https://doi.org/10.17221/153/2019-CJFS

Discrimination of *Thymus, Origanum, Satureja and Thymbra species fro*m the family Labiatae by untargeted metabolomic analysis

Nurcan Aysar Guzelsoy[1], Filiz Çavuş[1], Oya Kaçar[2]*

[1]Department of Food Authenticity, Central Research Institute of Food and Feed Control, Ministry of Agriculture and Forestry, Bursa, Turkey

[2]Department of Field Crops, Uludag University, Faculty of Agriculture, Bursa, Turkey

**Corresponding author: nurcan.aysarguzelsoy@tarimorman.gov.tr*

Citation: Guzelsoy A.N., Çavuş F., Kaçar O. (2020): Discrimination of *Thymus*, *Origanum*, *Satureja* and *Thymbra* species from family Labiatae by untargeted metabolomic analysis. Czech J. Food Sci., 38: 151–157.

Abstract: The term "thyme" does not refer to herbs that belong to a single species. The genera *Thymus*, *Origanum*, *Satureja* and *Thymbra* of the family Labiatae are traditionally named as thyme and locally known as 'kekik'. Unlike Turkey, these species are globally called differently. Spices made of *Origanum*, *Thymus* and *Satureja* are called oregano, thyme and savory, respectively. It is often difficult to differentiate them because of their similar smell and appearance. Most commercial products traded as a mixture of those genera and the mixing together of different species may lead to economically motivated adulteration and a product of reduced value. The species were analysed by LC-ESI-QTOF-MS and a comprehensive statistical workflow was designed. The data of methanolic extracts were assessed and an extraction algorithm was employed for the processing of raw data. Five species were discriminated using principal component analysis (PCA) and hierarchical cluster analysis (HCA). The results of PCA and HCA were consistent with each other. Twenty-one metabolites were determined for the discrimination.

Keywords: LC-QTOF; plant metabolomics; multivariate data analysis; thyme

Turkey is located in the gene centre of the family Labiatae and there are 45 genera, 556 species and 741 taxa belonging to this family in Turkey (Davis 1982). The term "thyme" does not refer to herbs that belong to a single species. The genera *Thymus*, *Origanum*, *Satureja*, *Thymbra* and *Coridothymus* that are members of the family Labiatae are traditionally named as thyme and commonly called as "kekik" in Turkey. They are marketed as the same due to their similar appearance and smell (Sancaktaroğlu & Bayram 2011). However, spices derived from the genus *Origanum, Thymus* and *Satureja* are called differently globally and they are known as oregano, thyme and savory, respectively, in the other countries.

There are 78 species, of which 39 are endemic, that belong to these five genera in Turkey (Başer & Arslan 2016). The genus *Thymus*, *Origanum*, *Satureja*, *Thymbra* and *Coridothymus* are represented by 38, 23, 14, 2 and 1 species, respectively, in Turkey (Davis et al. 1988). They smell similarly owing to the presence of thymol and carvacrol as main components in their essential oils. Carvacrol and thymol are monoterpenoid phenol compounds and are isomers of each other which have pharmacological properties including antibacterial, antifungal, anthelminthic, insecticidal, analgesic and antioxidant activities (Koparal & Zeytinoğlu 2003; Chishti et al. 2013; Nagoor Meeran et al. 2017; Bedoya-Serna et al. 2018).

Origanum onites L. is the most widely cultivated oregano species, followed by *O. vulgare* subsp. *hirtum.* They are used in the production of essential oils and they have a high economic value. Turkey is considered to be one of the largest oregano exporter countries (approximately 50% of world trade) in the world. Turkey's major export markets are USA, Canada, France, Hungary and Poland (Öztürk et al. 2012; Samet & Cikili 2015). Generally other species are mostly consumed as spices and herbal tea in local markets.

https://doi.org/10.17221/153/2019-CJFS

The similarity of the genus *Thymus*, *Origanum*, *Satureja* and *Thymbra* leads to confusions and economically motivated adulteration since their nutritional quality and economic values are very different (Gad et al. 2013). Studies have shown that a high level of heterogeneity and adulteration may be determined in thyme samples which are exported to the European Union (Marieschi et al. 2009). Detailed information and differentiation of metabolites are needed for their objective identification and quality control. In recent years, a metabolomics approach combined with multivariate data analysis has been widely used for discrimination, authentication, and quality assessment of plants, crops and herbal medicines (Sawada & Hirai 2013; Liu et al. 2016; Abu-Reidah et al. 2019). Spectroscopy, mass spectrometry (MS) and nuclear magnetic resonance (NMR) are the most popular analytical methods in metabolomics (Deborde et al. 2017; Sonibare et al. 2018; Awina et al. 2019). LC-MS-based untargeted metabolomics with high sensitivity provides an ideal procedure for the analysis of a wide range of non-volatile metabolites (Farag et al. 2014; Defernez et al. 2017; Tao et al. 2018).

