UNIVERSIDADE DE SÃO PAULO

FACULDADE DE CIÊNCIAS FARMACÊUTICAS DE RIBEIRÃO PRETO

Metabolomics studies of wild edible plants from Asteraceae family

Estudos metabolômicos de plantas alimentícias selvagens da família Asteraceae

JOLINDO ALENCAR FREITAS

RIBEIRÃO PRETO

2020

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE CIÊNCIAS FARMACÊUTICAS DE RIBEIRÃO PRETO

Metabolomics studies of wild edible plants from Asteraceae

Estudos metabolômicos de plantas alimentícias selvagens da família Asteraceae

Doctoral thesis presented to the Graduate Program of Pharmaceutical Sciences of School of Pharmaceutical Sciences of Ribeirão Preto/USP for the degree of Doctor in Sciences.

Concentration Area: Natural and synthetics products

Author: Jolindo Alencar Freitas

Supervisor**:** Prof**.** Dr. Fernando Batista da Costa

Revised version of doctoral thesis presented to the Graduate Program of Pharmaceutical Sciences in 06/05/2020. Original version available in School of Pharmaceutical Sciences of Ribeirão Preto/USP.

RIBEIRÃO PRETO
2020

FREITAS, J.A.	**Metabolomics studies of wild edible plants from Asteraceae family**		DOUTORADO FCFRPUSP 2020

Jolindo Alencar Freitas

Metabolomics studies of wild edible plants from Asteraceae family.

Ribeirão Preto, 2020.

78 p.

Doctoral thesis presented to the Graduate Program of School of Pharmaceutical Sciences of Ribeirão Preto/USP for the degree of Doctor in Sciences. Concentration Area: Natural and synthetic products

Supervisor: Prof. Dr. Fernando Batista da Costa

1. Wild edible plants 2. Asteraceae 3. Pyrrolizidine alkaloids 4. Chemotaxonomy

APPROVAL PAGE

Jolindo Alencar Freitas
Doctor

Doctoral thesis presented to the Graduate Program of School of Pharmaceutical Sciences of Ribeirão Preto/USP for the degree of Doctor in Sciences.

Concentration Area: Natural and synthetics products

Supervisor: Prof. Dr. Fernando Batista da Costa

Approved on: 06/05/2020

Examiners

Prof. Dr. _____

Institution: _____ Signature:_____

Prof. Dr. _____

Institution: _____ Signature:_____

Prof. Dr. _____

Institution: _____ Signature:_____

Prof. Dr. _____

Institution: _____ Signature:_____

“Eu dedico essa tese de doutorado à memória de minha mãe, que mesmo distante sempre me deu forças e que faz parte de quem eu sou”

“I dedicated this doctoral thesis to my mom (in memoriam), who always gave me strength and is part of who I am”

AKNOWLEDGEMENTS

I am grateful to my entire family: my parents Aparecida Silberschmidt Freitas (*in memoriam*) and Jolindo Bulgike de Alencar Freitas, my brother Joemar and my sister Juliane. I am not able to describe the support you all gave me. You all are great examples for me. I am grateful to you.

Thanks to my supervisor Prof. Dr. Fernando Batista da Costa for the opportunity, lab managing and guidance.

To my contemporaneous members of AsterBioChem research group: Annylory, Marcelo, Rosana, Ricardo, Daniele, Lucas, Federico, Amanda, Felipe, Tsvetelina, Yasmine, Álvaro, Beatriz, and to our two aggregated members: Camila and Marília. I am proud to say that I met funny, very smart and resilient people in this group. A special thanks to Marina for the friendship and support in collects and to Gari for all the knowledge exchanged, experiment advices and friendship.

To my lab mates from Pharmacognosy lab: Débora, Ingrid, Victor, Valdeline, Gabriel, Jennyfer, Adriany, Carol, João and Jonas. We had lunch together, we laughed, suffered and learned together.

To Prof. Jairo Kenupp Bastos and Prof. Niege Araçari Jacometti Cardoso Furtado for managing the lab and teaching us every time.

To Angélica, Mário and Waltinho for the technical support and to ensure the proper functioning of the laboratory. A special thanks to Dr. Luiz Fernando da Silva for the direct technical support and for being our trustful Orbitrap manager. To Dr. Nivaldo Boralle from Chemical Institute of UNESP for NMR acquisition. To Edimárcio from the Botany lab, Ricardo from Biology department of FFCLRP, Vanessa from Toxicology lab, and Tomaz from Organic Chemistry lab for technical support.

To people outside USP routine, a need to thanks Ovídio, Daniele Sayuri, Renata Fuga, and Suellen: you were very important friends in the last years. To all my pets which unfortunately left during these four years of doctoral. You all are in my thoughts and are part of me.

To Prof. Dr. Milton Groppo for helping with plants identification.

To Gersio and Nivaldo, gardeners of FCFRP, for initially assisting in the cultivation of my plants.

To every farmer and air-free market vendor which allowed me to harvest plants in their home. I learned a lot with you too.

Thanks to São Paulo Research Foundation (FAPESP, grants 2014/26866-7) by Orbitrap acquisition. Thanks to the Graduate Program in Pharmaceutical Sciences and to the School of Pharmaceutical Sciences of Ribeirão Preto, University of São Paulo. Thanks to CAPES for providing scholarships and financial support to research in Brazil. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

At last, I thank to all teachers and Professors which I had for knowledge gathered and to everyone else who support me directly or indirectly in these last years.

“*Let food be thy medicine, and let medicine be thy food.”*

attributed to *Hippocrates*

ABSTRACT

FREITAS, J. A. **Metabolomics studies of wild edible plants from Asteraceae family. 2020. 72 p. Thesis (Doctoral).** Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2020.

The plant family Asteraceae has important edible species, such as lettuce and chicory, which are widely cultivated around the world. Studies have demonstrated the importance to ingest wild and/or spontaneous plant species as salads or condiments, since they are sources of a wide variety of nutrients and secondary metabolites. Unfortunately, edible wild species ethnobotanical studies presented the use of species with chronic toxic compounds as pyrrolizidine alkaloids or unclear phytochemistry as *E. fosbergii*. This study aims to demonstrate, through LC-MS untargeted metabolomics analysis of plants, possible risks and advantages of including 11 Asteraceae wild edible leaves in population diet. Initially, *E. fosbergii* had its chemical profile studied, which showed for the first time in this species the occurrence of 28 compounds, including pyrrolizidine alkaloids, flavonoids and cinnamic acid derivatives. At second, *E. fosbergii* and *E. sonchifolia* had their metabolic profiling compared, revealing that *E. fosbergii* accumulates more pyrrolizidine alkaloids than *E. sonchifolia*, with main importance to emiline compound, present mainly in the first species. Then, the influence of soil composition and phenology were studied in both species, revealing the situation in which each plant accumulates certain classes of compounds. Apart from that, other nine species, including wild and domesticated crop species, had their LC-MS chemical profile compared. It was possible to observe that *Bidens pilosa*, *Galinsoga parviflora*, *Acmella oleraceae* samples had different chemical profiles from Cichorieae tribe samples. *Cichorium intybus* and *Youngia japonica* species had unique chemical profiles and were distinguished from other four species. *Lactuca canadensis*, *Sonchus oleraceus* and *Cichorium endivia* demonstrated a similar chemical profile among each other and with a few *L. sativa* samples. The main discriminant compounds were linked to the chemotaxonomic information of each tribe. In conclusion, these findings have not only demonstrated the importance of chemotaxonomy in Asteraceae but also presented that the studied wild edible species could bring a variety of secondary metabolites to population. At last, this study confirmed hypothesis of pyrrolizidine alkaloids presence in *E. fosbergii*, which is a risk to the population due to the toxicity of this class of secondary metabolites.

Keywords: Asteraceae, metabolomics, Wild edible plants, Unconventional edible plants, pyrrolizidine alkaloids

RESUMO

FREITAS, J. A. **Estudos metabolômicos de plantas alimentícias selvagens da família Asteraceae. 2020. 72 p. Tese (Doutorado).** Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2020.

A família Asteraceae possui importantes plantas alimentícias, como alfaces e chicórias, as quais são extensamente cultivadas pelo mundo todo. Estudos têm demonstrado a importância de se incluir espécies de plantas selvagens e/ou espontâneas na alimentação, nas formas de salada ou condimento, a fim de incrementar a ingestão de determinados nutrientes e metabólitos secundários. Infelizmente, alguns estudos de espécies alimentícias selvagens relatam o uso de plantas com presença de substâncias tóxicas, como alcaloides pirrolizidínicos, ou uso de espécies com fitoquímica pouco estudada, como *Emilia fosbergii*. Este estudo tem o objetivo de demonstrar, através de estudos metabolômicos não-direcionados por CL-EM possíveis riscos e vantagens de se incluir 11 espécies de plantas selvagens na alimentação da população. Inicialmente, *E. fosbergii* teve o perfil químico estudado e 28 substâncias foram descritas pela primeira vez nesta espécie, incluindo alcaloides pirrolizidínicos, flavonoides, e derivados do ácido cinâmico. Com isso, comparou-se o perfil metabólico de *E. fosbergii* com *E. sonchifolia,* revelando que *E. fosbergii* acumula mais alcaloides pirrolizidínicos que *E. sonchifolia*, com destaque ao alcaloide pirrolizidínico emilina, presente predominantemente na primeira espécie. Na sequência, ainda foi estudada a influência da composição do solo e do florescimento no perfil químico das duas espécies, que revelou condições em que cada espécie acumula certas classes de substâncias. Além disso, nove espécies, entre selvagens e domesticadas, tiveram o perfil químico obtido por CL-EM comparado. Foi possível observar que amostras das espécies *Bidens pilosa, Galinsoga parviflora e Acmella oleracea* tiveram perfil químico muito diferente de amostras da tribo Cichorieae. *Cichorium intybus* e *Youngia japonica* apresentam perfil químico únicos e distintos das outras quatro espécies. *Lactuca canadensis, Sonchus oleraceus* e *Cichorium endivia* demonstraram perfil químico similar entre si e com algumas amostras de *L. sativa*. Os principais discriminantes foram associados à informação quimiotaxonômica de cada tribo. Concluindo, estes resultados não apenas demonstram a importância da quimiotaxonomia em Asteraceae como também apresenta que o uso de algumas dessas espécies poderia contribuir com obtenção de uma variedade de metabólitos secundários para a população. Por último, este estudo confirmou a hipótese da presença de alcaloides pirrolizidínicos em *E. fosbergii*, o que é um alerta de risco à população devido à toxicidade dessa classe de metabólitos secundários.

Keywords: Asteraceae, metabolômica, plantas selvagens alimentícias, Plantas Alimentícias Não Convencionais (PANCs), alcaloides pirrolizidínicos.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

FAO	Food and Agriculture Organization of the United Nations
LC	Liquid Chromatography
UHPLC	Ultra High-Performance Liquid Chromatography
MS	Mass Spectrometry
PCA	Principal Component Analysis
HR	High resolution
PA	Pyrrolizidine alkaloids
CL-EM	Cromatografia líquida hifenada à espectrometria de massas
CCA	Canonical Correspondence Analysis

SUMMARY

1 INTRODUCTION

Food security and wild edible plants

The Food and Agriculture Organization (FAO) consider food security one of the biggest challenges for humanity, since it is estimated that over 820 million people worldwide were in chronic hunger situation in 2018 (FAO et al., 2019). That means that one in nine people have insufficient daily calorie intake and are uncapable to have a normal, active, and healthy life. Added to this number are 2 billion people, around a quarter of world population (FAO et al., 2019), whose daily access to food is uncertain or insufficient or who can only access food with compromised quality. In contrast to this information, the same document shows that world population suffering from obesity also reaches 2 billion people and contributes to 4 million of deaths yearly. The aforementioned numbers represent the current world situation on food security, however it could still be worsened in emergency situations, as conflicts, wars or environment disasters.

Despite seeming to be contradictory, undernourishment and obesity coexistence are very usual in poor regions and are deeply linked since both are signs of malnutrition (FAO et al., 2019; VILLENA-ESPONERA; MORENO-ROJAS; MOLINA-RECIO, 2019). Undernutrition during pregnancy or on infants could lead to stunt physical development or metabolism disorders that manifest as obesity or chronic diseases, as well as social, mental and cognition compromised levels (BLACK et al., 2004; LIU et al., 2017; VEENA et al., 2016). Malnutrition includes vitamin and minerals deficiency and is one of the biggest humanity challenges since it is generated as a consequence of health, educational, agriculture, social and economic politics (FAO et al., 2019). Obesity is one example of the connection between economy and food insecurity, since nutritive foods are expensive and processed foods (which are rich in fats, salt and sugar) are very cheap (FAO et al., 2019; VILLENA-ESPONERA; MORENO-ROJAS; MOLINA-RECIO, 2019).

Fighting against poverty is one approach, but there are other alternatives to combat malnutrition worldwide. Malnutrition could be a sign of an inadequate food system and alternatives should include diverse sectors as agriculture, transport, education and economy (FAO et al., 2019; KINUPP, 2007). Since meat-based food are expensive, and unaffordable to low-income families, an alternative to help general population to

reach good nutrition parameters could be the cultivation of native plants or collect of spontaneous foods, both considered as wild food. A huge variety of grains, fruits, vegetables and even mushrooms are used worldwide as food, but just a limited number of species are used as commercial domesticated crops (BALDERMANN et al., 2016; KINUPP, 2007; KUNKEL, 1984). There is no catalogue with all edible plant species but it is theorical estimated that around 10,000 plant species would be eatable (BALDERMANN et al., 2016; KUNKEL, 1984). Even with this variety of eatable plant species, most countries economy and food production rely in just a few domesticated crops used in all parts of the world, mainly potato, rice, wheat, and soy. Adaptation of these four species in different environments worldwide generates loss of characteristics and even facilitate propagation of diseases (FAO et al., 2019; KINUPP, 2007). Native and spontaneous plants are normally used in sustainable agriculture systems and could be an interesting alternative to food production because they are 1) resistant to diseases (reducing the use of pesticides), 2) adapted to different climates, 3) source of a variety of nutrients and bioactive compounds, which makes them a viable option for trades, leading to economic growth (BALDERMANN et al., 2016).

The collection of wild plants is part of humanity development and is still kept in a number of regions and communities (SCHULP; THUILLER; VERBURG, 2014). Indigenous and traditional plants from native communities are investigated in search of a variety of nutrients, which could be used to complement dietary intake or as supplement (KINUPP, 2007; SHIN et al., 2018). A deeper knowledge on native edible plants could be a difference to survive in scarcity times, as shown in an ethnobotanical survey with elderlies Dutch (VORSTENBOSCH et al., 2017). In this study, authors rescued elderly knowledge about wild plants eaten by population in Netherlands during famine caused by World War II and state that this knowledge seems restricted to elderly population, having been erased from youngers.

Not just used in scarcity periods, the use of native and wild food is a tradition in certain locals. In Europe, hunts, fruits, mushrooms, berries and vegetables collection are part of popular alimentation and even a form of population entertainment in a few regions (SCHULP; THUILLER; VERBURG, 2014). Wild food collection, beyond fulfilling alimentation needs, becomes a recreational activity, produce entertainment, also incentives physical activity, and promotes a sustainable and integrated community.

There is an estimate that in Europe around 65 million people collect wild foods and around 100 million people consume wild food (SCHULP; THUILLER; VERBURG, 2014).

These habits also were associated to Mediterranean population, which had lower risks of chronic diseases, as coronary-heart disorders, strokes (ROSATO et al., 2019; SCODITTI et al., 2012), development of cancer (SCHWINGSHACKL et al., 2017; TRICHOPOULOU; LAGIOU; KUPER, 2000), and even cognition and neurological diseases as depression and Alzheimer (PSALTOPOULOU et al., 2013; SERRA-MAJEM et al., 2019; SINGH et al., 2014). A Systematic review and meta-analysis study confirms association between Mediterranean Diet to lower levels of cancer mortality and risk of several cancer types, which includes colorectal, liver, gastric, breast, head and neck, gallbladder and biliary tract cancer (SCHWINGSHACKL et al., 2017). These effects were mainly attributable to higher intake of fruits, grains, and vegetables consumption. Another systematic review relates also evidences of Mediterranean Diet to lower risk of cardiovascular diseases, which include coronary diseases and ischemic stroke (ROSATO et al., 2019). Major cause to this healthier population characteristic would be the diet, which use big amounts of olive oils, a big variety of plants and fruits, and moderate amounts of wine. This behavior set – or healthy lifestyle – is known as Mediterranean Diet and it is considered an intangible cultural heritage by The United Nations Educational, Scientific and Cultural Organization (UNESCO) since 2010 (UNESCO, 2010).

Besides most known and studied dietary pattern in world, Mediterranean Diet is not the only project of knowledge rescue and ethnobotanic research of native species and foods to increase population's health. Another example is Ark of Taste project, from Slow Food foundation, which catalogued until now more than 5,000 products that includes native dishes, recipes, species and native products around the world. In Brazil, a term which got famous in the last decade was Unconventional Edible Plants, "Plantas Alimentícias não Convencionais" (PANCs) in Portuguese, which comprises native, regional, weed plants or use of dismissed parts of conventional plants as food (KINUPP, 2007; KINUPP; BARROS, 2008; PEISINO et al., 2019). Besides that, organizations and university, which includes Brazilian Network of Food Data Systems (Brasilfoods), Universidade of São Paulo (USP) e Food Research Center (FoRC/CEPID/FAPESP) developed a database to list species and food composition

from Brazil biodiversity, called “Tabela Brasileira de Composição de Alimentos” – TBCA (http://www.tbca.net.br/). Brazil possesses a huge unexplored biodiversity, but commercial food and agriculture is also based on large domesticated crops of potatoes, rice, beans, and soy. A major exception is cassava, which has processed foods, as tapioca and cassava flour and could be found in most markets. There are also famous regional plants such as “tucupi”, “jambu” and “açaí” at North, “pequi” and “baru nuts” in Midwest, and a huge variety of fruits in Northeast but their production still rarely reaches high scales, commercial exportation or even other parts of Brazil (KINUPP, 2007).

Most studies on wild species are related to nutritional value (MARTINS et al., 2011; RENNA et al., 2015), ethnobotanical surveys (GUARRERA; SAVO, 2016; LENTINI; VENZA, 2007; SHIN et al., 2018), antioxidant and polyphenols quantification (MARTINS et al., 2011; SAVO et al., 2019), but there are a few studies that carried out a detailed chemical profiling of edible species (GIAMBANELLI et al., 2018a, 2018b). Researchers then realized that most of native or wild plants used as food in ethnobotanic surveys were also used by population as phytomedicine to a huge diversity of illness, mainly to inflammatory conditions, fever, gastrointestinal disorders and infections (GUARRERA; SAVO, 2013; LENTINI; VENZA, 2007).

Use of plants as medicine or source of biological active compounds is embedded in humanity’s history. A lot of herbs was used as medicine to treat fever, wounds, gastrointestinal disorders, dermatitis and as painkillers, due to biological activities as analgesia, anti-inflammatory and antimicrobial. Biological activities usually are related to the presence of secondary metabolites, such as phenolics (flavonoids, anthocyanins, cinnamic acid derivatives, etc.), sesquiterpene lactones, coumarins, alkaloids and other classes (HARVEY, 2008; NEWMAN; CRAGG, 2007; WINK, 2003). So, wild plants could be used to combat hunger and malnutrition or used as an alternative to achieve a healthier life, with functional activity and even as prophylaxis to reduce risk and incidence of diseases. On the other hand, still there is a lack of studies about chemical profile, mainly related to secondary metabolites.

Antioxidant and anti-inflammatory activities are the biological activities most studied to edible plants (GUARRERA; SAVO, 2013; MARTINS et al., 2011; SAVO et al., 2019). One possible reason could be that oxidative damage is associated as the first step to

more complex conditions, as inflammation cascade triggering, and even development of chronic diseases as cardiovascular diseases, and neurodegenerative diseases as Alzheimer (KARSTENS et al., 2019) and tumorigenesis (VECCHIA, 2009). Antioxidant activity is usually related to the presence of flavonoids or other polyphenols compounds and most of articles in literature rely in just quantitative analysis of presence of flavonoids, total polyphenolic compounds, or antioxidant activity. On the other hand, high biological activity can be found even in extracts with low phenolic contents, which means that chemical composition is more important than phenolic amounts and indicates that profiling the compounds present in each species is necessary to better understand the biological activity (Conforti et al., 2009).

The family Asteraceae

Asteraceae is one of the largest plant families and comprises more than 20,000 species distributed worldwide in more than 1,500 genera. The current family classification is based on morphological and molecular data and comprehend 12 subfamilies and 43 tribes (FUNK et al., 2009).

Taxonomist Funk et al. (2009) describe Asteraceae flower morphology as following: “florets arranged on a receptacle in centripetally developing heads and surrounded by bracts, by anthers fused in a ring with the pollen pushed or brushed out by the style, and by the presence of achenes (cypselas) usually with a pappus”. Asteraceae comprehend herbs, shrubs, trees, and vines, which grow in almost every habitat, with annual or perennial behavior.

Hundreds of Asteraceae species are used as traditional remedies and their biological activities are intensely studied (FUNK et al., 2009). In a previous study, our group investigated the COX and LOX inhibition of 57 Asteraceae species (CHAGAS-PAULA et al., 2015) to demonstrate their anti-inflammatory potential. However, one of the most successful bioactive compounds in Asteraceae is the sesquiterpene artemisinin isolated from *Artemisia annua* L, which was discovered for malaria treatment and as precursor used to obtain more effective derivatives for malaria treatment, which lead Tu You-You to be awarded with the 2015 Nobel Prize in physiology or medicine (TU, 2011; TU et al., 1982).

Asteraceae biological diversity influences the chemical composition variety of species. The biosynthesis of secondary metabolites comprehends a broad diversity of

compounds in this family, including monoterpenes, diterpenes, triterpenes, sesquiterpenes and sesquiterpene lactones, polyacetylenes, flavonoids, phenolic acids, benzofurans, coumarins and pyrrolizidine alkaloids. Two of these classes, polyacetylenes and special sesquiterpene lactones, are found only in this family (FUNK et al., 2009). Pyrrolizidine alkaloids class, which will be discussed in a separated topic in this work, also is an unusual class restricted to a few plant families and, in Asteraceae, pyrrolizidine alkaloids production is restricted to the tribes Senecioneae and Eupatorieae (NORDENSTAM et al., 2009). The Asteraceae biodiversity associated with chemical diversity makes the family an interesting subject for chemotaxonomy studies (SAREEDENCHAI; ZIDORN, 2010; SHULHA; ZIDORN, 2019).

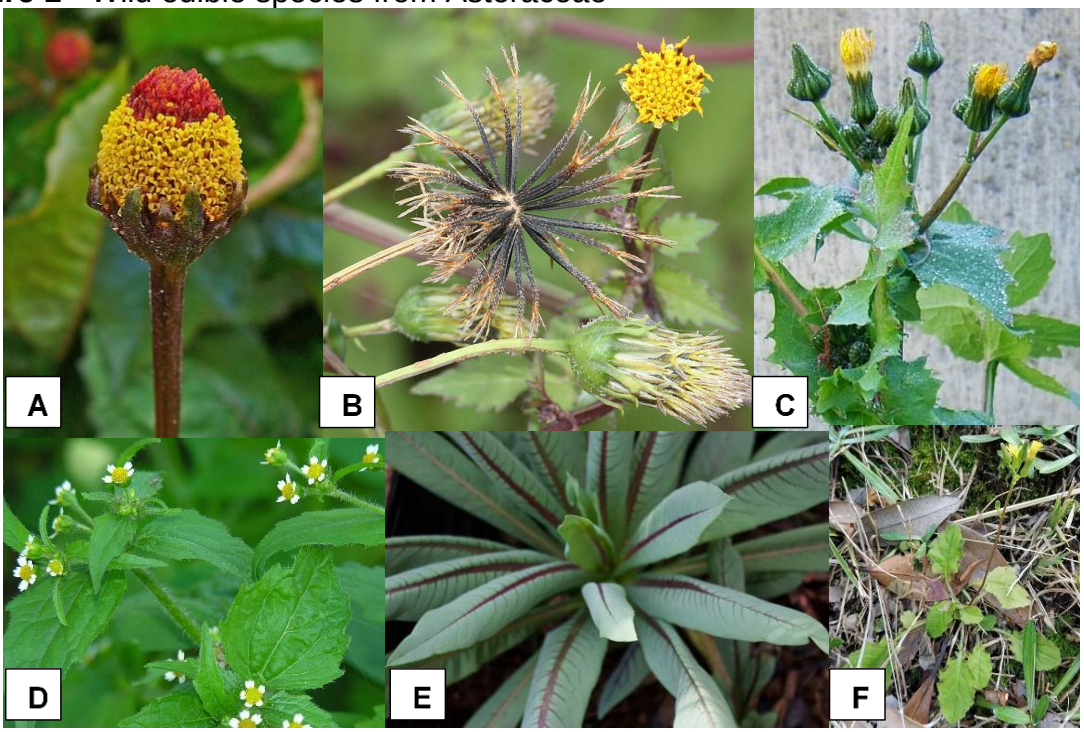
In opposition to the huge species variety, just a few Asteraceae species are used as domesticated crop. A few of the most known examples are sunflower, lettuce, and chicory (FUNK et al., 2009). Even so, these commodities are among most produced crops in the world. FAO data estimates that production of lettuce and chicory in 2018 in the world was over 42,000,000 tonnes (FAO, 2018). Except for sunflower, which seeds are used as well as for extraction of fixed oil, the edible Asteraceae plants described herein have its leaf used to be eaten, lettuce is eaten raw, most of other are eaten cooked, but raw consumption have also been reported (LORENZI; KINUPP, 2014).

The common lettuce (*Lactuca sativa* L.) is one of the major food commodities in the world and domesticated use dates from ancient Egypt 4,500 BC (HARLAN, 1986). It is considered the domesticated form of the wild species *L. serriola* L. (RYDER; WHITAKER, 1995). Chicory (*Cichorium intybus* L.) is native to Mediterranean region. *Cichorium intybus* L. comprises diverse varieties, such as chicory, Belgian endive, witloof, catalogna type, coffee chicory, and radicchio. *Cichorium endivia* L. have uncertain origin and comprises escarole, with broad-leaves, and endive, with narrow leaves (FUNK et al., 2009). Studies describe lettuce and chicory as good sources of glycosylated flavonoids, caffeic acid derivatives and sesquiterpene lactones, since they can be consumed raw in salads, avoiding degradation of these compounds (PEPE et al., 2015). Sesquiterpene lactones are described as the responsible by bitterness of lettuce and chicory (PRICE et al., 1990). Sesquiterpene lactones shown activity to a

few tumoral cell lines, which increased interest in these compounds as edible plants (ZIDORN, 2008).

On the other hand, Asteraceae diversity still could accomplish edible plants if we consider species not investigated, since several ethnobotanical studies reported that the majority of encountered wild edible species belongs to Asteraceae family (GUARRERA; SAVO, 2013; LICATA et al., 2016; RENNA et al., 2015). A few examples of Asteraceae wild edible species used around the world are *Acmella oleracea* (L.) R. K. Jansen, *Bidens pilosa* L., *Sonchus oleraceus* (L.) L., *Galinsoga parviflora* Cav., *Lactuca canadensis* L. and *Youngia japonica* (L.) DC (**Figure 1**).

Figure 1 - Wild edible species from Asteraceae



SOURCE: Wikimedia. Attributions: A) *Zell* B) *Alpsdake* C) Wildfeuer.
A) *A. oleracea*; ***B)*** *B. pilosa;* **C)** *S. oleraceus;* **D)** *G. parviflora;* **E)** *L. canadensis;* **F)** *Y. japonica*

The species *A. oleraceae,* popularly known in Brazil as “jambu”, have flowers and leaves used as condiment in North of Brazil, with presence of spilanthol and N-alkylamides in both parts of the plant (CHENG et al., 2015). *Acmella oleraceae* is used by population as local anesthetic, mainly to toothache, which studies report biological activity to spilanthol, and demonstrated anti-inflammatory and antinociceptive activities, by inhibition of proinflammatory mediators (RONDANELLI et al., 2019; WU et al., 2008).

