



Published in final edited form as:

*Food Chem.* 2017 August 01; 228: 270–278. doi:10.1016/j.foodchem.2017.01.118.

## A metabolomic evaluation of the phytochemical composition of tomato juices being used in human clinical trials

Morgan J. Cichon<sup>a,\*</sup>, Ken M. Riedl<sup>a,b</sup>, Steven J. Schwartz<sup>a,b</sup>

<sup>a</sup>Department of Food Science & Technology, The Ohio State University, Columbus, OH, USA

<sup>b</sup>The Ohio State University, Comprehensive Cancer Center, Columbus, OH, USA

### Abstract

Juices from the traditional red tomato and a unique *tangerine* tomato variety are being investigated as health promoting foods in human clinical trials. However, it is unknown how the *tangerine* and red tomato juices differ in biologically relevant phytochemicals beyond carotenoids. Here liquid-chromatography high-resolution mass spectrometry metabolomics was used to evaluate broadly the similarities and differences in carotenoids and other phytochemicals between red and *tangerine* tomato juices intended for clinical interventions. This untargeted approach was successful in the rapid detection and extensive characterization of phytochemicals belonging to various compound classes. The tomato juices were found to differ significantly in a number of phytochemicals, including carotenoids, chlorophylls, neutral lipids, and cinnamic acid derivatives. The largest differences were in carotenoids, including lycopene, phytoene, phytofluene, neurosporene, and  $\zeta$ -carotene. Smaller, but significant, differences were observed in polar phytochemicals, such as chlorogenic acid, hydroxyferulic acid, phloretin-di-*C*-glycoside, and isopropylmalic acid.

### Keywords

Metabolomics; Tomatoes; Phytochemicals; LC–MS; Carotenoids; Flavonoids

---

\*Corresponding author at: 2015 Fyffe Ct., 260 Parker Food Science and Technology Building, Columbus, OH 43210, USA, cichon.8@osu.edu (M.J. Cichon).

#### Conflict of interest

The authors declare no conflict of interest.

#### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.foodchem.2017.01.118>.

#### Chemical compounds studied in this article

Lycopene (PubChem CID: 446925)

Phytoene (PubChem CID: 5280784)

Phytofluene (PubChem CID: 6436722).

Neurosporene (PubChem CID: 5280789)

$\zeta$ -Carotene (PubChem CID: 5280788)

Rutin (PubChem CID: 5280805)

Naringenin (PubChem CID: 932)

Chlorogenic acid (PubChem CID: 1794427)

$\alpha$ -Tomatine (PubChem CID: 28523)

## 1. Introduction

Epidemiological research has shown a correlation between increased consumption of tomato products and a decreased risk of certain diseases, including prostate cancer (Giovannucci, 1999; Hadley, Miller, Schwartz, & Clinton, 2002). This relationship has also been observed in studies with animals fed diets supplemented with tomatoes. However, we still know little about the mechanism behind this observed protective effect. Much research has focused on the tomato carotenoid lycopene as a potential bioactive compound due to its antioxidant capacity and abundance in the tomato. Lycopene is an efficient singlet oxygen quencher and free radical scavenger (Böhm, Puspitasari-Nienaber, Ferruzzi, & Schwartz, 2002; Di Mascio, Kaiser, & Sies, 1989) and there is evidence suggesting that these properties could translate into a protective effect *in vivo*. Additionally, lycopene has been shown to accumulate in human tissues, such as the prostate, where it may have some biological effect (Clinton et al., 1996).

Research has traditionally focused on the red tomato, but in a recent human clinical study conducted by our group, lycopene from the juice of a unique *tangerine* tomato variety was found to be 8.5 times more bioavailable than lycopene from a red tomato juice (Cooperstone et al., 2015). *Tangerine* tomatoes are orange in color, which is a result of lycopene being – biosynthesized in a tetra-*cis* (7Z, 9Z, 7'Z, 9'Z) geometrical configuration rather the all-*trans* configuration found in red tomatoes. This conformational change causes a shift in the absorption spectrum of lycopene, resulting in a marked color change. The enhanced bioavailability of lycopene from the *tangerine* tomato has been attributed in part to the lipid dissolved state of tetra-*cis*-lycopene in the tomato compared to the crystalline form of all-*trans*-lycopene.