The aim of this study was to identify the diversity of some species from the family Labiatae through an untargeted metabolomics approach based on the use of LC-ESI-QTOF-MS followed by multivariate analyses. PCA and HCA were used as chemometric techniques. Moreover, the potential chemical markers for discrimination of the species were selected from the identified components.

MATERIAL AND METHODS

Plant material. In this study, a total of 60 samples from five different species which are called "kekik" (*Origanum onites* L., *Origanum vulgare* subsp. *hirtum*, *Thymus vulgaris* L., *Thymbra spicata* var. *spicata* and *Satureja hortensis* L.) were used as plant material. The samples were obtained from Medicinal and Aromatic Plants Research Area, Department of Field Crops, Agricultural Faculty, Bursa Uludağ University during the 2017 vegetation period. Taxonomic identification of the species was made according to the "Flora of Turkey" (Davis et al. 1988) at Biology Department, Science & Art Faculty, Bursa Uludağ University. *Origanum onites* L., *Origanum vulgare* subsp. *hirtum*, *Thymus vulgaris* L. and *Thymbra spicata* var. *spicata* are perennial, *Satureja hortensis* L. are annual plants. Harvesting dates changed according to species that were harvested during 50% of flowering period. 500 grams of each harvested plant were dried in the shade at 24 °C and 45% relative humidity, then the leaves and stems were separated by hand. Dried leaves were ground using a mill and homogenized before analysis.

Chemicals. MS quality methanol was purchased from Merck and ammonium acetate, a mobile phase additive, was obtained from Sigma Aldrich (Germany). Ultra-high purity water was prepared by a Milli-Q system (USA).

Preparation of samples. Five grams of homogenized sample mixture were weighed into 50 mL plastic tubes and 10 mL methanol/water (50/50) was added before homogenisation with Ultra-Turrax. The homogenized mixture was centrifuged at 1 699 × g for 5 minutes. Samples were filtered through a 0.22 micrometer filter, taken into a vial and injected into the LC-QTOF (Agilent 6550 iFunnel) device.

HPLC-Q/TOF-MS analysis. Chromatographic evaluation was carried out with an Agilent HPLC-1260 (USA) system equipped with a degasser, quaternary pump and automatic injector using a Poroshell 120 EC reverse phase C18 analytical column (4.6 × 100 mm, particle size 2.7 μm) (Agilent Technologies, USA). The sample injection volume was 5 μL, at a final flow rate of 0.6 mL min^{-1}. The mobile phases were composed of 5 μM ammonium acetate in water and methanol. The applied gradient was as follows: 0.5 min, 5% B; 0.5–25 min, 5–95% B; 25–28 min, 95% B; 28–29 min, 95–5% B; 29–33 min, 5% B. The analysis was performed in the negative ionization mode with the mass range setting at m/z 100–2 500.

Separated components were detected with liquid chromatography-quadrupole-time of flight system (LC-ESI-Q-TOF, Agilent 6550; Agilent Technologies, USA) equipped with an ESI interface. The instrument parameters were optimized automatically using the Agilent tune mix solution to maximize signal intensity and maintain acceptable resolution. The optimized instrument conditions were set as follows; drying N_2 gas flow rate, 14 L min^{-1}; temperature, 175 °C; nebulizer, 45 psig; capillary voltage of 3 500 V; fragmentor 300; skimmer, 65 V. The reference masses 121.0509 (Purine) and 922.0098 (HP-0921) were used for internal mass calibration. Each sample was analysed in triplicate.

