. J. 2010. Plant Phenolics: Extraction, Analysis and Their Antioxidant and Anticancer Properties. *Molecules,* **15,** 7313-7352.

4 Borras Linares, I., Arraez-Roman, D., Herrero, M., Ibáñez, E., Segura-Carretero, A. and Fernandez-Gutierrez, A. Comparison of different extraction procedures for the comprehensive characterization of bioactive phenolic compounds in Rosmarinus officinalis by reversed-phase high-performance liquid chromatography with diode array detection coupled to electrospray time-of-flight mass spectrometry. *Journal of Chromatography A,* **1218,** 7682-7690.

5 Gómez-Romero, M. A., Segura-Carretero, A. and Fernández-Gutiérrez, A. Metabolite profiling and quantification of phenolic compounds in methanol extracts of tomato fruit. *Phytochemistry,* **71,** 1848-1864.

6 Allwood, J.W., Parker, D., Beckmann, M., Draper J. and Goodacre R. 2012. In: N.W. Hardy and R.D. Hall (eds.), Plant Metabolomics: Methods and Protocols (Methods in Molecular Biology), vol. 860, Humana Press, pp. 157.

7 Rogachev, I. and Aharoni, A., in: J.M. Walker (Ed), Methods in Molecular Biology, Vol. 860, Humana press, pp. 129.

8 Iswaldi, I., Gómez-Caravaca, A. M. A., Lozano-Sanchez, J. S., Arrã¡Ez-Romã¡N, D., Segura-Carretero, A. and Fernandez-Gutierrez, A. Profiling of phenolic and other polar compounds in zucchini (Cucurbita pepo L.) by reverse-phase high-performance liquid chromatography coupled to quadrupole time-of-flight mass spectrometry. *Food Research International,* **50,** 77-84.

9 Abu-Reidah, I. M., Arraez-Roman, D., Quirantes-Pina, R., Fernandez-Arroyo, S., Segura-Carretero, A. and Fernandez-Gutiérrez, A. 2011. HPLC-ESI-Q-TOF-MS for a comprehensive characterization of bioactive phenolic compounds in cucumber whole fruit extract. *Food Research International,* **46,** 108-117.

10 Baslam, M., Pascual, I., Sanchez-Díaz, M., Erro, J., García-Mina, J. M. A. and Goicoechea, N. Improvement of Nutritional Quality of Greenhouse-Grown Lettuce by Arbuscular Mycorrhizal Fungi Is Conditioned by the Source of Phosphorus Nutrition. *Journal of Agricultural and Food Chemistry,* **59,** 11129-11140.

11 European Commission Heath & Consumer Protection Directorate-General, Guidelines on comparability, extrapolation, group tolerances and data requirements for setting MRLs. SANCO 7525/VI/95 - rev. 9 (2011).

12 Santos, J., Mendiola, J. A., Oliveira, M. B. P. P., Ibáñez, E. and Herrero, M. Sequential determination of fat- and water-soluble vitamins in green leafy vegetables during storage. *Journal of Chromatography A,* **1261,** 179-188.

13 Mou, B. 2009. Nutrient Content of Lettuce and its Improvement. *Current Nutrition & Food Science,* **5,** 242-248.

14 Blasa, M., Gennari, L., Angelino, D. and Ninfali, P. 2010. In: R.R. Watson, V.R. Preedy (Eds), Bioactive Foods in Promoting Health. Fruits and Vegetables, Academic Press, pp. 37.

15 Serafini, M., Bugianesi, R., Salucci, M., Azzini, E., Raguzzini, A. and Maiani, G. 2002. Effect of acute ingestion of fresh and stored lettuce (Lactuca sativa) on plasma total antioxidant capacity and antioxidant levels in human subjects. *British Journal of Nutrition,* **88,** 615-623.

16 Zidorn, C. 2008. Sesquiterpene lactones and their precursors as chemosystematic markers in the tribe Cichorieae of the Asteraceae. *Phytochemistry,* **69,** 2270-2296.

17 Chaturvedi, D., 2011. In: Tiwari, V. K. (Ed), *Opportunity, Challenge and Scope of Natural Products in Medicinal Chemistry*, Research Signpost, Kerala (India), pp. 313.

