Variation in secondary metabolites in a unique set of tomato accessions collected in Turkey

Food Chemistry

Bakir, Sena; Capanoglu, Esra; Hall, Robert D.; Vos, Ric C.H.

https://doi.org/10.1016/j.foodchem.2020.126406

Food Chemistry 317 (2020) 126406

Contents lists available at ScienceDirect

Food Chemistry

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

Variation in secondary metabolites in a unique set of tomato accessions collected in Turkey

Sena Bakir[a,b], Esra Capanoglu[a,*], Robert D. Hall[c,d], Ric C.H. de Vos[c,*]

[a] Istanbul Technical University, Faculty of Chemical and Metallurgical Engineering, Food Engineering Department, Maslak, Istanbul, Turkey
[b] Recep Tayyip Erdogan University, Faculty of Engineering, Merkez, Rize, Turkey
[c] Bioscience, Wageningen University and Research Centre (Wageningen-UR), PO Box 16, 6700 AA Wageningen, The Netherlands
[d] Laboratory of Plant Physiology, Wageningen University & Research, PO Box 16, 6700 AA, Wageningen, The Netherlands

ARTICLE INFO

Chemical compounds studied in this article:
LC-MS grade methanol (Pubchem CID: 887)
Acetonitrile (Pubchem CID: 6342)
Formic acid (Pubchem CID: 284) provided by Merck (Frankfurter, Germany)

Keywords:
Tomato
Semi-polar phytochemicals
Metabolomics
Multivariate analysis

ABSTRACT

In this study, 50 tomato landraces grown in Turkey were investigated in terms of their secondary metabolite profiles. Each accession was planted in 2016 and 2017 in 3 replicates in an open field. In this study, color, pH and brix of the fruit samples were measured and an unbiased LCMS-based metabolomics approach was applied. Based on Principal Components Analysis (PCA) and Hierarchical Cluster Analysis (HCA) of the relative abundance levels of > 250 metabolites, it could be concluded that fruit size was the most influential to the biochemical composition, rather than the geographical origin of accessions. Results indicated substantial biodiversity in various metabolites generally regarded as key to fruit quality aspects, including sugars; phenolic compounds like phenylpropanoids and flavonoids; alkaloids and glycosides of flavour-related volatile compounds. The phytochemical data provides insight into which Turkish accessions might be most promising as starting materials for the tomato processing and breeding industries.

1. Introduction

There is a general agreement that a healthy diet is important in preventing chronic diseases, and supporting the right energy balance of the body. Specific recognition has been given to the so-called 'Mediterranean diet' as being particularly healthy in terms of its balanced nutritious and health-promoting components (Trichopoulou et al., 2014). Tomato and its products are key components of the Mediterranean diet and are considered as healthy foods due to their low fat content and high levels of vitamins A and C as well as of a range of other health-promoting compounds such as phenolic acids, flavonoids and carotenoids. The health-related bioactive compounds in Mediterranean diet were previously investigated by several researchers (Mazzoni et al., 2016; Battino et al., 2019; Cianciosi, Simal-Gandara, & Forbes-Hernandez, 2019).

The (poly)phenolic compounds of tomato and tomato-based products have frequently been evaluated in relation to their potential effect against, for example, cardiovascular diseases (O'Kennedy et al., 2006), certain types of cancer (Ramos-Bueno et al., 2017), diabetes (Shidfar et al., 2011), and also regarding their general antioxidant activity (Kotíková, Lachman, Hejtmánková, & Hejtmánková, 2011) and anti-inflammatory activity (Burton-Freeman, Talbot, Park, Krishnankutty, & Edirisinghe, 2012).

Tomato is a member of the *Solanaceae* family, and part of the genus *Solanum* which includes many wild relatives as well as the cultivated tomato '*Solanum lycopersicum*'. It is one of the most widely consumed vegetables across the world and its consumption is still increasing. According to the data provided by FAOSTAT, tomato production in the world continually increased during the 1994–2016 period. Tomato is very important for the Turkish diet and Turkey is fourth in the world for tomato production (for fresh and processing goals) following China, USA, and India (Faostat, 2017).

Tomato fruit has been also a regular subject for metabolic profiling and has almost become a model organism for investigations relating to both the fresh fruit (Zhu et al., 2018) and their processed forms (Capanoglu, Beekwilder, Boyacioglu, De Vos, & Hall, 2010; Di Lecce, Martinez-Huelamo, Tulipani, Vallverdu-Queralt, & Lamuela-Raventos, 2013).

Metabolites are generally recognized as being the end products of the genome, and mirror the operation of the biological system (Herrero, Simo, Garcia-Canas, Ibanez, & Cifuentes, 2012) and metabolomics has become an application-driven science with a broad range of applications in various fields, including medical science, synthetic biology, medicine, food science and technology as well as nutrition (Putri et al.,

* Corresponding authors.
E-mail addresses: capanogl@itu.edu.tr (E. Capanoglu), ric.devos@wur.nl (R.C.H. de Vos).

https://doi.org/10.1016/j.foodchem.2020.126406
Received 24 November 2019; Received in revised form 10 February 2020; Accepted 11 February 2020
Available online 22 February 2020
0308-8146/

2013). The matrices of many food materials have been characterized in great detail, mainly by using comprehensive metabolomics approaches such as liquid chromatography or gas chromatography coupled to mass spectrometry (LC-MS and GC–MS, respectively) in order to investigate their biochemical composition in relation to e.g. phenotype, genotype, processing treatments etc. This approach can be used for food quality analysis (Capanoglu et al., 2010; Thissen et al., 2011; Wishart, 2008) and also, food safety aspects such as investigating the presence of pesticide residues, veterinary drugs, natural toxins, environmental contaminants and contaminants arising from food processing (Castro-Puyana, Perez-Miguez, Montero, & Herrero, 2017; Pinu, 2016). Adulteration and fraud in foods have also been revealed through metabolite profiling of materials using similar chromatographic techniques and, metabolites can be used as indicators for the provenance of food materials in terms of their origin (Tolstikov & Fiehn, 2002; Esteki et al., 2018; Esteki, Ahmadi, Heyden, & Simal-Gandara, 2019; Esteki, Regueiro, & Simal-Gandara, 2019; Esteki, Shahsavari, & Simal-Gandara, 2019a, 2019b) and post-harvest history (Thurman et al., 2005) as well as fraudulent deception of consumers regarding their purity and adulteration (Wishart, 2008; Esteki, Farajmand, Kolahderazi, & Simal-Gandara, 2017; Moldes, Mejuto, Rial-Otero, & Simal-Gandara, 2017). In the literature, there are few studies indicating the combined use of spectrophotometric techniques and multivariate analysis in order to identify food classification, authentication and adulteration (Esteki, Shahsavari, & Simal-Gandara, 2018; Astray et al., 2019; Gonzalez-Fernandez et al., 2019). Besides, nowadays blockchain technology is also becoming popular in the field of food traceability (Galvez, Mejuto, & Simal-Gandara, 2018).