Given the significant differences in lycopene bioavailability between the red tomato and the *tangerine* tomato, our group is interested in evaluating whether this translates into a difference in biological activity. We are actively using juices developed from red and *tangerine* tomatoes as health promoting foods in human clinical interventions with prostate cancer patients. While we know that these two tomato juices differ in their carotenoid profiles, it is unknown how they differ in other potentially bioactive phytochemicals. Tomatoes contain many phenolic compounds, including flavonoids, such as naringenin and kaempferol, and hydroxycinnamic acids, such as ferulic acid and coumaric acid (Moco et al., 2006). These compounds have been shown to possess considerable bioactivity in a variety of test systems (Erlund, 2004; Heim, Tagliaferro, & Bobilya, 2002; Meyer, Donovan, Pearson, Waterhouse, & Frankel, 1998) and therefore, it is reasonable to believe that they contribute to the health benefits associated with tomatoes. In fact, some research has demonstrated an enhanced protective effect when feeding whole tomatoes versus lycopene alone (Boileau et al., 2003; Canene-Adams et al., 2007). These data suggest a synergistic effect between the carotenoids and other phytochemicals in tomatoes. In order to understand better any biological effects observed in clinical trials with food products, it is paramount to have a more comprehensive understanding of the chemical differences between the products being evaluated.

The objective of this study is to use a liquid-chromatography mass spectrometry (LC–MS)-based metabolomics approach to identify phytochemical and metabolite differences in both the polar and lipophilic fractions of the red and *tangerine* tomato juices. Untargeted metabolomic profiling allows for the unbiased detection and differential analysis of thousands of phytochemical species belonging to a number of different compound classes in a single analysis. This approach has been used to characterize phenolic compounds and other secondary metabolites in tomatoes (Gómez-Romero, Segura-Carretero, & Fernández-Gutiérrez, 2010; Moco et al., 2006) and to evaluate the effects of thermal processing on tomato phytochemicals (Capanoglu, Beekwilder, Boyacioglu, Hall, & de Vos, 2008). We hypothesize that a metabolomics approach will allow us to rapidly characterize a broad array of phytochemical differences with potential biological relevance in food-based clinical trials.

## 2. Materials and methods

### 2.1. Chemicals and standards

All solvents were from Fisher Scientific (Pittsburgh, PA, USA). Methyl tert-butyl ether (MtBE) and acetone were HPLC grade, hexanes was Optima grade, and acetonitrile (ACN) was Optima LC/MS grade. Methanol (MeOH) and water were either HPLC grade (extraction solvents) or Optima LC/MS grade (LC/MS analysis). Ammonium acetate was from J.T. Baker (Phillipsburg, NJ, USA) and formic acid was from Fisher Scientific. Chlorogenic acid standard was purchased from Cayman Chemical (Ann Arbor, MI, USA), rutin standard was purchased from Sigma-Aldrich (St. Louis, MO, USA), and 2-isopropylmalic acid and  $\alpha$ -tomatine standards were purchased from Santa Cruz Biotechnology (Dallas, TX, USA). Pheophytin a standard was generated from chlorophyll *a* standard (Sigma-Aldrich) using the method described by Pumilia et al. (2014). All-*trans*-lycopene standard was isolated from tomato paste as previously reported (Kopec et al., 2010).

### 2.2. Tomato juices

*Tangerine* tomatoes (*Solanum lycopersicum* L., hybrid FG10–314) and red tomatoes (*Solanum lycopersicum* L., hybrid PS696) were grown in the field under the same soil and climate conditions in Fremont, OH at the OSU North Central Agricultural Research Station. Both tomato varieties were processed into juice in the OSU Food Industries Center Pilot Plant in the Department of Food Science & Technology (Columbus, OH, USA). Five cans of each juice were randomly sampled for the metabolomic analyses. In order to expand compound coverage, two separate extractions were performed on the juices to extract both lipophilic and polar/semi-polar phytochemicals.