18 Ribas-Agustí, A., Gratacós-Cubarsí, M., Sárraga, C., García-Regueiro, J.-A. and Castellari, M. 2011. Analysis of Eleven Phenolic Compounds Including Novel p-Coumaroyl Derivatives in Lettuce (Lactuca sativa L.) by Ultra-high-performance Liquid Chromatography with Photodiode Array and Mass Spectrometry Detection. *Phytochemical Analysis,* **22,** 555-563.

19 Cao, J., Chen, W., Zhang, Y., Zhang, Y. and Zhao, X. 2010. Content of Selected Flavonoids in 100 Edible Vegetables and Fruits. *Food Science and Technology Research*, **16,** 395-402.

20 Altunkaya, A. and Gokmen, V. 2009. Effect of various anti-browning agents on phenolic compounds profile of fresh lettuce (L. sativa). *Food Chemistry,* **117,** 122-126.

21 Llorach, R., Martinez-Sanchez, A. N., Tomas-Barberan, F. A., Gil, M. A. I. and Ferreres, F. 2008. Characterisation of polyphenols and antioxidant properties of five lettuce varieties and escarole. *Food Chemistry,* **108,** 1028-1038.

22 Sessa, R. A., Bennett, M. H., Lewis, M. J., Mansfield, J. W. and Beale, M. H. 2000. Metabolite Profiling of Sesquiterpene Lactones from Lactuca Species: Major Latex Components Are Novel Oxalate And Sulfate Conjugates Of Lactucin And Its Derivatives. *Journal of Biological Chemistry,* **275,** 26877-26884.

23 Loaiza-Velarde, J. G., Tomás-Barberá, F. A., and Saltveit, M. E. 1997. Effect of Intensity and Duration of Heat-shock Treatments on Wound-induced Phenolic metabolism in Iceberg Lettuce. *Journal of American Society for Horticultural Science*, **122,** 873-877.

24 Lee, J. G., Lee, B. Y. and Lee, H. J. 2006. Accumulation of phytotoxic organic acids in reused nutrient solution during hydroponic cultivation of lettuce (Lactuca sativa L.). *Scientia Horticulturae,* **110,** 119-128.

25 Sobolev, A. P., Brosio, E., Gianferri, R. and Segre, A. L. 2005. Metabolic profile of lettuce leaves by high-field NMR spectra. *Magnetic Resonance in Chemistry,* **43,** 625-638.

26 Krishnan, P., Kruger, N. J. and Ratcliffe, R. G. 2005. Metabolite fingerprinting and profiling in plants using NMR. *Journal of Experimental Botany,* **56,** 255-265.

27 Moco, S., Bino, R. J., Vorst, O., Verhoeven, H. A., de Groot, J., van Beek, T. A., Vervoort, J. and Ric de Vos, C. H. 2006. A Liquid Chromatography-Mass Spectrometry-Based Metabolome Database for Tomato. *Plant Physiology,* **141,** 1205-1218.

28 Lin, L.-Z., Harnly, J., Zhang, R.-W., Fan, X.-E. and Chen, H.-J. Quantitation of the Hydroxycinnamic Acid Derivatives and the Glycosides of Flavonols and Flavones by UV Absorbance after Identification by LC-MS. *Journal of Agricultural and Food Chemistry,* **60,** 544-553.

29 Jaiswal, R., Kiprotich, J. and Kuhnert, N. Determination of the hydroxycinnamate profile of 12 members of the Asteraceae family. *Phytochemistry,* **72,** 781-790.

30 Bremer, H. J., Anninos, A. and Schulz, B. 1996. Amino acid composition of food products used in the treatment of patients with disorders of the amino acid and protein metabolism. *European Journal of Pediatrics,* **155,** S108-S114.

31 Farrokhi, N., Whitelegge, J. P. and Brusslan, J. A. 2008. Plant peptides and peptidomics. *Plant Biotechnology Journal,* **6,** 105-134.

32 Larsen, P. O., Sørensen, F. T., Wieczorkowska, E. and Goldblatt, Biochem P. 1981. Meta-carboxy-substituted aromatic amino acids and v-glutamyl peptides: chemical characters for classification in the Iridaceae. *Biochemical Systematics and Ecology,* **9,** 313–323.

33 Jiang, Z., Rai, D. K., O'connor, P. M. and Brodkorb, A. Heat-induced Maillard reaction of the tripeptide IPP and ribose: Structural characterization and implication on bioactivity. *Food Research International,* **50,** 266-274.