There are several studies in which the metabolites of tomato fruit have been investigated by using LCMS methods (Moco et al., 2006; Moco et al., 2007; Iijima et al., 2008; Tohge & Fernie, 2015; Alseekh et al., 2015). On the other hand, there are yet few studies investigating different tomato accessions for their metabolite contents (Siracusa, Avola, Patane, Riggi, & Ruberto, 2013; Baldina et al., 2016; Naranjo et al., 2016; Di Paola Naranjo et al., 2016), and the number of accessions involved were relatively small and none applied an unbiased and comprehensive (untargeted) metabolomics approach. In this study, we collected a unique set of locally-grown accessions of tomato from all over Turkey. This set included many regional landraces and special accessions which were collected from hobby-growers in local villages and from farmers. Seeds from all lines were stored at the Turkish National Seed Gene Bank of the Aegean Agricultural Research Institute. The results presented here are therefore the first not only for this substantial collection of Turkish accessions, but also for exploring their natural biodiversity in an essentially untargeted manner. In this study we applied LC with high resolution MS to semi-polar aqueous-methanol extracts, which in tomato fruit, enables the detection of a large variety of secondary metabolites, including phenolic acids, phenylpropanoids, flavonoids, alkaloids, glycosylated volatiles, etc.

2. Materials and methods

2.1. Tomato samples

The origins of all Turkish tomato seed samples used in this screen were obtained from the National Seed Gene Bank of Aegean Agricultural Research Institute, Izmir, Turkey. All details are reported in Supplemental Table S1. Landraces were acquired from farmers who do not have access to commercial hybrid seeds to produce tomatoes and hence seed is retained on the farm from year to year. In total, 75 different landraces were initially included in this research. In addition, two commercial hybrids, Ayaş and Tosya, were included in the study of 2017 for comparison with the accessions. All accessions were labelled according to their origin, based on the name of their nearest city as indicated in Supplemental Table S1. All plants were cultivated in a field provided by the Plant Biodiversity, Geophyte Research and Training Center Directorate, Beykoz, Istanbul in 2016 and experiments were carried out following a uniform cultivation practice in an open field based on randomized block design with irrigation and fertilization being fully automated. Each accession was planted in 3 replicates. The same seed batch used in 2016 was used to grow the plants for a second season in 2017. The same cultivation procedure as in 2016 was applied in 2017, with the only exception being that the planting was carried out in a different open field provided by Ataturk Horticultural Central Research Institute, Yalova. Both in 2016 and 2017, the planting time in the field was at the end of March and harvesting was performed at the end of August. However, not all accessions yielded sufficient fruits: 51 accessions could be harvested in 2016 and 50 in 2017. In the end, 37 accessions were common across the two years. After harvesting at the ripe stage, fruits were immediately brought to the Food Engineering Department Laboratories at Istanbul Technical University, Istanbul, Turkey, for general fruit assessments (size, brix, pH and colour), sample preparation and storage. Representative fruits per accession were selected by eliminating damaged tomatoes and those with black spots. Cherry-type tomatoes were grouped as being small-sized (< 3.5 cm), beef type tomatoes were grouped as large-sized (> 6.5 cm) and the tomatoes sized between 3.5 and 6.5 cm were grouped as being medium. Three to ten intact fruits (depending on the size of the fruit) per accession were pooled. The whole fresh tomatoes were subsequently blended for 10 *sec* in order to obtain pulp for both pH and brix measurements. After this short blending step the main part of the pulp was immediately frozen in liquid nitrogen for metabolite analysis. Frozen pulps were powdered using a precooled IKA A11 grinder (Staufen, Germany), powders were dried using a freeze dryer (ALPHA 1–2 LDplus, Osterode am Harz, Germany) and these dried powders were stored at -20 °C until transport to WUR, Wageningen, The Netherlands, for metabolome analysis.

2.2. Colour, pH and brix measurements

Color analyses were performed on the sliced surfaces of freshly-cut ripe fruits, using a colorimeter (Chroma Meter CR-410, Konica-Minolta, Milan, Italy) in the CIE 1976 (L*, a*, b*) color space. The instrument was calibrated against a white tile before measurements were performed. Chroma and hue values of samples were calculated from L*, a*, b* using the equation ($\sqrt{a^{*2} + b^{*2}}$).

An Abbe refractometer (PCE Instruments, UK) was used for the determination of brix values and a pH meter (HI 2211 Hanna Instruments, The Netherlands) for pH values of the freshly blended tomatoes prepared as explained in Section 2.1.

2.3. Metabolomics

Metabolite extraction and LCMS-based metabolomics was basically carried out as originally described in De Vos et al. (2007). Briefly, 30 mg of freeze dried samples were accurately weighed and mixed with 1.2 mL 75% MeOH containing 0.1% formic acid. Samples were sonicated in a water bath (Branson 3510, Hampton) at maximum frequency for 15 min at 20 °C, and then centrifuged at 16,000 rpm (Eppendorf Centrifuge 5810R and 5415R-Ontario, Canada) before transferring the clear supernatant into HPLC glass vials. Quality control samples (QC's) were made using an equal mixture of all tomato powders and preparing five independent extractions as technical replicates. In addition, 2 tomato samples were selected randomly to prepare four replicate extractions as additional individual technical replicates.

Chromatographic separation of extracts was performed with an HPLC system (Waters Acquity, Milford, MA, USA) fitted with a C_{18} column (Phenomenex Luna 150 × 2 mm i.d., 3 μm-Torrence, USA). 0.1% formic acid in Milli-Q water (Purelab Ultra water purification system, High Wycombe, UK) (eluent A) and 0.1% formic acid in acetonitrile (eluent B) were prepared for the mobile phase. Gradient phase

started with 95% eluent A and 5% eluent B; at 45 min 65% eluent A and 35% eluent B: at 47 min 25% eluent A and 75% eluent B; at 52 min 25% eluent A and 75% eluent B and at 54 min returns to the initial conditions with a 0.19 mL/min flow rate. The column temperature was set to 40 °C and detection was accomplished with both a photodiode array (PDA) detector (Waters) at 210–600 nm and an LTQ-Orbitrap FTMS hybrid mass spectrometer (Thermo Scientific, Bremen, Germany) in negative ionization mode at an m/z range of 90–1350 D and a mass resolution of 60,000 FWHM (Mokochinski et al., 2018).

The MetAlign software package (Lommen, 2009) was used for mass signal extraction and alignment, with a peak detection threshold value (Metalign parameter 8B) of 10,000 ions/scan. The data obtained from MetAlign were then filtered for mass peaks having a minimum intensity of 20,000 ions/scan in at least 2 samples. The resulting filtered peak list was imported into MSClust software (Mokochinski et al., 2018) in order to remove metabolite signal redundancy by clustering signals derived from the same metabolite, based on their similar retention time and relative abundance pattern across all samples, and creating an in-source mass spectrum consisting of e.g. natural isotopes, fragments and adducts created in the ion source. In addition, for each metabolite the MSClust tool calculated a single representative intensity value (total ion counts) based on clustered representative mass signals. The resulting list with relative intensities of each putative metabolite detected in each sample was used for multivariate analysis after log transformation of the metabolite intensity data; to enable log-transformation, non-detects with intensity value 0 were randomized between 45% and 55% of the Metalign detection threshold. Principal components analysis (PCA) was made with Simca v. 14 (Umetrics; Umea; Sweden) using pareto scaling, while hierarchical clustering analysis (HCA) was performed with GeneMaths XT (Applied Maths; 1.6 software, Belgium) using mean-centering.