The species *S. oleraceus* is among the most gathered wild edible plant in Mediterranean area (GUARRERA; SAVO, 2013). In Brazil it is known as “serralha” and its first official report of medicinal use dates from 1874 (RICARDO et al., 2017) as diuretic and to treat nephritis. Crude hydroethanolic (50%) extract of *S. oleraceus* aerial parts demonstrated anti-inflammatory, antipyretic, anxiolytic, and antinociceptive activities in rodent models. The chemical composition reports occurrence of sesquiterpene lactones from the guaianolide and eudesmanolide and mainly derivatives from lactucin (ZIDORN, 2008), and the flavonoids apigenin and luteolin glycosides (SAREEDENCHAI; ZIDORN, 2010).

Bidens pilosa, popularly known as “picão-preto”, has traditional use to several illnesses and has anti-inflammatory, anti-pyretic, and hepatoprotective effects (RICARDO et al., 2017). Phytochemical studies comprehend almost 200 compounds described to this species, with main importance to acetylenes and flavonoids from aurone, chalcone and flavanone classes (SILVA et al., 2011).

The species *G. parviflora,* known as “fazendeiro” or “picão-branco” in Brazil or “guasca” in Colombia, Peru and Bolivia, are traditionally used as anti-inflammatory to treat dermatological problems as eczema and it has in its composition flavonoids, flavanones, caffeic acid derivatives, diterpenes and steroids and high level of vitamin C (ALI; ZAMEER; YAQOOB, 2017; KINUPP, 2007).

Lactuca canadensis, known as “almeirão-roxo” in Brazil, has high levels of minerals Ca, Zn, B and Mn (SILVA et al., 2018). The presence of guaianolides and eudesmanolides sesquiterpene lactones was described previously (MICHALSKA; SZNELER; KISIEL, 2013).

Youngia japonica is distributed throughout the temperate zone and it is used as antipyretic and for detoxification (YAE et al., 2009). Its chemistry comprehends flavonoids, cinnamic acid derivatives, and phenylpropanoids as well as guaianolides and eudesmanolides sesquiterpene lactones (YAE et al., 2009; ZIDORN, 2008).

Pyrrolizidine alkaloids and the genus Emilia

Pyrrolizidine alkaloids (PAs) are ester alkaloids composed by a necine base, which can be saturated (platynecine base) or unsaturated (otonecine and retronecine bases). PAs with unsaturation in the necine moiety are transformed into highly reactive pyrroles, which are able to bind to proteins and even DNA cross-links, the cause of

genotoxicity potential which could cause tumorigenesis (CHEN; HUO, 2010; MATTOCKS, 1968). Main affected organs in mammals are the liver, where bioactivation occurs, and the lung (CULVENOR et al., 1976). PA poisoning initial symptoms include upper abdominal discomfort, reduction of urine excretion, painful hepatomegaly, and abdominal and feet swelling. The clinical manifestation most frequently caused by PA poisoning is called veno-occlusive disease, which cause obstruction of blood vessels, mainly from liver to heart, which could have fast progression, causing blood vomiting and has high mortality rate. Even after recovery, patients could develop cirrhosis (CULVENOR et al., 1976; MOREIRA; PEREIRA; ANDRADE, 2018; WHO, 1989). The Brazilian Health Regulatory Agency (in Portuguese, “Agência Nacional de Vigilância Sanitária”, ANVISA) restricts PAs maximum daily intake at 1 ppm in phytomedicines or contaminated food (ANVISA, 2014), same restriction limit is reported by MOREIRA & PEREIRA (2018) to the German Health Government. However, the same author cites that Food and Drug Administration (FDA) and Austria government had banned all products containing PAs from markets, since a few vigilance departments consider that no tolerable daily intake could be established.

PAs comprehend a very important information to chemotaxonomy, since PAs are an unusual class and its subclasses pattern is different to each tribe or plant family (HARTMANN; OBER, 2000). PAs are present only in Asteraceae (Eupatorieae and Senecioneae tribes) and Boraginaceae families and a few genera from Fabaceae and Orchidaceae families (HARTMANN; OBER, 2000).

Senecioneae is one of the most diverse tribes in Asteraceae family and comprehend around 150 genera and more than 3,400 species spread worldwide (NORDENSTAM et al., 2009). In Brazil, Senecioneae represents 97 species classified in eight genera. The genus *Emilia* (Cass.) Cass. has currently more than 110 species (The Plant List), however, a global taxonomic review is needed because of the current classification only includes species from Central Africa and the *E. coccinea* complex (JEFFREY, 1997; LISOWSKY, 1990). In Brazil only two species are present: *E. sonchifolia* and *E. fosbergii* (**Figure 2**) (HIND, 1993).

Figure 2 – Aerial parts of *E. sonchifolia* (**A**) and *E. fosbergii* (**B**).

SOURCE: Wikimedia. Attributions: A) Rose; B) Starr.

Misidentification is a serious problem to these *Emilia* species. *Emilia fosbergii* is sometimes misidentified as *E. coccinea* or *E. sonchifolia* var. *javanica*, present only in Oceania and Asia (NICOLSON, 1980). Worldwide known as tassel flower, the most usual popular name used for both species in Brazil is “emília”, “serralhinha” or “serralha vermelha”. However, both *Emilia* species are sometimes identified in Brazil as “serralha”, the popular name of *S. oleraceus*, and even “dente-de-leão”, the popular name of *Taraxacum officinale* L.; both these plants are used as phytomedicine, food and/or condiment (KINUPP, 2007; LEITÃO et al., 2014). The misidentification of these species becomes an aggravated problem due to described presence of hepatotoxic PAs in *Emilia* genus.

E. sonchifolia is originated from Asia (NICOLSON, 1980) and used a long time ago in Traditional Chinese Medicine and in Ayurveda Indian medicine to treat a lot of issues such as fever, wounds, respiratory affections, and inflammatory diseases. Some studies confirmed the anti-inflammatory and antinociceptive activities (COUTO et al., 2011; MUKO; OHIRI, 2000). Were also evaluated antitumoral and cytotoxic activity (SHYLESH; PADIKKALA, 2000) with a potential inhibition of metastasis in mice caused by B16F10 melanoma (GEORGE; KUTTAN, 2016). *Emilia sonchifolia* is also used as food by populations from Asian countries, such as Bangladesh, Taiwan and Hong Kong (GEORGE; KUTTAN, 2016; KUMAR et al., 2015) and even Brazil (KINUPP, 2007).

Previous phytochemical studies of the genus *Emilia* have generally been restricted to *E. coccinea* and *E. sonchifolia*. In these two species, flavonoids, chlorogenic acid derivatives, acetylenes, and ketoesters were described, which could explain the anti-inflammatory medicinal use (BOHLMANN; KNOLL, 1978; RAJ, 2012; SHEN et al., 2013; SRINIVASAN; SUBRAMANIAN, 1980). Most health-related issues and chemotaxonomic significance reports for this genus are associated with the presence of toxic PAs. A total of 11 macrocyclic rings PAs were reported until now in *E. sonchifolia* including N-oxides, retronecine and otonecines bases (HSIEH et al., 2015). The GC-MS analysis of *E. sonchifolia* plant extracts allowed the identification of PAs containing both retronecine (senecionine, seneciphylline, and integerrimine) and otonecine bases (senkirkine, otosenine, neosenkirkine, petasitenine, acetylsenkirkine, acetylpetasitenine, desacetyldoronine, and doronine) (HSIEH et al., 2015). In *E. coccinea* extract, LC-MS analysis detected platyphylline-N-oxide, three stereoisomers of the non-toxic platyphylline, ligularidine, neoligularidine, neosenkirkine, and senkirkine (MROCZEK et al., 2004). Emiline was isolated from *E. flammea*, a synonym of *E. coccinea* (Barbour and Robins, 1987; The Plant List, 2013), and until now, it was described only in this species.

The origin of *E. fosbergii* Nicolson is still undefined, but the main hypothesis regarding its origin suggests that this species is a neotropical descendant of *E. sonchifolia* (MORAES; GUERRA, 2010; NICOLSON, 1980). Peisino et al. 2019 reports the use of *E. fosbergii* as an edible plant with antioxidant and anti-inflammatory properties, probably due to the presence of a few flavonoids and chlorogenic acid. But pyrrolizidine alkaloids were not investigated until now in *E. fosbergii*.

Metabolomics and multivariate analysis

Metabolomics is the youngest in the “omics” research area, which includes genomics, proteomics and transcriptomics (FIEHN, 2002). An acceptable definition to metabolomics is a quantitative or qualitative analysis of a wide number of metabolites, compounds with low molecular weight, from a biological system, as a specific cell, tissue or organism (GOODACRE, 2005; ROCHFORT, 2005). However, there is no strategy able to extract or to analyze the entire range of metabolites present in a biological system yet, since metabolome is composed by a wide range of compound classes with very broad physicochemical properties related to extraction as polarity, solubility and properties directly related to analysis, as UV absorption, fluorescent

emission, or mass ionization (GOODACRE et al., 2004; KOPKA et al., 2004; VILLAS-BÔAS; RASMUSSEN; LANE, 2005).

One of the simplest metabolomics subdivisions splits metabolomics in two approaches: targeted and untargeted analysis. Targeted analysis aims to identification and absolute quantification of metabolites selected before analysis. Untargeted analysis also known as metabolite/metabolic profiling is a rapid and qualitative analysis of a large number of different metabolites which aims to identify a specific metabolic profile or trends of metabolites variation that characterize a determined sample, and often absolute quantification it is not necessary (VILLAS-BÔAS et al., 2004).

Metabolomics analysis is possible because of the advance of analytical instruments and machine learning knowledge. The most used techniques to metabolomics are nuclear magnetic resonance (NMR) and mass spectrometry (MS), usually hyphenated or indirectly associated to chromatography. Association between gas or liquid chromatography to mass spectrometry (GS-MS and LC-MS, respectively) is one of the most used technique due to high sensitivity, which allows analysis of metabolites even in very low concentration and comprehensiveness of data, since fragmentation (MS2 or MS/MS) provide structural information and are relatively easy to understanding.

Different methods are necessary to understand the huge amount of information generated by metabolome dataset, for this, multivariate analysis and chemometrics are used to simplify the information and consolidate the knowledge. Chemical measurements obtained by analytical tools, usually called raw data, are treated and analyzed by multivariate analysis to search for patterns or trends able to characterize samples. In unsupervised multivariate analysis, which include principal component analysis (PCA) and hierarchical clustering analysis (HCA), it is possible to observe samples grouping based only in quantitative measurement of data information, which means that samples are grouped by chemical profile similarity (HASTIE; TIBSHIRANI; FRIEDMAN, 2009; VILLAS-BÔAS et al., 2004). PCA is an unconstrained ordination method which aims to represent the main trends of variation of the data (BORCARD; GILLET; LEGENDRE, 2018). In supervised analysis a previous information must be stablished before analysis, for example “active and inactive plant extracts”, and samples information are constrained ordered by these groups sets. Supervised analysis usually needs a training set to build a model, a validation set to ensure quality

of the model and a test set, which usually is the set to have information to be analyzed. A few examples of popular supervised analysis tools are partial least square discriminant analysis (PLS-DA) and its orthogonal variety (OPLS-DA), and neural networks (HASTIE; TIBSHIRANI; FRIEDMAN, 2009; VILLAS-BÔAS et al., 2004).

The identification of metabolites is essential to ensure the information and to be integrated by other researches (CREEK et al., 2014; SUMNER et al., 2007). Metabolite identification still is one of the major problems in natural products metabolomics. Secondary metabolism generates outnumbered metabolites from a huge variety of classes and comprehends hundreds or thousands of molecular skeletons (ERNST et al., 2014). In 2007, minimum information to assure identification accuracy level was proposed by the Metabolomics Standard Initiative, where four confidence levels of identification were established (from 1 to 4) (SUMNER et al., 2007). This system was latter adapted (BLAŽENOVI et al., 2018) and the level 0 of identification was added, resulting in the system described in Erro! Fonte de referência não encontrada..

Table 1 - Confidence level of identification.

Confidence level[1]	**Compound [2]**	**Description**
Level 0	Isolated	Isolated compound. Require 3D structure, including stereochemistry.
Level 1	Identified	Standard or reference physical compound match. At least two orthogonal techniques, as MS/MS, or MS and RT.
Level 2	Annotated	Probable structure. Compound matches to literature information at least in two orthogonal information, as MS and MS2 fragmentation.
Level 3	Characterized	Possible structure or class. One or several candidates possible (including isomers), requires at least one information to support propose identification, as MS, MS2 fragmentation or UV profile.
Level 4	Unknown	Unknown feature. Presence in sample.

Adapted from [1] Blaženovi et al. (2018) and [2] SUMNER et al., (2007).

1.1.1 Plant metabolomics

Metabolomics is a crescent field and calls attention of researchers from the areas of natural products and plant research (ERNST et al., 2014). A few applications of

metabolomics approaches include drug discovery from medicinal plants, quality control of food and medicinal plants, biogeographic and tracking, environmental metabolomics and chemotaxonomy (CASOTI et al., 2018; PADILLA-GONZÁLEZ; DIAZGRANADOS; COSTA, 2017).

Plant chemotaxonomy can be shortly described as taxonomical organization based on chemical constituents, usually secondary metabolites, in which biosynthetic capacity to produce a set of compounds could be used to distinguish species, genus, tribes or families (SPRING, 2000). It is used as an alternative to understand or as a complementary information to taxonomic classification. Chemotaxonomy is a consistent field in Asteraceae since this wide family have a complex taxonomic classification (DA COSTA; TERFLOTH; GASTEIGER, 2005; SAREEDENCHAI; ZIDORN, 2010; SHULHA; ZIDORN, 2019).

Metabolomics is an approach that can be used to understand the entire plant mechanisms to interact with ambient. Resource availability can influence plant growth rate as well as stressed conditions can influence secondary metabolites biosynthesis and accumulation, that have a wide range of functions in plants (COLEY; BRYANT; CHAPIN, 1985; WINK, 2003). So, plant metabolomics can be used to understand the influence of cultivation, diseases, climates and/or environmental in production of secondary metabolism as well as primary metabolism injures (CASOTI et al., 2018). In a previous study of our research group, it was possible to infer main climate and soil nutrients which induce production of each main class of secondary metabolites from *Tithonia diversifolia* (Hemsl.) A. Gray (SAMPAIO; EbDRADA-EBEL; DA COSTA, 2016).

2 HYPOTHESES

Use of wild species could have vantages and risks. Use of species with presence of toxic compounds or with unknown chemical composition could put population in risk. On the other hand, the use of wild species could be an alternative to poor population and/or a rich source of bioactive secondary metabolites.

In this context, this study presents two hypotheses:

1) Can *E. fosbergii* be a potential risk to population if used as food?

2) Are there remarkable differences in the chemical profiles of domesticated crops when compared with wild species from Asteraceae used as food?

Objectives

To answer these two main questions, this study has the following specific aims:

1a) to dereplicate leaf extract and investigate the presence of pyrrolizidine alkaloids in *E. fosbergii*;

1b) to compare the metabolic profiles of *E. fosbergii* and *E. sonchifolia;*

1c) to verify the influence of cultivation and flowering in the chemistry profile of *E. fosbergii* and *E. sonchifolia*;

2) to verify differences and similarities between nine edible species of Asteraceae, including three domesticated crops and six wild species.

3 MATERIAL AND METHODS

Because this study has two independent main objectives, with the aim to be clearer, the experiments will follow the numbering described in “Objectives” section.

Experiment 1a

3.1.1 Plant material

To obtain alkaloids enriched fractions and isolated compounds, *E. fosbergii* was collected by Jolindo Freitas in the “Parque das Artes”, Ribeirão Preto, SP, Brazil (geographic coordinates S 21º 12’ 55” W 47º 49’ 06”) in April and May 2018. A voucher specimen (EF 16,534) was deposited in the SPFR Herbarium (Department of Biology, Faculty of Philosophy, Science and Letters of Ribeirão Preto, University of São Paulo, FFCLRP-USP).

3.1.2 Pyrrolizidine alkaloids isolation and enriched fractions obtention

After collection, *E. fosbergii* plants were separated into four parts: roots (235 g), stems (500 g), leaves (1 kg), and flowers (235 g), then dried in an oven (50 ºC). Dried leaves were extracted with 10 L of ethanol alkalinized with NH_4OH (1%) and the extract was concentrated in a rotary evaporator. The resulting dried extract was solubilized in HCl 0.5 M (400 mL) and further partitioned with CH_2Cl_2 (2 x 200 mL) to remove lipophilic components. Zinc pellets (900 mg) were added to the aqueous phase and this solution

was subject to ultrasonic bath (40 kHz) for 1h and heated for 1h to reduce pyrrolizidine alkaloids N-oxides. The solution was adjusted to pH 11 with NH_4OH (10%). Dichloromethane (3 x 200 mL) was used to extract pyrrolizidine alkaloids and the resulting solution was concentrated to obtain 300 mg leaf fraction (EFLD) enriched in alkaloids. This procedure was carried out in the same weight/volume proportion to obtain enriched fractions from the roots (200 mg, EFRD), stems (35 mg, EFSD), and flowers (136 mg, EFFD). These four fractions were further analyzed by LC-MS (1 mg/mL), in conditions as described below in Item 3.2.4.

EFLD (300 mg) was submitted to centrifugal thin-layer chromatography (Chromatotron) in silica gel 60 with CH_2Cl_2 (100 mL) and further eluted (200 mL) with a solution of CH_2Cl_2:MeOH:NH_4OH (85:14:1), providing 11 fractions which were analyzed by LC-MS. Emiline (2 mg) was isolated from fractions 6 and 7 by semipreparative HPLC (C18 Onyx monolithic column, 100 x 10 mm), monitored in UV (216 nm) in gradient elution mode using water (A) and MeCN (B), both with 0.1% formic acid at a flow rate of 4 mL/min. The elution program was 5% to 13% B in 5 min, isocratic 13% until 17 min and 13 to 100% B until 23 min.

3.1.3 NMR analysis

NMR spectra of compound 3 were acquired on a Bruker Avance III HD of 14.1 T spectrometer (Bruker, Germany) equipped with a Triple Inverse TCI CryoProbe head (3.0 mm). Chloroform-d_1 was used as solvent. Chemical shifts (δ) are given in ppm and coupling constants (J) in hertz (Hz), 600.13 MHz to ^{1}H and 150.9 MHz to ^{13}C. ^{13}C chemical shifts were obtained by DEPTQ, HSQC and HMBC experiments.

Experiment 1b and 1c

3.1.4 Plant material - Sampling of *E. fosbergii* and *E. sonchifolia*

Individuals of *E. fosbergii* and *E. sonchifolia* were cultivated on pots (3.5 L) in a greenhouse in School of Pharmaceutical Sciences of Ribeirão Preto, FCFRP-USP, Ribeirão Preto, SP, Brazil, in February 2017. A voucher specimen of each species (ES 16,536 and EF 16,534) were deposited in the SPFR Herbarium (Department of Biology, Faculty of Philosophy, Science and Letters of Ribeirão Preto, University of São Paulo, FFCLRP-USP).

Plants were cultivated under three different soil conditions: A) non-treated soil; B) soil treated with horse manure; C) soil treated with chicken manure. Four replicates of each treatment (200g) were collected in labelled plastic bags and sent for soil analysis of macro and micronutrients (SILVA et al., 2019). Soil analysis is attached in **Appendix A**.

Two harvest were made: before (1) and after flowering (2), 50 and 100 days after seedling, respectively. A total of 29 samples of *E. sonchifolia* and 27 samples of *E. fosbergii* were collected. All plants were cultivated in same conditions of irrigation, wind, and light exposure.

3.1.5 Samples harvest

Health leaf from health individuals were collected, frozen in liquid nitrogen bath and kept in dry ice during transport. In the laboratory, liquid nitrogen was used to assist samples powdering in mortar and pestle. Then, powdered samples were freeze-dried

3.1.6 Preparation of plant extracts

Dried leaves powder (20 mg) of each sample were weighted in analytical balance in Eppendorf tubes of 1.5 mL, where 1 mL of aqueous ethanol (80 %) was added. Chrysin (Sigma-Aldrich ®) was used as internal standard to a solution of final concentration of 10 µg/mL. Samples were vortexed for a few seconds followed by ultrasonication (40 kHz, Nova Instruments) for 10 min at room temperature. After extraction, samples were centrifuged (Nova Instruments) for 15 min at 11,000 rpm and hexane (300 µL, LC-MS grade) was added to clean up. Samples were vortexed a few seconds again and then centrifuged (10 min, 11,000 rpm). The hexane supernatant phase was discarded, and aqueous ethanol extract was filtered through a 0.2 µm PTFE filter.

3.1.7 Metabolic profiling

Metabolic fingerprints of each sample (4 µL) were obtained in a UHPLC equipment (Thermo Scientific, USA), composed by two Accela 1250 quaternary pumps coupled to an Exactive Plus mass spectrometer (Orbitrap mass analyzer) synchronized by the Xcalibur 2.2 software (Thermo Scientific). Chromatographic analyses were performed in a Kinetex XB-C18 (1.7 µm, 150 × 2.1 mm) column coupled to a compatible guard column in oven temperature of 35 °C. Water (A) and acetonitrile (B), both with 0.1% formic acid, were used as mobile phases (400 µL/min flow rate) with the following

elution program: 5%→50% B in 15 min, isocratic 50% until 20 min, 50→100% B until 30 min. Mass spectrometer operated in both negative and positive modes using Fullscan, All-ion-fragmentation (resolution setting of 70,000) to MS and MS/MS modes. Electrospray ionization mode (ESI) conditions were the following: spray voltage of 3.6 kV in positive and 3.2 kV in negative mode, capillary temperature of 320 °C, and a scan window of 100 to 1,500 *m*/*z*. Higher energy collisional dissociation (HCD) fragmentation were conducted for MS/MS in negative and positive modes using normalized collision energy at 35%.

3.1.8 Data preprocessing

Chromatographic ".raw" data were split in negative and positive modes by Proteowizard software. All samples of each mode were uploaded and processed by MZMine 2.37, where peaks where extracted, chromatogram rebuild, peaks were deconvoluted, isotopes from same compound were grouped and duplicate peaks were excluded. The MZmine parameters used for data preprocessing it is described in **Appendix B**. After MZmine preprocessing, data for each ionization mode were exported as tables in ".csv" format representing peak areas of each sample and secondary peak identification, as retention time and *m/z* value of each peak.

Peak areas detected in the blank (extraction solvent) were subtracted from the original matrix in Excel spreadsheet.

3.1.9 Compounds identification and annotation

Structures described previously in literature for each species were extracted in ".sdf" format from databases SciFinder Scholar and PubChem or draw with assistance of MarvinSketch (version 18.24). Chemical structure databases were built with assistance of JChem for Excel 18.8.0, Chemaxon (www.chemaxon.com) plugin to transform structures to SMILES format and calculate exact mass. Structure information were included in AsterDB, *in house* database that contains around 2,500 unique chemical structures of terpenoids, flavonoids, trans-cinnamic acid derivatives and other minor chemical classes of natural products reported in Asteraceae family (http://www.asterbiochem.org/asterdb). Mass accuracy was calculated in Excel and tolerance stablished to 5.0 ppm (BRENTON; GODFREY, 2010).

The process of metabolite identification acquired by LC-MS was carried out following guidelines of Metabolomics Society (BLAŽENOVI et al., 2018; SUMNER et al., 2007). Following data information were used to identify compounds:

- Level 1: detected peaks had retention time and exact mass compared to previously isolated compounds and commercial standards (CHIBLI et al., 2018) by MzMine2. Retention time was confirmed manually, mass accuracy calculated to confirm MS data, and MS2 fragments were compared manually to standards and to literature to confirm proposed identity.
- Level 2: detected peaks had accurate mass compared to AsterDB database by MZMine2. Mass accuracy was calculated to confirm MS data and MS2 fragments were compared manually to literature.
- Level 3: exact mass or MS2 fragments match with literature.
- Level 4: exact mass and MS2 fragments with no match to databases.

3.1.10 Multivariate analysis

Peak areas detected in the blank (extraction solvent with internal standard) were subtracted in other samples from the original matrix by Excel. Principal Component Analysis (PCA) was performed to observe the distribution of samples by similarity. Analyses were carried out in the software SIMCA 13.0.3.0 (Umetrics, Sweden). The obtained data matrix was scaled by the Pareto method, which keeps the data structure partially intact (major peaks in each sample remain) while reducing the relative importance of large values (BERG et al., 2006). In Pareto method, scaling is made by subtracting average value of all peak areas divided by square of standard deviation, as demonstrated in formula $\tilde{x}_{ij} = \frac{x_{ij} - \bar{x}_i}{\sqrt{s_i}}$ in which x_{ij} represents peak area value i in sample j; $\bar{x}_i$ represents average value of peak area i; s_i represents standard deviation of peak area i; and $\tilde{x}_{ij}$ represents the data after pretreatment (BERG et al., 2006).

Canonical Correspondence Analysis (CCA) was carried out in the software R 3.5.1, using *vegan* package to correlate flowering and the soil composition data to metabolic fingerprints of the samples. R scripts are attached in **Appendix C**.

Experiment 2

3.1.11 Plant material - Sampling domesticated crop and wild edible species

Plant samples (leaves) were collected from six farms in region of Ribeirão Preto city and cities around (state of São Paulo, Brazil), in December 2018 and January 2019. Geographic coordinates are available in **Appendix D**.

A total of 57 samples were collected, which belong to species *Lactuca sativa* (two cultivars of *L. sativa* var. *crispa* and two cultivars of *L. sativa* var. *capitata*), *L. canadensis*, *Cichorium endivia*, *C. intybus* (three cultivars), *A. oleracea*, *Y. japonica*, *B. subalternans*, *G. parviflora* and *S. oleraceus*. Samples were collected in farms which cultivate domesticated crop plants, as *L. sativa* and *C. intybus*, meanwhile wild species had grown spontaneously. Geographic coordinates and popular name of each species are listed in **Appendix E**.

Vouchers for botanical identification of each species were deposited in the SPFR Herbarium (Department of Biology, Faculty of Philosophy, Science and Letters of Ribeirão Preto, University of São Paulo, FFCLRP-USP) under codes, LS 17679-17681, CI 1673-17674, LC17683, CE 17678, YJ 17671, AO 17675, BP 16537, GP 16533, SO 16535.

3.1.12 Samples harvest

Health leaf samples from health individuals were collected, frozen and kept in dry ice during transport. In the laboratory, liquid nitrogen was used to assist samples powdering in mortar and pestle. Then, powdered samples were freeze-dried.

3.1.13 Metabolic profiling

Metabolic profiling procedures follows as described before to Experiments 1b and 1c to Preparation of extracts (item 3.2.3), Metabolic profiling (item 3.2.4), Data preprocessing (3.2.5), and compounds identification (item 3.2.6).

3.1.14 Multivariate analysis

Principal Component Analysis (PCA) and Orthogonal Partial Least Square Discriminant Analysis (OPLS-DA) were carried out in the software SIMCA 13.0.3.0 (Umetrics, Sweden).

PCA was performed to observe the distribution of samples by similarity. After initial PCA analysis by SIMCA, loadings with low significance (near zero) were excluded to recalculate a new model until the model get contribution over 0.50 in three first PCs. PCA analyses were used to verify robustness of triplicate samples, since samples near each other means chemical profiling similarity.

Orthogonal Partial Least Square Discriminant Analysis (OPLS-DA) was used to assist discriminant determination of each group.