### 2.3. Preparation of lipophilic extracts

To extract lipid soluble phytochemicals, 5 mL of MeOH was added to 2 g of tomato juice, probe sonicated for 8 s, and centrifuged for 5 min at 2000 $\times$ g. The supernatant was removed and saved. The remaining pellet was then extracted with 5 mL hexane/acetone (1:1, v/v), probe sonicated for 8 s, and centrifuged for 5 min at 2000 $\times$ g. The supernatant was again removed and added to the previously saved supernatant. The hexane/acetone extraction was repeated two more times. To the pooled supernatants, 5 mL of water was added to drive

phase separation and 1 mL of the upper organic layer was dried under nitrogen. Dried extracts were stored at  $-80^{\circ}\text{C}$  for no more than 24 h before analysis by LC-QTOF-MS.

#### 2.4. Preparation of polar extracts

Polar and semi-polar phytochemicals were extracted using a method adapted from Moco et al. (2006). In summary, 3 mL of MeOH was added to 1 g of tomato juice to yield an extract of approximately 75% MeOH and 25% water. The sample was then vortexed for 10 s, bath sonicated for 15 min, and centrifuged for 10 min at  $2000\times g$ . The supernatant was subsequently removed and immediately analyzed by LC-QTOF-MS.

#### 2.5. Lipophilic phytochemical analysis by LC-QTOF-MS (APCI+)

Both tomato juice extracts were analyzed using a 1290 Infinity UHPLC system with a diode array detector (DAD) coupled to an iFunnel 6550 QTOF-MS (Agilent, Santa Clara, CA). Data were acquired using Agilent MassHunter Acquisition software (version B.06.01).

Lipophilic tomato juice extracts were redissolved in 1 mL MtBE followed by 1 mL of MeOH. Extracts were filtered through a  $0.2\ \mu\text{m}$  nylon filter and  $5\ \mu\text{L}$  were injected onto a C30 reversed phase column ( $250\ \text{mm} \times 4.6\ \text{mm}$  i.d.,  $3\ \mu\text{m}$  particle size) (YMC America, Inc., Allentown, PA, USA) maintained at  $40^{\circ}\text{C}$ . Compounds were eluted using a mobile phase of A = MeOH/MtBE/2% ammonium acetate (60:35:5, v/v/v) and B = MtBE/MeOH/2% ammonium acetate (78:20:2, v/v/v). A gradient was applied at 1.3 mL/min starting at 0% B and increasing to 35% B over 9 min, then increasing to 100% B over 6.5 min, and returning to 0% B over the final 6 min. The UHPLC system was coupled to the QTOF-MS via atmospheric pressure chemical ionization (APCI) operated in positive ion mode. The following MS parameters were used: gas temperature,  $290^{\circ}\text{C}$ ; vaporizer,  $500^{\circ}\text{C}$ ; gas flow, 13 L/min; nebulizer, 20 psig; VCap, 3500 V; corona current,  $5\ \mu\text{A}$ . Data were acquired in 2 GHz extended dynamic range (EDR) mode with approximately 20K resolution in the range of 100–1700  $m/z$  at a scan rate of 1 spectra/s. Prior to each experiment, the TOF mass axis was calibrated with ESI-L solution (Agilent). Reference solution (Agilent) containing purine and HP-0921 (hexakis(1H, 1H, 3H-tetrafluoropropoxy)phos phazine) was infused concurrently with the analytical nebulizer through the dedicated reference sprayer providing mass calibration correction for each scan.