2.4. Metabolite annotation

The annotation of selected metabolites was primarily based on existing databases like the MotoDB (Sofia Moco, Raoul J Bino, Oscar Vorst, Harrie A Verhoeven, Joost de Groot, Teris A van Beek, et al., 2006), the KOMICS database http://www.kazusa.or.jp/komics/en/database-en.html and the Dictionary of Natural Products, as well as on annotations reported in other papers dealing with tomato fruits and other tomato plant parts (e.g. Gómez-Romero, Segura-Carretero, and Fernández-Gutiérrez (2010), Tikunov, de Vos, Paramás, Hall, and Bovy (2010), Vallverdú-Queralt Jauregui Medina-Remón Andrés-Lacueva & Lamuela-Raventós (2010), Ferreres, Taveira, Pereira, Valentao and Andrade (2010), Schmidt, Li, Shi, Jones, & Pichersky (2011)). Annotation was based on a number of complementary features – their corresponding accurate mass (using a threshold of 5 ppm), chromatographic retention time, UV/Vis absorbance spectrum (when available) and any in-source fragments that were detected. New metabolites (the so-called 'unknowns') were putatively annotated based on their elemental formula calculated from the observed accurate mass also using additional information such as UV absorption spectra and MS/MS fragmentation data.

3. Results and discussion

3.1. Geographical distribution

In this study, we collected 75 tomato accessions from Turkey and 2 commercial hybrids. As can be seen from Fig. 1, the accessions collected represent all geographical regions in Turkey. Collection areas were at diverse altitudes (2–1566 m above sea level), and with different climate conditions ranging from Adana with a relatively hot and dry climate to Erzurum which is relatively cold and dry (Supplementary metadata: Table S1). There were significant phenotypic differences between the fruit samples (Fig. 2): some were large beef type tomatoes, others were small cherry-type tomatoes or intermediate sized with round or elliptic shapes. Among all accessions there was a single yellow accession (number 37; Fig. 2e) while all others had the typical red color.

3.2. Color, pH and brix values

The color, pH and brix values of the tomato accessions were measured and are reported in Supplementary data Table S2 and demonstrated with box plot in Supplementary data Fig. S1. For the color analyses, in the CIELAB color space the a* value represents the redness and greenness from positive to negative, the b* value represents the yellowness and blueness from positive to negative and the L* value represents the lightness (Singh, Chauhan, Vatsa, & Singh, 2003). Tomatoes used in this study were harvested after reaching its maturity level which was decided visually (spherical, uniform red color formation), so a* values of accessions were always positive and there were no negative values which is related with the unripe stages. Chroma is the reflection of purity or saturation of light and is a good indicator of consumer acceptance (López Camelo & Gómez, 2004). The chroma values of the Turkish tomato accessions were calculated based on measured a* and b* values and varied between 30 and 60. On the other hand, the L, a*, b* values of the samples ranged between 36.9–46.1, 8.1–34.9, and 21.1–55.7, respectively. The yellow-colored accession had the lowest a* and the highest b* values, as expected, and thus its chroma ($\sqrt{a^{*2} + b^{*2}}$) is in the middle. Brix values of fruits ranged between 5.5 and 9.42 (Fig. 3) and pH values were found to be between 4.02 and 4.47. These pH and Brix values were observed to be in the range of other published commercial hybrids/landraces (Bottino, Capannelli, Turchini, Della Valle, & Trevisan, 2002; Tiwari, O' Donnell, Brunton, & Cullen, 2009).

3.3. Metabolic diversity

Fruits of 48 accessions and the 2 hybrids from the 2017 harvest were examined for their metabolic diversity by analyzing semi-polar (aqueous-methanol) extracts with a high-resolution LC-MS system and processing the data in an unbiased manner.

To estimate the overall technical variation of the untargeted analysis, appropriate reference extracts were prepared both from a pool of all samples (QC's, n = 5, spread evenly across the real samples and used to check system stability during the entire analysis series; and 2 random accessions (n = 4 each).

We first checked whether the metabolite intensity data generated by the untargeted metabolomics workflow, as applied in the present study, are representative of data obtained by targeted data analysis. For 3 well-known phenolic compounds in tomato fruit i.e. rutin, naringenin chalcone and chlorogenic acid, we plotted the peak intensities derived from the untargeted processing workflow against those derived from targeted LCMS data analysis based on retrieving the peak height of their corresponding molecular ion-specific mass signals from the same raw LCMS data files (Suppl Fig. S2). For all three, the targeted data were linearly and highly ($R^2 > 0.99$) correlated with the untargeted data. In addition, the correlations between the data from this untargeted LCMS analysis and those from targeted LC-PDA analysis, based on integrating chromatographic peak areas of UV/Vis traces – a common approach in phenolic compound analysis – across the tomato extracts analyzed by the LC-PDA-MS system were also high, with regression values of 0.987, 0.988, and 0.94 for rutin, naringenin chalcone, and chlorogenic acid, respectively (data not shown). These high correlations across samples indicate that the untargeted LC-MS approach provides similar relative abundance data for metabolites as targeted analysis using LC-MS or LC-PDA, while additionally providing insight into the relative abundance of many other known and yet unknown compounds in the fruit samples.

Using our untargeted workflow, we were able to obtain relative abundance values for 251 putative compounds present in the ripe fruits

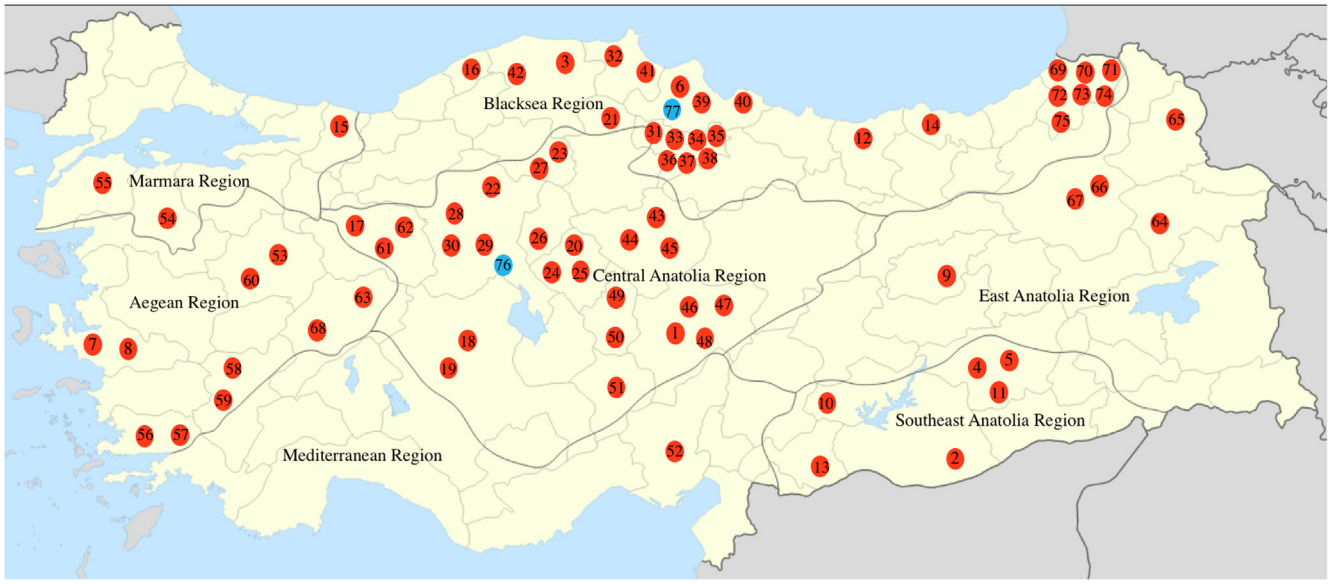

Fig. 1. Map of Turkey with indicated sampling areas of 75 tomato landraces and 2 hybrids. The red points on the map represent locations of the landraces and the blue points of the hybrids.