Data processing and statistical analysis. The data obtained from the LC-QTOF has been collected and evaluated with the aid of Agilent MassHunter software (Version B 05.00 Qualitative Analysis; Agilent Technologies, USA) and then converted to compound exchange format (CEF). CEF file format is used to exchange data between Agilent MassHunter software and MassHunter Mass Profiler Professional (MPP) software package 2.0 (Agilent Technologies, USA). The data was charac-

https://doi.org/10.17221/153/2019-CJFS

terized by retention time (RT), abundance, and accurate mass. MPP software was used in differential analysis to determine relationships between different species. Principal component analysis (PCA) and hierarchical clustering analysis (HCA) were performed by MPP software.

RESULTS AND DISCUSSION

To assess the differences in metabolite composition of five species from the family Labiatae, an untargeted metabolite profiling of extracts derived from dried leaves was conducted. Metabolomics workflow was applied in the following order: Sample preparation, separation, data collection, feature finding, statistical analysis and compound identification. Q-TOF MS with an electrospray ionization source was used because of its high selectivity, high resolution and great mass accuracies, which makes it suitable for fast fingerprinting of complex plant extracts (Mauri & Pietta 2000). During the method development some settings were applied to the LC mass system in order to obtain as many compounds as possible for metabolite profiling; ionization mode, mobile phases and flow rate were tested. The negative ion mode was chosen for the analysis due to relatively higher intensities of most metabolites compared with those in the positive mode and the maximum separation for polar metabolites was achieved by using 5mM ammonium acetate at 0.6 mL min^{-1}. Representative total ion chromatograms (TIC) of *Origanum onites* L., *Thymbra spicata* var. *spicata, Satureja hortensis* L., *Origanum vulgare* subsp. *hirtum* and *Thymus vulgaris* L. in the negative ation mode are given in Figure 1. To check the reproducibility of metabolite retention times (RT), repeated measurements were performed, and the variability of RT was found less than 7%.

The data files were processed, and compounds were extracted from the raw data using the Molecular Feature Extraction (MFE) algorithm in MassHunter qualitative analysis software.

The MFE algorithm is an easy way to mine information from complex data and it is designed to find the true ion signals in the total ion chromatogram (Sana et al. 2008). This chromatographic deconvolution step cleans data background ion noise and generates a list of MFs consisting of retention times (RT) and molecular masses. All compounds were characterized by accurate mass, their abundance and retention time. The initial MFs were filtered with the following MFE parameters; peak height higher than 5 000 counts, mass tolerance 5 ppm, negative ions; compound quality score of > 80, isotope grouping peak spacing tolerance 0.0025 m/z plus 7.0 ppm. This step is useful for reducing the dimensionality of the data set prior to PCA and to select the most characteristic marker compounds for species discrimination.

The list of molecular features was converted to compound exchange format (.CEF file) and exported to MPP software for differential analysis to determine relationships between samples. The data on the retention time-molecular mass of approximately 2 552 compounds, which was obtained as a result of MS data of different species, was prepared for analysis; data preparation includes grouping, filtering, alignment, normalization,