#### 2.6. Polar phytochemical analysis by LC-QTOF-MS (ESI-)

The methanol extracts were filtered and  $2\ \mu\text{L}$  were injected onto a Zorbax Eclipse Plus RRHD C18 column ( $150\ \text{mm} \times 2.1\ \text{mm}$  i.d.,  $1.8\ \mu\text{m}$  particle size) (Agilent) maintained at  $40^{\circ}\text{C}$ . Compounds were eluted with A = water (0.1% formic acid) and B = ACN (0.1% formic acid). A 20 min gradient was applied at 0.6 mL/min starting at 5% B, increasing to 35% B over 10 min, holding at 35% B for 2 min, increasing to 75% B over 3 min, holding at 75% B for 2 min, and then returning immediately to 5% B and re-equilibrating for 3 min. The UHPLC system was coupled without flow splitting to the QTOF-MS via electrospray ionization (ESI) operated in negative ion mode. The following MS parameters were used: gas temperature,  $290^{\circ}\text{C}$ ; gas flow, 13 L/min; nebulizer, 30 psig; sheath gas temperature,  $400^{\circ}\text{C}$ ; sheath gas flow, 12 L/min; VCap, 4000 V; nozzle voltage, 2000 V. Data were acquired in 2 GHz extended dynamic range (EDR) mode with 20K resolution in the range of 50–1700

$m/z$  at a scan rate of 3 spectra/s. Mass calibration and reference sprayer operation was performed as described under Section 2.5.

## 2.7. Data processing and statistical analysis

Raw LC–MS data were processed using the Batch Recursive Feature Extraction option in the Agilent MassHunter Profinder software (version B.06.00). This option extracts mass spectral features and collapses related isotopes and common adducts ( $[M+Na]^+$ ,  $[M+K]^+$ ,  $[M+HCOO]^-$ ) into one feature. This is followed by mass and retention time alignment of all features. Only features detected in at least two juice replicates were retained. The software then implements a recursive feature extraction on the raw data where the mass and retention time results from the untargeted feature extraction in the first step are used in a targeted search. This improves the accuracy of the feature extraction by reducing the number of false negatives and false positives in the dataset, thereby increasing the quality of the data exported for differential analysis.

Extracted compounds, comprised of a neutral mass, retention time, and abundance, were exported as compound exchange files (.CEF) for further analysis using the Agilent Mass Profiler Professional (MPP) chemometrics software (version 13.1). Data were evaluated and filtered to remove low quality and inconsistent mass spectral features. Only those compounds that were present in at least 80% of the samples in the red and/or *tangerine* tomato juice groups, had a coefficient of variation <25% within a group, and an average ion abundance >100,000 were retained. Statistically significant differences between the red and *tangerine* tomato juices were determined using an unpaired *t*-test ( $P < 0.05$ ; Benjamini-Hochberg false discovery rate (FDR) multiple testing correction). Compounds differing by a fold change greater than two between the red tomato juice and the *tangerine* tomato juice were considered for MS/MS identification.

## 2.8. Compound identification

Compounds were identified using a combination of UV–vis spectra, accurate mass, isotope ratios, MS/MS fragmentation patterns, and authentic standards when available. This information was compared against published literature on the chemical composition of tomatoes and publicly available online metabolite databases Metlin (<https://metlin.scripps.edu>) and PlantCyc through the Plant Metabolic Network ([www.plantcyc.org](http://www.plantcyc.org)). Targeted MS/MS experiments for structural identification were conducted using the same LC-QTOF-MS systems and mobile phase gradients described in Sections 2.5 and 2.6. MS/MS data were collected by isolating precursor ions with a quadrupole resolution of 1.3 amu and using fixed collision energies of 10, 20, and 40 eV in the mass range of 50–1700  $m/z$  at an acquisition rate of 2 spectra/s. Molecular Structure Correlator (MSC) software (version B.05.00, Agilent) was used as needed to correlate collected MS/MS fragmentation data with candidate chemical structures. Major carotenoid, flavonoid, phenolic acid, and glycoalkaloid phytochemical species in the tomato juices were quantitated by external calibration curves using authentic standards.