Fig. 2. Demonstration of some contrasting tomato accessions. [a] accession 16 TR68513 Bartin, [b] accession 21 TR69785 Corum, [c] accession 24 TR69805 Kirsehir, [d] accession 29 TR69817 Ankara, [e] accession 37 TR70708 Amasya, [f] accession 42 TR70740 Kastamonu, [g] accession 44 TR71370 Yozgat, [h] accession 57 TR61697 Mugla and [i] accession 72 TR75261 Artvin. Pictures are taken after harvest 2017.

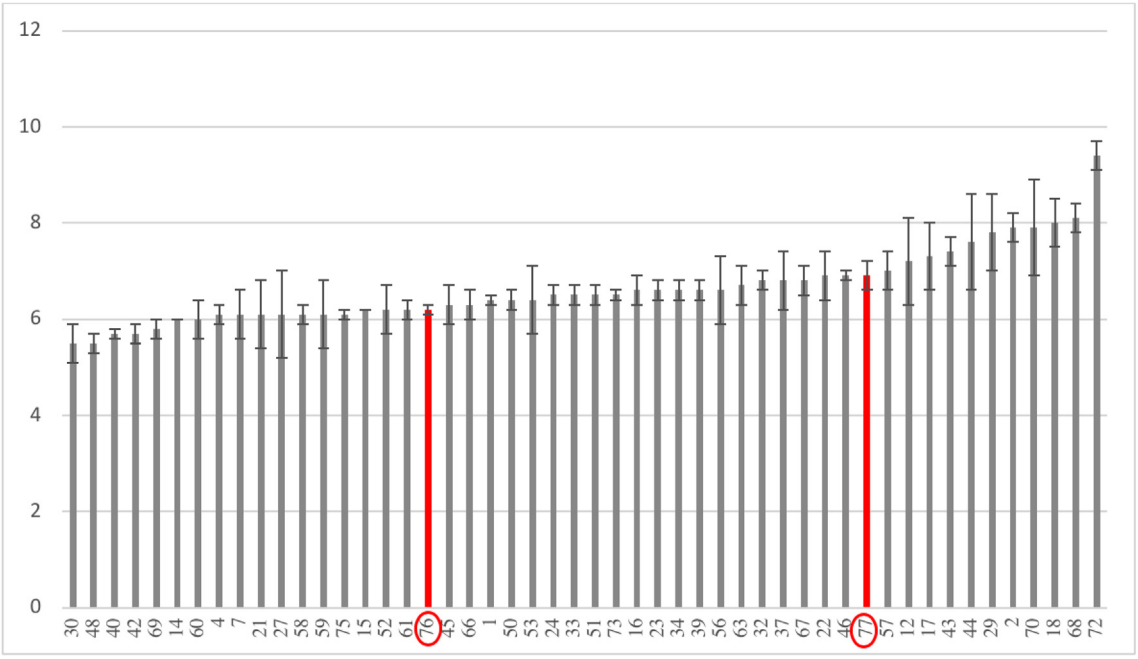

Fig. 3. Brix values of 48 tomato accessions and 2 hybrids (red columns with red circles) after harvest in 2017. Error bars indicate the standard deviation of 3 measurements from the tomato surface. The genotypes associated with the accession codes are to be found in Supp. Table S1.

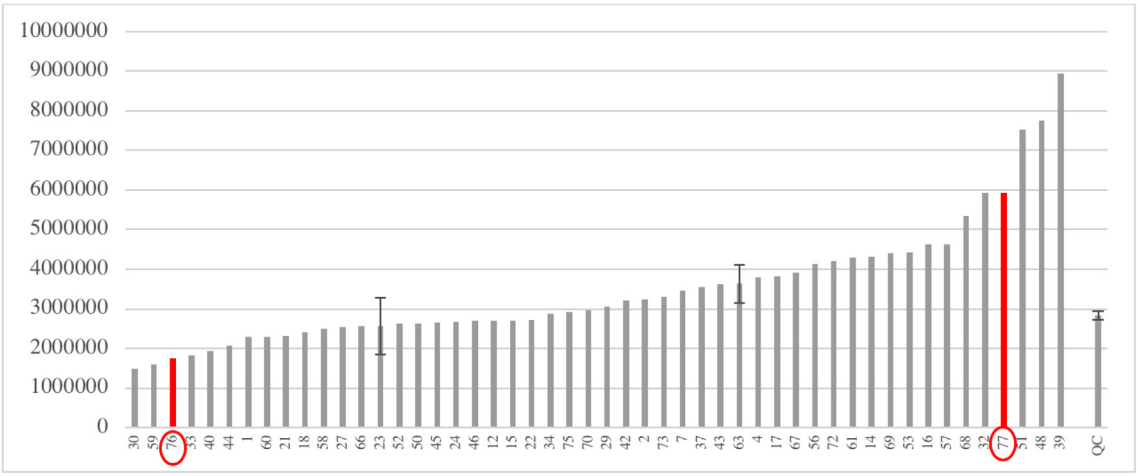

Fig. 4. Relative level of the main flavonol-glycoside rutin (quercetin-3-*O*-rutinoside) in ripe fruits of 48 Turkish tomato accessions and 2 hybrids (red columns with red circles) from the 2017 harvest based on untargeted LC-MS analysis. Samples are ordered from lowest to highest level. To estimate technical error, accessions 23 and 63 were analysed 4 times while the quality control samples (QC), consisting of a mix of all genotypes, was analysed 5 times during the entire sample series; error bars represent the standard deviations of their means.

of the 48 Turkish accessions and 2 hybrids harvested in 2017 (Suppl Excel Table 1, sheet 2017) and 265 compounds present in the ripe fruits of the 37 accessions in 2016 which were common to the accessions measured in 2017 (Suppl Excel Table 1, Sheet 2016). We identified substantial biodiversity among these accessions in metabolites representing various biochemical classes, including flavonoids, regarded as potentially healthy dietary compounds, phenylpropanoids, alkaloids and glycosides of flavour-related volatile compounds like benzyl alcohol and methyl salicylate. As an example, results for rutin (quercetin-3-*O*-rutinoside) obtained with untargeted metabolite analysis of the fruits in 2017 are presented in Fig. 4, and those of some other compounds are to be found in the supplementary data, Fig. S3. Results indicate that the metabolite content of samples with the lowest to the highest levels varied for instance 6 fold for rutin, 8 fold for naringenin chalcone, 5 fold for ferulic acid hexose and 2.5 fold for citric acid, while several other compounds, including glycosides of flavour-related compounds, were only detected in specific accessions (Fig. 4 and Suppl data Fig. S3).