HCA - Heatmap was built by R software, using “heatmap.2” function present in “gplots” package. Only annotated and identified mass peaks obtained in positive ionization mode were used to this analysis, considering only one injection. R scripts are attached in **Appendix C**.

4 RESULTS AND DISCUSSION

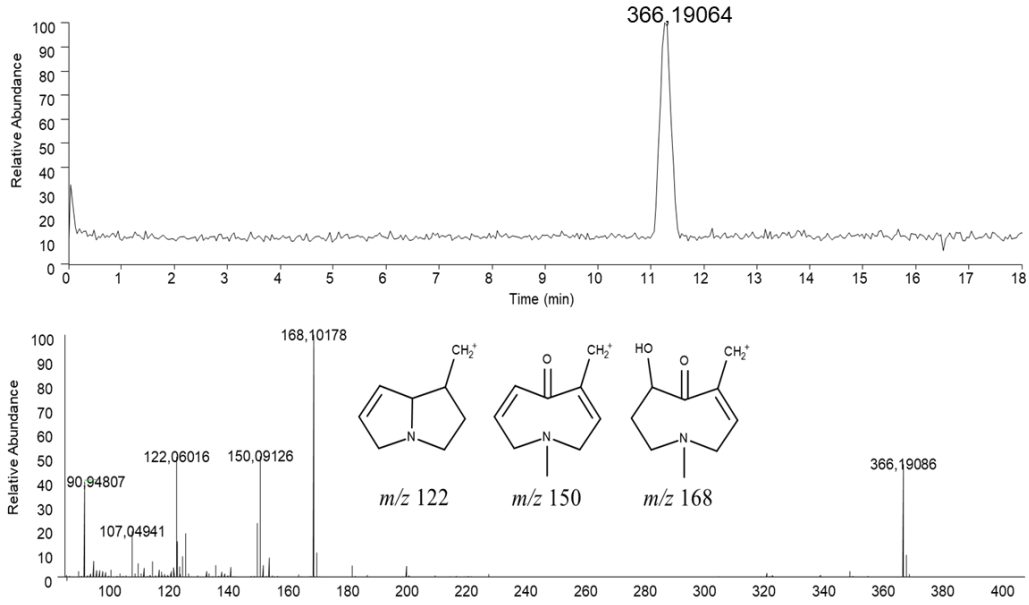
Experiment 1

4.1.1 Dereplication and investigation of PAs occurrence in *E. fosbergii* (Experiment 1a)

In this study, a total of 28 compounds were described in *E. fosbergii* as following: one PA was isolated, three flavonoids and three caffeoylquinic acid derivatives were identified; five PAs, four PA-NO, and one flavonoid were annotated; four PAs (considering isomers), one flavonoid, and one caffeoylquinic acid derivative were characterized in *E. fosbergii*. Except for quercetin and 3-*O*-caffeoylquinic acid, other compounds are described by first time in this species.

Metabolite 3 was isolated as a yellow oil. LC-MS analysis of 3 presented *m/z* $[M+H]^+$ = 366.19064, and MS2 fragments at *m/z* 122, 150, 168 (**Figure 3**), which are diagnostic fragments of pyrrolizidine alkaloids with otonecine base (HSIEH et al., 2015). To confirm the identity of emiline, NMR signals were compared with published data (**Table 2)** (BARBOUR; ROBINS, 1987). NMR spectra are attached in **Appendix F.**

Figure 3 - LC-MS chromatogram (A) and MS2 spectra (B) of compound emiline in positive ionization mode.



The compound presented *m/z* $[M+H]^+$ 366.19052. Analysis was obtained in a column ACE3 C18 (3.0 x 150 mm, 3 µm), using water and acetonitrile solvents as eluent, both added formic acid 0.1% with the following gradient program: 15% acetonitrile (15min), 15→100% acetonitrile (18min), 100% acetonitrile (23min). Fragments structure based on literature (HSIEH et al., 2015; RUAN et al., 2012).

The ^{1}H NMR spectrum of **3** showed characteristic signals for olefinic protons at δ 6.05 (1H, br s, H-2) and at δ 5.11 (2H, m, H-19). In addition, ^{13}C signals are correspondent to four olefinic carbons at δ 134.2 (C-1), 135.6 (C-2), 146.6 (C-13), and 118.1 (C-19), three carbonyl groups at δ 192.4 (C-8), 177.1 (C-11), and 173.7 (C-16), three aliphatic oxygen-bearing carbons at δ 77.1 (C-7), 74.9 (C-12), and 66.2 (C-9), and a aliphatic methyl carbon at δ 12.1 (C-21).

TOCSY experiment were realized to confirm three sets of spin network present in emiline structure. Selective excitation performed at δ 6.05 ppm (H-2) correlated to chemical shift signal at 3.44 (H-3a) and 3.24 (H-3b) and allylic coupling to signals at 5.10 (H-9a) and 4.48 (H-9b) ppm (**Figure F5**). Allylic coupling between signals 4.48 and 5.10 ppm was shown also by change in residual multiplet structure after Homonuclear decoupling experiment by irradiating signal at δ 4.48 ppm (H-9) (**Figure F6**). Selective excitation performed at 2.84 ppm (proton bound to C-14) correlated to chemical shift signal at δ 2.29 (H-15), 2.15 (H-14b), 1.61 (H-20a), 1.51 (H-20b), 0.86 (H-21) and allylic coupling to signal at 5.11 (H-19) ppm (**Figure F7**). Selective excitation performed at δ 2.49 ppm (H-6a) correlated to chemical shift signal at 4.84 (H-7), 2.93 (H-5a), 2.73 (H-5b) and 2.22 (H-6b) ppm (**Figure F8**).

Figure 4 - Emiline chemical structure

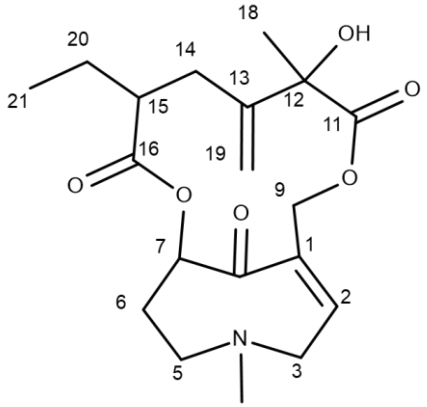


Table 2 – Comparative experimental NMR signals of 3 (ppm) to literature emiline

	^{13}C exp.[ab]	^{13}C lit.[c]	^{1}H exp.[b]	^{1}H lit. [c]	HMBC
1	134.2	135.7	-	-	
2	135.6	131.8	6.05 (br s, 1H)	6.02 (br s, 1H)	
3	59.2	66.6	3.49 (d, *J*=18 Hz, 1-H)	3.44 (br d, *J*= 18Hz, 1H)	C-1, C-5
			3.29 (d, *J* = 18 Hz, 1H)	3.20 (br d, *J*= 18Hz, 1H)	C-1, C-5
5	53.9	53.2	2.93 (m, 1H)	2.85 (m, 1H)	
			2.73 (m, 1H)	2.65 (m, 1H)	
6	35.7	37.5	2.49 (m, 1H)	2.43 (m, 1H)	
			2.22 (br s, 1H)	2.2 (m, 1H)	C-8
7	77.1	77.2	4.84 (t, *J*= 3.0 Hz, 1H)	4.81 (t, *J*= 3Hz, 1H)	
8	192.4	191.5	-	-	
9	66.2	58.6	5.10 (d, 1 H)	5.09 (d, *J*= 11Hz, 1H)	C-2, C-11
			4.48 (d, *J*= 11 Hz, 1H)	4.46 (br d, *J*= 11, 1H)	C-2, C-11
11	177.1	174.7	-	-	
12	74.9	75.2	-	-	
13	146.6	146.5	-	-	
14	37.7	36.1	2.84 (dd, *J*= 13, 3 Hz, 1H)	2.80 (m, 1H)	C-16,C-19
			2.15 (m, 1H)	2.10 (m, 1H)	C-12, C-19
15	47.0	46.9	2.27 (m, 1H)	2.25 (m, 1H)	
16	173.7	177.6	-	-	
18	28.6	18.5	1.55 (s, 3H)	1.53 (s, 3H)	C-11,C-13
19	118.1	117.9	5,11 (m, 2 H)	5.1 (d, *J*= 6, 2H)	C-11, C-12, C-14
20	26.6	26.5	1.61 (m, 1H)	1.53 (m, 2H)	
			1.50 (m, 1H)		C-16
21	12.1	12.0	0.86 (t, *J*= 7.4 Hz, 3H)	0.85 (t, *J*= 7, 3H)	C-15
NMe	40.5	40.2	2.16 (s, 3H)	2.06 (s, 3H)	C-3, C-5

[a] Spectra recorded in $CDCl_3$ at 600 MHz for ^{1}H and 150 MHz for ^{13}C.
[b] Signals obtained from DEPTQ, HSQC and HMBC experiments.
[c] Barbour & Robins, 1987. Phytochemistry 26, 2430–2431. Spectra recorded in $CDCl_3$ at 200 MHz for ^{1}H and 50 MHz for ^{13}C.

In addition to the isolated compound emiline (**3**), fractions of all *E. fosbergii* parts were analyzed by LC-MS and 14 other PAs were annotated or characterized (**Table 3**). It was possible to see PAs in all plant parts of *E. fosbergii* (Erro! Fonte de referência não e ncontrada. and **Figure 5)**. Pyrrolizidine alkaloids structure of annotated compounds are present in **Figure 6.** PAs with otonecine base were identified by the diagnostic ions *m/z* 122, 150, 168 (HSIEH et al., 2015): two isomers of emiline (**4** and **9**), probable senkirkine and neosenkirkine; an acetylated derived of senkirkine was annotated as ligularidine (**7**); and a putative PA with *m/z* $[M+H]^+$ 410.2162, which we proposed to be a hydrogenated ligularidine derivative (**8**).

Retronecine base PAs were also annotated in *E. fosbergii* as seneciphylline (**1**), senecionine (**2**), and acetylseneciphylline (**6**) , as well as correspondent N-oxide of these three pyrrolizidine alkaloids: seneciphylline N-oxide (**1-NO)**, senecionine N-oxide (**2-NO**), and acetylseneciphylline N-oxide (**6-NO)**. Fragments of *m/z* $[M+H]^+$ 140 and 158 indicate the presence of nontoxic PAs from platynecine base (RUAN et al., 2012), such as platyphylline (**10**), as well as platyphylline N-oxide (**10-NO**), and a putative PA with *m/z* $[M+H]^+$ 340.2113, which we proposed to be a hydrogenated platyphylline derivative (**5**).

Just one mass peak different of **1-10** annotated as PA peak was observed in the metabolic profiling of leaf hydroethanolic extract (**Figure 5E**). Peak (**11**) presented *m/z* $[M+H]^+$ 382.1861, at retention time 4.67 min and annotation was confirmed by presence of fragments at *m/z* 122, 150, and 168. Peak **11** was annotated as petasitenine or its isomer otosenine.

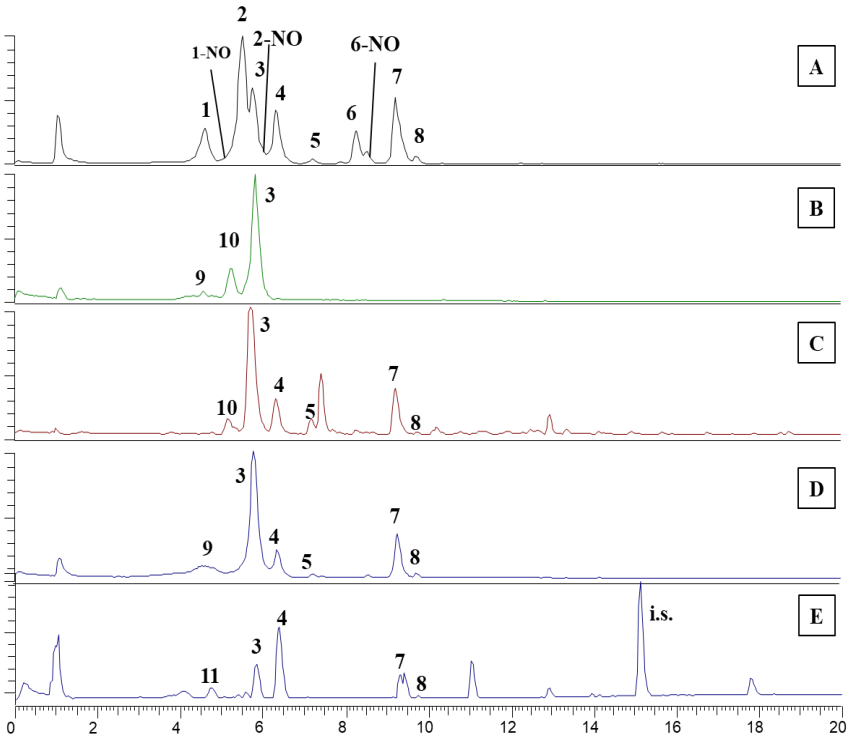
Putative pyrrolizidine alkaloids 5, in low amount, and 8 were found also in hydroethanolic leaves extract (**Figure 5E**), which reduces hypothesis of been adducts formed by zinc reduction step. MS2 fragments and probable structure of putative pyrrolizidine alkaloids are present in **Appendix G**.

Table 3 - Pyrrolizidine alkaloids identified in *E. fosbergii*.

ID	*m/z* Value $[M+H]^+$	RT (min)	Fragments	Error (ppm)	Identity	EFFD	EFLD	EFSD	EFRD
1	334.1638	4.64	120.08; 138.09; 94.06; 306.17	-5.0	seneciphylline	+	+	-	+
1-NO	350.1599	5.15	118.07; 120.08; 136.07; 138.09	-3.7	seneciphylline N-oxide	+	lc	lc	+
2	336.1795	5.44	120.08; 138.09; 94.06; 308.17	-4.8	senecionine	+	+	-	+
3	366.1899	5.72	168.10; 150.09; 122.06; 348.18	-4.6	emiline	+	+	+	+
2-NO	352.1750	6.00	118.07; 120.08; 136.07; 138.09	-3.0	senecionine N-oxide	+	lc	lc	+
4	366.1903	6.35	168.10; 150.09; 122.06; 338.19	-3.8	senkirkine or neosenkirkine	-	+	+	+
5	340.2113	7.18	122.10; 140.10; 155.10; 176.10; 158.11	-3.3	putative base platynecine	+	+	Lc	+
6	376.1743	8.28	120.08; 138.09; 316.15	-4.6	acetylseneciphylline	-	+	-	+
6-NO	392.1695	8.59	118.07; 120.08; 136.07; 138.09	-3.8	acetylseneciphylline N-oxide	lc	lc	Lc	+
7	408.2006	9.24	122.08; 150.09; 168.10; 348.18	-3.9	ligularidine	-	+	+	+
8	410.2162	9.74	122.05; 150.08; 168.09; 350.18	-4.1	putative base otosenine	-	+	lc	lc
9	366.1903	4.55	122.05; 150.08; 158.09; 168.09	-3.8	senkirkine or neosenkirkine	lc	-	+	lc
10	338.1953	5.36	122.09; 140.11; 310.20; 158.11; 240.16	-4.2	platyphylline	lc	+	-	-
10-NO	354.1907	5.74	118.07; 120.08; 136.07; 138.09; 140.11	-2.7	platyphylline N-oxide	lc	lc	lc	-
11 [1]	382.1861	4.67	122.08; 150.08; 168.10	1.3	petasitenine or otosenine	-	-	-	-

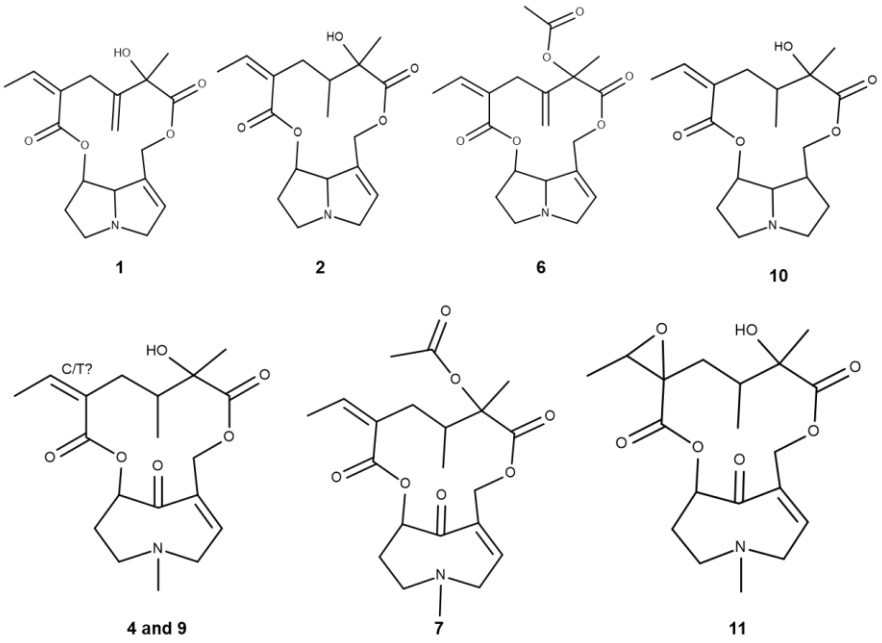
EFFD: flowers CH_2Cl_2 fraction; EFLD: leaves CH_2Cl_2 fraction; EFSD: stems CH_2Cl_2 fraction; EFRD: roots CH_2Cl_2 fraction. + indicates presence, - indicates absence, lc indicates low concentration. [1] Peak 11 was present only in leaves hydroethanolic (80%) extract.

Figure 5 - Base peak LC-ESI-MS chromatograms in positive mode of *E. fosbergii.*



PA-enriched fractions of roots (**A**), flowers (**B**), leaves (**C**), and stems (**D**), leaves hydroethanolic (80%) extract (**E**). i.s.: internal standard (chrysin). Peak numbering is according to Erro! Fonte de referência não encontrada.. 1- seneciphylline; 2- s enecionine; 3- emiline; 4- senkirkine; 5- putative platinecine; 6- acetylseneciphylline; 7- ligularidine; 8- putative otonecine; 9- neosenkirkine; 10- platyphylline; 11- petasitenine or otosenine.

Figure 6 - Chemical structure of annotated pyrrolizidine alkaloids in *E. fosbergii*.



1- seneciphylline; 2- senecionine; 4- (*Z*) senkirkine; 6- acetylseneciphylline; 7- ligularidine; 9- (*E*) neosenkirkine; 10- platyphylline; 11- petasitenine or otosenine.

Thus far, after *E. coccinea*, *E. fosbergii* is the second species reported to contain emiline (BARBOUR; ROBINS, 1987), which could indicate a close phylogenetic relationship between these two species. Nicholson (1980) hypothesized that *E. fosbergii* originated recently from the hybridization of *E. sonchifolia* var. *sonchifolia* and a somewhat different version of *E. coccinea*. Moraes and Guerra, (2010) confirmed *E. fosbergii* as tetraploid, an indicative of hybrids, and cytological similarity with *E. sonchifolia*. *Emilia fosbergii* different capacity for emiline production from that of *E. sonchifolia* could indicate descended traces from *E. coccinea*. All pyrrolizidine alkaloids found in *E. fosbergii* were described previously in *E. sonchifolia* and/or *E. coccinea*.

Pyrrolizidine alkaloids are important chemotaxonomic markers at the intra-tribe level in Asteraceae and the genera level as well, as previously described in *Senecio* and *Crotalia* (FLORES; TOZZI; TRIGO, 2009; TRIGO et al., 2003).

Biosynthesis of PAs in *Senecio* species was determined to start in the roots, and senecionine is the precursor of most PAs. Senecionine is converted to senkirkine, considered the precursor of otonecine base PAs (PELSER et al., 2005). In our results, senecionine and seneciphylline, both retronecine base PAs, were found mainly in roots, while trace amounts were found in leaves and stems. Meanwhile, the otonecine base PAs emiline and ligularidine were found in the whole plant. That could indicate that PA transformation of senecionine to senkirkin and emiline in *E. fosbergii* occurs in roots.

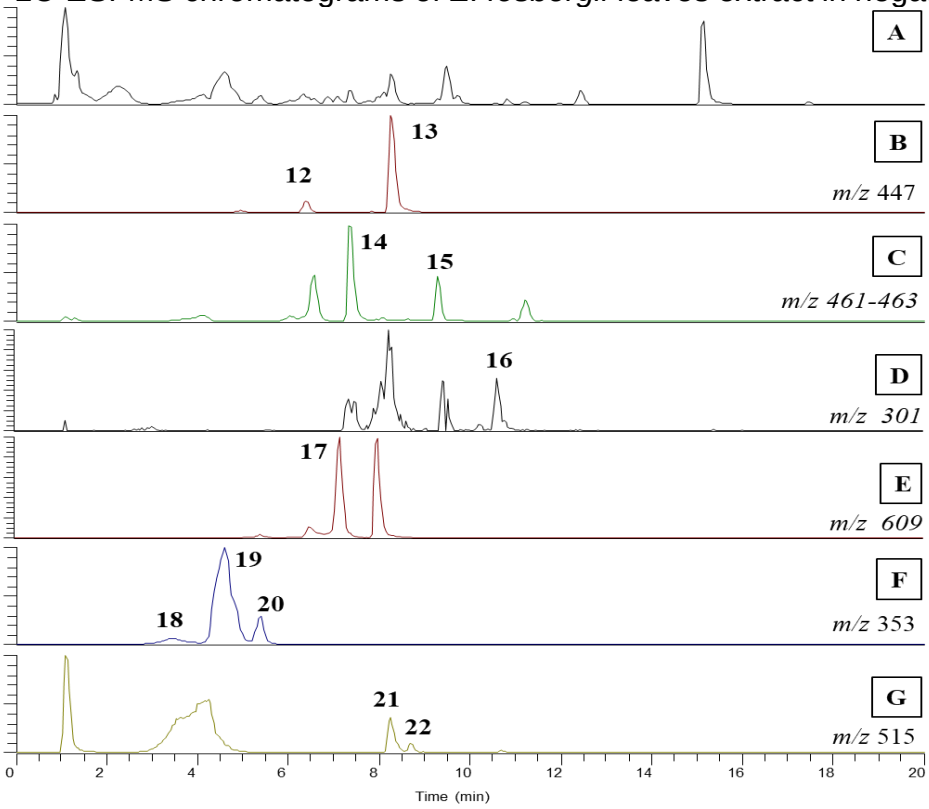
Additionally, chemical profiles of *E. fosbergii* leaf hydroethanolic extracts (Samples described in item 3.2.3 and metabolic profiling in item 3.2.4) were analyzed by LC-MS in both positive and negative ionization modes to reveal the presence of isoorientin (**12**), quercitrin (**13)**, isoquercitrin (**14),** and isorhamnetin-3-*O*-rhamnoside (**15**). Chlorogenic acid derivatives present in the extracts were 3-*O*-caffeoylquinic acid (**18**), 5-*O*-caffeoylquinic acid (**19**), 4,5-dicaffeoylquinic acid (**22**), a putative caffeoylquinic acid (**20**), and a putative dicaffeoylquinic acid (**21**). The presence of compounds **12−24** in *E. fosbergii* hydroethanolic leaf extracts (**Figure 7**) are presented in **Table 4** and structures in **Figure 8.**

To the best of our knowledge, with exception of quercetin and 3-*O*-caffeoylquinic acid related by PEISINO et al., (2019) this is the first report of all compounds (**1-24)** in *E. fosbergii*.

Table 4 - Flavonoids and chlorogenic acid derivatives present in E. fosbergii hydroethanolic leaf extracts.

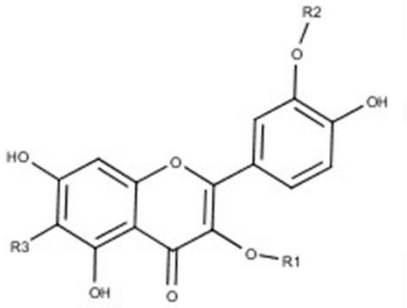
ID	*m/z* Value	RT (min)	Fragments	Error (ppm)	ID level	Identity
12	447.0932	6.38	327.05; 167.03; 357.06; 173.04; 191.06; 447.09; 297.04	1.0	1	isoorientin
13	447.0949	8.25	300.03	4.8	1	quercitrin
14	463.0891	7.39	300.03	3.0	1	isoquercitrin
15	461.1098	9.27	314.04; 96.96; 461.11; 271.03; 285.04	3.1	2	isorhamnetin-3-*O*-rhamnoside
16	301.0358	10.68	245,05	-0.3	1	quercetin
17	609.1469	7.17	300.03	-1.2	1	rutin
18	353.0881	3.37	191.06; 135.04; 179.03; 173.05; 167.03; 108; 123.04	2.4	1	3-*O*-caffeoylquinic acid
19	353.0880	4.57	353.09; 191.06; 173.05; 179.03; 135.04	2.1	1	5-*O*-caffeoylquinic acid
20	353.0884	5.38	191.06; 179.03; 135.04	3.2	3	putative caffeoylquinic acid
21	515.1215	8.25	353.08; 255.03; 191.05; 179.00; 173.05; 135.04	4.9	3	putative dicaffeoylquinic acid
22	515.1207	8.63	353.08; 255.03; 191.05; 179.00; 173.05; 135.04	3.4	1	4,5-dicaffeoylquinic acid
23	191.0555	1.06	191.05	-3.0	1	Quinic acid
24	115.0026	1.17		-4.5	1	Maleic acid

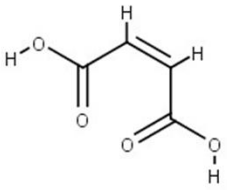
Figure 7 - LC-ESI-MS chromatograms of *E. fosbergii* leaves extract in negative mode.