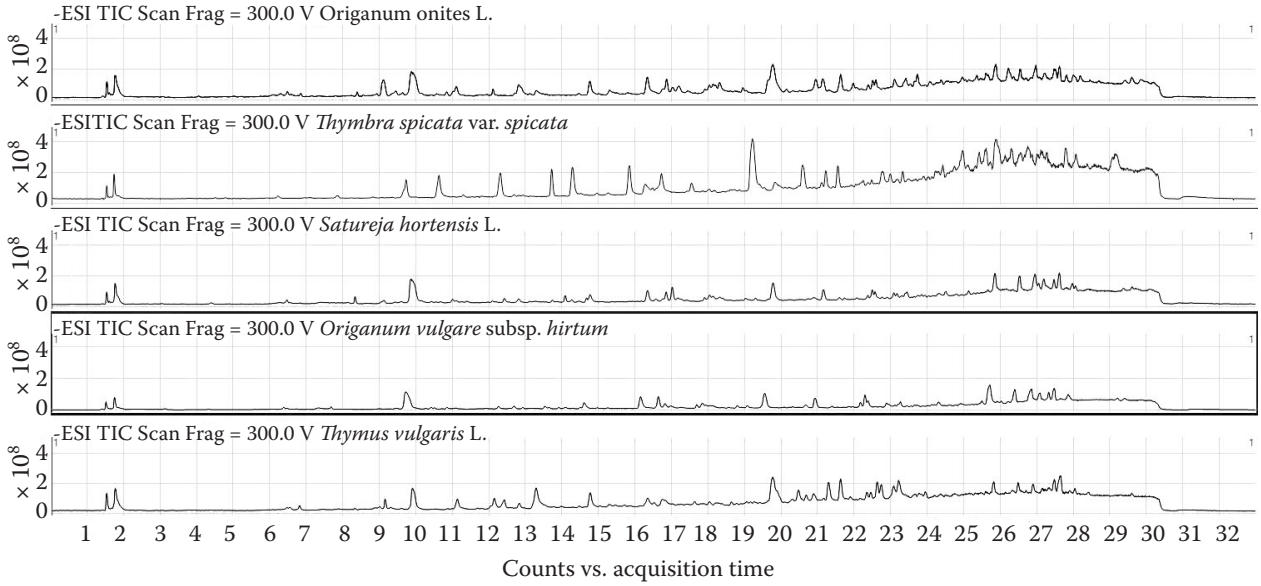

Figure 1. Representative total ion chromatograms (TIC) of species in the negative ion mode of methanol/water extracts (-ESI TIC Scan Frag)

https://doi.org/10.17221/153/2019-CJFS

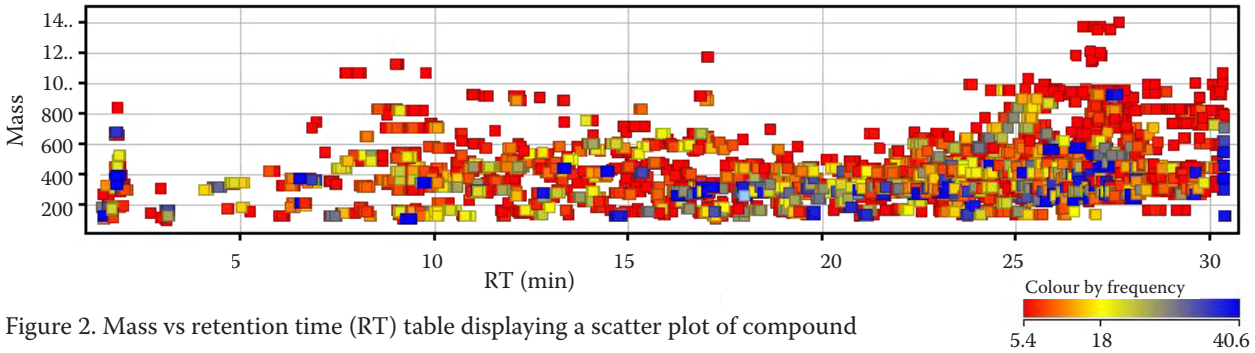

Figure 2. Mass vs retention time (RT) table displaying a scatter plot of compound

and baselining. To reject low intensity data during filtering, minimum absolute abundance was determined as 5 000 counts. Tolerance values were 0.15 min for retention time and 2.0 mDa for mass spectral similarity to align the molecular features across the samples. According to selected RT and mass tolerances, the mass versus RT table (Figure 2) has been obtained which displays scatter plots of compounds. In the mass vs. RT table, replicate samples are expected to have a similar number of present and absent compounds. Compounds with a low incidence in samples are shown in red, on the other hand, compounds with a high incidence are shown in blue.

The number of features contributing to the differentiation of the species was very high (2 552 features), and therefore it was reduced by selecting the most abundant features discriminative between different species. The frequency of these compounds was determined according to the number of detections in the samples. Entities were filtered based on their frequency values. Retained entities that appeared in at least 100% of samples in at least one condition were defined for filtering conditions. As it is shown in Figure 3 after the filtering process, 560 entities were retained in which 100% of the values in any 1 out of 5 conditions had acceptable values and were used in subsequent steps.

After pretreatment of data, principal component analysis (PCA) and hierarchical clustering analysis (HCA) were performed in MPP software to evaluate sample clustering according to the different species and to find groups of data. Principal component analysis (PCA) is a powerful statistical tool commonly performed in metabolomics for identification and grouping of samples due to similarity in the metabolic profile (Kirkwood et al. 2013). Considering the analytical data, PCA was able to discriminate among varieties. As a result