In order to provide insight into the overall differences and similarities in the metabolite composition between accessions, the metabolite data of the 2017 harvest year were subjected to multivariate analyses. The Principal Components Analysis (PCA) results (Fig. 5) showed that the technical replicates (of both the pooled and 2 individual accessions) were mostly relatively closely grouped, indicating that the observed distribution of accessions is due to the biological variation (accession differences) rather than the analytical variation, as expected. In addition, this PCA plot indicates that the main metabolite variation between fruits (PC1) could be explained by the differences in their approximate fruit sizes (color coded in Fig. 5): small sized (cherry type) tomatoes are generally located opposite to the bigger (beef type) fruits. Within the metabolites (loadings) underlying this separation, we could identify several flavonoids, including naringenin and quercetin derivatives, as

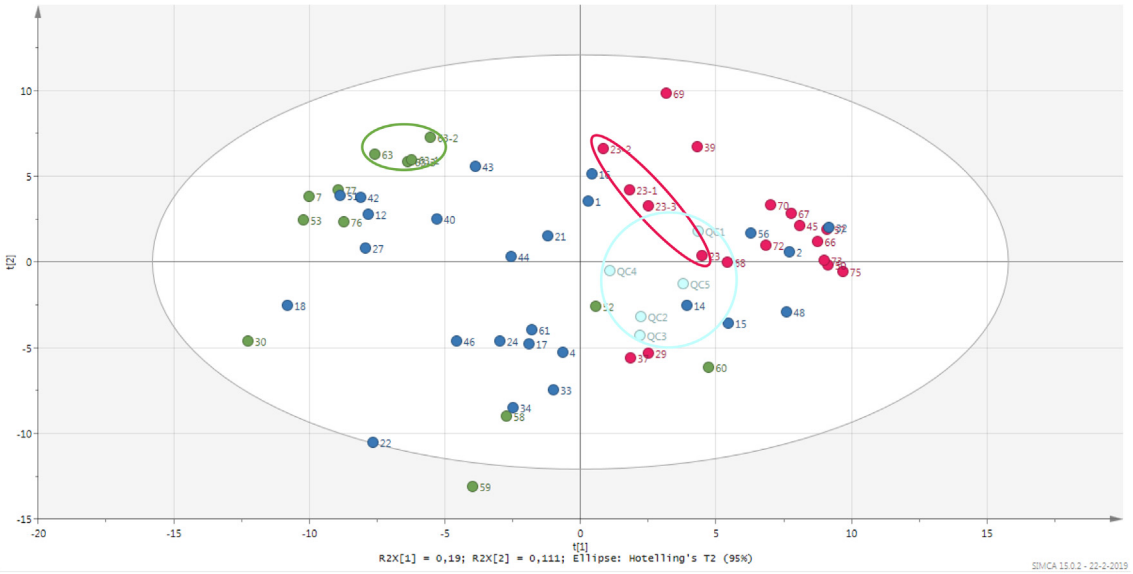

Fig. 5. PCA plot of ripe tomato fruits, based on their metabolite variation analysed using LCMS, from 48 Turkish landraces harvested in 2017. Samples are colour-coded based on their approximate fruit size, with green, blue and magenta coloured markers representing large, medium and small sized tomatoes, respectively. Extracts of the pooled fruit sample (QC1-5) are represented with turquoise markers. Green, magenta and turquoise circles comprise the technical replicates of accessions 63 and 23 and of the pooled sample, respectively. PC1 explains 19% and PC2 11% of the total metabolite variation.

being higher in the small fruited accessions. This is likely related to the fact that tomato flavonoids are known to specifically accumulate in the fruit peel (Moco et al., 2007) and the relative proportion of the peel to the whole fruit is larger in smaller fruits. On the other hand, the geographical origin of fruits did not have a clear effect on the occurrence of the metabolites in this 2017 harvest (Suppl data Fig. S4). For example, accessions coded as 22, 29, and 30 which were all originated from the same city (Ankara) were clearly separated in the PCA diagram, and indeed accessions coded with 43, 44, and 45 originating from a relatively close city, Yozgat, were also distributed separate to each other and also from the Ankara accessions.

For comparing the two harvest years (2016 and 2017), we subsequently performed LCMS analysis on fruit samples of those 37 accessions for which sufficient material had been collected in both years (Supplementary data, Table 1) and performed PCA for both years separately (Suppl data, Fig. 5a and b). A similar trend was observed between the two years, i.e. the distribution of accessions in both PCA plots was mostly related to the fruit size, which suggests that within this population, the fruit phenotype (here in terms of fruit size) has a strong influence on the overall biochemical profile, at least based on the semi-polar metabolites considered here. Similar results have also been obtained in a previous study (Adato et al., 2009). In both years, the geographic origin of accessions was less important to the overall variation of metabolite composition. On the other hand, Anton et al. (2017) investigated the polyphenols of tomatoes during ripening of fruits. They emphasized that accumulation patterns of these compounds were the same in standard-type fruits while it may change in a few situations in cherry-type cultivars and they pointed out that changes in the content of polyphenolic compounds during ripening mostly depend on cultivar. In our study, we only focused on red-ripe fruits, so including fruits at different ripening stages may change the results with respect to the variation of metabolites between different varieties.

To evaluate further the metabolic diversity and identify possible common metabolic pathways between specific accessions (Etalo et al., 2013), we applied a Hierarchical Cluster Analysis (HCA) on both accessions and metabolites, based on the variation in metabolites in fruits of the 2017 harvest (Fig. 6). Conform the results in the PCA (Fig. 5), the technical replicates of individual accessions cluster tightly together, indicating similar metabolite profiles, as expected. Based on the common variation across the accessions, several metabolite clusters were observed. For instance, glycoalkaloids such as esculeoside A, esculeoside B and acetoxy-hydroxytomatine, dihydroxytomatine as well as some other alkaloids, clustered together. The flavonoid glycosides, including quercetin 3-*O*-rutinoside, quercetin-hexose-deoxyhexose, quercetin 3-*O*-rutinoside-7-*O*-glucoside, kaempferol 3-*O*-rutinoside also clustered together as a separate group (Fig. 6). Additionally, phenylpropanoids including 5-caffeoylquinic acid, 3,4,5-tricaffeoylquinic acid, dicaffeoylquinic acid, 3,4 (or 3,5)-dicaffeoylquinic acid and 4,5-dicaffeoylquinic acid also clustered together. Furthermore, we observed a group of previously-described compounds which are known to define the contrast between so-called 'smoky' and 'non-smoky' fruit flavours (Tikunov, Molthoff, de Vos, Beekwilder, van Houwelingen, van der Hooft, et al., 2013). Thus, the non-smoky glycoside form 2-*O*-β-d-glucopyranosyl-(1 → 2)-[*O*-β-d-xylopyranosyl-(1 → 6)]-*O*-β-d-glucopyranoside (GXG) of both methyl salicylate (MeSA-GXG) and eugenol (Eugenol-GXG), as well as the malonylated form of MeSA-GXG clustered together due to their relative high abundance in 11 of the 50 accessions analyzed (highlighted by a yellow box in Supplementary data, Fig. 5; i.e. accessions TR70701 Amasya, TR61870 Denizli, TR61967 Denizli, TR69796 Ankara, TR69163 Konya, TR69818 Ankara, TR68526 Eskisehir, TR70702 Amasya, TR69817 Ankara, TR66043 Kütahya and TR71389 Kayseri). The relative high abundance of these specific glycosides suggests that these accessions contain the fully functional gene controlling the key glycosylation step that prevents the release of the smoky-flavor related volatiles methyl salicylate and eugenol upon cell disruption of the fresh fruit (Tikunov et al., 2013) as usually happens on cutting (e.g. food processing) or chewing (directly consuming) the fruits. Interestingly, in the HCA these known non-smoky compounds clustered together with some as yet so-called 'unknown compounds' (Suppl. data, Fig. 6S). Based on their elemental composition and MSMS data, we could putatively identify these new compounds as being malonylated glycosides of the flavor-related volatiles phenylethanol and benzyl alcohol, namely phenylethanol malonyl pentosyl-dihexoside (*m/z* 663.214), phenylethanol malonyl-dihexoside (*m/z* 531.172), benzyl alcohol malonyl-pentosyl-dihexoside (*m/z* 649.198) and benzyl alcohol malonyl-dihexoside (*m/z* 517.156). All the other varieties did not show the presence of these specific glycosylated forms of volatiles.