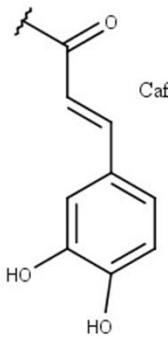
Full scan of *E. fosbergii* leaves extract (**A**). Selected ion monitoring chromatograms of *E. fosbergii* leaf extract to demonstrate presence of flavonoids at m/z 447 (**B**) and at *m/z* 461-463 (**C**), quercetin at *m/z* 301 (**D**), rutin at *m/z* 609 (**E**), caffeoylquinic acids at m/z 353 (**F**), and dicaffeoylquinic acids at m/z 515 (**G**). Constituents numbering according to **Table 4.**

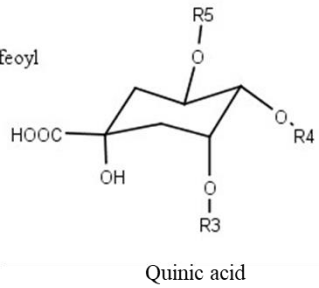
Figure 8 - Chemical structures of flavonoids and caffeoylquinic acid derivatives found in *E. fosbergii* leaf extracts





24
Maleic acid





Quinic acid

		R1	R2	R3
Isoorientin	**12**	H	H	Glucosyl
Quercitrin	**13**	Rhamnosyl	H	H
Isoquercitrin	**14**	Glucosyl	H	H
Isorhamnetin-3-O-rhamnoside	**15**	Rhamnosyl	CH_3	H
Quercetin	**16**	H	H	H
Rutin	**17**	Glucosyl-Rhamnosyl	H	H

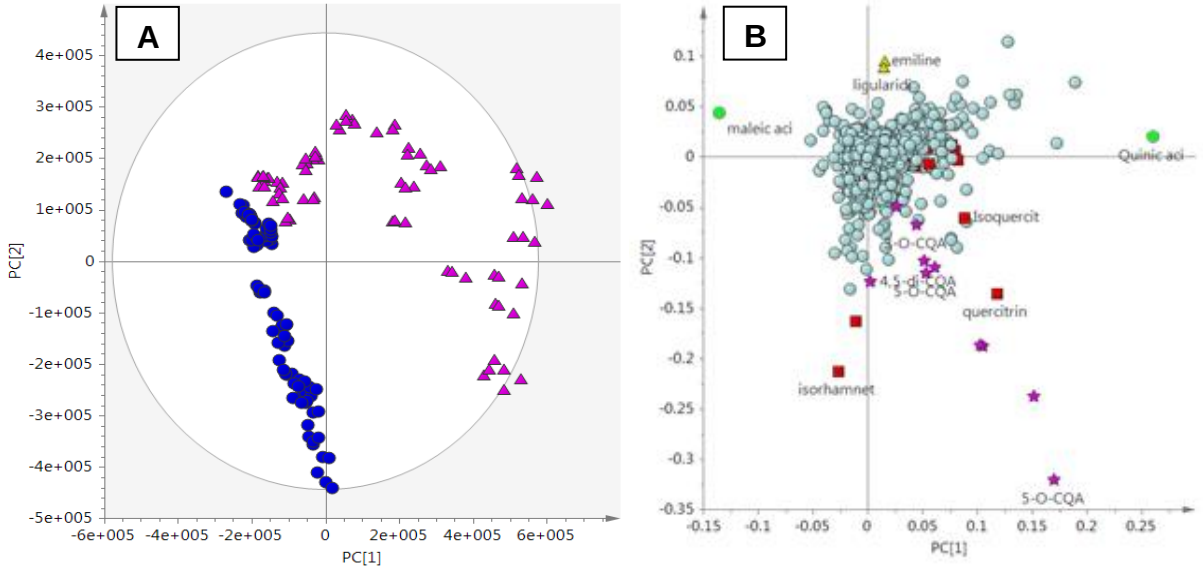
		R3	R4	R5
3CQA	**18**	Caffeoyl	H	H
5CQA	**19**	H	H	Caffeoyl
4,5diCQA	**22**	H	Caffeoyl	Caffeoyl
Quinic acid	**23**	H	H	H

4.1.2 Comparison of metabolic profiling of *E. fosbergii* and *E. sonchifolia* (Experiment 1b)

In this topic, it was possible to observe that there is chemical difference between *E. sonchifolia* and *E. fosbergii* chemical profiling, with highlights to PAs, quercitrin, isoquercitrin and caffeoylquinic acid derivatives as chemical markers.

Metabolomic fingerprints of 27 samples of *E. fosbergii* and 29 samples of *E. sonchifolia* resulted in obtention of a total of 672 features in negative mode and 746 features in positive mode. Both modes were used in multivariate analysis. In PCA, a model was built with R^2= 0.93, Q^2 = 0.808 in 28 PCs, which Erro! Fonte de referência não e ncontrada. shows PC1 (0.27) x PC2 (0.16). In scores plot (Erro! Fonte de referência não encontrada.**A)**, it was possible to observe sample distribution was separated by species, with most of *E. sonchifolia* samples clustered in the third quadrant and a subgroup of *E. sonchifolia* in fourth quadrant. Species clustered separately indicates chemical composition different between each other. In loadings plot (Erro! Fonte de r eferência não encontrada.**B**), it is possible to identify higher peak area of following compounds in *E. fosbergii* leaves: quinic acid, senkirkin, emiline and acetylsenkirkine presents in the first quadrant, and quercitrin, isoquercitrin and caffeoylquinic acids derivatives in the second quadrants. This result indicates that leaves of *E. fosbergii* accumulate higher amounts of PAs, mainly senkirkine, emiline and ligularidine, than *E. sonchifolia*. PAs investigation in PCA and chromatograms revealed that senkirkine is predominant in *E. sonchifolia*, meanwhile isomer emiline is predominant in *E. fosbergii* and absent in *E. sonchifolia* samples, except for two *E. sonchifolia* samples. Even so, these two *E. sonchifolia* samples are nearest *E. sonchifolia* cluster than *E. fosbergii*, which means that solely emiline presence it is not enough to differentiate species. Ligularidine was absent in *E. sonchifolia* also. All peaks found in *E. sonchifolia* were also found in *E. fosbergii*. Peaks information annotated in *E. sonchifolia* are described in Erro! Fonte de referência não encontrada..

Figure 9 - PCA analysis of *E. sonchifolia* and E. *fosbergii*.



A) Score plot PC1 (0.27) x PC2 (0.16) represents samples of *E. fosbergii* (pink triangles) and *E. sonchifolia* (blue circles). **B**) Loading plot represents peaks area. It is possible to visualize pyrrolizidine alkaloids (yellow triangles), flavonoids (red box), chlorogenic acid derivatives (pink stars), organic acids (green circles), and unknown metabolites (blue circles) in loadings plot. R^2= 0.929, Q^2 = 0.808 (28 PCs).

Both species presented a similar profile in richness (variety) of compounds present in leaves extracts, but compounds peak area variation were enough to differentiate *E. sonchifolia* from *E. fosbergii* in PCA analysis. This similarity of compounds variety could be because of both species are adapted to the same climate conditions, herbivores or diseases (KIRK et al., 2010). The capacity to biosynthesize same variety of compounds could also be due very close taxonomy relationship and even fortify the hypothesis of *E. fosbergii* origin as descendant from *E. sonchifolia* (MORAES; GUERRA, 2010; NICOLSON, 1980). Chromatograms visualization of one single sample from each group to both species is available in **Appendix H.**

PAs are studied as chemotaxonomic markers for long time. Presence in just Boraginaceae, Asteraceae (Senecioneae and Eupatorieae tribes) and a few genera of Fabaceae and Orchidaceae indicates important chemotaxonomy information. Infrageneric studies evaluated PA as important chemotaxonomic markers for the Asteraceae genera *Crotalia* and *Senecio* (FLORES; TOZZI; TRIGO, 2009; TRIGO et al., 2003). Emiline and ligularidine could be potential chemical markers to differentiate *E. fosbergii* and *E. sonchifolia* leaves, mainly in earlier development stages, when morphology discrimination is still unclear. In 29 *E. sonchifolia* leaves extract samples, emiline was found in just two adult plants as a minority compound (**Figure H3** present

in **Appendix H**), which corroborates with literature, since emiline was not described in *E. sonchifolia* samples so far.

Table 5 - Peaks information detected in *E. sonchifolia*.

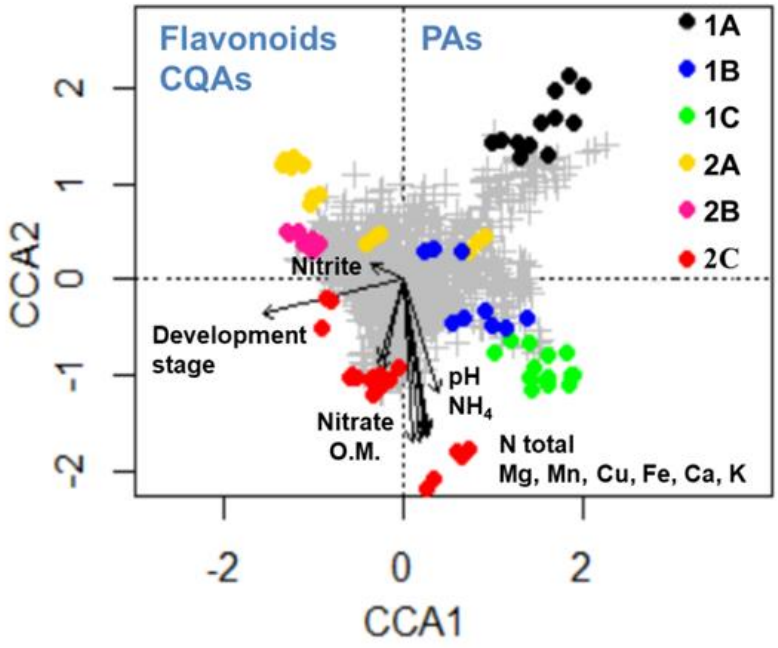
Ionization	*m/z value*	RT (min)	Identity	ID level*	Error (ppm)
$[M-H]^-$	447.0949	8.25	Quercitrin	1	4.8
$[M-H]^-$	463.0891	7.39	Isoquercitrin	1	3.0
$[M-H]^-$	461.1098	9.30	Isorhamnetin-3-O-rhamnoside	2	3.1
$[M-H]^-$	353.0880	4.54	5-*O*-caffeoylquinic acid	1	2.1
$[M-H]^-$	353.0884	5.38	Putative caffeoylquinic acid	3	3.2
$[M-H]^-$	515.1215	8.29	Putative dicaffeoylquinic acid	3	4.9
$[M-H]^-$	115.0026	1.14	Maleic acid	3	-9.2
$[M-H]^-$	191.0556	1.03	Quinic acid	1	-2.7
$[M-H]^-$	301.0354	10.7	Quercetin	1	-1.7
$[M-H]^-$	315.0514	11.25	3-Methylquercetin	1	-1.4
$[M-H]^-$	609.1459	7.25	Rutin	1	-2.9
$[M+H]^+$	366.1917	5.85	Emiline	3	0.0
$[M+H]^+$	366.1926	6.39	Senkirkine	2	2.6

4.1.3 Influence of cultivation and flowering in the chemical profile of *E. fosbergii* and *E. sonchifolia* (Experiment 1c)

Canonical correspondence analysis (CCA) is a multivariate analysis that associate two data sets and allows to visualize relationship between them. In concept, CCA is a multivariate multiple linear regression followed by an unconstrained ordination by an unconstrained ordination of the matrix of fitted values. In this study, CCA search for a series of linear combinations of the environmental variables that best explain variation of metabolic profiling (BORCARD; GILLET; LEGENDRE, 2018). Visualization of relationship between the two datasets is possible by CCA triplot, in which arrows represent environmental variables, crosses represent metabolites, and points represent plant samples.

Canonical Correspondence Analysis of *E. fosbergii* (Erro! Fonte de referência não e ncontrada.), demonstrate that development stage was the factor which most influences *E. fosbergii* chemical profile samples, so plants collected before flowering are separated by CCA1, which means those samples are clustered in first and second quadrants, with exception of a few samples of the group 2A (green) collected after flowering. Group 2A represent samples cultivated in non-treated soil. Size growth of this samples was stunted and flowering stage of a few plants was retarded, and a few samples did not even bloomed, probably this is the reason to chemical profile of these samples looks like samples of plants collected in earlier development stage from other groups, presents in third and fourth quadrants.

Figure 10 – CCA triplot of *E. fosbergii* leaves chemical profile related to cultivation.



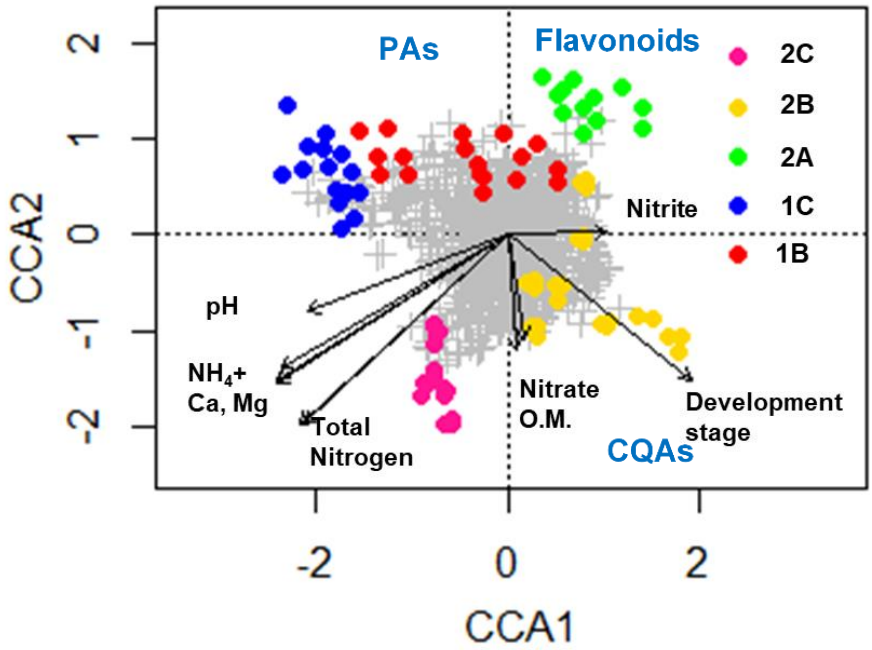
Numbers means harvest time 1- non-flowering (50 days); 2- flowering (100 days). Capital letters represent soil planted: A- non treated soil; B- soil treated with horse manure; C- soil treated with chicken manure. CQA – caffeoylquinic acid derivatives. Blue words represent identified chemical classes of loadings. CCA1 (0.435) x CCA2 (0.240), R^2= 0.504, Inertia: 0.447 (constrained) and 0.441 (unconstrained).

Canonical analysis indicates that development stage is the most significative parameter analyzed to PA variation, since samples present in first square presented positive correlation to PAs emiline and ligularidine. Samples cultivated in soil A and collected before flowering (group 1A) had the highest peak area of PAs in study, mainly emiline, senkirkine and ligularidine. Samples of *E. fosbergii* collected after flowering presented higher peak area of flavonoids and caffeic acids derivatives. Nitrogen

species (nitrite, ammonium, and nitrate) have different importance for PAs content. In *E. fosbergii*, higher contents of nitrate and nitrite (soil treated with horse manure) produced lower amounts of PA in comparison to other groups.

Between most important parameters which influences *E. sonchifolia* chemical profile, canonical analysis (Erro! Fonte de referência não encontrada.) indicates development, f ollowed by total nitrogen, ammonium, calcium, and magnesium ions. Higher contents of flavonoids and caffeic acid derivatives were markers to distinguish chemical profile of plants after flowering from plants in earlier development stage. Detected PAs had no significative variance in *E. sonchifolia*.

Figure 11 – CCA triplot of *E. sonchifolia* leaves chemical profile related to cultivation



Numbers means harvest time **1**- non-flowering (50 days); **2**- flowering (100 days). Capital letters represent soil planted: **A**- non treated soil; **B**- soil treated with horse manure; **C**- soil treated with chicken manure. CQA – caffeoylquinic acid derivatives. Blue words represent identified chemical classes of loadings. CCA1 (0.453) x CCA2 (0.287), R^2= 0.498, Inertia: 0.329 (constrained) and 0.331 (unconstrained).

High content of PAs in plants cultivated in non-treated soil corroborates with resource availability hypothesis. This theory postulates that plants produce more defense metabolites in stressed conditions and compromised development (COLEY; BRYANT; CHAPIN, 1985; WINK, 2003). Meanwhile plants with high grown rate compensate herbivore attacks, so plant in good conditions do not need to produce so much defense mechanisms (COLEY; BRYANT; CHAPIN, 1985). Similar results were found in

Senecio jacobeae, *Senecio aquaticus* and derivate hybrids, when plants cultivated in poor soil or in water privation had higher PAs content (HOL; VRIELING; VAN VEEN, 2003; KIRK et al., 2010). In an experiment where nitrogen and phosphor availability were controlled in *S. jacobeae*, no effect was found on PAs concentration under N or P limitation in soil (Vrieling, 1994).

Biosynthesis of PAs was studied in *Senecio* species and relates that PAs production starts in roots with production of senecionine and PAs transformations could occur in different plant parts. N-oxide retronecine is formed to increase polarity and deliver these compounds to other plant parts. Senecionine would be converted to senkirkine, considered the precursor of otonecine base PAs and then another reactions occur to form derivatives, for example, acetylation to produce ligularidine (PELSER et al., 2005). Plant nutrition affects flowering, which can be associated with PA transformation (HOL; VRIELING; VAN VEEN, 2003). If derivatives compounds were formed on flowers, would be expected that plants grown under nourishing soil, which flowering had been retarded, would produce lower variety of PAs. The results did not follow this theory, which indicates that PAs transformations in *Emilia* does not occur in flowers, but in other parts of the plant.

In a few *E. fosbergii* leaves samples collected before flowering, emiline (**3**) and senkirkine (**4**) are majority compounds (**Figure H3** in **Appendix H**). The observed PAs variation in both species is an important health security information, and misidentification use of *E. fosbergii* instead of *E. sonchifolia* could be riskier since the first specie shown higher amounts of toxic PA. Meanwhile grown plants, usually used as edible and medical plants, produced lower PAs contents, so population probably is already using plants with lower PA contents.

To summarize, *E. fosbergii* has different chemical profile in comparison to *E. sonchifolia*, mainly in concentration of compounds such as caffeoylquinic acids derivatives, isoquercitin and the PAs emiline and ligularidine. PAs contents were higher in *E. fosbergii* than *E. sonchifolia* and emiline and ligularidine presences are potential chemical markers to identify *E. fosbergii,* mainly in plants before flowering, when morphology differences are not clear yet. Chemical profile varies in function of development stage and soil conditions. Both species presented higher contents of flavonoids and caffeoylquinic acid derivatives in samples collected after flowering. In

E. fosbergii, higher contents of PAs were found in plants grown in soil with lower contents of nitrate and nitrite.

Experiment 2

4.1.4 Comparison metabolic profiling of wild and domesticated crop species

Asteraceae is one of the biggest plant families and have as representative common lettuce as one of the most famous edible plants consumed around the world. Lettuce and a few varieties of chicory, which also belongs to Asteraceae family, are accessible and cheap leafy vegetables which have a vantage of being consumed raw in salads, since cooking can cause degradation of thermolabile phytochemicals (PEPE et al., 2015). These species belong to *Cichorieae* tribe which phytochemical composition had been described mainly by presence of glycosylated flavonoids, cinnamic acids derivatives, anthocyanin in a few blue/purple color varieties and sesquiterpene lactones (SAREEDENCHAI; ZIDORN, 2010; ZIDORN, 2008). Besides antioxidant and anti-inflammatory potential, evaluation of sesquiterpene lactones to antitumoral activity against a diversity of tumoral cell lines improved importance of Asteraceae species as edible plants (ZIDORN, 2008). Sesquiterpene lactones are usually responsible for the bitterness present in chicory and in lettuce (PRICE et al., 1990).

UHPLC-HRMS analysis of 57 plant extracts from nine species, in a total of 171 injections, resulted in 303 and 509 mass features obtained in negative and positive ionization modes, respectively, after pretreatment. Two multivariate analysis were performed with detected peaks: unsupervised PCA with all extracted peak areas and heatmap with identified compounds peak areas. Comparison to previous isolated and standard compounds (CHIBLI et al., 2018) literature and AsterDB allowed us to propose the identity of 49 compounds: 3 alkylamides, 7 amino acids, 1 aurone, 3 chalcones, 2 anthocyanidins, 2 flavanones, 7 flavonoids, 2 coumarins, 12 organic acids (cinnamic acid derivatives included), 2 phenols, 8 sesquiterpene lactones and 14 putative sesquiterpene lactones distributed in these species. Putative compounds were not identified due to existence of more than one peak with same *m/z* value or lack of information in literature about fragmentation in ESI. All identified compounds and detailed chromatogram and mass spectrometer information are listed in Erro! F onte de referência não encontrada.. Chemical structures of annotated compounds and putative compounds are presented in **Appendix I**.

Table 6 - Compounds identity, chromatographic and spectrometric information

ID	Ionization	*m/z*	RT	Identity	ID level[1]	Compound class	Exact Mass	Error (ppm)
25	$[M+H]^+$	222.1861	17.02	Spilanthol	3	alkylamide	221.1780	1.5
26	$[M+H]^+$	236.2021	19.11	(2E,6Z,8E)-N-(2-methylbutyl)-2,6,8-decatrienamide	3	alkylamide	235.1936	2.7
27	$[M+H]^+$	230.1550	15.71	(2E,5Z)-N-isobutylundeca-2,5-diene-8,10-diynamide	3	alkylamide	229.1467	2.3
28	$[M+H]^+$	147.0770	1.03	L-glutamine	1	amino acid	146.0691	0.0
29	$[M+H]^+$	132.1025	1.26	L-isoleucine;L-leucine	1	amino acid	131.0946	0.4
30	$[M+H]^+$	166.0869	1.52	L-phenylalanine	1	amino acid	165.0790	0.5
31	$[M+H]^+$	116.0713	1.03	L-proline	1	amino acid	115.0633	1.2
32	$[M+H]^+$	120.0662	1.04	L-threonine	1	amino acid	119.0582	1.1
33	$[M+H]^+$	205.0978	3.16	L-tryptophan	1	amino acid	204.0899	0.6
34	$[M+H]^+$	118.0870	1.10	L-valine	1	amino acid	117.0790	1.2
35	$[M+H]^+$	535.1101	8.79	cyanidin 3-*O*-(6'-*O*-malonylglucoside)	3	anthocyanidin	534.1010	2.6
36	$[M-H]^-$	463.1236	11.30	(2E)-1-[3-(β-D-Glucopyranosyloxy)-2,4-dihydroxyphenyl]-3-(3-hydroxy-4-methoxyphenyl)-2-propen-1-one	3	chalcone	464.1319	-1.0
37	$[M-H]^-$	491.1185	9.32	Okanin 4'-*O*-(6"-*O*-acetyl-β-D-glucopyranoside)	3	chalcone	492.1268	-1.0
	$[M+H]^+$	493.1359	9.34					2.6

Continue...

38	$[M+H]^+$	575.1414	12.20	(2Z)-2-[(3,4-Dihydroxyphenyl)methylene]-7-hydroxy-6-[(2,4,6-tri-*O*-acetyl-β-D-glucopyranosyl)oxy]-3(2H)-benzofuranone	3	aurone	574.1323	2.2
39	$[M-H]^-$	533.1288	10.67	(2E)-1-[4-[(4,6-Di-*O*-acetyl-β-D-glucopyranosyl)oxy]-2,3-dihydroxyphenyl]-3-(3,4-dihydroxyphenyl)-2-propen-1-one	3	flavanone	534.1373	-1.3
	$[M+H]^+$	535.1463	10.73					2.2
40	$[M+H]^+$	577.1564	12.31	(2S)-2-(3,4-Dihydroxyphenyl)-2,3-dihydro-8-hydroxy-7-[(2,4,6-tri-*O*-acetyl-β-D-glucopyranosyl)oxy]-4H-1-benzopyran-4-one	3	flavanone	576.1479	1.1
	$[M-H]^-$	575.1395	12.32					-0.9
14	$[M+H]^+$	465.1037	7.47	Isoquercitrin	1	flavonol	464.0955	0.9
	$[M-H]^-$	463.0872	7.44					-1.0
41	$[M-H]^-$	447.0921	7.59	kaempferol-3-*O*-glucopyranoside	1	flavonol	448.1006	-1.5
	$[M+H]^+$	449.1092	7.78					1.8
42	$[M-H]^-$	477.0660	7.53	quercetin-3-*O*-glucuronide	1	flavonol	478.0747	-1.8
	$[M+H]^+$	479.0836	7.52					2.1
17	$[M-H]^-$	609.1448	7.16	Rutin	1	flavonol	610.1534	-1.2
	$[M+H]^+$	611.1628	7.17					2.6
43	$[M-H]^-$	593.1287	10.89	Tiliroside	1	flavonol	594.1373	-1.3
	$[M+H]^+$	595.1455	10.95					0.6
44	$[M-H]^-$	431.0962	8.38	Vitexin	1	flavone	432.1056	-3.7

Continue...

45	$[M-H]^-$	461.0715	7.70	luteolin or kaempferol glucuronide	3	flavonoid	462.0798	-1.1
	$[M+H]^+$	463.0885	7.75					1.9
46	$[M-H]^-$	549.0874	7.96	quercetin 3-*O*-malonylglucoside	3	flavonol	550.0959	-1.2
	$[M+H]^+$	551.1050	7.96					2.4
47	$[M+H]^+$	163.0394	7.56	7-hydroxycoumarin	1	coumarin	162.0317	-0.5
48	$[M+H]^+$	147.0446	9.19	coumarin	1	coumarin	146.0368	0.2
49	$[M+H]^+$	517.1360	8.25	3,5-*O*-dicaffeoylquinic acid	1	organic acid	516.1268	2.7
	$[M-H]^-$	515.1180	8.26					-1.9
50	$[M-H]^-$	353.0870	3.50	3-*O*-caffeoylquinic acid	1	organic acid	354.0951	-0.7
51	$[M-H]^-$	353.0869	5.37	4-*O*-caffeoylquinic acid	1	organic acid	354.0951	-0.9
52	$[M+H]^+$	355.1035	4.52	5-*O*-E-caffeoylquinic acid	1	organic acid	354.0951	1.6
	$[M-H]^-$	353.0869	4.52		1	organic acid		-1.0
53	$[M-H]^-$	163.0390	6.96	coumaric acid	1	organic acid	164.0473	-3.2
54	$[M-H]^-$	179.0337	5.03	caffeic acid	1	organic acid	180.0423	-4.2
55	$[M-H]^-$	193.0492	7.56	ferulic acid	1	organic acid	194.0579	-4.4
56	$[M-H]^-$	191.0547	1.01	quinic acid	1	organic acid	192.0634	-4.4
57	$[M-H]^-$	133.0131	1.07	malic acid	3	organic acid	134.0215	-4.4
58	$[M-H]^-$	115.0026	1.11	maleic acid	1	organic acid	116.0110	-4.9
59	$[M-H]^-$	191.0185	1.11	citric acid	1	organic acid	192.0270	-3.5

Continue...

60	$[M-H]^-$	473.0712	7.36	chicoric acid	3	organic acid	474.0798	-1.7
61	$[M+Na]^+$	395.1320	4.36	syringin	3	Phenol	372.1420	0.6
62	$[M+H]^+$	209.0815	4.78	sinapaldehyde	3	Phenol	208.0736	0.5
63	$[M+H]^+$	279.1241	4.65	11β,13-dihydrolactucin	3	sesquiterpene lactone	278.1154	3.0
64	$[M+H]^+$ $[M-H]^-$	341.0700 339.0533	8.73 8.67	8-deoxylactucin-15-sulfate	3	sesquiterpene lactone	340.0617	1.4 -1.7
65	$[M+H]^+$	409.1853	9.48	8-epi-Deacylcinaropicrin glucoside	3	sesquiterpene lactone	408.1760	3.6
66	$[M+Na]^+$	463.1586	4.57	cichorioside B	3	sesquiterpene lactone	440.1682	1.4
67	$[M+H]^+$	409.1848	5.25	glucozaluzanin C	3	sesquiterpene lactone	408.1760	2.4
68	$[M+Na]^+$	433.1256	8.50	lactucopicrin	3	sesquiterpene lactone	410.1366	-1.7
69	$[M+H]^+$	243.1024	9.33	lettucenin B	3	sesquiterpene lactone	242.0943	1.1
70	$[M+H]^+$	397.1670	11.36	youngiajaponicol A	3	sesquiterpene lactone	396.1573	4.8
71	$[M+H]^+$	241.0866	11.74	putative 1: lettucenin A analogue	3	sesquiterpene lactone	240.0786	0.7
72	$[M+H]^+$	241.0867	15.69	putative 2: lettucenin A analogue	3	sesquiterpene lactone	240.0786	1.0

Continue...