34 Schräder, I., Eichner, K. 1996. Zeitschrift fur Lebensmittel-Untersuchung und -Forschung, **202,** 474–480.

35 Hui, P. D. Y. H., Chandan, R. C., Clark, S., Cross, N. A., Dobbs, J. C., Hurst, W. J., Leo M. L. Nollet, P. D., Shimoni, E., Nirmal Sinha, P. D. and Smith, E. B. 2007. *Handbook of Food Products Manufacturing: Health, Meat, Milk, Poultry, Seafood, and Vegetables* John Wiley & Sons.

36 Flores, P., Hellin, P. and Fenoll, J. Determination of organic acids in fruits and vegetables by liquid chromatography with tandem-mass spectrometry. *Food Chemistry,* **132,** 1049-1054.

37 Al-Achi, A. 2008. *An Introduction to Botanical Medicines: History, Science, Uses, and Dangers: History, Science, Uses, and Dangers* ABC-CLIO.

38 Kitajima, J., Kamoshita, A., Ishikawa, T., Takano, A., Fukuda, T., Isoda, S. and Ida, Y. 2003. Glycosides of *Atractylodes lancea. Chemical and Pharmaceutical Bulletin,* **51,** 673-678.

39 Liu, R., Ye, Y., Qiang, L., Liao, X., and Zhao, Y. 2008. The fragmentation pathway of the nucleosides under the electrospray ionization multi-stage mass spectrometry. *Life Science Journal,* **5,** 37-40.

40 Leu, Y.-L., Shi, L.-S. and Damu, A. G. 2003. Chemical Constituents of *Taraxacum formosanum*. *Chemical and Pharmaceutical Bulletin,* **51,** 599-601.

41 Herraiz, T. and Galisteo, J. 2003. Tetrahydro-Î²-carboline Alkaloids Occur in Fruits and Fruit Juices. Activity as Antioxidants and Radical Scavengers. *Journal of Agricultural and Food Chemistry,* **51,** 7156-7161.

42 Herraiz, T. and Galisteo, J. 2002. Tetrahydro-ß-carboline Alkaloids that Occur in Foods and Biological Systems Act as Radical Scavengers and Antioxidants in the ABTS Assay. *Free Radical Research,* **36,** 923-928.

43 Wu, S.-Y., Wang, G.-F., Liu, Z.-Q., Rao, J.-J., Lu, L., Xu, W., Wu, S.-G. and Zhang, J.-J. 2009. Effect of geniposide, a hypoglycemic glucoside, on hepatic regulating enzymes in diabetic mice induced by a high-fat diet and streptozotocin. *Acta Pharmacol Sin,* **30,** 202-208.

44 Mahmoud, Z. F., Kassem, F. F., Abdel-Salam, N. A. and Zdero, C. 1986. Sesquiterpene lactones from Lactuca sativa. *Phytochemistry,* **25,** 747-748.

45 Han, Y.-F., Cao, G.-X., Gao, X.-J. and Xia, M. Isolation and characterisation of the sesquiterpene lactones from Lactuca sativa L var. anagustata. *Food Chemistry,* **120,** 1083-1088.

46 Xia, Z., Qu, W., Lu, H., Fu, J., Ren, Y. and Liang, J. Sesquiterpene lactones from Sonchus arvensis L. and their antibacterial activity against Streptococcus mutans ATCC 25175. *Fitoterapia,* **81,** 424-428.

47 Liu, Y., Chen, Y., Momin, A., Shaner, R., Wang, E., Bowen, N. J., Matyunina, L. V., Walker, L. D., McDonald, J. F., Sullards, M. C. and Merrill Jr., A. H. 2010. Elevation of sulfatides in ovarian cancer: an integrated transcriptomic and lipidomic analysis including tissue-imaging mass spectrometry, *Molecular Cancer,* **9,** 186.

48 Michalska, K., and Kisiel, W. 2007. A Guaianolide Sulfate Conjugate and Other Constituents from Taraxacum alpinum. *Polish Journal of Chemistry,* **81,** 515-519.

49 Wesołowska, A., Nikiforuk, A., Michalska, K., Kisiel, W. and Chojnacka-Wójcik, E. 2006. Analgesic and sedative activities of lactucin and some lactucin-like guaianolides in mice. *Journal of Ethnopharmacology,* **107,** 254-258.