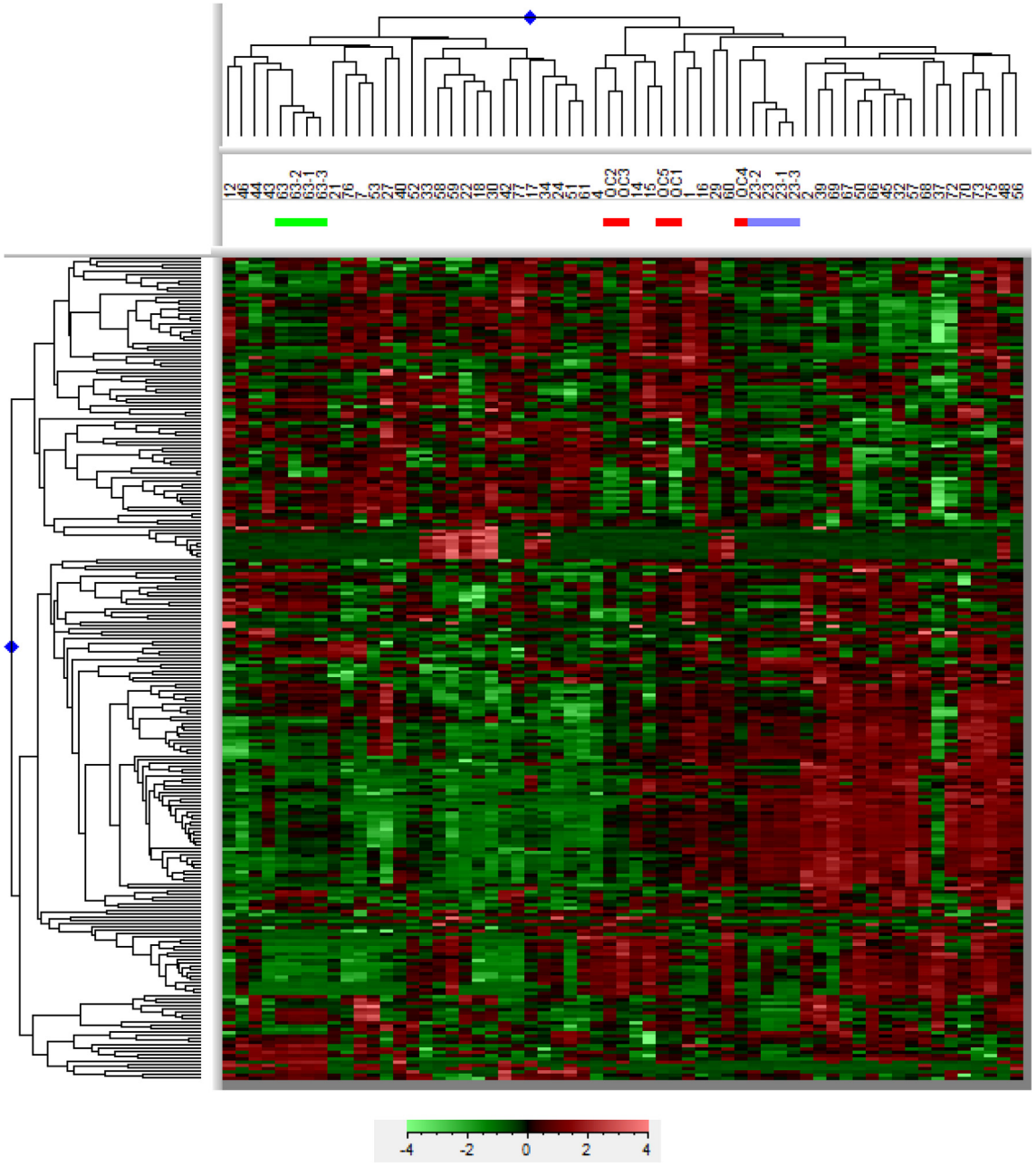

Fig. 6. Hierarchical clustering analysis (HCA) diagram and heatmap of 48 Turkish landraces and 2 hybrids (columns) from the 2017-harvest, based on similarities in their profiles of metabolites (rows) as obtained by untargeted LCMS. Green, purple and red squared genotypes represent the replications of accessions 63, 23 and quality control sample, respectively. Color scale of the heatmap represents relative intensity values after log2-transformation and normalisation across samples by mean-centring.

4. Conclusions

In this study, untargeted comprehensive LCMS-based metabolomics was used to determine biodiversity in phytochemical composition, regarding mainly semi-polar secondary metabolites, in ripe fruits of 50 unique accessions originating in Turkey (landraces). Our results of unsupervised multivariate analyses based on the variation in 265 and 251 compounds detected in harvests from 2016 and 2017, respectively, indicated that there is no clear metabolome distinction between these tomato accessions based on their geographic origin, while fruit type appears most influential to their metabolite composition. Small cherry-type fruits generally contained higher levels of peel-specific compounds, including flavonoids, than did large beef-type fruits. Within the accessions analyzed a large biodiversity was detected in various classes of compounds related to tomato fruit quality, such as flavonoids, alkaloids and smoky-flavor related compounds. This study is the first comprehensive screen of secondary metabolites in Turkish tomato landraces on this large scale.

CRediT authorship contribution statement

Sena Bakir: Methodology, Formal analysis, Investigation, Resources, Writing - original draft, Writing - review & editing, Visualization. **Esra Capanoglu:** Conceptualization, Methodology, Formal analysis, Investigation, Resources, Writing - original draft, Writing - review & editing, Supervision, Project administration. **Robert D. Hall:** Conceptualization, Resources, Writing - review & editing, Supervision, Project administration. **Ric C.H. de Vos:**

Conceptualization, Methodology, Investigation, Resources, Writing - review & editing, Supervision, Project administration.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was financially supported by The Scientific and Technological Research Council of Turkey (TUBITAK) with 2214-A-International Research Fellowship Programme for PhD student (application number 1059B141700390) and by the Istanbul Technical University, Scientific Research Projects (BAP) Unit (Project ID number: 41359).

The authors kindly express their gratitude to cultivation teams of Plant Biodiversity, Geophyte Research and Training Center Directorate, Beykoz, Istanbul in 2016 with the leadership of Halil İbrahim Tuzlacı, MSc., and Ataturk Horticultural Central Research Institute, Yalova with the leadership of Ibrahim Sönmez, PhD. The authors also thank Bert Schipper, Bioscience Wageningen-UR, for operating the HPLC-PDA-LTQ-Orbitrap FTMS system and his excellent help in sample extractions and analysis.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodchem.2020.126406.

References

Adato, A., Mandel, T., Mintz-Oron, S., Venger, I., Levy, D., Yativ, M., ... Jetter, R. (2009). Fruit-surface flavonoid accumulation in tomato is controlled by a SlMYB12-regulated transcriptional network. *PLoS Genetics, 5*(12), e1000777.

Alseekh, S., Tohge, T., Wendenberg, R., Scossa, F., Omranian, N., Li, J., ... Fernie, A. R. (2015). Identification and mode of inheritance of quantitative trait loci for secondary metabolite abundance in tomato. *The Plant Cell, 27*(3), 485–512.