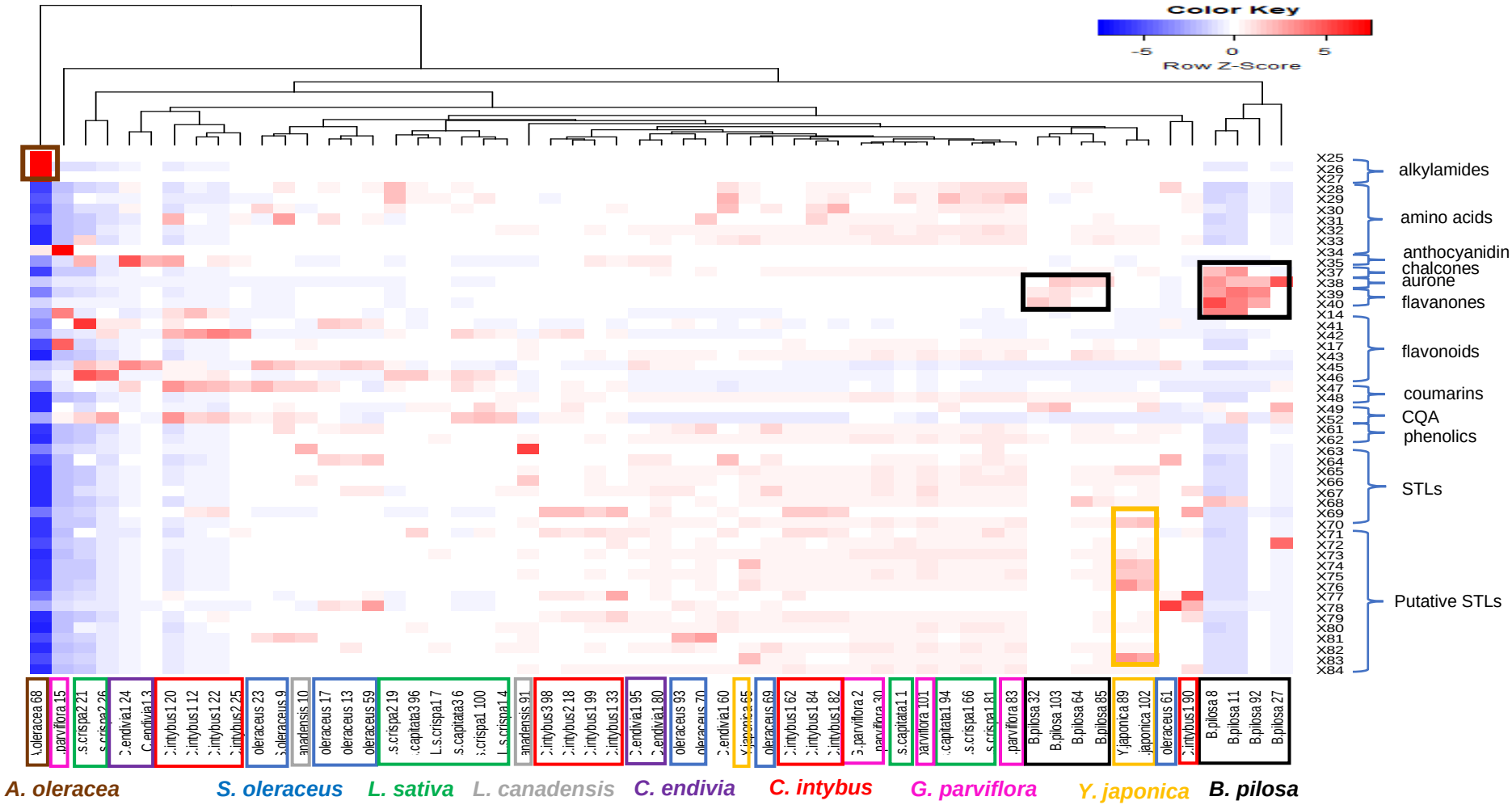
73	$[M+Na]^+$	263.0684	12.97	putative 3: lettucenin A analogue	3	sesquiterpene lactone	240.0786	-0.2
74	$[M+H]^+$	245.1179	6.13	putative 4: Dehydrozaluzanin C	3	sesquiterpene lactone	244.1099	0.4
75	$[M+H]^+$	245.1184	11.30	putative 5: Dehydrozaluzanin C	3	sesquiterpene lactone	244.1099	2.4
76	$[M+H]^+$	247.1334	9.49	putative 6: zaluzanin C;3-epizaluzanin C; Estafiatone	3	sesquiterpene lactone	246.1256	-0.1
77	$[M+H]^+$	263.1285	8.07	putative 7: armexifolin; grosheimin; jacquinelin	3	sesquiterpene lactone	262.1205	0.6
78	$[M+H]^+$	261.1127	7.92	putative 8: 8-deoxylactucin analogue	3	sesquiterpene lactone	260.1049	0.1
79	$[M-H]^-$	259.0969	9.35	putative 9: 8-deoxylactucin analogue	3	sesquiterpene lactone	260.1049	-0.7
	$[M+H]^+$	261.1127	9.12					0.1
80	$[M-H]^-$	409.1285	11.68	putative 10: lactucopicrin analogue	3	sesquiterpene lactone	410.1366	-0.5
	$[M+H]^+$	411.1445	11.62		3	sesquiterpene lactone	410.1366	0.3
81	$[M+H]^+$	411.1447	9.92	putative 11: lactucopicrin analogue	3	sesquiterpene lactone	410.1366	0.9

Continue...

82	$[M+H]^+$	411.1999	5.59	putative 12: picriside B or 11β,13-dihydro-glucozaluzanin C	3	sesquiterpene lactone	410.1941	-4.9
83	$[M+Na]^+$	447.1639	5.97	putative 13: crepiside C, or E; crepidiaside B;lactuside C; macrocliniside A	3	sesquiterpene lactone	424.1733	1.8
84	$[M+Na]^+$	447.1642	8.06	putative 14: crepiside C, or E; crepidiaside B; lactuside C	3	sesquiterpene lactone	424.1733	2.6
85	$[M-H]^-$	611.1606	6.94	okanin 3',4'-di-O-β-D-glucoside	3	chalcone	612.1690	-0.9
86	$[M+H]^+$	449.1087	7.47	cyanidin 3-O-glucoside	3	anthocyanidin	449.1084	0.8
87	$[M-H]^-$	261.1124	9.56	putative 15: armexifolin; grosheimin; jacquinelin	3	sesquiterpene lactone	262.1205	-1.2

In the HCA-Heatmap (**Figure 12**), samples are displayed horizontally related to chemical profile similarity. It is possible to observe the difference in the chemical profile of species A. *oleracea*, *Y. japonica* and *B. pilosa*. *Acmella oleracea* was the only specie which presented alkylamides compounds (**25-28**) in chemical profile (brown square), including spilanthol (**25**), one of the most studied compounds in this plant (CHENG et al., 2015; RONDANELLI et al., 2019). *Bidens pilosa* (black squares) main discriminants were an aurone (**38**), two chalcones (**36-37**) and two flavanones (**39-40**), all described before in the specie (SILVA et al., 2011). Heatmap demonstrates that *Y. japonica* presented high peak areas of sesquiterpene lactones (yellow squares), with highlight to compound youngiajaponicol A (**70**), present in all *Y. japonica* samples, but just in one sample of *S. oleraceus* and one sample of *L. canadensis*. All annotated or putative sesquiterpene lactones present in *Y. japonica* would be from pseudoguaianolide subclass, with presence of peak **65**, possible crepisides (**83-84**) and zalulazin derivatives (**74-75**). The species *C. intybus* is spread by HCA, with three main groups (red squares), but no distinction between cultivars. Also, no distinction between varieties or cultivars were found in *L. sativa or C. intybus* samples.

Figure 12 – HCA - Heatmap based on metabolic fingerprint of 57 samples from nine species.



Color rectangles in heatmap highlight major discriminant compounds. Samples are ordered in horizontal by chemical similatiry based on peak area obtained by LC-MS in positive mode. CQA – caffeoylquinic and dicaffeoylquinic acids. STLs – sesquiterpene lactones.

HCA-Heatmap used just identified metabolite peak areas obtained by positive ionization mode (**Figure 12**). Similar results were obtained using triplicates and both ionization modes using successive PCA modelling (**Appendix J**) showing that HCA-Heatmap summarizes the same information. At first, it was not possible to modelling all samples by just one single PCA analysis, since all *A. oleracea* samples were outlier to the model (**Figure J1A**), which were excluded to build a new PCA model. In second PCA model, *B. pilosa* samples also had different chemical profile, been outlier to this model (**Figure J1C**). In a third PCA model, *Y. japonica* samples were outlier to the model due to a very different chemical profile from others species (**Figure J1E**). Also, it is possible to visualize a few discriminant metabolites in loading plots present in metabolic profiling of each species, mainly to alkylamides as *A. oleracea* discriminants (**25** and **27**, **Figure J1B**), chalcones and aurones as discriminants to *B. pilosa* (**36**-**40**, **Figure J1D**) and a few sesquiterpene lactones as discriminants to *Y. japonica* (**65**, **70**, **74**-**76**, **83**, **Figure J1F**).

In this study, comparisons of plant samples by metabolomics were more robust in higher taxonomy clades. Zidorn, 2008 says that chemical variation within genera is smaller than chemical variation between genera, and the extent of phytochemical variation is generally unknown. In other words, even in similar conditions of cultivation or harvest, higher clades in taxonomy would have different chemical profile, otherwise lower clades would have more chance to present similar chemical profile. Our results demonstrated a direct comparison of species, with wild species grown in the same places, soil, moist, luminosity, among other conditions in which conventional lettuce and chicory were cultivated. Also, samples were collected from more than one place, which could generate a more robust result about difference between species. This approach aimed to minimize cultivation conditions effect in chemical profile fluctuation caused by phenology variation or cultivation conditions and allow us to see if there is a robust chemical difference between studied species.

In Heatmap (**Figure 12**), it is possible to distinguish chemical profile of species from the subtribe Cichorieae to *B. pilosa*, from the tribe Coreopsideae, *A. oleracea* from the tribe Heliantheae and *G. parviflora*, from the tribe Milleriae (FUNK et al., 2009). In this case, even samples grown in same conditions had chemical profile very different to other species and were characterized by a few chemical markers, which means that these species have a robust difference in chemical profile comparison to other species.

On the other hand, it was not possible to differentiate completely species inside Cichorieae tribe in the model and looks like local of cultivation influenced strongly samples metabolic profile.

Chalcones, flavanones and aurone were good chemical markers of *B. pilosa* samples. Since chalcones is present in only two subtribes from Cichorieae tribe (Crepidinae and Microseridinae), aurone is present just in Hypochaeridinae subtribe, and flavanones are present in these three previous mentioned subtribes and Hieraciinae. However, these three flavonoids classes are absent in Cichoriineae, Lactuceae and Hyoseridinae subtribes, which belongs *Cichorium*, *Lactuca* and *Sonchus* genera, respectively (SAREEDENCHAI; ZIDORN, 2010).

It is possible to see in Heatmap plot (**Figure 12**) that two groups of *C. intybus* samples (red rectangles) in center have higher amounts of sesquiterpene lactones than group from same species located in left side of heatmap, which have higher amounts of flavonoids, coumarins and caffeoylquinic acid derivatives. Probably this difference in chemical profile between these groups was caused by different locations cultivation, since samples present in left red rectangle were harvested in farms A and C, while samples present in other two red rectangles were obtained in farms B, D and E. A similar result was found in a chemical profile investigation of hydroxycinnamic acid derivatives and flavonoids in 12 wild species collected in North Italy, where phenolic levels variated by locality, mainly in higher hydroxycinnamic acid derivatives in samples from Monterchi and higher contents of flavonoids in samples from Bertinoro (GIAMBANELLI et al., 2018b). Our group demonstrate previously how cultivation can influence metabolic profile of plants using *Tithonia diversifolia* as example (SAMPAIO; EDRADA-EBEL; DA COSTA, 2016) and as demonstrated in this document to both Emilia species studied.

Since initial chemical profiles of evaluated samples were related to taxonomy difference among them, a second multivariate analysis was made by PCA restricted to samples from Cichorieae tribe. The chemical profile of wild species *L. canadensis* and *S. oleraceus* were compared to domesticated crop species *S. sativa*, *C. intybus* and *C. endivia.* Samples of *Y. japonica* were positioned out of ellipse hotelling confidence level (**Appendix K**), so those samples are considered outlier to the model, what means a very different chemical profile from other samples, most by higher contents of

sesquiterpene lactones, as described before. Samples of *Y. japonica* were excluded to build a new model with other species.

The species *Y. japonica* belongs to Cichorieae tribe, same of *Lactuca* and *Cichorium* genera, but it is classified in the subtribe Crepidinae, while *Lactuca* belongs to the subtribe Lactucinae and the genus *Cichorium* to the subtribe Cichoriineae (FUNK et al., 2009). Our results corroborate to chemotaxonomy importance of sesquiterpene lactones in Cichorieae tribe, since the literature describe predominance of costus guaianolide lactones in *Y. japonica*, which include crepisides C-E, youngiajaponicol A and zaluzanin derivatives (SHULHA; ZIDORN, 2019), main discriminant compounds to distinguish *Y. japonica* in our results.

The PCA model with five species from Cichorieae tribe allowed to observe samples dispersion (**Figure 13)**, R^2= 0.877, Q^2= 0.567, 14 PCs) which can be correlated to chemical profile differences and similarities. In PC2 (0.128) x PC3 (0.113) score plot (**Figure 13A**) it is possible to observe a tendency of separation of *C. intybus* samples (diamond format) in the top right quadrant and a tendency in *L. sativa* samples (triangles) separated by PC2 in bottom quadrants. Otherwise, *C. endivia* (4-point blue stars) is spread in top left in score plot (**Figure 13A)** grouped with *S. oleraceus* (5-points red stars) and *L. canadensis* (yellow circles) samples.

Since PC2 and PC3 together contributes with just 24.1% of information, 3D score plot was built with PC1 (0.266) x PC2 (0.128) x PC3 (0.113), and tendencies described before are not changed by presence of PC1 (**Figure 13C**). It is possible to see in 3D PCA score plot a tendency to form three major groups: one with *C. intybus* samples (yellow cone), the second with *L. sativa* samples (green snowflakes) and the third with other three species near each other: *L. canadensis* (pink circles), *S. oleraceus* (aqua circles), and *C. endivia* (red boxes). That means *C. endivia* have a chemical profile more similar to *S. oleraceus* and *L. canadensis* than to *L. sativa* and/or *C. intybus*.

Loading plot (**Figure 13B, PC2 x PC3**) indicates compounds discriminants for each species. *Cichorium intybus* samples have positive correlation to compounds disposed at left-top side of the plot, which comprehend mainly a putative lactone (**79**), chicoric acid (**60**), 7-hydroxycoumarin (**47**) and quercetin-3-*O*-glucorunide (**42**). Otherwise, *C. intybus* have negative correlation to compounds disposed at right side, which comprehends citric acid (**59**), 8-deoxylactucin-sulfate (**64**), kaempferol-3-*O*-

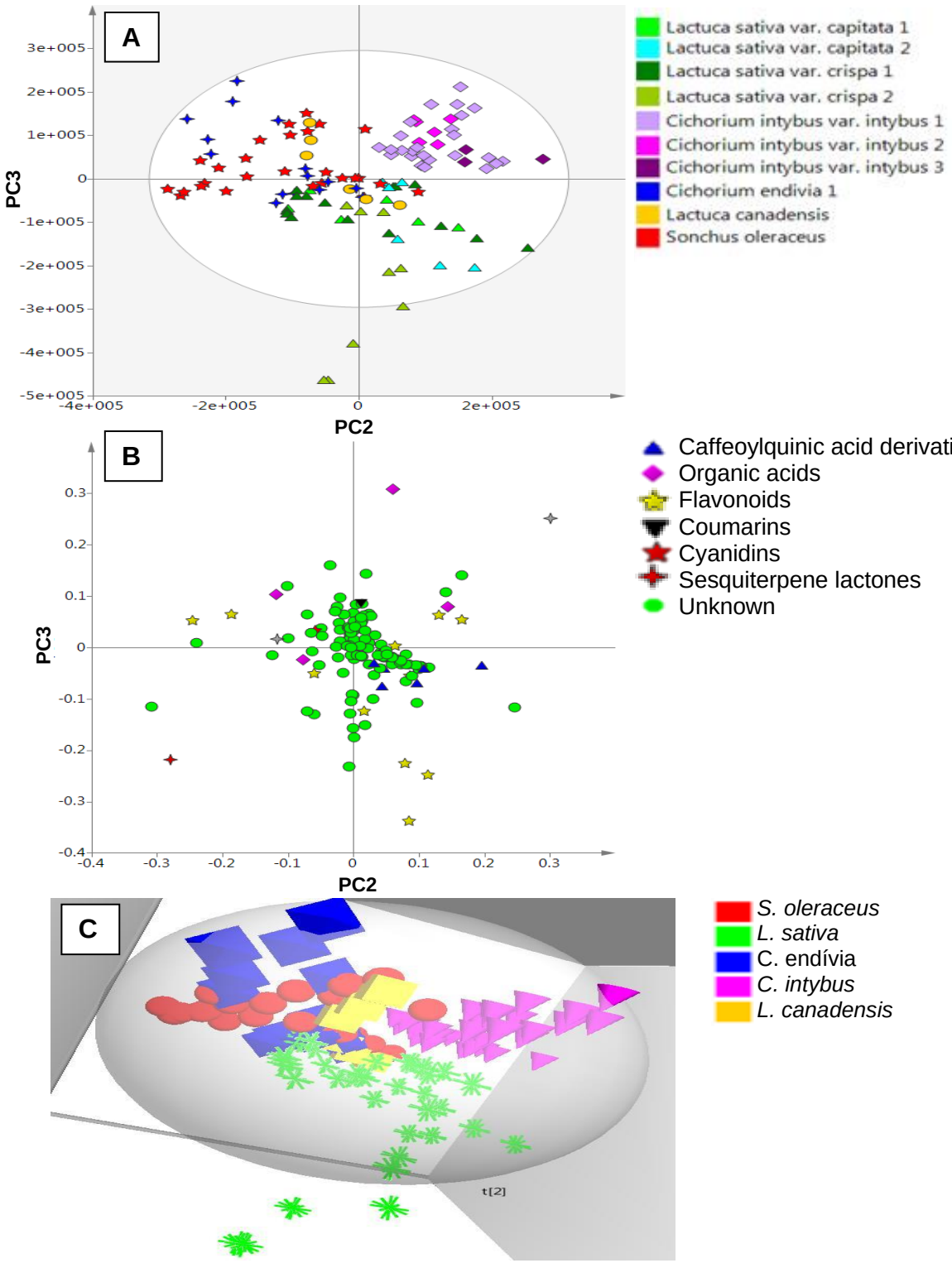
glucopyranoside (**41**) and luteolin-3-*O*-glucorunide (**23**). *Lactuca sativa* have positive correlation to caffeoylquinic acid derivatives (**49-52**), quinic acid (**56**) and isoquercitrin (**14**) and a flavonoid annotated as quercetin-3-*O*-malonylglucoside (**46**), disposed at bottom-right, which indicates higher contents of these compounds in *L. sativa* than in other species. Peak **46** also formed a dimer complex with *m/z* 1099.179 and produced a fragment in negative mode with *m/z* 505.098, both were marked as flavonoids in **Figure 13B**. Samples of *L. canadensis*, *S. oleraceus* and *C. endivia* present positive correlation to citric acid (**59**), a putative lactone (**78**) and (kaempferol or luteolin)-glucuronide (**45**), and negative correlation to caffeoylquinic acid derivatives (**49-52**) and isoquercitrin (**14**). In none of *C. endivia* samples were found isoquercitrin (**14**).

It was noticed a horizontal distribution in PC1, influenced by local of harvest, which demonstrates cultivation influence. PC1 x PC2 plot, present in **Appendix L** shows PC1 x PC2, from the same model presented in **Figure 13**, but colors are based on local instead species. Samples obtained from farms A and C clustered in the right side of the plot. On the other hand, samples from B are clustered in top left quadrant, samples collected in E clustered in bottom left quadrant and samples from D and F farms spread along left side. This result shows that sites and climatic conditions can influence chemical profile enough to differentiate samples from same species.

In samples of Cichorieae tribe were found flavonoids (**14**,**17-41-46**), cinnamic acid derivatives (**48-54**), two coumarins (**47-48**), sesquiterpene lactones derived from lactucin (**63-84**) and two anthocyanidins (**35, 86**). Coumarin and 7-hydroxycoumarin (**47,48** respectively) compounds were found in higher amounts mainly in *C. intybus* samples, and higher amounts of anthocyanidins were found in red lettuce extracts (*L. sativa* var. *crispa* 2).

A previous chemotaxonomy study about flavonoids in Cichorieae reports predominance of quercetin derivatives in *Lactuca* genus. Meanwhile kaempferol, luteolin and apigenin derivatives are predominant in *C. intybus*, as well as presence of cyanidins in samples of purple/red lettuce or in radicchio, classified as *C. intybus* var. *foliosum* (SAREEDENCHAI; ZIDORN, 2010).

Figure 13 - PCA analysis of edible species from Cichorieae tribe.



A- Score plot PC2 (0.128) x PC3 (0.113) demonstrates the disposal of samples by chemical profile. Colors indicate varieties and cultivars. Geometric forms represent species: triangles - *L. sativa*; diamonds - *C. intybus*; 5-points stars - *S. oleraceus*; 4-points stars – *C. endivia*; circles – *L. canadensis*. **B** – Loadings plot PC2 x PC3 represent mass features detected in samples. **C** – 3D score plot PC1 (0.266) x PC2 (0.128) x PC3 (0.113)

In our results, metabolite **45** (negative and positive modes) presented negative correlation to *C. intybus* as demonstrated in PCA (**Figure 13**), indicating lower peak area in *C. intybus* than in other species. Peak **45** presented *m/z* 461.0715 and 463.0885 in negative and positive mode, respectively, and MS2 fragments at *m/z* 285 and 287 in negative and positive mode respectively, which indicates presence of glucuronide moiety attached to a kaempferol or a luteolin aglycone. Literature described presence of both compounds in *C. intybus* (SAREEDENCHAI; ZIDORN, 2010) and kaempferol glucuronide was related as the most abundant kaempferol derivative in *C. intybus* (GIAMBANELLI et al., 2018b). Since just one peak was detected at this mass, presence of kaempferol-3-glucuronide could be more plausible than a luteolin derivative.

Cichorium intybus and *L. sativa* species usually have predominant lactucin derivatives as sesquiterpene lactones constituents (SHULHA; ZIDORN, 2019). Giambanelli et al., 2018 described low contents of lactucopicrin in *C. intybus,* in contrast to study of Graziani et al., 2015 where lactucopicrin were the predominant lactone in different chicory varieties. Variation of lactucopicrin content is expected, since a previous study demonstrated that lactucopicrin content levels variate even after storage (PETER; AMERONGEN, 1996). Sesquiterpene lactone is associated to bitterness and that variation in lactucopicrin could even be a problem in chicory production since it could influence in taste and people consumption (PETER; AMERONGEN, 1996).

Presence of malic acid as discriminant in top-right quadrant in PCA loading plot could indicate plants with higher vegetative growth rates or nutrient deficiency. Organic acids as malic and citric acids have several functions in plants, including assistance of organic acids in iron and phosphorus absorption in stressed conditions, since excretion of organic acids can turn these elements presents in soil in a more soluble and available form to roots. Besides that, organic acids accumulation could be used to balance pH and charges generates by nitrate (NO_3) reduction to NH_4^+, utilized in amino acid biosynthesis. In rapid vegetative growth, rates of nitrate reduction are higher and could generate toxic residues, so organic acids could be produced in higher scale to balance (LÓPEZ-BUCIO et al., 2000).

Proximity of samples in score plot of PCA (**Fig. 13**) indicates that *Cichorium endivia* showed a chemical profile more similar to *L. canadensis* and *S. oleraceus* than to *L.*

sativa and *C. intybus* samples. A few samples of *L. sativa* are near of these three species, which indicates that chemical profile of *L. sativa* have a few similarities with these species.

Presented result supports the use of studied wild edible species as a source of different phytochemicals. Use of these species as edible plants could open new opportunities in rural economy if these species were managed and/or cultivated as domesticated crop in higher scale. Secondary metabolites could make these species more attractive since public is attracting the attention to new sources of flavonoids and others antioxidant compounds, linked to idea for a healthier life. Metabolic fingerprint comparison demonstrates chemical difference mainly between species from different tribes. Use of *B. pilosa*, *Y. japonica*, *G. parviflora* and A. *oleracea* as edible plants could bring taste diversity and a variety of phytochemicals to population. Species *L. canadensis* and *S. oleraceus* had shown a similar chemical profile to *C. endivia* and a few samples of *L. sativa*, and secondary metabolites should not influence in chooses between these species, but other characteristics could be interesting to population, as odor, taste, and texture. Our study does not comprehend effect of cooking these species, which could cause degradation of some secondary metabolites.

5 CONCLUSION

In order to have a good health, a balanced diet is recommended, and it can be achieved by increasing the consumption of high amounts of vegetables. It could be done by promoting the collection and/or domestication of wild species as a public health politics measure together with the stimulation of rural economy, bringing together improvement of population's general health and economy growth. However, there is a risk to the population when considering that toxic plants could be ingested, by disinformation or by misidentification.

With regards to the first question, routine consumption of *E. fosbergii* for long periods is a risk to population, since it was found toxic pyrrolizidine alkaloids in this species, including otonecine base, which is considered the most toxic between pyrrolizidine alkaloids. In controlled cultivation experiment, it was possible to see influence of cultivation variables and collect time in amounts of metabolites present in leaves. *Emilia fosbergii* presented higher contents of PAs than *E. sonchifolia*, which increases the risk to population, since there is a misidentification problem in this genus, even with

species from other genera. Still, chemical profiles of both species were highly influenced by flowering, which influenced flavonoids and caffeoylquinic acid derivatives accumulation. In *E. fosbergii*, higher contents of PA were found in plants grown in soil with lower contents of nitrate and nitrite, which could increase toxicity.

To answer the second hypothetical question, this study shows a comparison between wild and domesticated edible species from Asteraceae family. Metabolic fingerprint comparison demonstrates chemical difference mainly between species from different tribes. The use of *B. pilosa*, *Y. japonica*, *G. parviflora* and *A. oleracea* as edible plants could bring taste diversity, new sensations, and a variety of secondary metabolites to population. Species *L. canadensis* and *S. oleraceus* had shown similar chemical profile to *C. endivia* and a few samples of *L. sativa*, which could be an alternative to population with no access to domesticated crop species.

6 REFERENCES

ALI, S.; ZAMEER, S.; YAQOOB, M. Ethnobotanical, phytochemical and pharmacological properties of *Galinsoga parviflora* (Asteraceae): A review. **Tropical Journal of Pharmaceutical Research**, v. 16, n. 12, p. 3023–3033, 2017.

ALPSDAKE. *Bidens pilosa* picture. [s.d.].

ANVISA. **Resolução da Diretoria Colegiada - RDC nº 26, 13 de maio de 2014.** Brasil. Ministério da Saúde, 2014.

Ark of Taste. Disponível em: <https://www.fondazioneslowfood.com/en/ark-of-taste-slow-food/>.

BALDERMANN, S. et al. Critical reviews in plant sciences are neglected plants the food for the future? **Critical Reviews in Plant Sciences**, v. 35, n. 2, p. 106–119, 2016.

BARBOUR, H.; ROBINS, D. J. Structure revision of emiline, a pyrrolizidine alkaloid from *Emilia flammea*. **Phytochemistry**, v. 26, n. 8, p. 2430–2431, 1987.

BERG, R. A. VAN DEN et al. Centering, scaling, and transformations: improving the biological information content of metabolomics data. v. 15, p. 1–15, 2006.

BLACK, M. M. et al. Iron and zinc supplementation promote motor development and exploratory behavior among Bangladeshi infants. **Am J Clin Nutr2**, v. 80, p. 903–910, 2004.

BLAŽENOVI, I. et al. Software Tools and Approaches for Compound Identification of LC-MS / MS Data in Metabolomics. **Metabolites**, v. 8, n. 31, p. 1–23, 2018.

BOHLMANN, F.; KNOLL, K. New acetylenic compounds from *Emilia* species. **Phytochemistry**, v. 17, p. 557–558, 1978.

BORCARD, D.; GILLET, F.; LEGENDRE, P. **Numerical Ecology with R**. [s.l.] Springer, 2018.