50 Bischoff, T. A., Kelley, C. J., Karchesy, Y., Laurantos, M., Nguyen-Dinh, P. and Arefi, A. G. 2004. Antimalarial activity of Lactucin and Lactucopicrin: sesquiterpene lactones isolated from Cichorium intybus L. *Journal of Ethnopharmacology,* **95,** 455-457.

51 Sánchez-Patán, F., Saì¤Nchez-Pataì,¤N. F., Monagas, M. A., Moreno-Arribas, M. V. and Bartolomeì, B. 2011. Determination of Microbial Phenolic Acids in Human Faeces by UPLC-ESI-TQ MS. *Journal of Agricultural and Food Chemistry,* **59,** 2241-2247.

52 Albayrak, S., Aksoy, A., Sagdic, O. and Hamzaoglu, E. Compositions, antioxidant and antimicrobial activities of Helichrysum (Asteraceae) species collected from Turkey. *Food Chemistry,* **119,** 114-122.

53 Mukhtar, N., Iqbal, K. and Malik, A. 2002. Novel Sphingolipids from *Conyza canadensis*. *Chemical and Pharmaceutical Bulletin,* **50,** 1558-1560.

54 Kokoska, L. and Janovska, D. 2009. Chemistry and pharmacology of Rhaponticum carthamoides: A review. *Phytochemistry,* **70,** 842-855.

55 Schütz, K., Carle, R. and Schieber, A. 2006. ***Taraxacum*—**A review on its phytochemical and pharmacological profile. *Journal of Ethnopharmacology,* **107,** 313-323.

56 Wang, S.-Y., Chang, H.-N., Lin, K.-T., Lo, C.-P., Yang, N.-S. and Shyur, L.-F. 2003. Antioxidant Properties and Phytochemical Characteristics of Extracts from Lactuca indica. *Journal of Agricultural and Food Chemistry,* **51,** 1506-1512.

57 Yaoita, Y. and Kikuchi, M. 1997. New Phenylpropanoid and New Eremophilenolides from Rhizomes of Petasites japonicus MAXIM. *Natural medicines,* **51,** 372-375.

58 Gómez-Romero, M. A., Zurek, G., Schneider, B., Baessmann, C., Segura-Carretero, A. and Fernández-Gutiérrez, A. Automated identification of phenolics in plant-derived foods by using library search approach. *Food Chemistry,* **124,** 379-386.

59 Matilla, P. and Hellstrom, J. 2007. Phenolic acids in potatoes, vegetables, and some of their products. *Journal of food composition and analysis: an official publication of the United Nations University, International Network of Food Data Systems,* **20,** 152-160.

60 Hussein, S. A. M., Hashim, A. N. M., El-Sharawy, R. T., Seliem, M. A., Linscheid, M., Lindequist, U. and Nawwar, M. A. M. 2007. Ericifolin: An eugenol 5-O-galloylglucoside and other phenolics from Melaleuca ericifolia. *Phytochemistry,* **68,** 1464-1470.

61 Fraisse, D., Felgines, C., Texier, O. and Lamaison, J.-L. 2011. Caffeoyl Derivatives: Major Antioxidant Compounds of Some Wild Herbs of the *Asteraceae* Family. *Food Nutritional Sciences*, **2,** 181-192.

62 Leu, Y.-L., Wang, Y.-L., Huang, S.-C., and Shi, L.-S., 2005. Chemical Constituents from Roots of Taraxacum formosanum, *Chemical and Pharmaceutical Bulletin,* **53,** 853-855.

63 Xu, F., Zou, G. A., Liu, Y. Q. and Aisa, H. A. 2012. Chemical constituents from seeds of Lactuca sativa. *Chemistry of Natural Compounds,* **48,** 574-576.

64 Clifford, M. N., Zheng, W. and Kuhnert, N. 2006. Profiling the chlorogenic acids of aster by HPLC–MSn. *Phytochemical Analysis,* **17,** 384-393.

65 Winter, M. and Herrmann, K. 1986. Esters and glucosides of hydroxycinnamic acids in vegetables. *Journal of Agricultural and Food Chemistry,* **34,** 616-620.

66 Stalmach, A. L., Mullen, W., Barron, D., Uchida, K., Yokota, T., Cavin, C., Steiling, H., Williamson, G. and Crozier, A. 2009. Metabolite Profiling of Hydroxycinnamate Derivatives in Plasma and Urine after the Ingestion of Coffee by Humans: Identification of Biomarkers of Coffee Consumption. *Drug Metabolism and Disposition,* **37,** 1749-1758.