Anton, D., Bender, I., Kaart, T., Roasto, M., Heinonen, M., Luik, A., & Pussa, T. (2017). Changes in Polyphenols Contents and Antioxidant Capacities of Organically and Conventionally Cultivated Tomato (Solanum lycopersicum L.) Fruits during Ripening. International Journal of Analytical Chemistry.

Astray, G., Mejuto, J. C., Martinez-Martinez, V., Nevares, I., Alamo-Sanza, M., & Simal-Gandara, J. (2019). Prediction models to control aging time in red wine. *Molecules, 24*(5).

Baldina, S., Picarella, M. E., Troise, A. D., Pucci, A., Ruggieri, V., Ferracane, R., ... Mazzucato, A. (2016). Metabolite profiling of italian tomato landraces with different fruit types. *Frontiers in Plant Science, 7*, 664.

Battino, M., Forbes-Hernandez, T. Y., Gasparrini, M., Afrin, S., Cianciosi, D., Zhang, J. J., ... Giampieri, F. (2019). Relevance of functional foods in the Mediterranean diet: The role of olive oil, berries and honey in the prevention of cancer and cardiovascular diseases. *Critical Reviews in Food Science and Nutrition, 59*(6), 893–920.

Bottino, A., Capannelli, G., Turchini, A., Della Valle, P., & Trevisan, M. (2002). Integrated membrane processes for the concentration of tomato juice. *Desalination, 148*(1–3), 73–77.

Burton-Freeman, B., Talbot, J., Park, E., Krishnankutty, S., & Edirisinghe, I. (2012). Protective activity of processed tomato products on postprandial oxidation and inflammation: A clinical trial in healthy weight men and women. *Molecular Nutrition & Food Research, 56*(4), 622–631.

Capanoglu, E., Beekwilder, J., Boyacioglu, D., De Vos, R. C. H., & Hall, R. D. (2010). The effect of industrial food processing on potentially health-beneficial tomato antioxidants. *Critical Reviews in Food Science and Nutrition, 50*(10), 919–930.

Castro-Puyana, M., Perez-Miguez, R., Montero, L., & Herrero, M. (2017). Reprint of: Application of mass spectrometry-based metabolomics approaches for food safety, quality and traceability. *TRaC-Trends in Analytical Chemistry, 96*, 62–78.

Cianciosi, D., Simal-Gandara, J., & Forbes-Hernandez, T. Y. (2019). The importance of berries in the human diet. *Mediterranean Journal of Nutrition and Metabolism, 12*(4), 335–340.

De Vos, R. C. H., Moco, S., Lommen, A., Keurentjes, J. J. B., Bino, R. J., & Hall, R. D. (2007). Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. *Nature Protocols, 2*(4), 778–791.

Di Lecce, G., Martinez-Huelamo, M., Tulipani, S., Vallverdu-Queralt, A., & Lamuela-Raventos, R. M. (2013). Setup of a UHPLC-QqQ-MS method for the analysis of phenolic compounds in cherry tomatoes, tomato sauce, and tomato juice. *Journal of Agricultural and Food Chemistry, 61*(35), 8373–8380.

Di Paola Naranjo, R. D., Otaiza, S., Saragusti, A. C., Baroni, V., Carranza, A. V., Peralta, I. E., ... Asis, R. (2016). Data on polyphenols and biological activity analyses of an Andean tomato collection and their relationships with tomato traits and geographical origin. *Data Brief, 7*, 1258–1268.

Esteki, M., Ahmadi, P., Heyden, Y. V., & Simal-Gandara, J. (2019). Fatty acids-based quality index to differentiate worldwide commercial pistachio cultivars. *Molecules, 24*(1).

Esteki, M., Farajmand, B., Kolahderazi, Y., & Simal-Gandara, J. (2017). Chromatographic fingerprinting with multivariate data analysis for detection and quantification of apricot kernel in almond powder. *Food Analytical Methods, 10*(10), 3312–3320.

Esteki, M., Regueiro, J., & Simal-Gandara, J. (2019). Tackling fraudsters with global strategies to expose fraud in the food chain. *Comprehensive Reviews in Food Science and Food Safety, 18*(2), 425–440.

Esteki, M., Shahsavari, Z., & Simal-Gandara, J. (2018). Use of spectroscopic methods in combination with linear discriminant analysis for authentication of food products. *Food Control, 91*, 100–112.

Esteki, M., Shahsavari, Z., & Simal-Gandara, J. (2019a). Food identification by high performance liquid chromatography fingerprinting and mathematical processing. *Food Research International, 122*, 303–317.

Esteki, M., Shahsavari, Z., & Simal-Gandara, J. (2019b). Gas chromatographic fingerprinting coupled to chemometrics for food authentication. *Food Reviews International*.

Esteki, M., Simal-Gandara, J., Shahsavari, Z., Zandbaaf, S., Dashtaki, E., & Heyden, Y. V. (2018). A review on the application of chromatographic methods, coupled to chemometrics, for food authentication. *Food Control, 93*, 165–182.

Etalo, D. W., Stulemeijer, I. J. E., van Esse, H. P., de Vos, R. C. H., Bouwmeester, H. J., & Joosten, M. H. A. J. (2013). System-wide hypersensitive response-associated transcriptome and metabolome reprogramming in tomato. *Plant Physiology, 162*(3), 1599–1617.

Faostat, F. (2017). *Statistical data.* Rome: Food and Agriculture Organization of the United Nations.

Ferreres, F., Taveira, M., Pereira, D. M., Valentao, P., & Andrade, B. P. (2010). Tomato (Lycopersicon esculentum) seeds: New flavonols and cytotoxic effect. *Journal of Agricultural and Food Chemistry, 58*(5), 2854–2861.

Galvez, J. F., Mejuto, J. C., & Simal-Gandara, J. (2018). Future challenges on the use of blockchain for food traceability analysis. *TRaC-Trends in Analytical Chemistry, 107*, 222–232.

Gómez-Romero, M., Segura-Carretero, A., & Fernández-Gutiérrez, A. (2010). Metabolite profiling and quantification of phenolic compounds in methanol extracts of tomato fruit. *Phytochemistry, 71*(16), 1848–1864.

Gonzalez-Fernandez, I., Iglesias-Otero, M. A., Esteki, M., Moldes, O. A., Mejuto, J. C., & Simal-Gandara, J. (2019). A critical review on the use of artificial neural networks in olive oil production, characterization and authentication. *Critical Reviews in Food Science and Nutrition, 59*(12), 1913–1926.

Herrero, M., Simo, C., Garcia-Canas, V., Ibanez, E., & Cifuentes, A. (2012). Foodomics: MS-based strategies in modern food science and nutrition. *Mass Spectrometry Reviews, 31*(1), 49–69.

Iijima, Y., Nakamura, Y., Ogata, Y., Tanaka, K., Sakurai, N., Suda, K., ... Shibata, D. (2008). Metabolite annotations based on the integration of mass spectral information. *Plant Journal, 54*(5), 949–962.

Kotíková, Z., Lachman, J., Hejtmánková, A., & Hejtmánková, K. (2011). Determination of antioxidant activity and antioxidant content in tomato varieties and evaluation of mutual interactions between antioxidants. *LWT-Food Science and Technology, 44*(8), 1703–1710.

Lommen, A. (2009). MetAlign: Interface-driven, versatile metabolomics tool for hyphenated full-scan mass spectrometry data preprocessing. *Analytical Chemistry, 81*(8), 3079–3086.