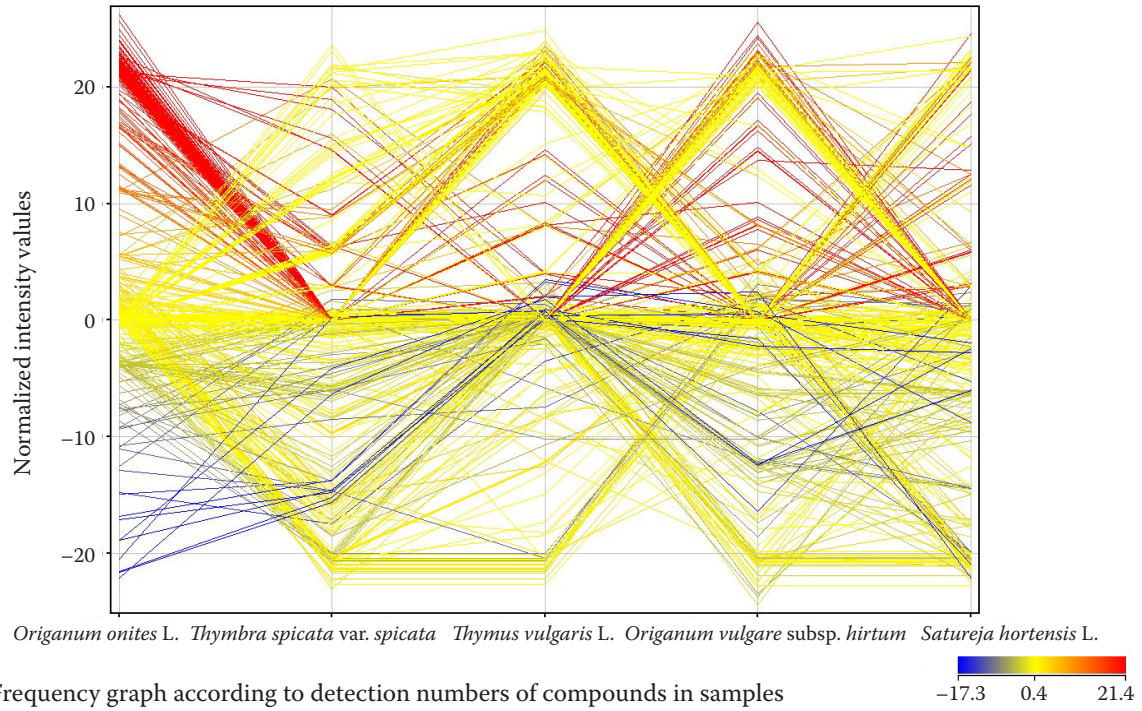

Figure 3. Frequency graph according to detection numbers of compounds in samples

https://doi.org/10.17221/153/2019-CJFS

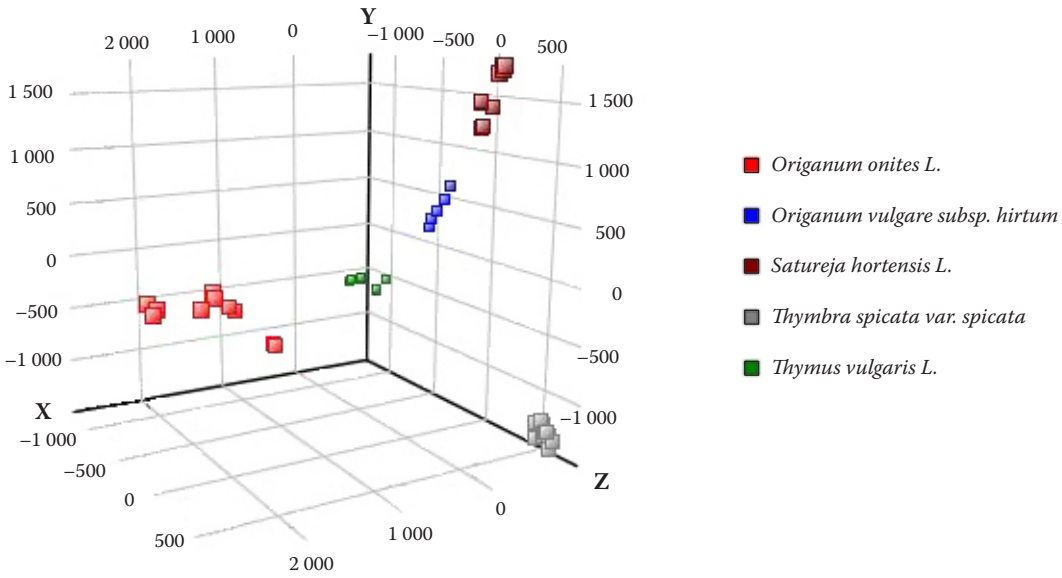

Figure 4. Results of PCA analysis conducted with MS data using negative electrospray ionization of species within the three principal components (PCs)