67 Dupont, M. S., Mondin, Z., Williamson, G. and Price, K. R. 2000. Effect of Variety, Processing, and Storage on the Flavonoid Glycoside Content and Composition of Lettuce and Endive. *Journal of Agricultural and Food Chemistry,* **48,** 3957-3964.

68 Rodrigues, E. D., Da Silva, D. B., De Oliveira, D. C. R. and Da Silva, G. V. J. 2009. DOSY NMR applied to analysis of flavonoid glycosides from Bidens sulphurea. *Magnetic Resonance in Chemistry,* **47,** 1095-1100.

69 Jian-gong, S., Zhong-jian, J. and Yu, L. 1991. Studies on the Chemical Constituents of Saussurea Japonica (Thunb) DC., *Chemical Research In Chinese Universities*, **7,** 906.

70 Katta, V., Chowdhury, S.K., Chait, B.T. 1991. Use of a Single-Quadrupole Mass Spectrometer for Collision- Induced Dissociation Studies of Multiply Charged Peptide Ions Produced by Electrospray Ionization. *Analytical Chemistry*, **63,** 174.

71 López-Lázaro, M. 2009. Distribution and Biological Activities of the Flavonoid Luteolin. *Mini-Reviews in Medicinal Chemistry,* **9,** 31-59.

72 Romani, A., Pinelli, P., Cantini, C., Cimato, A. and Heimler, D. 2006. Characterization of Violetto di Toscana, a typical Italian variety of artichoke (Cynara scolymus L.). *Food Chemistry,* **95,** 221-225.

73 Fabre, N., Rustan, I., Hoffmann, E. and Quetin-Leclercq, J. L. 2001. Determination of flavone, flavonol, and flavanone aglycones by negative ion liquid chromatography electrospray ion trap mass spectrometry. *Journal of the American Society for Mass Spectrometry,* **12,** 707-715.

74 Belboukhari, N., Cheriti, A., Roussel, C. and Vanthuyne, N. Chiral separation of hesperidin and naringin and its analysis in a butanol extract of Launeae arborescens. *Natural Product Research,* **24,** 669-681.

75 Lu, Y., Song, W., Liang, X., Wei, D. and Zhou, X. 2009. Chemical Fingerprint and Quantitative Analysis of Cirsium setosum by LC. *Chromatographia,* **70,** 125-131.

76 Xie, H., Wang, T., Matsuda, H., Morikawa, T., Yoshikawa, M. and Tani, T. 2005. Bioactive Constituents from Chinese Natural Medicines. XV. Inhibitory Effect on Aldose Reductase and

Structures of Saussureosides A and B from *Saussurea medusa*. *Chemical and Pharmaceutical Bulletin,* **53,** 1416-1422.

77 Pinelli, P., Agostini, F., Comino, C., Lanteri, S., Portis, E. and Romani, A. 2007. Simultaneous quantification of caffeoyl esters and flavonoids in wild and cultivated cardoon leaves. *Food Chemistry,* **105,** 1695-1701.

78 Rees, S. and Harborne, J. 1984. Flavonoids and other phenolics of Cichorium and related members of the Lactuceae (Compositae). *Botanical Journal of the Linnean Society,* **89,** 313-319.

79 Plouvier, V. 1989. Recherche et répartition du coniféroside, de la skimmine, d'hétérosides coumariniques voisins et de l'aspéruloside dans quelques groupes botaniques. *Bulletin du Muséum national d'histoire naturelle. Section B, Adansonia*, **11,** 217-232.

80 Murata, T., Arai, Y., Miyase, T. and Yoshizaki, F., 2009. An alkaloidal glycoside and other constituents from *Leucosceptrum japonicum, Journal of Natural Medicines*, **63,** 402-407.

81 Paraschos, S., Magiatis, P., Kalpoutzakis, E., Harvala, C. and Skaltsounis, A.-L. 2001. Three New Dihydroisocoumarins from the Greek Endemic Species Scorzonera cretica1. *Journal of Natural Products,* **64,** 1585-1587.

82 Shi, P., Zhang, Y., Qu, H. and Fan, X. Systematic characterisation of secondary metabolites from Ixeris sonchifolia by the combined use of HPLC-TOFMS and HPLC-ITMS. *Phytochemical Analysis,* **22,** 66-73.