López Camelo, A. F., & Gómez, P. A. (2004). Comparison of color indexes for tomato ripening. *Horticultura Brasileira, 22*(3), 534–537.

Mazzoni, L., Perez-Lopez, P., Giampieri, F., Alvarez-Suarez, J. M., Gasparrini, M., Forbes-Hernandez, T. Y., ... Battino, M. (2016). The genetic aspects of berries: From field to health. *Journal of the Science of Food and Agriculture, 96*(2), 365–371.

Moco, Capanoglu, E., Tikunov, Y., Bino, R. J., Boyacioglu, D., Hall, R. D., Vervoort, J., & de Vos, R. C. H. (2007). Tissue specialization at the metabolite level is perceived during the development of tomato fruit. *Journal of Experimental Botany, 58*(15–16), 4131–4146.

Moco, S., Bino, R. J., Vorst, O., Verhoeven, H. A., de Groot, J., van Beek, T. A., ... de Vos, R. C. H. (2006). A liquid chromatography-mass spectrometry-based metabolome database for tomato. *Plant Physiology, 141*(4), 1205–1218.

Mokochinski, J. B., Mazzafera, P., Sawaya, A., Mumm, R., de Vos, R. C. H., & Hall, R. D. (2018). Metabolic responses of Eucalyptus species to different temperature regimes. *Journal of Integraive Plant Biology, 60*(5), 397–411.

Moldes, O. A., Mejuto, J. C., Rial-Otero, R., & Simal-Gandara, J. (2017). A critical review on the applications of artificial neural networks in winemaking technology. *Critical Reviews in Food Science and Nutrition, 57*(13), 2896–2908.

Naranjo, R. D. D., Otaiza, S., Saragusti, A. C., Baroni, V., Carranza, A. D., Peralta, I. E., ... Asis, R. (2016). Hydrophilic antioxidants from Andean tomato landraces assessed by their bioactivities in vitro and in vivo. *Food Chemistry, 206*, 146–155.

O'Kennedy, N., Crosbie, L., Whelan, S., Luther, V., Horgan, G., Broom, J. I., ... Duttaroy, A. K. (2006). Effects of tomato extract on platelet function: A double-blinded crossover study in healthy humans. *The American Journal of Clinical Nutrition, 84*(3), 561–569.

Pinu, F. R. (2016). Early detection of food pathogens and food spoilage microorganisms: Application of metabolomics. *Trends in Food Science & Technology, 54*, 213–215.

Putri, S. P., Nakayama, Y., Matsuda, F., Uchikata, T., Kobayashi, S., Matsubara, A., & Fukusaki, E. (2013). Current metabolomics: Practical applications. *Journal of Bioscience and Bioengineering, 115*(6), 579–589.

Ramos-Bueno, R. P., Romero-González, R., González-Fernández, M. J., & Guil-Guerrero, J. L. (2017). Phytochemical composition and in vitro anti-tumour activities of selected tomato varieties. *Journal of the Science of Food and Agriculture, 97*(2), 488–496.

Schmidt, A., Li, C., Shi, F., Jones, A. D., & Pichersky, E. (2011). Polymethylated myricetin in trichomes of the wild tomato species Solanum habrochaites and characterization of trichome-specific 3′/5′-and 7/4′-myricetin O-methyltransferases. *Plant Physiology, 155*(4), 1999–2009.

Shidfar, F., Froghifar, N., Vafa, M., Rajab, A., Hosseini, S., Shidfar, S., & Gohari, M. (2011). The effects of tomato consumption on serum glucose, apolipoprotein B, apolipoprotein A-I, homocysteine and blood pressure in type 2 diabetic patients. *International Journal of Food Sciences and Nutrition, 62*(3), 289–294.

Singh, S. K., Chauhan, D. S., Vatsa, M., & Singh, R. (2003). A robust skin color based face detection algorithm. *Tamkang Journal of Science and Engineering, 6*(4), 227–234.

Siracusa, L., Avola, G., Patane, C., Riggi, E., & Ruberto, G. (2013). Re-evaluation of traditional Mediterranean foods. The local landraces of 'Cipolla di Giarratana' (Allium cepa L.) and long-storage tomato (Lycopersicon esculentum L.): Quality traits and polyphenol content. *Journal of the Science of Food and Agriculture, 93*(14), 3512–3519.

Thissen, U., Coulier, L., Overkamp, K. M., Jetten, J., van der Werff, B. J. C., van de Ven, T., & van der Werf, M. J. (2011). A proper metabolomics strategy supports efficient food quality improvement: A case study on tomato sensory properties. *Food Quality and Preference, 22*(6), 499–506.

Thurman, E. M., Ferrer, I., Zweigenbaum, J. A., Garcia-Reyes, J. F., Woodman, M., & Fernandez-Alba, A. R. (2005). Discovering metabolites of post-harvest fungicides in citrus with liquid chromatography/time-of-flight mass spectrometry and ion trap tandem mass spectrometry. *Journal of Chromatography A, 1082*(1), 71–80.

Tikunov, Y. M., de Vos, R. C. H., Paramás, A. M. G., Hall, R. D., & Bovy, A. G. (2010). A role for differential glycoconjugation in the emission of phenylpropanoid volatiles from tomato fruit discovered using a metabolic data fusion approach. *Plant Physiology, 152*(1), 55–70.

Tikunov, Y. M., Molthoff, J., de Vos, R. C. H., Beekwilder, J., van Houwelingen, A., van der Hooft, J. J., ... van de Geest, H. (2013). Non-smoky glycosyltransferase1 prevents the release of smoky aroma from tomato fruit. *The Plant Cell, TPC, 113*, 114231.

Tiwari, B. K., O'Donnell, C. P., Brunton, N. P., & Cullen, P. J. (2009). Degradation kinetics of tomato juice quality parameters by ozonation. *International Journal of Food Science and Technology, 44*(6), 1199–1205.

Tohge, T., & Fernie, A. R. (2015). Metabolomics-inspired insight into developmental, environmental and genetic aspects of tomato fruit chemical composition and quality. *Plant and Cell Physiology, 56*(9), 1681–1696.

Tolstikov, V. V., & Fiehn, O. (2002). Analysis of highly polar compounds of plant origin: Combination of hydrophilic interaction chromatography and electrospray ion trap mass spectrometry. *Analytical Biochemistry, 301*(2), 298–307.

Trichopoulou, A., Martinez-Gonzalez, M. A., Tong, T. Y., Forouhi, N. G., Khandelwal, S., Prabhakaran, D., ... de Lorgeril, M. (2014). Definitions and potential health benefits of the Mediterranean diet: Views from experts around the world. *BMC Medicine, 12*, 112.

Vallverdú-Queralt, A., Jauregui, O., Medina-Remón, A., Andrés-Lacueva, C., & Lamuela-Raventós, R. M. (2010). Improved characterization of tomato polyphenols using liquid chromatography/electrospray ionization linear ion trap quadrupole Orbitrap mass spectrometry and liquid chromatography/electrospray ionization tandem mass spectrometry. *Rapid Communications in Mass Spectrometry, 24*(20), 2986–2992.

Wishart, D. S. (2008). Metabolomics: Applications to food science and nutrition research. *Trends in Food Science & Technology, 19*(9), 482–493.

Zhu, G., Wang, S., Huang, Z., Zhang, S., Liao, Q., Zhang, C., ... Huang, S. (2018). Rewiring of the fruit metabolome in tomato breeding. *Cell, 172*(1–2), 249–261 e212.