PCs – 74.4% of the total variance, with 33.13% in the first dimension (X-Axis), 26.26% in the second dimension (Y-Axis) and 15.02% in the third dimension (Z-Axis)

of PCA analysis, it has been observed that samples of *Origanum onites* L., *Origanum vulgare* subsp. *hirtum*, *Satureja hortensis* L., *Thymbra spicata* var. *spicata* and *Thymus vulgaris* L. formed different groups. The result of the PCA in a 3D scatter plot is shown in Figure 4.

The 3D scatter plot shows each species in different colour and samples from the same species are very close to each other in each group. Entities were filtered on their *P*-values calculated from statistical analysis. To assess the differential significance among the species the analysis of variance (ANOVA) test was applied with *P*-value cut-off ≤ 0.05 and fold change cut-off 2.0 using the Benjamini-Hochberg multiple testing correction. As a result of PCA conducted with the use of MS data, the first component (X-axis) has been determined as 33.13%, the second component (Y-axis) as 26.26% and the third component (Z-axis) as 15.02%. Good resolution among samples representing the respective species was obtained for the negative ionization mode. In the next step HCA was conducted with normalized intensity values. HCA is a powerful method to uncover subgroups within a dataset, permitting observations with similar abundance profiles to merge into clusters. In HCA the "Euclidean distance" was selected for measurement. The result is displayed as a dendrogram in Figure 5. Five species were divided into three main clusters in accordance with their metabolic profiles. Cluster 2 was subdivided into two subclusters of *T. spicata* and *T. vulgaris* while cluster 3 was subdivided into those of *O. vulgare* and *S. hortensis*. Cluster 1 comprised the single species *O. onites*. Results of HCA were consistent with those of PCA, suggesting that the analysis methods and data processing in this study are reliable.

As a result of chemometric analysis, differential metabolites were identified among the species. Twenty-one compounds were selected to be the potential chemical markers for discriminating thyme species because of their characteristic distribution (relative higher abundance, %) in the specific thyme species. The list of marker compounds for *O. onites*, *T. spicata*, *S. hortensis*, *O. vul-*

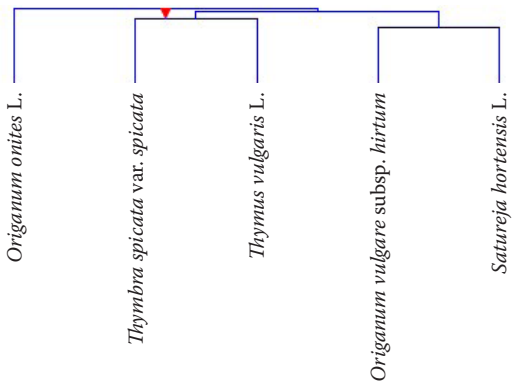

Figure 5. Results of HCA analysis conducted with MS data on the samples.