83 Yang, M. C., Choi, S. U., Choi, W. S., Kim, S. Y. and Lee, K. R. 2008. Guaiane Sesquiterpene Lactones and Amino Acid-Sesquiterpene Lactone Conjugates from the Aerial Parts of Saussurea pulchella. *Journal of Natural Products,* **71,** 678-683.

84 Cha, M.-R., Choi, C. W., Lee, J. Y., Kim, Y.S., Yon, G. H., Choi, S.-U., Kim, Y. H., and Ryu, S. Y. 2012. Two New Amino Acid-Sesquiterpene Lactone Conjugates from Ixeris dentate. *Bulletin of the Korean Chemical Society*. **33,** 337-340.

85 Roepstorff, P., Fohlman, J. 1984. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomedical Mass Spectrometry,* **11,** 601.

Conclusions

CONCLUSIONS

1. The phytochemical composition of three different varieties of green beans (*Phaseolus vulgaris L)* namely, Perona, Helda, and Strike, have been characterized using HPLC-ESI-TOF-MS. For this purpose, methanol:H_2O 80% (v:v) has been used as a distinctive extraction solvent of the polar fraction of green bean samples. The separation capacity of the HPLC together with the inherent characteristics such as the accurate mass data, elevated sensitivity and resolving power provided by the ESI-TOF-MS, have allowed the characterization of 72 phytochemical compounds from different families (5 hydrobenzoic acids, 5 hydrocinnamic acids, 41 flavonols, 6 flavones, 1 dihydroflavonol, 6 flavanones, 4 isoflavones, 1 isoflavanone, 2 lignans, and 1 iridoid) present in the three varieties of green beans in a short time. Interestingly, 54 of them have been reported in green beans for the first time. According to the results, it bears highlighting that the variety Strike is qualitatively more abundant in flavonoids than in the other varieties analyzed.

2. In the analysis of cucumber fruits, HPLC–ESI-QTOF-MS has been confirmed to be a reliable analytical technique for the separation and detecting of phenolic compounds. The hydro-methanolic extract of cucumber has been separated using a reversed-phase C18 analytical column (4.6×150 mm, 1.8 µm particle size; ZORBAX Eclipse Plus) and maintained at 25 °C. Thus, the proposed method offers the tentative identification of 73 compounds established relying on the MS data and MS/MS fragmentation pattern which are generated by the QTOF-MS as a powerful detection tool. The characterized compounds can be divided into several classes, such as 22 phenolic acids, 49 flavonoids, and 2 lignans. This work marks the first comprehensive study of the phenolic composition of cucumber whole fruit extract and the obtained results demonstrate that cucumber as an abundant source of phenolic compounds. These results help define the past and current usage of *Cucumis sativus L.* as food, in folk medicine; also boost a further use of this vegetable in nutrition, cosmetology and health.

3. The coupling of HPLC with ESI-QTOF-MS has demonstrated to be a powerful analytical technique for the characterization of the phenolic and other phytochemical compounds in watermelon. The mobile phase, consisting acetic acid (0.5%) in deionized water (A) and acetonitrile (B), was pumped at 0.8 mL/min into the HPLC system with an injection volume of 5 µL with the following gradient elution program: 0–2 min, 1–3% B; 2–15 min, 3–9% B; 15–32 min, 9–35% B; and 32–39 min, 35–100% B. Thus, by using the proposed method a total of 71 phenolic and phytochemical compounds were tentatively identified using the accurate mass determination of the deprotonated ions (MS data and MS/MS fragmentation pattern). In this sense, a wide variety of phytochemical components have been detected in this report, namely, 11 hydrobenzoic acids, 21 hydrocinnamic acids, 23 flavonoids, 1 lignan, 4 Iridoids, 4 coumarins, and other 6 phenolic compounds. This report represents a preliminary study for the phytochemicals in watermelon and marks the first characterization of these compounds in watermelon flesh. In this regard, the data reported indicate richness of watermelon in flavonoids.