BRENTON, A. G.; GODFREY, A. R. Accurate mass measurement: Terminology and treatment of data. **Journal of the American Society for Mass Spectrometry**, v. 21, n. 11, p. 1821–1835, 2010.

CASOTI, R. et al. Topics on metabolomics. In: DINIZ, M. F. F. M. et al. (Eds.). . **Natural products and drug discovery: from pharmacochemistry to pharmacological approaches**. João Pessoa: UFPB, 2018. p. 83–126.

CHAGAS-PAULA, D. A. et al. Outstanding anti-inflammatory potential of selected asteraceae species through the potent dual inhibition of cyclooxygenase-1 and 5-lipoxygenase. **Planta Medica**, v. 81, n. 14, p. 1296–1307, 2015.

CHEN, Z.; HUO, J. Hepatic veno-occlusive disease associated with toxicity of pyrrolizidine alkaloids in herbal preparations. **The Netherlands Journal of Medicine**, v. 68, n. 6, p. 252–260, 2010.

CHENG, Y. et al. Alkylamides of *Acmella oleracea*. v. 1101, p. 6970–6977, 2015.

CHIBLI, L. A. et al. Natural products as inhibitors of Leishmania major dihydroorotate

dehydrogenase. **European Journal of Medicinal Chemistry**, v. 157, p. 852–866, 2018.

COLEY, P. D.; BRYANT, J. P.; CHAPIN, F. S. Resource Availability and Plant Antiherbivore Defense. **Science**, v. 230, n. 4728, p. 895–899, 1985.

CONFORTI, F. et al. The protective ability of Mediterranean dietary plants against the oxidative damage: The role of radical oxygen species in inflammation and the polyphenol, flavonoid and sterol contents. **Food Chemistry**, v. 112, p. 587–594, 2009.

COUTO, V. M. et al. Antinociceptive effect of extract of Emilia sonchifolia in mice. **Journal of Ethnopharmacology**, v. 134, p. 348–353, 2011.

CREEK, D. J. et al. Metabolite identification: are you sure? And how do your peers gauge your confidence? **Metabolomics**, v. 10, n. 3, p. 350–353, 2014.

CULVENOR, C. C. J. et al. Hepato and Pneumotoxicity of pyrrolizidine alkaloids and derivatives in relation to molecular structure. **Chem. Biol. interactions**, v. 12, p. 299–324, 1976.

DA COSTA, F. B.; TERFLOTH, L.; GASTEIGER, J. Sesquiterpene lactone-based classification of three Asteraceae tribes: A study based on self-organizing neural networks applied to chemosystematics. **Phytochemistry**, v. 66, n. 3, p. 345–353, 2005.

ERNST, M. et al. Mass spectrometry in plant metabolomics strategies : from analytical platforms to data acquisition and processing. **Nat. Prod. Rep.**, v. 33, n. di, p. 784–806, 2014.

FAO. **FAOStats**. Disponível em: <http://www.fao.org/faostat/en/#data/QC>. Acesso em: 12 jan. 2020.

FAO et al. **The state of food security and nutrition in the world**. Rome: FAO, 2019.

FIEHN, O. Metabolomics--the link between genotypes and phenotypes. **Plant molecular biology**, v. 48, n. 1–2, p. 155–71, 2002.

FLORES, A. S.; TOZZI, A. M. G. DE A.; TRIGO, J. R. Pyrrolizidine alkaloid profiles in *Crotalaria* species from Brazil : Chemotaxonomic significance. **Biochemical Systematics and Ecology**, v. 37, p. 459–469, 2009.

FUNK, V. A. et al. **Systematics, Evolution, and Biogeography of Compositae**. Viena: International Association for Plant Taxonomy, 2009.

GEORGE, G. K.; KUTTAN, G. Inhibition of pulmonary metastasis by *Emilia sonchifolia* (L.) DC: An in vivo experimental study. **Phytomedicine**, v. 23, p. 123–130, 2016.

GIAMBANELLI, E. et al. Sesquiterpene lactones and inositol 4-hydroxyphenylacetic acid derivatives in wild edible leafy vegetables from Central Italy. **Journal of Food Composition and Analysis**, v. 72, n. June, p. 1–6, 2018a.

GIAMBANELLI, E. et al. Identification and quantification of phenolic compounds in edible wild leafy vegetables by UHPLC / Orbitrap-MS. **J Sci Food Agric**, v. 98, p. 945–954, 2018b.

GOODACRE, R. et al. Metabolomics by numbers : acquiring and understanding global

metabolite data. **Trends in Biotechnology**, v. 22, n. 5, p. 245–252, 2004.

GOODACRE, R. Metabolomics – the way forward. **Metabolomics**, v. 1, n. 1, p. 1–2, 2005.

GRAZIANI, G. et al. Profiling chicory sesquiterpene lactones by high resolution mass spectrometry. **Food Research International**, v. 67, p. 193–198, 2015.

GUARRERA, P. M.; SAVO, V. Perceived health properties of wild and cultivated food plants in local and popular traditions of Italy: A review. **Journal of Ethnopharmacology**, v. 146, n. 3, p. 659–680, 2013.

GUARRERA, P. M.; SAVO, V. Wild food plants used in traditional vegetable mixtures in Italy. **Journal of Ethnopharmacology**, v. 185, p. 202–234, 2016.

HARLAN, J. R. Lettuce and the sycamores: sex and romance in ancient Egypt. **Economic Botany**, v. 40, p. 4–15, 1986.

HARTMANN, T.; OBER, D. Biosynthesis and metabolism of pyrrolizidine alkaloids in plants and specialized insect herbivores. In: LEEPER, F. J.; VEREDAS, J. C. (Eds.). . **Biosynthesis**. Berlim: Springer, 2000. p. 207–239.

HARVEY, A. L. Natural products in drug discovery. **Drug discovery today**, v. 13, n. 19–20, p. 894–901, 2008.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The elements of statistical leraning: Data mining, inference, and prediction**. 2. ed. New York: Springer Science & Business Media, 2009.

HIND, D. J. N. A Checklist of the Brazilian Senecioneae (Compositae). **Kew Bulletin**, v. 48, n. 2, p. 279–295, 1993.

HOL, W. H. G.; VRIELING, K.; VAN VEEN, J. A. Nutrients decrease pyrrolizidine alkaloid concentrations in *Senecio jacobaea*. **New Phytologist**, v. 158, n. 1, p. 175–181, 2003.

HSIEH, C. H. et al. Hepatotoxic pyrrolizidine alkaloids in *Emilia sonchifolia* from Taiwan. **Journal of Food Composition and Analysis**, v. 42, p. 1–7, 2015.

JEFFREY, A. C. What Is Emilia coccinea (Sims) G. Don (Compositae)? A revision of the large-headed *Emilia* species of Africa. **Kew Bulletin**, v. 52, n. 1, p. 205–212, 1997.

KARSTENS, A. J. et al. Associations of the Mediterranean diet with cognitive and neuroimaging phenotypes of dementia in healthy older adults. p. 361–368, 2019.

KINUPP, V. F. **Plantas alimentícias não-convencionais da região de Porto Alegre, RS**. [s.l.] Universidade Federal do Rio Grande do Sul, 2007.

KINUPP, V. F.; BARROS, I. B. I. DE. Teores de proteína e minerais de espécies nativas, potenciais hortaliças e frutas. **Ciênc. e Tecnol. de Aliment.**, v. 28, n. 4, p. 846–857, 2008.

KIRK, H. et al. Species by environment interactions affect pyrrolizidine alkaloid expression in *Senecio jacobaea*, *Senecio aquaticus*, and their hybrids. **Journal of Chemical Ecology**, v. 36, p. 378–387, 2010.

KOPKA, J. et al. Metabolite profiling in plant biology : platforms and destinations. **Genome biology**, v. 5, n. 6, p. 1–9, 2004.

KUMAR, D. G. et al. Traditional uses, phytochemical and pharmacological aspects of *Emilia sonchifolia* (L.) DC. **International Journal of Research in Ayurveda and Pharmacy**, v. 6, n. 4, p. 551–556, 2015.

KUNKEL, G. **Plants for human consumption**. Koenigstein: Koeltz Scientific Books, 1984.

LEITÃO, F. et al. Medicinal plants traded in the open-air markets in the State of Rio de Janeiro, Brazil: an overview on their botanical diversity and toxicological potential. **Brazilian Journal of Pharmacognosy**, v. 24, p. 225–247, 2014.

LENTINI, F.; VENZA, F. Wild food plants of popular use in Sicily. v. 12, p. 1–12, 2007.

LICATA, M. et al. A survey of wild plant species for food use in Sicily (Italy) – results of a 3-year study in four Regional Parks. **Journal of Ethnobiology and Ethnomedicine**, 2016.

LISOWSKY, S. Le genre *Emilia* (Asteraceae, Senecioneae) en Afrique Centrale (Congo, Zaire, Rwanda, Burundi). **Polish Botanical Studies**, v. 1, p. 67–116, 1990.

LIU, L. et al. Exposure to famine in early life and the risk of obesity in adulthood in Qingdao: evidence from the 1959-1961 Chinese Famine. **Nutrition, Metabolism and Cardiovascular Diseases**, v. 27, n. 2, p. 154–160, 2017.

LÓPEZ-BUCIO, J. et al. Organic acid metabolism in plants : from adaptive physiology to transgenic varieties for cultivation in extreme soils. **Plant Science**, v. 160, p. 1–13, 2000.

LORENZI, H.; KINUPP, V. F. **Plantas alimentícias não convencionais (PANC) no Brasil**. Nova Odessa: Plantarum, 2014.

MARTINS, D. et al. Nutritional and in vitro antioxidant properties of edible wild greens in Iberian Peninsula traditional diet. **Food Chemistry**, v. 125, n. 2, p. 488–494, 2011.

MATTOCKS, A. R. Toxicity of pyrrolizidine alkaloids. **Nature**, v. 217, p. 723–728, 1968.

MICHALSKA, K.; SZNELER, E.; KISIEL, W. Phytochemistry Sesquiterpene lactones from *Lactuca canadensis* and their chemotaxonomic significance. **Phytochemistry**, v. 90, p. 90–94, 2013.

MORAES, A. P.; GUERRA, M. Cytological differentiation between the two subgenomes of the tetraploid *Emilia fosbergii* Nicolson and its relationship with *E. sonchifolia* (L.) DC . (Asteraceae). **Plant Syst Evol**, v. 287, p. 113–118, 2010.

MOREIRA, R.; PEREIRA, D. M.; ANDRADE, P. B. Pyrrolizidine alkaloids: chemistry, pharmacology, toxicology and food safety. **International journal of molecular sciences**, v. 19, n. 1668, p. 1–22, 2018.

MROCZEK, T. et al. On-line structure characterization of pyrrolizidine alkaloids in *Onosma stellulatum* and *Emilia coccinea* by liquid chromatography-ion-trap mass spectrometry. **Journal of Chromatography A**, v. 1056, n. 1- 2 SPEC.ISS., p. 91–97, 2004.

MUKO, K. N.; OHIRI, F. C. U. A preliminary study on the anti-inflammatory properties of *Emilia sonchifolia* leaf extracts. **Fitoterapia**, v. 71, p. 65–68, 2000.

NEWMAN, D. J.; CRAGG, G. M. Natural products as sources of new drugs over the last 25 years. **Journal of Natural Products**, v. 70, n. 3, p. 461–477, 2007.

NICOLSON, D. H. Summary of cytological information on *Emilia* and the taxonomy of four Pacific taxa of *Emilia* (Asteraceae: Senecioneae). **Systematic Botany**, v. 5, n. 4, p. 391–407, 1980.

NORDENSTAM, B. et al. Senecioneae. In: FUNK, V. A. et al. (Eds.). . **Systematics, evolution and biogeography of Compositae**. Vienna: International Association for Plant Taxonomy, 2009. p. 503–525.

PADILLA-GONZÁLEZ, G. F.; DIAZGRANADOS, M.; COSTA, F. B. DA. Biogeography shaped the metabolome of the genus Espeletia: a phytochemical perspective on an Andean adaptive radiation. **Scientific Reports**, v. 7, n. 1, p. 1–11, 2017.

PEISINO, C. M. O. et al. Health - Promoting properties of brazilian unconventional food plants. **Waste and Biomass Valorization**, v. 1, p. 1–10, 2019.

PELSER, P. B. et al. Frequent gain and loss of pyrrolizidine alkaloids in the evolution of *Senecio* section *Jacobaea* (Asteraceae). **Phytochemistry**, v. 66, p. 1285–1295, 2005.

PEPE, G. et al. Evaluation of anti-inflammatory activity and fast UHPLC – DAD – IT-TOF profiling of polyphenolic compounds extracted from green lettuce (*Lactuca sativa* L .; var . Maravilla de Verano). **Food Chemistry**, v. 167, p. 153–161, 2015.

PETER, A. M.; AMERONGEN, A. VAN. Sesquiterpene lactones in chicory (*Cichorium intybus* L.) Distribution in chicons and effect of storage. **Food Research International**, v. 29, n. 5–6, p. 439–444, 1996.

PRICE, K. R. et al. Relationship between the chemical and sensory properties of exotic salad crops - coloured lettuce (*Lactuca sativa*) and chicory (*Cichorium intybus*). **J Sci Food Agric**, v. 53, p. 185–192, 1990.

PSALTOPOULOU, T. et al. Mediterranean diet and stroke, cognitive impairment, depression: a meta-analysis. **Annals of neurology**, v. 74, n. 4, p. 580–91, 2013.

RAJ, M. Natural antioxidant (flavone glycoside) from *Emilia sonchifolia* DC. and its potential activity. **International Journal of Pharmaceutical Science**, v. 4, n. 3, p. 159–162, 2012.

RENNA, M. et al. Elemental characterization of wild edible plants from countryside and urban areas. **Food Chemistry**, v. 177, p. 29–36, 2015.

RICARDO, L. M. et al. Plants from the Brazilian traditional medicine: Species from the books of the polish physician Piotr Czerniewicz (Pedro Luiz Napoleão Chernoviz, 1812–1881). **Brazilian Journal of Pharmacognosy**, v. 27, n. 3, p. 388–400, 2017.

ROCHFORT, S. Biology and implications for natural products research. **J. Nat. Prod.**, v. 68, p. 1813–1820, 2005.

RONDANELLI, M. et al. *Acmella oleraceae* for pain management. **Fitoterapia**, p.

104419, 2019.

ROSATO, V. et al. Mediterranean diet and cardiovascular disease: a systematic review and meta-analysis of observational studies. **European Journal of Nutrition**, v. 58, n. 1, p. 173–191, 2019.

ROSE, H. ***Emilia sonchifolia*** **picture**. Disponível em: <https://commons.wikimedia.org/wiki/File:Emilia_sonchifolia_plant6_(14047479524).jpg>.

RUAN, J. et al. Characteristic ion clusters as determinants for the identification of pyrrolizidine alkaloid N-oxides in pyrrolizidine alkaloid-containing natural products using HPLC-MS analysis. **Journal of Mass Spectrometry**, v. 47, n. 3, p. 331–337, 2012.

RYDER, E. J.; WHITAKER, T. W. Lettuce. In: SMARTT, J.; SIMMONDS, N. W. **Evolution of Crop Plants**. Essex: Longman Scientific and Technical, 1995. p. 53–56.

SAMPAIO, B. L.; EDRADA-EBEL, R.; DA COSTA, F. B. Effect of the environment on the secondary metabolic profile of Tithonia diversifolia: A model for environmental metabolomics of plants. **Scientific Reports**, v. 6, p. 1–11, 2016.

SAREEDENCHAI, V.; ZIDORN, C. Flavonoids as chemosystematic markers in the tribe Cichorieae of the Asteraceae q. **Biochemical Systematics and Ecology**, v. 38, n. 5, p. 935–957, 2010.

SAVO, V. et al. Traditional salads and soups with wild plants as a source of antioxidants : a comparative chemical analysis of five species growing in central Italy. v. 2019, 2019.

SCHULP, C. J. E.; THUILLER, W.; VERBURG, P. H. Wild food in Europe : A synthesis of knowledge and data of terrestrial wild food as an ecosystem service. **Ecological Economics**, v. 105, p. 292–305, 2014.

SCHWINGSHACKL, L. et al. Adherence to mediterranean diet and risk of cancer: An updated systematic review and meta-analysis. **Nutrients**, v. 9, n. 10, p. 1063, 2017.

SCODITTI, E. et al. Mediterranean diet polyphenols reduce inflammatory angiogenesis through MMP-9 and COX-2 inhibition in human vascular endothelial cells : A potentially protective mechanism in atherosclerotic vascular disease and cancer. **Archives of Biochemistry and Biophysics**, v. 527, n. 2, p. 81–89, 2012.

SERRA-MAJEM, L. et al. Molecular Aspects of Medicine Benefits of the Mediterranean diet : Epidemiological and molecular aspects. **Molecular Aspects of Medicine**, v. 67, n. April, p. 1–55, 2019.

SHEN, S. et al. A new cyclohexylacetic acid derivative from the aerial parts of *Emilia sonchifolia*. **Natural Product Research**, v. 27, n. 15, p. 1330–1334, 2013.

SHIN, T. et al. Traditional knowledge of wild edible plants with special emphasis on medicinal uses in Southern Shan State, Myanmar. **Journal of Ethnobiology and Ethnomedicine**, v. 14, n. 48, p. 1–13, 2018.

SHULHA, O.; ZIDORN, C. Phytochemistry Sesquiterpene lactones and their precursors as chemosystematic markers in the tribe Cichorieae of the Asteraceae

revisited: An update (2008 – 2017). **Phytochemistry**, v. 163, n. March, p. 149–177, 2019.

SHYLESH, B. S.; PADIKKALA, J. In vitro cytotoxic and antitumor property of *Emilia sonchifolia* (L.) DC in mice. **Journal of Ethnopharmacology**, v. 73, p. 495–500, 2000.

SILVA, C. S. et al. Manual de análises químicas de solos, plantas e fertilizantes. In: **EMBRAPA**. [s.l: s.n.].

SILVA, F. L. et al. Compilation of Secondary Metabolites from *Bidens pilosa* L. **Molecules**, v. 16, p. 1070–1102, 2011.

SILVA, L. F. L. et al. Nutritional Evaluation of Non-Conventional Vegetables in Brazil. **Annals of the Brazilian Academy of Sciences**, v. 90, n. 2, p. 1775–1787, 2018.

SINGH, B. et al. Association of Mediterranean diet with mild cognitive impairment and Alzheimer's disease: A systematic review and meta-analysis. **Journal of Alzheimer's Disease**, v. 39, n. 2, p. 271–282, 2014.

SPRING, O. Chemotaxonomy based on metabolites from glandular trichomes. **Advances in Botanical Research**, v. 31, n. 4, p. 153–169, 2000.

SRINIVASAN, K. K.; SUBRAMANIAN, S. S. Chemical investigation of *Emilia sonchifolia*. **Fitoterapia**, v. 51, n. 5, p. 241–243, 1980.

STARR, F. & K. ***Emilia fosbergiii*** **picture**. Disponível em: <https://commons.wikimedia.org/wiki/File:Starr_040330-0145_Emilia_fosbergii.jpg>.

SUMNER, L. W. et al. Proposed minimum reporting standards for chemical analysis. **Metabolomics**, v. 3, p. 211–221, 2007.

THE PLANT LIST. **The Plant List**. Disponível em: <www.theplantlist.org>. Acesso em: 12 set. 2019.

TRICHOPOULOU, A.; LAGIOU, P.; KUPER, H. Cancer and Mediterranean Dietary Traditions. v. 9, n. September, p. 869–873, 2000.

TRIGO, J. R. et al. Chemotaxonomic value of pyrrolizidine alkaloids in southern Brazil Senecio (Senecioneae : Asteraceae). **Biochemical Systematics and EcologySystematics and**, v. 31, p. 1011–1022, 2003.

TU, Y. The discovery of artemisinin (qinghaosu) and gifts from Chinese medicine. **Nature Medicine**, v. 17, n. 10, p. 1217–1220, 2011.

TU, Y. Y. et al. Studies on the Constituents of *Artemisia annua* Part II. **Planta Medica**, v. 44, p. 143–145, 1982.

UNESCO. **Convention for the safeguarding of the intangible cultural heritage**, 2010.

VECCHIA, C. LA. Association between Mediterranean dietary patterns and cancer risk. **Nutrition Reviews**, v. 67, n. 1, p. 126–129, 2009.

VEENA, S. R. et al. Association between maternal nutritional status in pregnancy and offspring cognitive function during childhood and adolescence; a systematic review. **BMC Pregnancy and Childbirth**, v. 16, n. 1, p. 220, 2016.

VILLAS-BÔAS, S. G. et al. Mass spectrometry in metabolome analysis. **Mass Spectrometry Reviews**, v. 24, n. 5, p. 613–646, 2004.

VILLAS-BÔAS, S. G.; RASMUSSEN, S.; LANE, G. A. Metabolomics or metabolite profiles? **Trends in Biotechnology**, v. 23, n. 8, p. 385–386, 2005.

VILLENA-ESPONERA, M. P.; MORENO-ROJAS, R.; MOLINA-RECIO, G. Food Insecurity and the Double Burden of Malnutrition of Indigenous Refugee Épera Siapidara. **Journal of Immigrant and Minority Health**, v. 21, p. 1035–1042, 2019.

VORSTENBOSCH, T. et al. Famine food of vegetal origin consumed in the Netherlands during World War II. p. 1–15, 2017.

VRIELING, K.; VAN WIJK, C. A. M. Cost assessment of the production of pyrrolizidine alkaloids in ragwort (*Senecio jacobaea* L.). **Oecologia**, v. 97, p. 541–546, 1994.

WHO. **Pyrrolizidine alkaloids health and safety guide**. Genova: World Health Organization, 1989.

WILDFEUER. ***Sonchus oleraceus* picture**. Disponível em: <https://upload.wikimedia.org/wikipedia/commons/4/4e/2006-11-16Sonchus_oleraceus03-03.jpg>.

WINK, M. Evolution of secondary metabolites from an ecological and molecular phylogenetic perspective. **Phytochemistry**, v. 64, p. 3–19, 2003.

WU, L.-C. et al. Anti-inflammatory effect of spilanthol from *Spilanthes acmella* on murine macrophage by down-regulating LPS-induced inflammatory mediators. **Journal of Agricultural and Food Chemistry**, v. 56, p. 2341–2349, 2008.

YAE, E. et al. Studies on the constituents of whole plants of *Youngia japonica*. **Chem. Pharm. Bull.**, v. 57, n. 7, p. 719–723, 2009.

ZELL, H. ***Acmella oleracea* picture**. Disponível em <https://upload.wikimedia.org/wikipedia/commons/f/f6/Acmella_oleracea_003.JPG>.

ZIDORN, C. Phytochemistry sesquiterpene lactones and their precursors as chemosystematic markers in the tribe Cichorieae of the Asteraceae. v. 69, p. 2270–2296, 2008.

7 APPENDIX

Appendix A – Soil composition parameters

Table A1 - Soil composition.

Samples	pH	O.M.[1]	Ca[2]	Mg[2]	K[2]	SO_4[3]	B[3]	Cu[3]	Fe[3]	Mn[3]	Zn[3]	Na[2]	Si[4]	Total Nitrogen[4]	Amonium[3]	Nitrate[3]	Nitrite[3]
A_1	7.5	15	69	5	1.7	63	0.12	1.5	4	2.9	0.4	0.1	3	603	3.9	4.7	0.1
A_2	7.5	15	67	5	1.5	62	0.01	1.6	4	2.8	0.5	0.1	4	642	7.1	4.3	0.1
A_3	7.4	15	66	5	1.5	68	0.12	1.6	3	2.6	0.4	0.1	3	666	6.4	2.8	0.1
A_4	7.5	15	69	5	1.8	69	0.14	1.8	4	3.2	0.4	0.1	3	564	6.7	4.3	0.1
B_1	7.2	31	72	17	18.7	87	0.24	2	16	6.1	2.1	0.7	7	1400	4.9	143.2	3.90
B_2	7.2	30	73	16	21.4	76	0.23	1.9	13	5.6	1.9	0.8	7	1365	4.6	140.4	2.10
B_3	7.1	36	74	19	24.1	77	0.24	2	17	6.7	2.3	0.9	7	1429	6.4	153.9	1.70
B_4	7.2	32	70	22	22.5	80	0.25	2.1	18	6.6	2.4	0.9	7	1471	9.6	152.5	2.50
C_1	7.9	23	110	29	63.6	251	1.13	3.7	21	13.2	11.5	8.4	6	1957	78.7	88.6	0.1
C_2	7.9	25	110	30	63	257	1.21	3.6	22	12.3	11.4	8.7	6	1968	83.7	90.1	0.1
C_3	7.9	25	104	31	65.7	273	1.25	4	26	14	13.7	8.9	7	2227	79.8	88.6	0.1
C_4	7.8	24	100	29	62.7	254	1.24	3.6	23	12.3	11.3	7.9	6	2085	85.5	82.9	0.1

[1] Soil Organic Matter (g/dm³), related to soil texture, bulk density, microbial biomass and water holding capacity; Units: [2] mmol/dm³; [3] mg/dm³; [4] mg/kg. Plants were cultivated under three different soil conditions: A) non-treated soil; B) soil treated with horse manure; C) soil treated with chicken manure. Four replicates of each treatment were collected.