### 3. Results and discussion

An untargeted metabolomics approach was taken to broadly evaluate and characterize carotenoid and other phytochemical differences between red and *tangerine* tomato juices designed for clinical interventions. Metabolomics has been used as a tool for studying food composition in a variety of studies, including those focused on cultivar variation (Dobson et al., 2008; Gómez-Romero et al., 2010), food quality (Johanningsmeier & McFeeters, 2011), and product adulteration (Jandri et al., 2014; Vaclavik, Schreiber, Lacina, Cajka, & Hajslova, 2012). Here metabolomics was used to enhance food-based cancer preventative research at a molecular level. It is important to comprehensively understand the phytochemical differences between health promoting foods, in this case red and *tangerine* tomato juices, in order to better interpret clinical outcomes from dietary interventions with these food products.

#### 3.1. Validation of untargeted metabolomics method with carotenoid results

While targeted methods exist for the analysis of carotenoids in foods (Kopec, Cooperstone, Cichon, & Schwartz, 2012), it was important to determine if these compounds and their isomers could be detected and evaluated along with other lipophilic phytochemicals using untargeted data acquisition and analysis methods. Carotenoids and other lipophilic phytochemicals were analyzed using APCI+. Of the 423 compounds detected in the lipophilic fraction following quality control filtering of the data, 352 (83%) were found to be significantly different with a corrected  $P < 0.05$  (Benjamini-Hochberg FDR multiple testing correction) between the red tomato juice and the *tangerine* tomato juice. Focusing on those compounds differing by a fold change of at least two when comparing the two juices, 299 compounds remained with 157 being significantly higher in the red tomato juice and 142 being significantly higher in the *tangerine* tomato juice.

A number of lipophilic compounds highly correlated with either the red or *tangerine* tomato juices were identified as carotenoids based on authentic standards, accurate mass, and characteristic UV-vis spectra as reported previously (Cooperstone et al., 2015). As shown in Table 1, the red tomato juice and the *tangerine* tomato juice differed significantly in their carotenoid profiles. The all-*trans* configuration of lycopene was 50-fold higher in the red tomato juice than the *tangerine* tomato juice, while the lycopene precursors phytoene and phytofluene were 4.5 and 3.6 times higher in the *tangerine* tomato juice compared to the red, respectively. The oxidative metabolite, lycopene 1,2-epoxide, was identified by UV-vis/MS and was only detected in the red tomato juice. On the other hand, three carotenoids were found to be unique to the *tangerine* tomato juice-  $\zeta$ -carotene, neurosporene, and tetra-*cis*-lycopene. Several isomers of both  $\zeta$ -carotene and neurosporene were detected in our metabolomics analysis. *Tangerine* tomatoes are rich in carotenoid *cis* isomers due to the lack of a functional carotenoid isomerase to enzymatically convert poly-*cis*-lycopene isomers to all-*trans*-lycopene. This results in the accumulation of *cis*-isomers of lycopene and its precursors  $\beta$ -carotene and neurosporene in the *tangerine* tomato. For this reason, chromatography was particularly important for our analysis and we chose a C30 analytical column to aid in the chromatographic separation of these carotenoid species. Additionally, many of these carotenoids only differ by one double bond and therefore have overlapping



first and third isotopes that require high mass resolution or chromatographic separation to distinguish. As many of the carotenoids detected and identified in this analysis have been previously reported in red and *tangerine* tomatoes (Clough & Pattenden, 1979), they served as validation of the untargeted metabolomics approach. Concentrations of the predominant carotenoids, all-*trans* lycopene, tetra-*cis* lycopene, phytoene, phytofluene, neurosporene, and  $\zeta$ -carotene, are reported in Table S1 and are in the ranges expected for these tomato varieties.