4. The analysis of artichoke (*Cynara scolymus*) composition has been established by HPLC coupled to two successive detection techniques; DAD and ESI-QTOF-MS which has been shown to be a powerful tool in screening phenolics in a hydro-methanolic extract of artichoke in less than 36 min. The QTOF detector operated in the following parameters in negative ion mode: nebulizer 2 bars; dry gas flow 8 L/min and temperature of 210 °C; endplate offset −500 V; capillary voltage +4 kV; mass range 50–1100 m/z. In this way, 61 phenolic compounds were characterized in artichoke, of which 33, namely, 3 hydroxybenzoic acids, 17 hydroxycinnamic acids, 4 lignans, 7 flavones, 2 flavonols, and one phenol derivative, have been reported for the first time in this plant. The characterized compounds provide more about the chemical composition of artichoke, which may help consumers and food scientists to know the benefits of using this traditional plant in contemporary diets as potential sources of antioxidants. Moreover, the findings of this report may also offer a better understanding of the bioactive components in artichoke that contribute to human health.

5. The use of UPLC-ESI-QTOF-MS has been shown a particular role in the 'non-targeted' phytochemical profiling despite of the high diversity of metabolites that are contained in three varieties of lettuce (cv. baby, romaine and iceberg). The use of Bruker Compass DataAnalysis™ containing the Data Analysis 4.0 software, provided a list of possible elemental formulas for the phytochemical compounds by combining accurate mass and isotopic distribution, reflected in their error and mSigma values, respectively to reduce the high number of molecular formula hits acting as a very helpful practice. In this regard, the method used has enabled the tentative characterization of 171 compounds belonging to various structural classes: 95 phenolic compounds belonging to different families, 42 amino acids and peptides, 11 organic acids, 3 nucleosides, 2 alkaloids, 92 phenolic compounds, 17 sesquiterpene lactones and 1 iridoid. Among the phytochemicals characterized, the most abundant were the phenolic compounds. Furthermore, all the varieties studied, besides their nutritional value, were qualitatively rich in phenolic compounds, particularly the cultivar baby. It bears mentioning that methodology applied was useful to characterize known phytochemicals as well as to discover new metabolites. In this work, two tryptophan-derived alkaloids were also identified and another 10 novel structures of amino acids and sesquiterpene lactones have been also proposed. These results could be useful in epidemiological studies and as well for dietary and nutritional recommendations, aiding in the selection of salad ingredients with abundant bioactive phytochemicals.

Conclusiones finales

CONCLUSIONES

1. La composición fitoquímica de tres variedades de la judía verde (*Phaseolus vulgaris L*), Perona, Helda y Strike, se ha caracterizado mediante HPLC-ESI-TOF-MS. Para este propósito, se ha utilizado una mezcla de metanol-H_2O 80:20 (v:v) como disolvente de extracción de los compuestos polares presentes en las muestras estudiadas. La capacidad de separación de la HPLC junto con la exactitud de masa, la sensibilidad así como el poder de resolución proporcionado por el ESI-TOF-MS ha permitido la caracterización de 72 compuestos fitoquímicos pertenecientes a diferentes familias (5 ácidos hidrobenzoicos, 5 ácidos hidrocinámicos, 41 flavonoles, 6 flavonas, 1 dihidroflavonol, 6 flavanonas, 4 isoflavonas, 1 isoflavanone, 2 lignanos, y 1 iridoide) en las tres variedades de judía verde en un corto período de tiempo siendo 54 de ellos caracterizados por primera vez. Según los resultados obtenidos, es interesante destacar que la variedad Strike es cualitativamente más abundante en flavonoides que las otras variedades analizadas.

2. En cuanto al análisis de la parte comestible de pepino (*Cucumis sativus L.*), HPLC-ESI-QTOF-MS ha demostrado ser una técnica analítica fiable para la separación y detección de los compuestos fenólicos presentes en esta matriz. Para ello, el extracto hidro-metanólico de pepino ha sido separado usando una columna analítica de fase reversa C18 (con dimensiones 4,6 x 150 mm y 1,8 micras ZORBAX Eclipse Plus) la cual se mantuvo a 25 °C durante todo el análisis. Así, el método propuesto permitió la identificación tentativa de 73 compuestos de diferentes familias, 22 ácidos fenólicos, 49 flavonoides y 2 lignanos, basándonos en la información de MS y MS/MS generada por el QTOF-MS. Este trabajo de investigación es el primer estudio exhaustivo sobre la composición fenólica de la parte comestible de un extracto de pepino, demostrando el interés de este vegetal como una fuente abundante en compuestos fenólicos. Además, estos resultados podrían definir el pasado y actual uso del pepino como alimento así como en la medicina tradicional y los datos obtenidos se podrían utilizar para impulsar el mayor uso de esta hortaliza en la nutrición y en la salud.