https://doi.org/10.17221/153/2019-CJFS

Table 1. The characteristics of marker compounds

Species	RT (min.)	Mass-to-charge ratio (m/z)	Ion	Probable ion formula	Mass error (mDa)
Origanum onites L.	4.28	191.0562	$[MH]^-$	$C_5H_{12}N_4O_2S$	–0.94
	6.25	337.0932	$[MH]^-$	$C_{12}H_{14}N_4O_2S$	0.36
	8.78	367.1091	$[MH]^-$	$C_{13}H_{16}N_4O_3S$	0.65
	10.83	445.2076	$[MH]^-$	$C_{18}H_{22}O_{11}S$	–0.87
Thymbra spicata var. *spicata*	8.22	329.0875	$[MH]^-$	$C_6H_{15}N_4O_6P$	0.62
	20.64	419.1492	$[MH]^-$	$C_{25}H_{24}O$	–0.67
	22.91	303.1969	$[MH]^-$	$C_{16}H_8N_4O_3$	0.43
	25.52	333.2074	$[MH]^-$	$C_{20}H_{30}O_4$	–0.87
	27.49	291.2321	$[MH]^-$	$C_{23}H_{48}O_2S_2$	0.41
Satureja hortensis L.	11.95	493.0989	$[MH]^-$	$C_{22}H_{22}O_{13}$	–0.87
	12.07	299.1084	$[MH]^-$	$C_{12}H_{18}N_2O_2S$	0.38
	13.19	593.1571	$[MH]^-$	$C_{27}H_{30}O_{15}$	0.89
	14.23	259.0976	$[MH]^-$	$C_{15}H_{16}O_4$	–0.28
	17.01	283.0608	$[MH]^-$	$C_{16}H_{12}O_5$	0.92
Origanum vulgare subsp. *hirtum*	11.74	138.0193	$[MH]^-$	$C_6H_5NO_3$	0.43
	13.58	261.1135	$[MH]^-$	$C_{15}H_{18}O_4$	–0.35
	25.45	303.2302	$[MH]^-$	$C_{17}H_{36}O_2S$	0.26
	25.67	736.3837	$[MH]^-$	$C_{26}H_{48}N_{20}O_4S$	0.41
Thymus vulgaris L.	12.08	180.0791	$[MH]^-$	$C_{10}H_{12}O_3$	–0.35
	16.54	332.1990	$[MH]^-$	$C_{15}H_{28}N_2O_6$	0.62
	19.11	164.0838	$[MH]^-$	$C_{10}H_{12}O_2$	0.37

RT – retention time

gare and T. vulgaris with m/z and monoisotopic mass are given in Table 1. However, molecular mass alone cannot unambiguously assign the molecular formula because of analytical constraints, so further confirmation by the characteristic MS/MS spectra or by other techniques such as NMR or isotope ratio is needed for the confirmation of compounds. Clear separation of the samples in PCA and uniform clustering results in HCA both show that LC–QTOFMS coupled with chemometric analysis is a valid and accurate approach to discriminate different species from the family Labiatae.

CONCLUSIONS

The goal of this metabolic fingerprinting study was to determine the differences in the metabolomes of five species from the family Labiatae. The LC–QTOFMS technique together with chemometric approach was successfully used for discrimination of five different species of the family Labiatae. As a result of multivariate analysis, several key metabolites were determined for the discrimination of species. The selected metabolites could be used for authentication and quality assessment of different species from the family Labiatae. The HPLC-QTOF/MS-based metabolomics approach is demonstrated to be an effective method for differentiation of different species from the family Labiatae which are important export products. As a perspective of this work, it has been planned to make further research on confirmation and targeted quantification of markers selected in this paper to improve authentication and quality control of thyme species on a routine basis.

REFERENCES

Abu-Reidah I.M., Arráez-Román D., Al-Nuri M., Warad I., Segura-Carretero A. (2019): Untargeted metabolite profiling and phytochemical analysis of *Micromeria fruticosa* L. (Lamiaceae) leaves. Food Chemistry, 279: 128–143.

https://doi.org/10.17221/153/2019-CJFS

Awina T., Mediani A., Maulidiani, Leonga S.W., Faudzia S.M.M., Shaaria K., Abas F. (2019): Phytochemical and bioactivity alterations of Curcuma species harvested at different growth stages by NMR-based metabolomics. Journal of Food Composition and Analysis, 77: 66–76.

Başer K.H.C., Arslan N. (2016): Oregano. In: Ambrose D.C.P., Manickavasagan A., Naik V. (eds): Leafy Medicinal Herbs: Botany, Chemistry, Postharvest Technology and Uses. CAB International, India: 170–188.

Bedoya-Serna G.M., Dacanal G.C., Fernandes A.M., Pinho S.C. (2018): Antifungal activity of nanoemulsions encapsulating oregano (*Origanum vulgare*) essential oil: in *vitro study* and application in Minas Padrão cheese. Brazilian Journal of Microbiology, 49: 929–935.

Chishti S., Kaloo Z.A., Sultan P. (2013): Medicinal importance of genus Origanum: A review. Journal of Pharmacognosy and Phytochemistry, 5: 170–177.

Davis P.H. (1982): Flora of Turkey and the East Aegean Islands. Edinburgh University Press, Edinburgh, UK; 7: 321.

Davis P.H., Mill R.R., Tan K. (1988): Flora of Turkey and the East Aegean Islands. Edinburgh University Press: Edinburgh, UK: 10: 201–209.

Deborde C., Moing A., Roch L., Jacob D., Rolin D., Giraudeau P. (2017): Plant metabolism as studied by NMR spectroscopy. Progress in Nuclear Magnetic Resonance Spectroscopy, 102–103: 61–97.