Appendix B – MZmine2 preprocessing parameters

B1- Experiment 1 parameters

1. **Scan by scan filtering:**
   - Filter: Savitzky-Golay
   - Number of datapoints: 5
2. **Mass detection:**
   - MS level 1
   - Mass detector: Exact mass
   - Noise level at 1.0E6 *
3. **Baseline correction*:**
   - Smoothing: 10^5
   - Asymmetry: 0.1
   - *m/z* bin 1.0
4. **FTMS shoulder peaks:**
   - Mass resolution of 70,000
   - Peak model function: Lorentzian extended
5. **Chromatogram builder:**
   - Min height: 5.0E6 *
   - Min time spam: 0.1 min
   - *m/z* tolerance: 0.001 *m/z* or 5.0 ppm
6. **Chromatogram deconvolution:**
   - Chromatographic threshold: 0.5
   - Minimum RT range: 0.1 min
   - Minimum height: 15 %
   - Min ratio of peak/edge: 5
   - Peak duration range: 0.10 – 4 min
7. **Isotopic peak group:**
   - Retention time tolerance: 0.9 min
   - *m/z* tolerance at 0.001 *m/z* or 5.0 ppm
   - Maximum charge: 2
   - Representative isotope: Most intense
8. **Join aligner:**
   - *m/z* tolerance 0.001 *m/z* or 5.0 ppm
   - Weight for *m/z* 15

Retention time tolerance 0.9 min

Weight for RT 10

Compare isotope pattern Setup

Isotope *m/z* tolerance: 0.001 *m/z* or 5.0 ppm

Minimum absolute intensity: 5.0E6

Minimum score: 65.0%

9. **Same RT and *m/z* range gap filler**

   m/z tolerance 0.001 *m/z* or 5.0 ppm

10. **Fragment search:**

    Retention time tolerance of 0.9 min

    Max fragment peak height 50.0%

    Minimum MS2 peak height at 5.0E6 *

    m/z tolerance at 0.001 *m/z* or 5.0 ppm

11. **Adduct search:**

    Retention time tolerance of 0.9 min

    m/z tolerance at 0.001 *m/z* or 5.0 ppm

    Max relative adducts peak height 50.0%

12. **Complex search:**

    Retention time tolerance of 0.9 min

    m/z tolerance at 0.001 *m/z* or 5.0 ppm

    Max complex peak height 50.0%

13. **Custom database search:**

    Retention time tolerance of 0.9 min

    m/z tolerance at 0.003 *m/z* or 5.0 ppm

B2 - Experiment 2 used same parameters described above, except for marked items (*), which parameters are described as following:

1. **Mass detection:** Noise level at 1.5E7
2. **Baseline correction:** Smoothing: 10^5; Asymmetry: 0.1; *m/z* bin 1.0
3. **Chromatogram builder:** Min height: 1.5E7
4. **Join aligner:** Minimum absolute intensity: 5.0E7
5. **Fragment search:** Minimum MS2 peak height at 5.0E7

Appendix C – R scripts

C1 - Script for canonical analysis

```
essoil<- read.csv("es_soil.csv", sep=";")          # read soil spreadsheet

rownames(essoil) <- essoil[,1]                     # exclude samples name

auto.escaled <- scale(essoil[,-c(1)], center=T, scale=T)          # normalize, if necessary

teste.env<- data.frame(auto.escaled)    # To transform spreadsheet in data frame (properties of matrices and of lists)

teste <- read.csv("es_metabolomics.csv", h=T, sep=";")          #read metabolomics spreadsheet

rownames(teste) <- teste[,1]                # exclude samples name

colvec2<- c("deeppink", "gold", "green", "blue", "black")

with(teste, levels(Groups))                # color related to groups. Groups is the row name

library(vegan)

small.cca <- cca(teste2~., teste.env, na.omit)          # canonical analysis

plot(small.cca, type= "n", scaling= 3)

points(small.cca, display="spec", pch= 3, col= "gray")          # to plot loadings

with(teste,points(small.cca,display="sites",col=colvec2[Groups2],pch=21,bg=colvec2[Groups2]))                # to plot scores with color based on groups

text(small.cca, display= "bp", col= "darkblue")          # to plot soil information as arrows

with(teste, legend("topright", legend= levels(Groups), bty= "n", col= colvec, pch= 21, pt.bg= colvec))                # shows legend

summary(small.cca)                # show model statistical parameters

# "sites" related to samples/scores; "spec" related to loadings or mass peak area
```

C2- Script for heatmap

```
library("gplots")

data<- t(read.csv("heatmap_somentepos2.csv",header = T, row.names=1, sep=";"))
        #read spreadsheet

centered<-scale(data,center=TRUE, scale=FALSE)          # centering

pareto.matrix<- apply(centered,1,function(x) x/sqrt(sd(x)))    # pareto scaling

hv3<-heatmap.2 (t(pareto.matrix), scale="row",col=bluered(50), dendrogram="both", Colv= TRUE, Rowv= FALSE, distfun= dist, reorderfun=function(d, w) reorder(d, w, agglo.FUN = mean), hclustfun=hclust, symkey=TRUE, key=TRUE, keysize=1, density.info="none", trace="none")                    # heatmap
```

Appendix D– Location of collect

Table D1 - Location of collect.

Local	Farm name	City	Geographic coordinates
A	“Augusto”	Ribeirão Preto urban area	S 21º 10’ 44” W 47º 51’ 20”
B	“Mário Lago” settlement community	Ribeirão Preto rural area	S 21º 8’ 8” W 47º 41’ 56”
C	“Santa Fé” Community-supported agriculture model	Cravinhos rural area	S 21º 19’ 48” W 47º 43’ 12”
D	“Santa Esília”	Bonfim Paulista rural area	S 21º 17’ 27” W 47º 48’ 45”
E	“Biagio”	Ribeirão Preto urban area	S 21º 19’ 48” W 47º 43’ 12”
F	“São Domingos”	Jardinópolis rural area	S 21º 11’ 31” W 47º 52’ 01”

Appendix E – Samples description and source

Table E1 - Location and description of harvested samples.

Total of samples	Species/variety	Popular name in Brazil	Harvest location
1	*Acmella oleracea*	jambu	D
8	*Bidens pilosa*	picão preto	A, B, C, D, F
5	*Cichorium endivia*	chicória	A, B, C, D, F
9	*Cichorium intybus* 1	almeirão	A-F
2	*Cichorium intybus* 2	catalônia	A, C
1	*Cichorium intybus* 3	pão de açúcar	F
5	*Galinsoga parviflora*	fazendeiro	A, C, E, F
2	*Lactuca canadensis*	almeirão roxo	C, D
2	*Lactuca sativa* var. *capitata* 1	alface americana	C, F
2	*Lactuca sativa* var*. capitata 2*	alface mimosa	A, F
5	*Lactuca sativa* var. *crispa* 1	alface crespa	A, B, C, D, F
3	*Lactuca sativa* var. *crispa* 2	alface roxa	C, F
9	*Sonchus oleraceus*	serralha	A-F
3	*Youngia japonica*	youngia, crepe	B, D, F

Appendix F – NMR spectrums of compound 1 (emiline)

Figure F1 - ^{1}H spectra of emiline ($CDCl_3$).

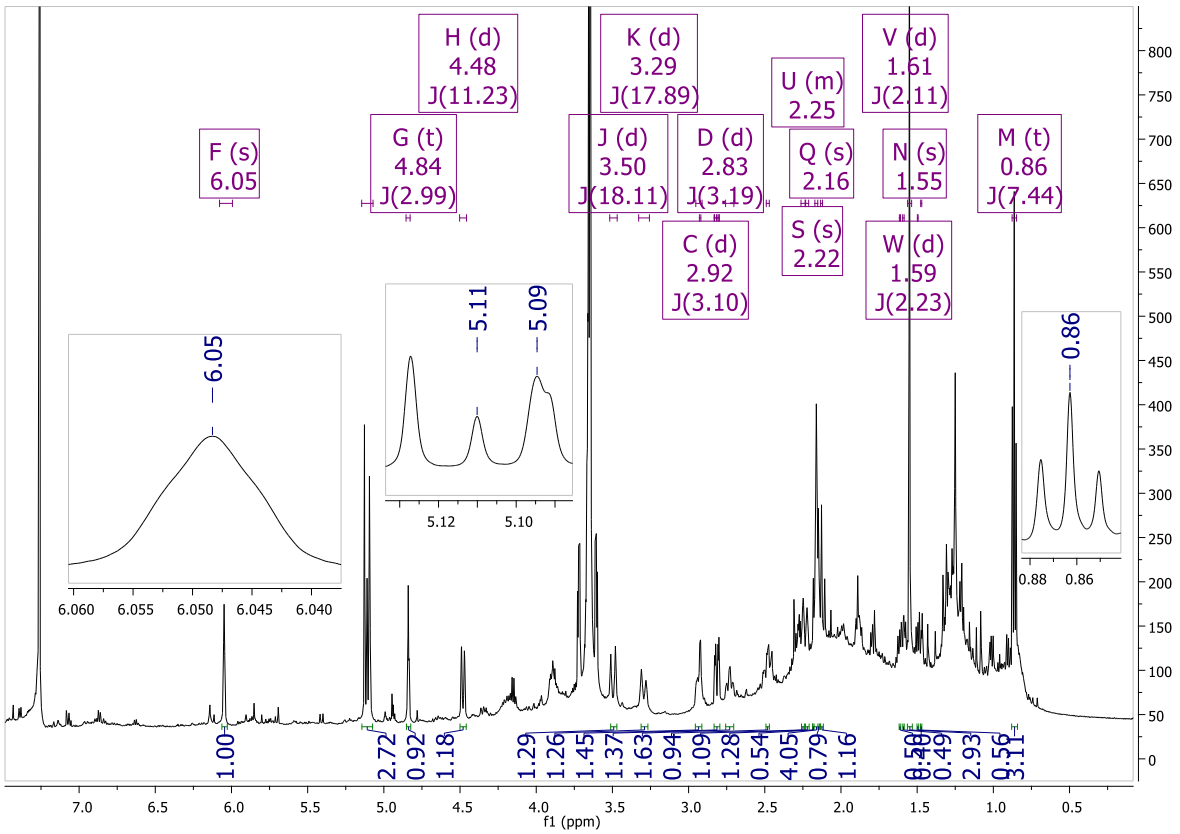


Figure F2 – DEPTQ spectra of emiline ($CDCl_3$).

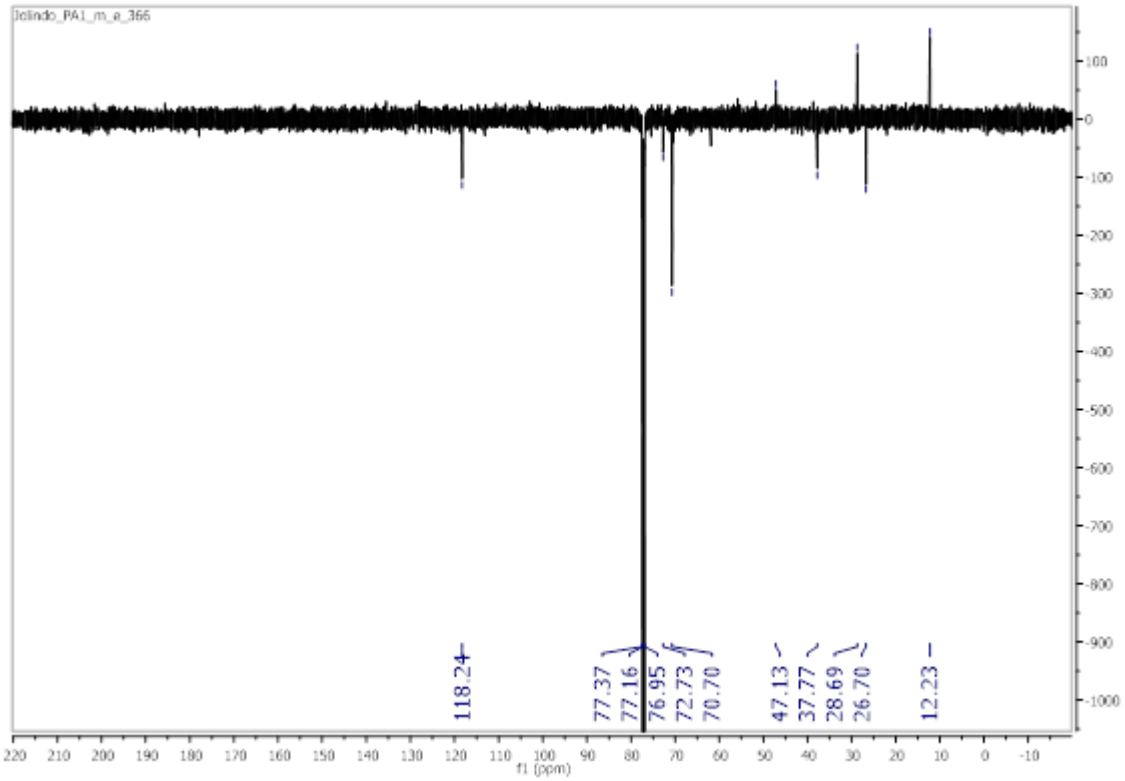


Figure F3 - HSQC spectra of emiline ($CDCl_3$).

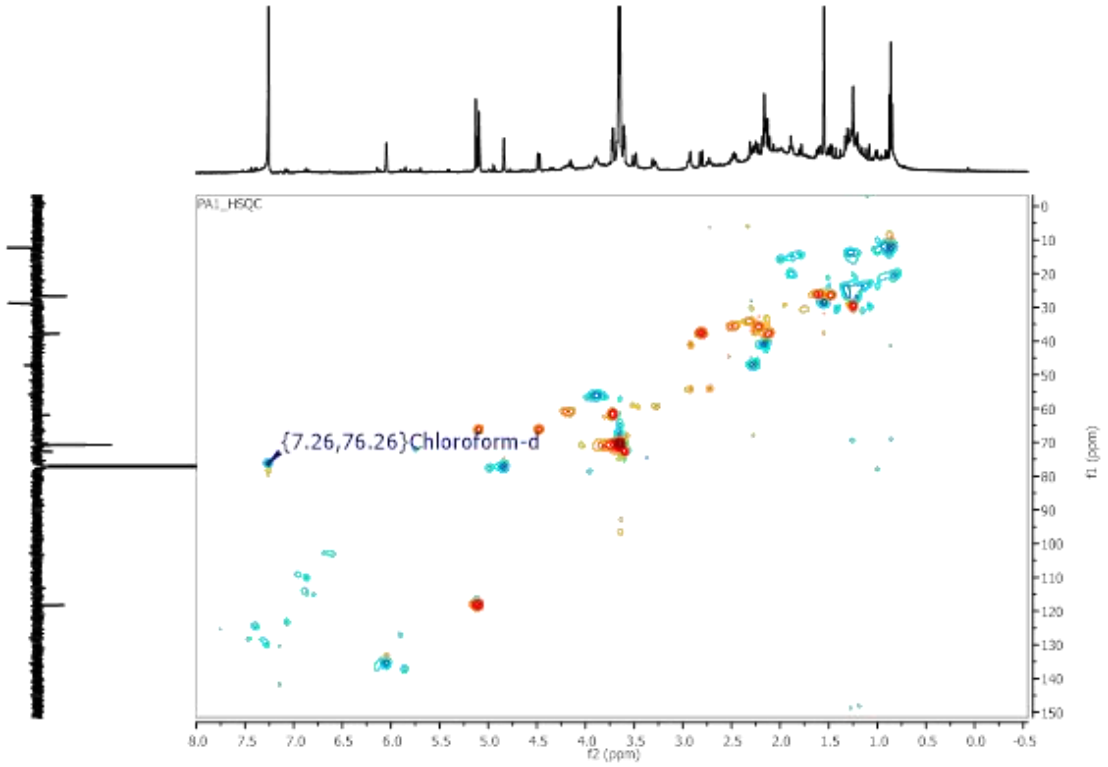


Figure F4 - HMBC spectra of emiline ($CDCl_3$).

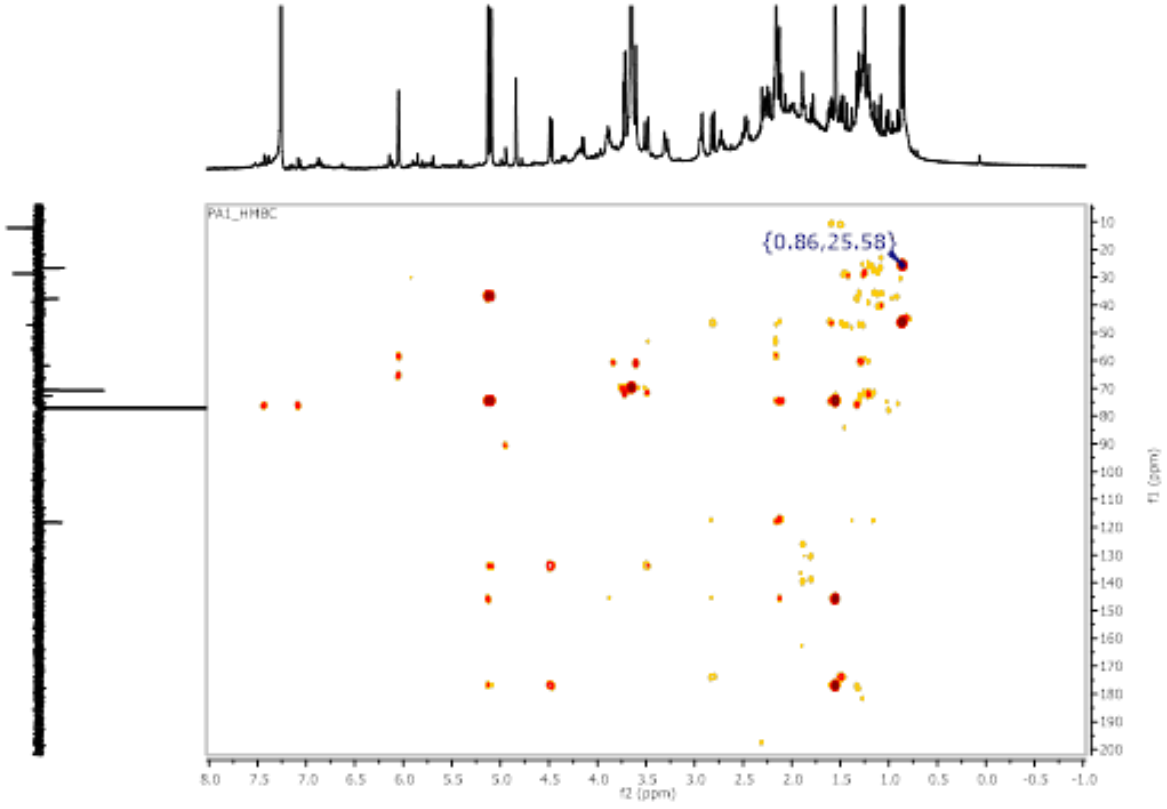


Figure F5 - TOCSY 1D spectra of emiline ($CDCl_3$). Selective excitation performed at δ 6.05 ppm.

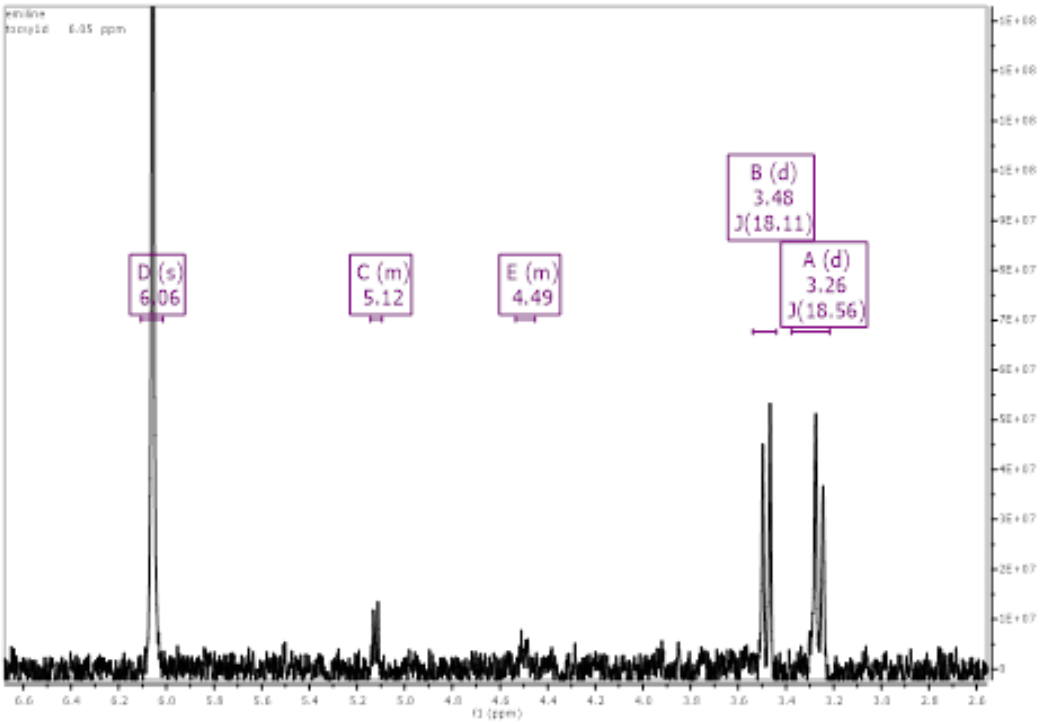
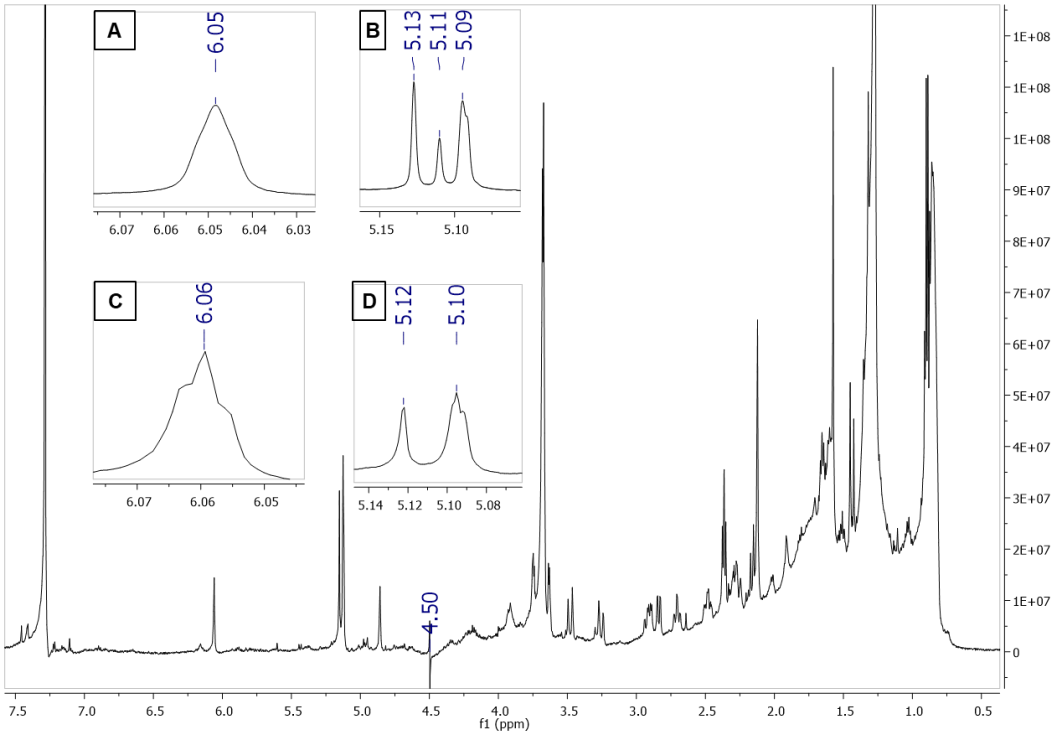


Figure F6 – Homonuclear decoupling ^1H spectra. Irradiation signal at δ 4.48 ppm.



A and **B**- Multiplet structures of chemical shifts at δ 6.05 and δ 5.10 ppm in ^1H original spectra. **C** and **D**- Spectra demonstrate residual multiplet structures of chemical shifts at δ 6.05 and δ 5.10 ppm in Homonuclear decoupling spectrum irradiated at δ 4.48 ppm.

Figure F7 - TOCSY 1D spectra of emiline ($CDCl_3$). Selective excitation performed at δ 2.84 ppm

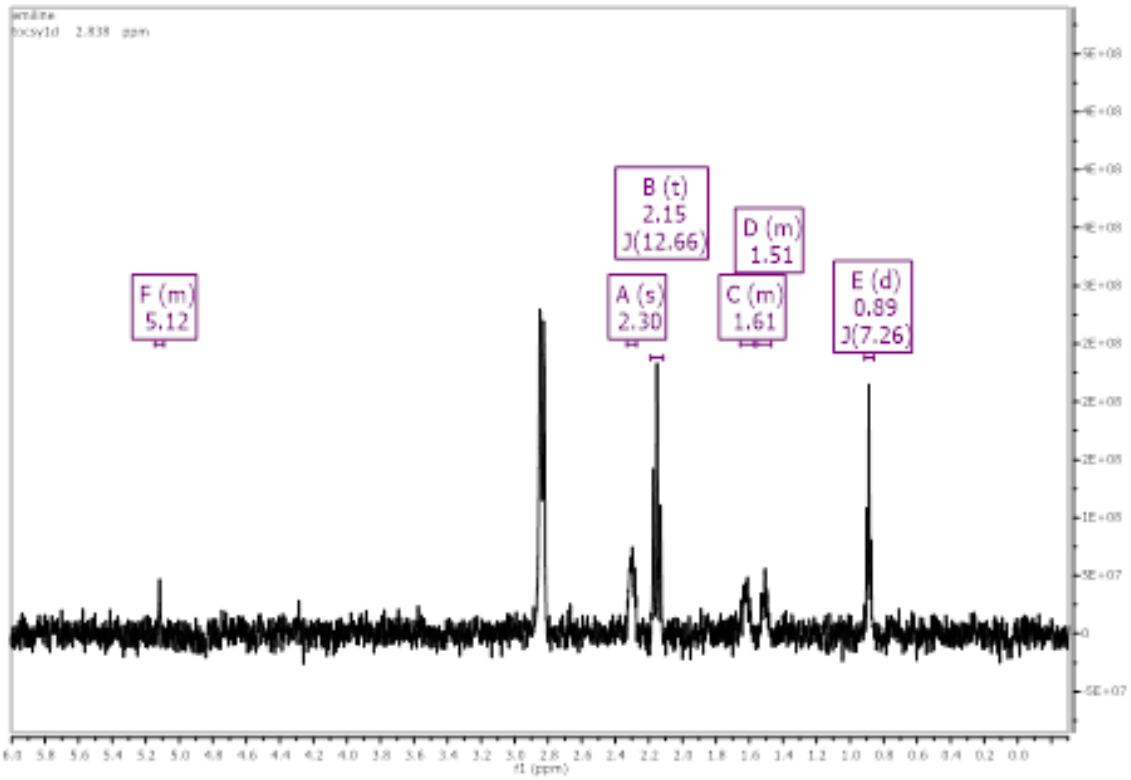
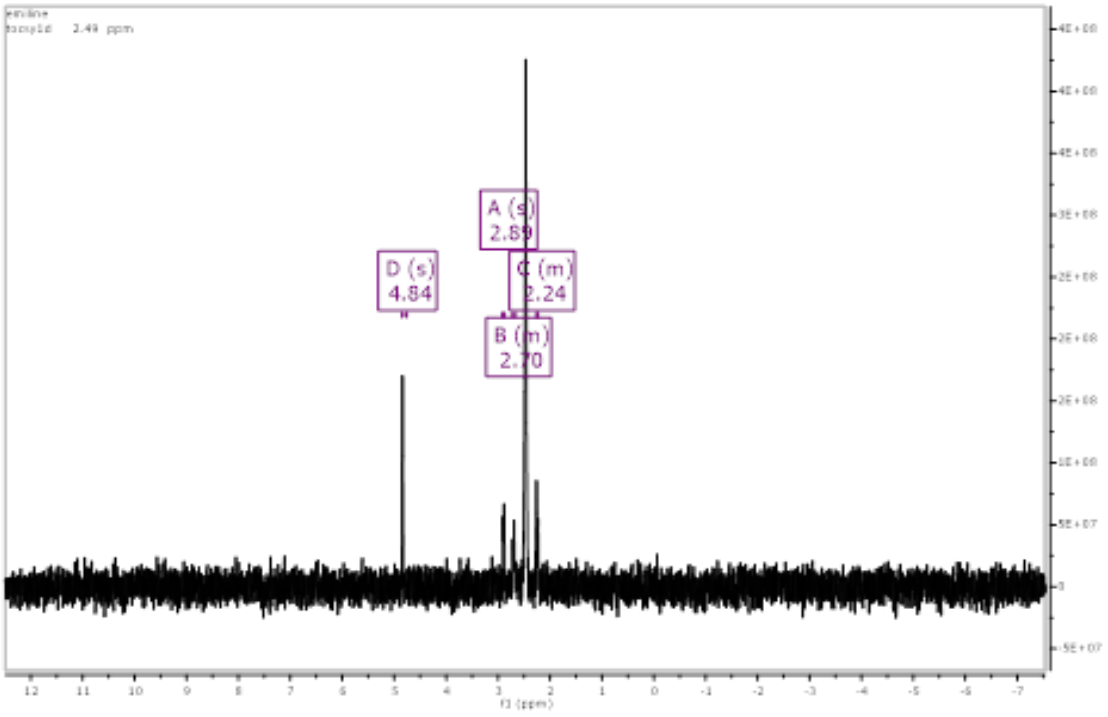
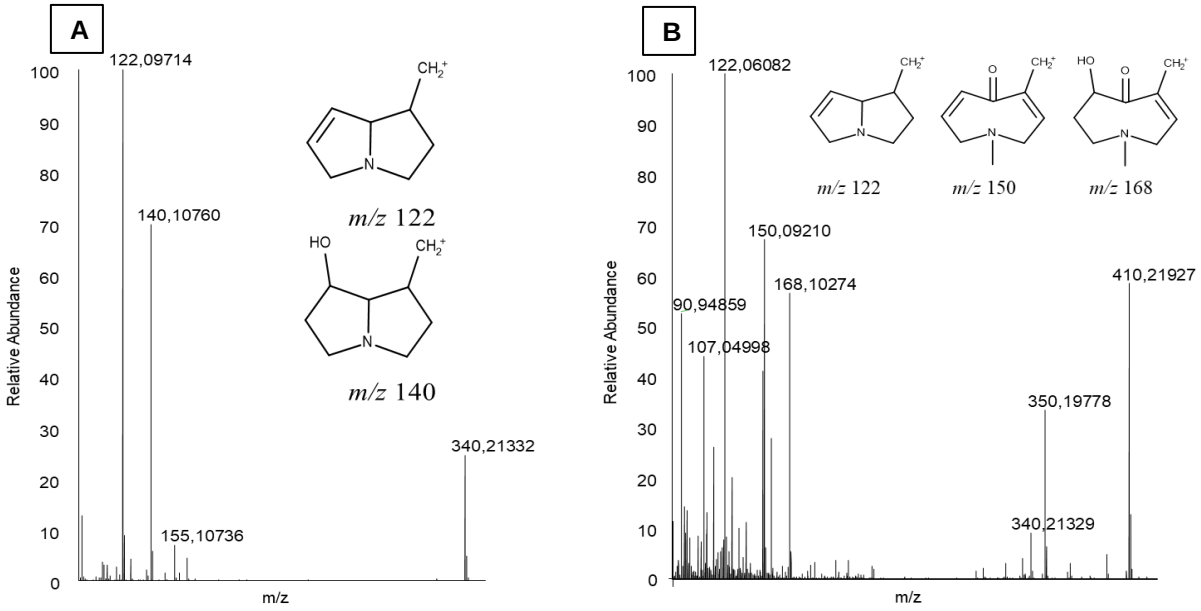


Figure F8 - TOCSY 1D spectra of emiline ($CDCl_3$). Selective excitation performed at δ 2.49 ppm.