3. El acoplamiento de HPLC con ESI-QTOF-MS ha demostrado ser una poderosa técnica analítica para la caracterización de los compuestos fenólicos así como otros compuestos fitoquímicos en la parte comestible de muestras de sandía. Para este propósito, se utilizó una fase móvil compuesta por agua y ácido acético (0,5%) como fase móvil (A) y acetonitrilo (B) la cual fue mantenida a un flujo de 0,8 ml/min durante todo el análisis empleando el siguiente gradiente: 0-2 min, 1-3% de B; 2-15

min, 3-9% de B; 15-32 min, 9-35% de B, y 32-39 min, 35-100% B. Así, un total de 71 compuestos fenólicos además de otros fitoquímicos polares, tales como 11 ácidos hidrobenzoicos, 21 ácidos hidrocinámicos, 23 flavonoides, 1 lignano, 4 cumarinas, 4 iridoides y otros 6 compuestos fenólicos, pudieron ser tentativamente identificados haciendo uso de los datos de MS y de MS/MS. Esta investigación se presenta como un estudio preliminar de los fitoquímicos presentes en sandía y marca la primera caracterización de estos compuestos en esta fruta, indicando una gran riqueza en flavonoides.

4. El análisis de la composición fitoquímica en muestras de alcachofa (*Cynara scolymus*) ha sido llevado a cabo mediante HPLC-DAD-ESI-QTOF-MS, la cual ha demostrado ser una herramienta poderosa para el análisis de estos compuestos presentes en un extracto hidro-metanólico de alcachofa en menos de 36 minutos. Para conseguir los mejores resultados en cuanto a resolución y sensibilidad en el sistema de detección QTOF se propusieron los siguientes parámetros en modo de ionización negativo: 2 bares de presión del gas de nebulización, un flujo de gas de secado de 8 L/min a una temperatura de 210 ° C y un voltaje del capilar de +4 kV trabajando en una rango de masas de 50-1100 m/z. De esta manera, 61 compuestos fenólicos pudieron ser caracterizados de los cuales 34 de ellos, 3 ácidos hidroxibenzoicos, 17 ácidos hidroxicinámicos, 4 lignanos, 7 flavonas, 2 flavonoles y un derivado de fenol, han sido caracterizados por primera vez en alcachofa. Los compuestos caracterizados podrían proporcionar más información sobre la composición química de la alcachofa, los cuales puede ayudar a los consumidores y a los científicos a conocer los beneficios del uso de este vegetal en la dieta como potencial fuente de antioxidantes. Por otra parte, los resultados de este estudio también pueden ofrecer una mejor comprensión de los componentes bioactivos de la alcachofa que contribuyen a mejorar la salud.

5. El uso de UPLC-ESI-QTOF-MS ha jugado un papel crucial en el registro de perfiles metabólicos, a pesar de la gran diversidad de componentes, en tres variedades de lechuga (Baby, Romaine e Iceberg). Para ello, el empleo de la plataforma Bruker DataAnalysis Compass™ y del software DataAnalysis 4.0 proporcionó una lista de posibles fórmulas moleculares de los compuestos mediante la combinación de la masa exacta y su distribución isotópica la cual permite reducir el número de posibles fórmulas moleculares protestas. En este sentido, el método utilizado ha permitido la caracterización de 171 compuestos pertenecientes a diferentes clases estructurales:

95 compuestos fenólicos pertenecientes a diferentes familias, 42 aminoácidos y péptidos, 11 ácidos orgánicos, 3 nucleósidos, 2 alcaloides, 17 lactonas de sesquiterpeno y 1 iridoide. En este sentido podemos afirmar que todas las variedades estudiadas, además de su alto valor nutricional, eran cualitativamente ricas en compuestos fenólicos, particularmente la variedad Baby. Así, la metodología utilizada fue útil tanto para llevar a cabo la caracterización de compuestos conocidos como desconocidos, logrando proponer la presencia de nuevos metabolitos por primera vez: dos alcaloides derivados de triptófano además de otras 10 nuevas estructuras de aminoácidos y lactonas de sesquiterpenos. Estos resultados podrían ser útiles tanto en los estudios epidemiológicos como para las recomendaciones dietéticas y nutricionales.