Defernez M., Wren E., Watson A.D., Gunning Y., Colquhoun I.J., Le Gall G., Kemsley E.K. (2017): Low-field 1H NMR spectroscopy for distinguishing between arabica and robusta ground roast coffees. Food Chemistry, 216: 106–113.

Farag M.A., Mohsen M. Heinke R., Wessjohann L.A. (2014): Metabolomic fingerprints of 21 date palm fruit varieties from Egypt using UPLC/PDA/ESI-qTOF-MS and GC-MS analysed by chemometrics. Food Research International, 64: 218–226.

Gad H.A., El-Ahmady S.H., Abou-Shoer M.I., Al-Azizi M.M. (2013): A modern approach to the authentication and quality assessment of thyme using UV spectroscopy and chemometric analysis. Phytochemical Analysis, 24: 520–526.

Kirkwood, J.S., Maier C., Stevens J.F. (2013): Simultaneous, untargeted metabolic profiling of polar and nonpolar metabolites by LC-Q-TOF mass spectrometry. Current Protocols in Toxicology, 4: 4–39.

Koparal A.T., Zeytinoğlu M. (2003): Effects of carvacrol on a human non-small cell lung cancer (NSCLC) cell line, A549. Cytotechnology, 43: 149–154.

Liu Y., Li L., Xiao Y.Q., Yao J.Q., Li P.Y., Yu D.R., Ma Y.L. (2016): Global metabolite profiling and diagnostic ion filtering strategy by LC–QTOF MS for rapid identification of raw and processed pieces of *Rheum palmatum* L. Food Chemistry, 192: 531–540.

Marieschi M., Torelli A., Poli F., Sacchetti G., Bruni R. (2009): RAPD-Based method for the quality control of Mediterranean oregano and its contribution to pharmacognostic techniques. Journal of Agricultural and Food Chemistry, 57: 835–1840.

Mauri P., Pietta P. (2000): Electrospray characterization of selected medicinal plant extracts. Journal of Pharmaceutical and Biomedical Analysis, 23: 61–68.

Nagoor Meeran M.F., Javed H., Al Taee H., Azimullah S., Ojha S.K. (2017): Pharmacological properties and molecular mechanisms of thymol: Prospects for its therapeutic potential and pharmaceutical development. Frontiers in Pharmacology, 8: 380–414.

Öztürk M., Temel M., Tinmaz A.B., Kil L. (2012): The position of medicinal and aromatic plants in our foreign trade. In: Medicine and Aromatic Plants Symposium, 13–15 September 2012, Tokat, Turkey: 33–44. (In Turkish)

Samet H., Cikili Y. (2015): Importance of medicinal and aromatic plants as an alternative crop in the rural development of Turkey. The Journal of Rural and Community Development, 10: 75–84.

Sana T.R., Roark J.C., Li X., Waddel K.L, Fischer S.M. (2008): Molecular formula and METLIN personal metabolite database matching applied to the identification of compounds generated by LC/TOF-MS. Journal of Biomolecular Techniques, 19: 258–266.

Sancaktaroğlu S., Bayram E. (2011): Research on the determination of the yields and quality characteristics of Istanbul Oregano (*Origanum vulgare* subsp. *hirtum* L.) populations of the different origin. Ege Ünivesitesi Ziraat Fakültesi Dergisi (Ege Journal of Agricultural Research), 48: 265–276. (In Turkish)

Sawada Y., Hirai M.Y. (2013): Integrated LC-MS/MS system for plant metabolomics. Computational and Structural Biotechnology, 4: 1–6.

Sonibare M.A., Ayoola I.O., Gueye B., Abberton M.T., D'Souza R., Kuhnert N. (2018): Leaves metabolomic profiling of *Musa acuminate* accessions using UPLC–QTOF–MS/MS and their antioxidant activity. Food Measurement, 12: 1093–1106.

Tao H.X., Xionga W., Zhao G.D., Penga Y., Zhonga Z.F., Xu L., Duand R., Tsim K.W.K., Yu H., Wanga Y.T. (2018): Discrimination of three *Siegesbeckiae Herba* species using UPLC-QTOF/MS-based metabolomics approach. Food Chemical Toxicology, 119: 400–406.

Received: May 31, 2019
Accepted: May 25, 2020