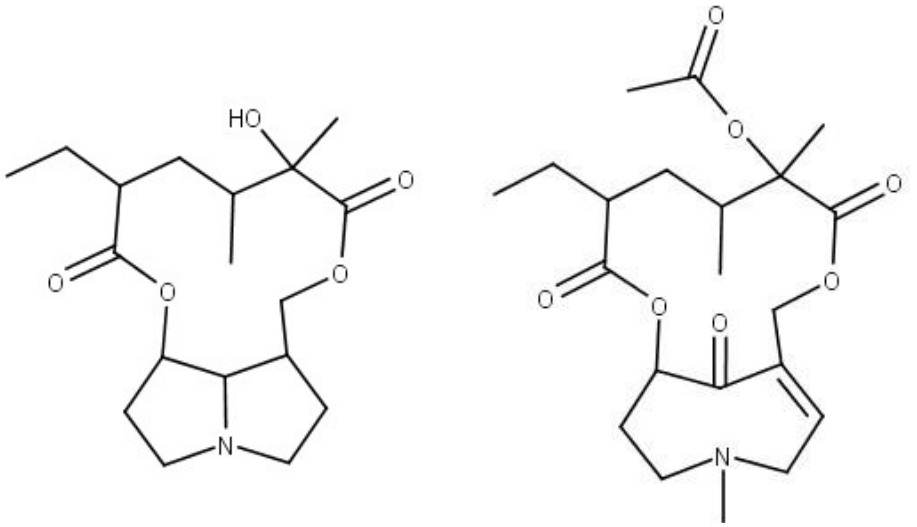
Appendix G – Putative pyrrolizidine alkaloids

Figure G1 - MS2 spectra of putative pyrrolizidine alkaloids



A- MS2 spectra of metabolite 5, with *m/z* $[M+H]^+$ 340.2113; B- MS2 spectra of metabolite 8, with *m/z* $[M+H]^+$ 410.2162. Structure of pyrrolizidine alkaloids diagnostic fragments are based on HSIEH et al. (2015) and RUAN et al. (2012)

Figure G2 – Proposed structures to metabolites 5 and 8.

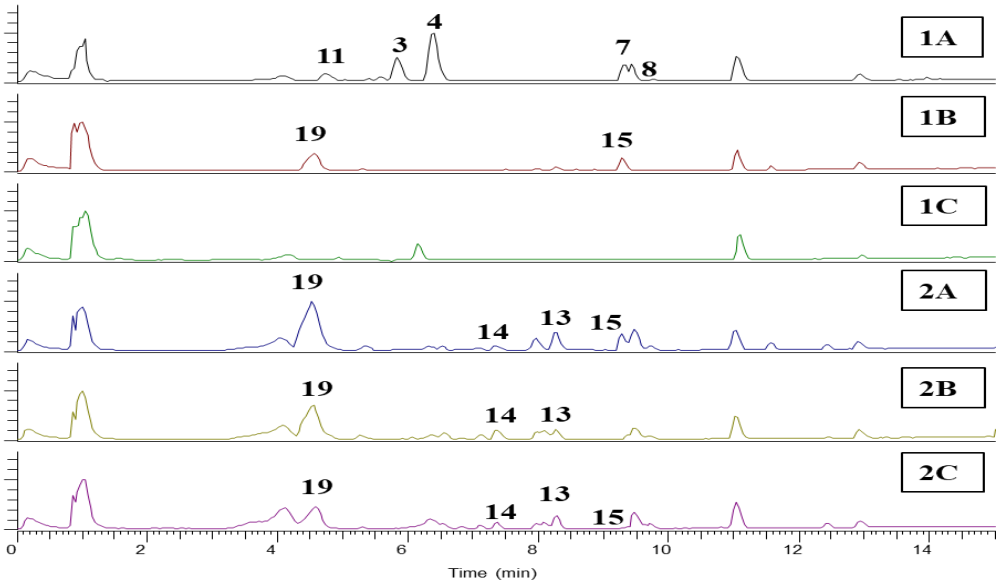


m/z $[M+H]^+$ = 340
5 – hydrogenated platyphylline

m/z $[M+H]^+$ = 410
8 – hydrogenated ligularidine

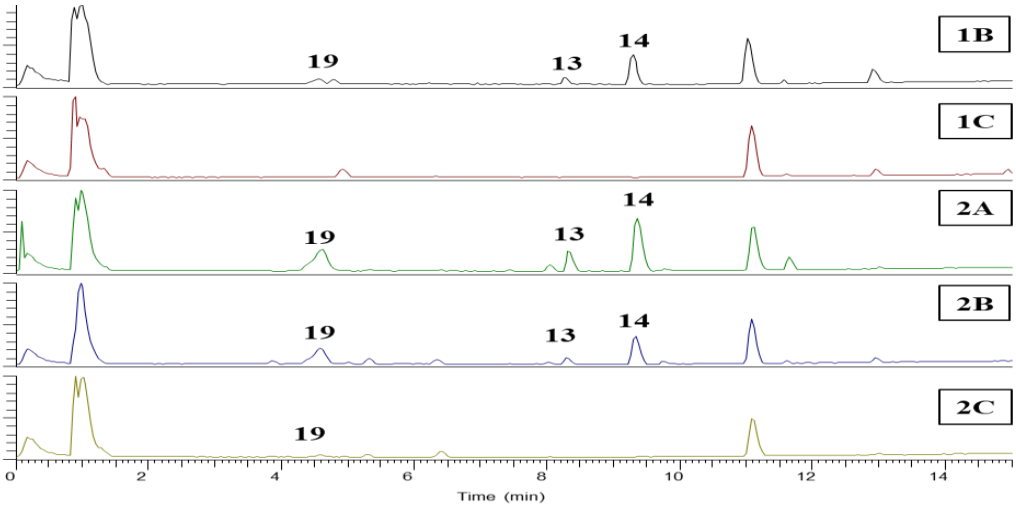
Appendix H – Chromatogram of leaves hydroethanolic extracts of *Emilia*

Figure H1 – Base peak LC-MS chromatogram (positive mode) of *E. fosbergii* hydroethanolic (80%) extracts.



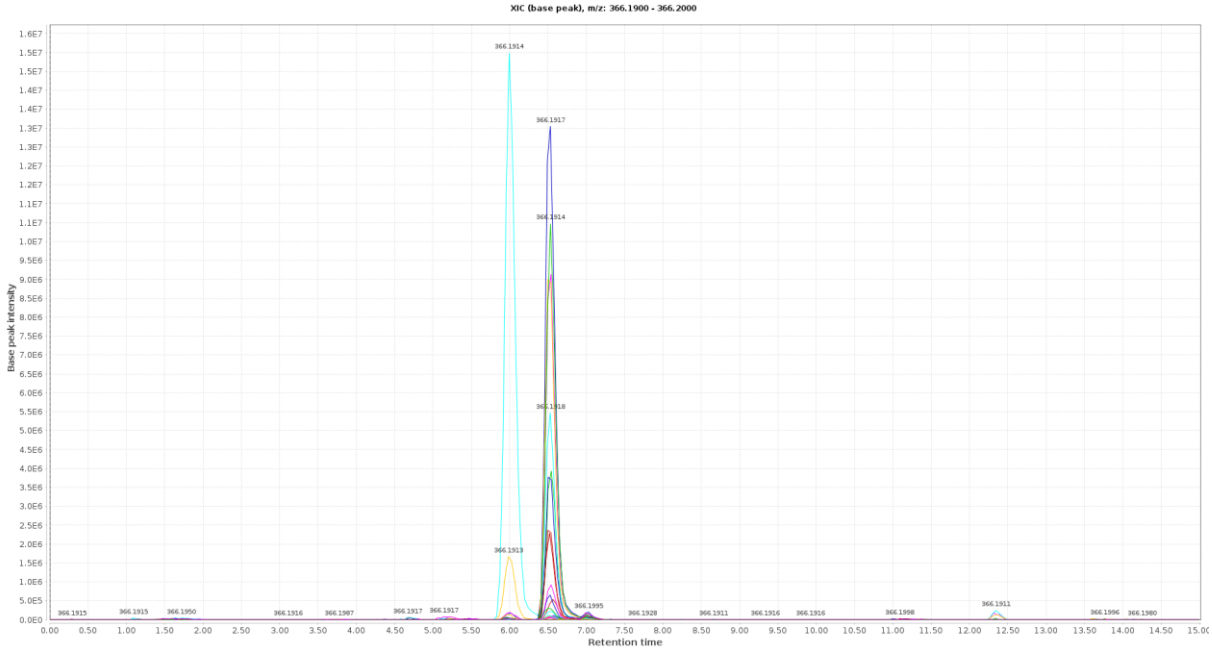
Capital letters represent soil cultivation as described in item 3.2.1: **A)** non-treated soil; **B)** soil treated with horse manure; **C)** soil treated with chicken manure. Numbers in legend represent time of harvest: **1)** before flowering; **2)** after flowering. Figure represents one chromatogram by group. Peak numbering is according to Tables 2 and 3. 3- emiline, 4- senkirkine, 7- ligularidine, 8- putative otonecine, 11- petasitenine, 13- quercitrin, 14- isoquercitrin, 15- isorhamnetin-3-*O*-rhamnoside, 19- 5-*O*-CQA.

Figure H2 – Base peak LC-MS chromatogram (positive mode) of *E. sonchifolia* hydroethanolic (80%) extracts.



Capital letters represent soil cultivation as described in item 3.2.1: **A)** non-treated soil; **B)** soil treated with horse manure; **C)** soil treated with chicken manure. Numbers in legend represent time of harvest: **1)** before flowering; **2)** after flowering. Figure represents one chromatogram by group. Peak numbering is according to Table 3. 13- quercitrin, 14- isoquercitrin, 19- 5-*O*-CQA.

Figure H3 – LC-MS chromatograms (base peak in positive mode, mass range monitoring at *m/z* 366) of all *E. sonchifolia* samples overlaid



Appendix I – Chemical structures annotated in nine edible species studied

Figure I1 – Structures described in **Table 6**.

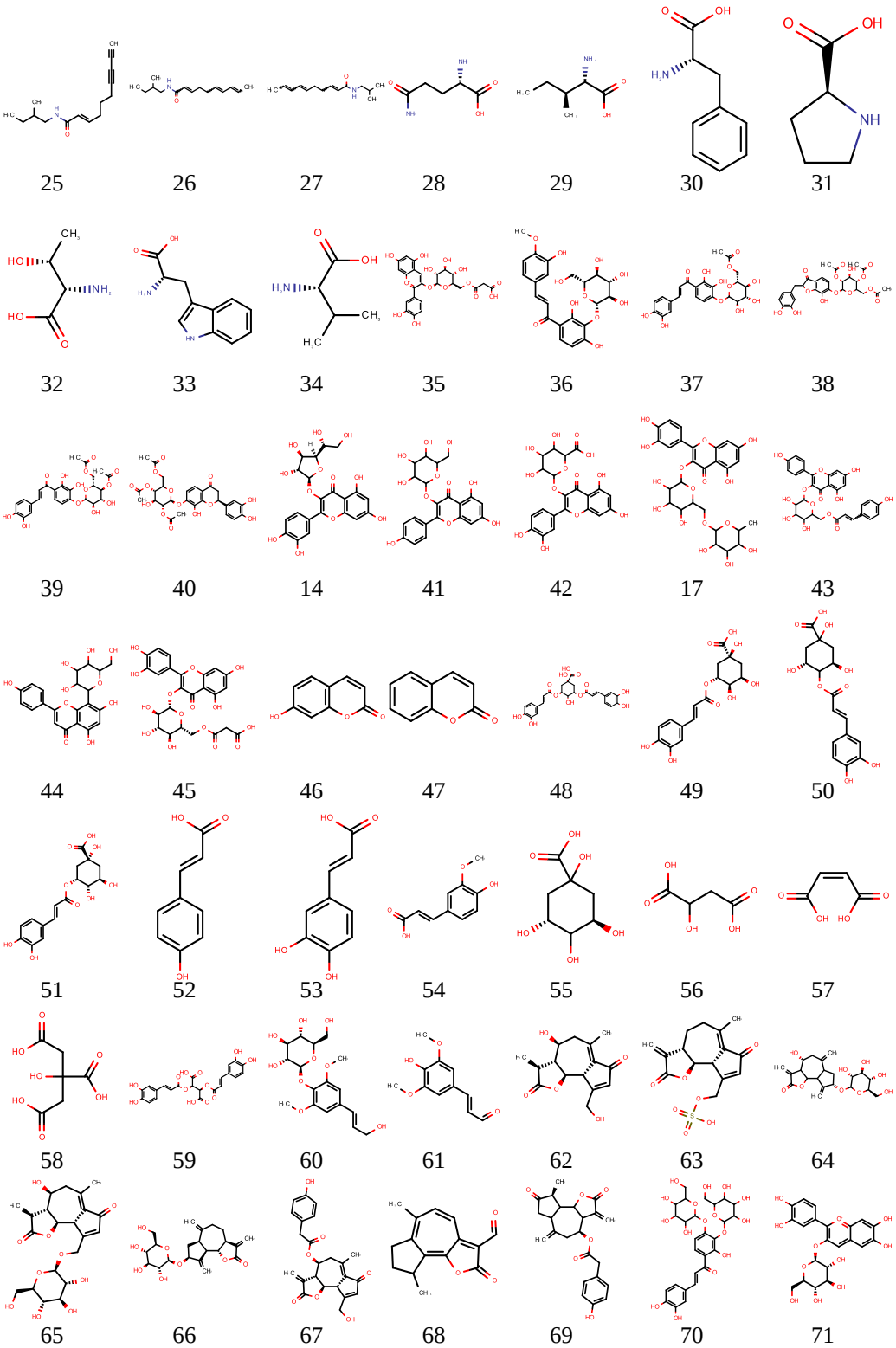
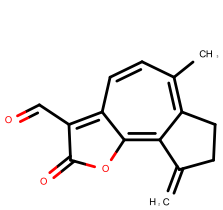
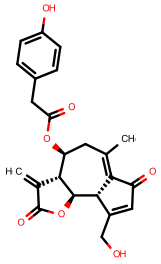


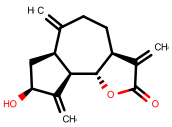
Figure I2 – Chemical structures of putative sesquiterpene lactones



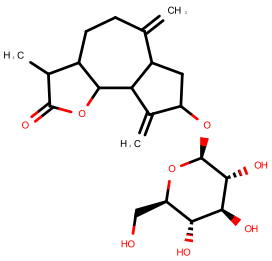
lettucenin



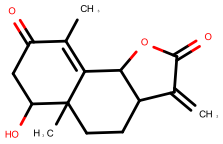
lactupicrin



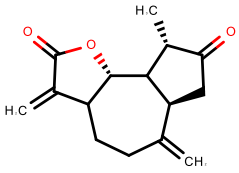
zaluzanin C



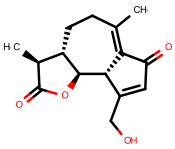
11β,13-dihydro-glucozaluzanin C



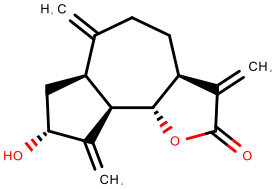
armexifolin



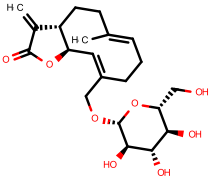
estafiatone



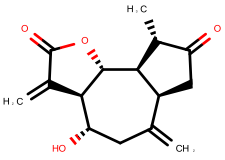
jacquinelin



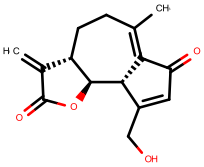
3-epizaluzanin C



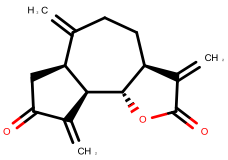
picriside B



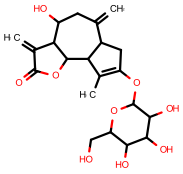
grossheimin



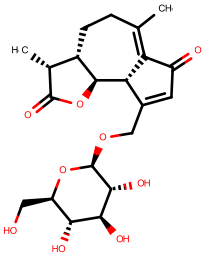
deoxylactucin



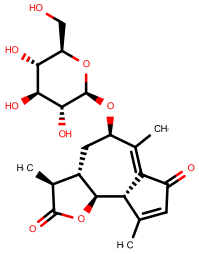
dehydrozaluzanin C



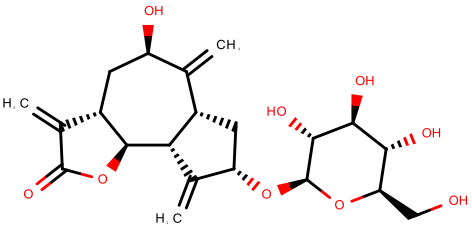
crepiside C



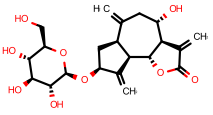
crepidiaside B



lactuside C



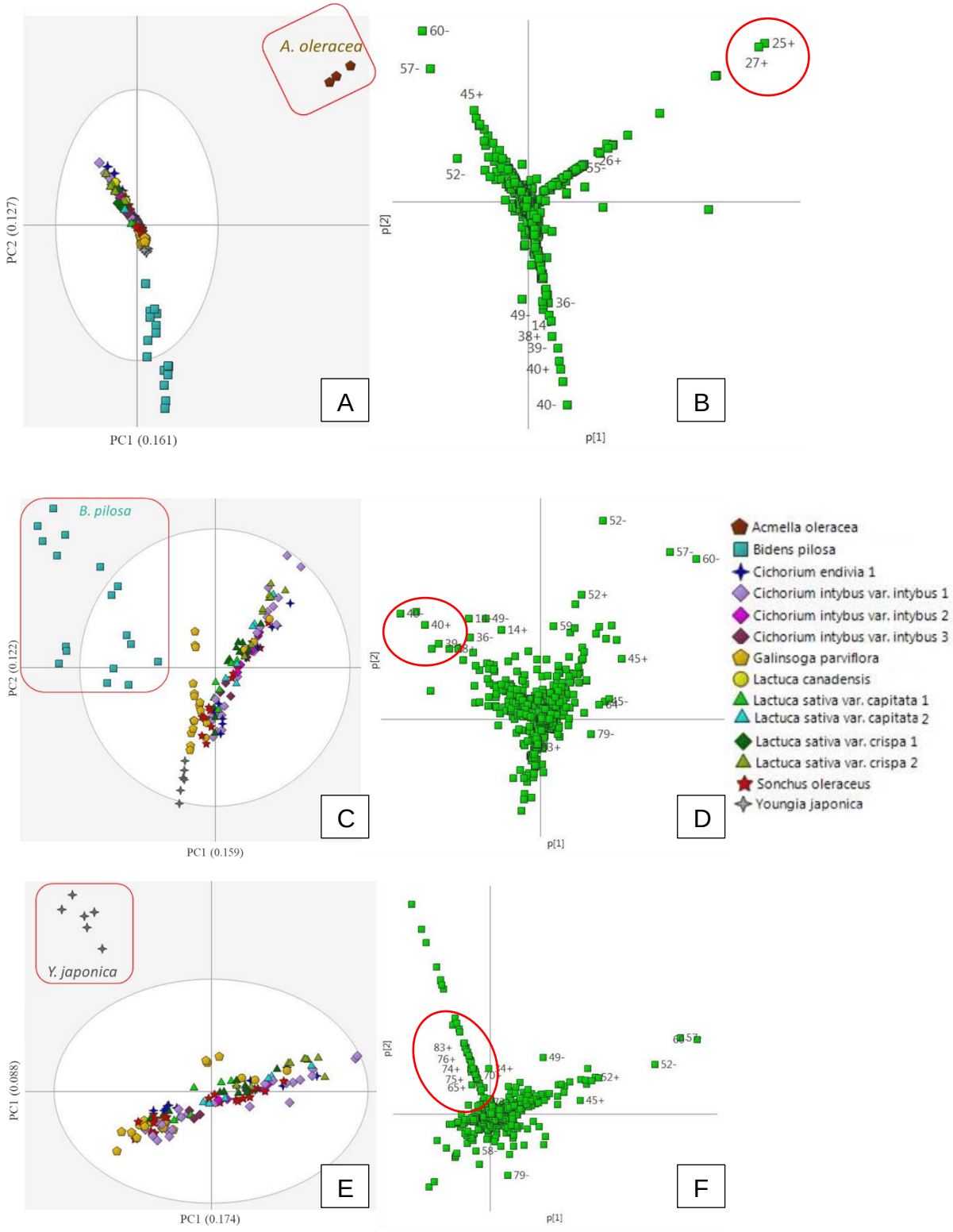
macrocliniside A



crepiside E

Appendix J – PCA modelling of Asteraceae edible plants metabolic profile

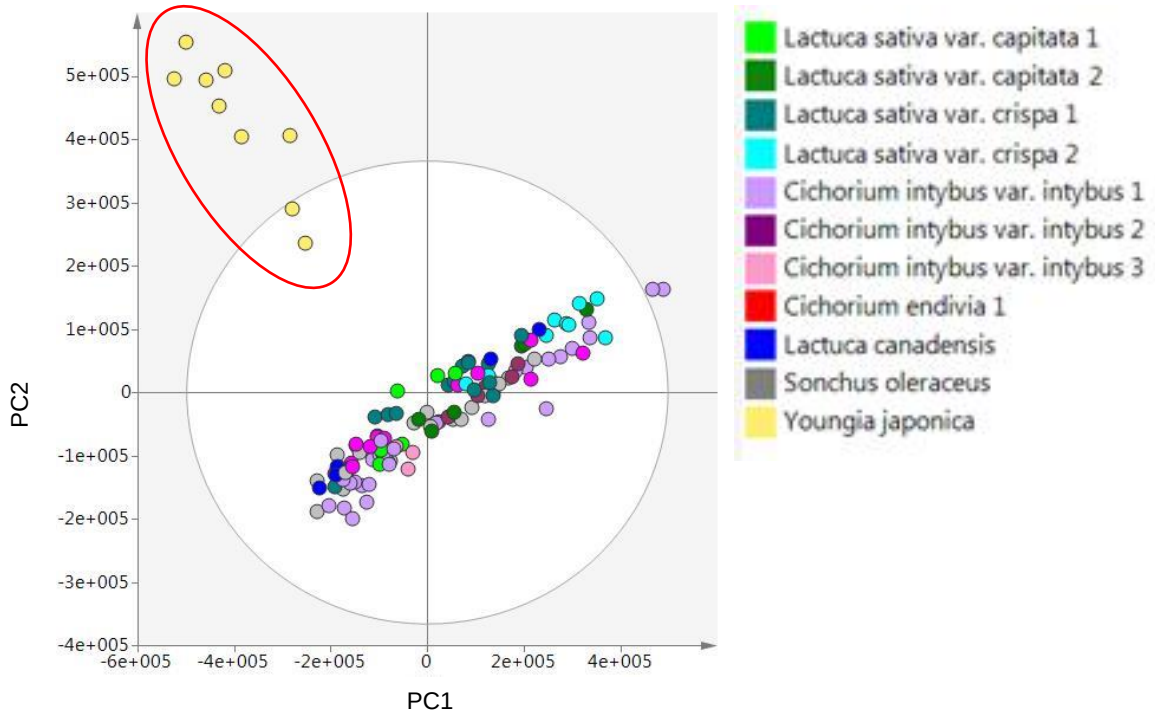
Figure J1 – Sequential PCA analysis of Asteraceae edible plants



Red squares in score plots mark species outliers to the model due to different chemical profile. Red circles in loading plots mark a few discriminant metabolites.

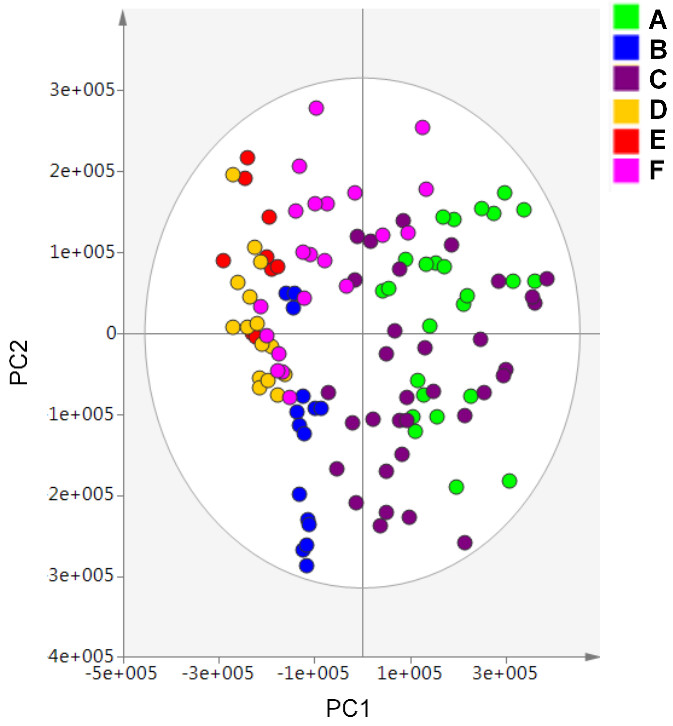
Appendix K – *Youngia japonica* samples as outlier to model

Figure K1 - PCA analysis of Cichorieae samples**.** Red ellipse indicates all *Y. japonica* samples. PC1 (0.185) x PC2 (0.0998), R^2 = 0.829, Q^2 = 0.56, 17 PCs.



Appendix L – Influence of cultivation in plants metabolic profile

Figure L1 - PCA analysis to demonstrate influence of cultivation**.**



PCA analysis reveals that samples distribution in horizontal (PC1) could be influenced by location of cultivation, since samples are divided in two major groups: samples from A and C at right side and samples B, D, E, F at left side.
A previous study of our research shown how cultivation can influence in metabolic profile (SAMPAIO, 2016).