Unsupervised hierarchical clustering analysis was performed on the normalized abundance values of the 299 significantly different compounds with a fold change greater than two between the tomato varieties using Euclidean distances and Ward's linkage rule (Fig. 1A). As is shown by the heat map coloring, many of these lipophilic compounds were present in high abundance (red) in one of the juices and present in much lower abundance or absent (blue) in the other juice. These results were unexpected as it was assumed that the aforementioned carotenoids were the major differentiating lipophilic phytochemicals. In an attempt to understand and identify these other phytochemical differences, we plotted them by retention time versus mass and colored by fold change regulation (Fig. 2). In doing so, we observed clusters of compounds in the data with the same retention times, but unique masses. Interestingly, these clusters lined up with the predominant carotenoids in the tomato juices. For example, the vertical streaks at 7.9, 8.2, 9.3, 9.6, and 9.9 min appear to correspond with phytoene, phytofluene, tetra-*cis*-lycopene, neurosporene, and  $\zeta$ -carotene, respectively, in the *tangerine* tomato juice, while the streak at 15.2 min appears to correspond with all-*trans*-lycopene in the red tomato juice. These are the most abundant carotenoids in the two juices. As carotenoids are highly conjugated structures, they are fairly labile and susceptible to degradation once extracted from the plant or food matrix (Kopeck et al., 2012). Therefore, we hypothesize that these clusters are arising in-source from complex gas-phase chemistry of carotenoids and carotenoid fragments. These "metabolite streaks" would not pose a problem in a targeted analysis, but present a unique challenge when analyzing carotenoids using an untargeted metabolomics approach and we suggest researchers be aware of this issue when data mining. Here we collapsed each "metabolite streak" into one feature, which greatly reduced the complexity of the data.

### 3.2. Differences in other pigments and lipids

In addition to carotenoids, a number of other lipophilic compounds were identified and found to significantly differentiate the red and *tangerine* tomato juices. Plant pigments belonging to the chlorophyll class of compounds were identified as differentiating phytochemicals in the tomato juices. Pheophytin *a*, identified by accurate mass and confirmed by authentic standard, was found to be 5 times higher in the *tangerine* tomato juice compared to the red tomato juice. Pheophytin is a chlorophyll degradation product resulting from the chemical displacement of magnesium from the porphyrin ring of the molecule. This compound is formed with heat and is commonly found in thermally processed vegetables (Schwartz, Woo, & Von Elbe, 1981). Another chlorophyll degradation product, pyropheophytin *a*, was also found to be higher in the *tangerine* tomato juice (4-fold difference). Pyropheophytin is characterized by the loss of a carbomethoxy group from pheophytin and is also formed during thermal treatment of chlorophyll-containing foods

(Pumilia et al., 2014; Schwartz et al., 1981). All immature tomatoes contain the green pigment chlorophyll, which is enzymatically degraded during ripening (Kozukue & Friedman, 2003). As the tomatoes were thermally processed into juice for this study, the presence of both pheophytin *a* and pyropheophytin *a* can be rationalized. It is unclear if the differences in chlorophyll derivatives can be attributed to variety or the degree of maturity of the tomatoes at the time of harvest. Still, we were intrigued to learn that using an untargeted metabolomics approach we were able to identify other differentiating plant pigments beyond the more obvious carotenoids.

Our analysis also revealed significant differences in neutral lipids, namely triglycerides, between the red and *tangerine* tomato juices. The fatty acid compositions of the triglycerides in Table 1 were inferred based on their characteristic in-source fragmentation patterns (Holcapek, Jandera, Zderadicka, & Hrubá, 2003). All identified triglycerides were comprised of some combination of linoleic (C18:2), linolenic (C18:3), and palmitic (C16:0) acids, which corresponds with the fatty acids that have been previously reported in tomatoes (Lenucci et al., 2012). The positions of the fatty acids on the glycerol backbone (sn-1, -2, or -3) were not confirmed for this analysis and a number of positional isomers are possible for some of the identified triglycerides.

Carotenoids need to be incorporated into lipid micelles to be absorbed and research has demonstrated that dietary fat is important for lipid micelle formation (Erdman, Bierer, & Gugger, 1993). While tomatoes alone are not a rich source of dietary lipids, the 2–3-fold increase in triglycerides in the *tangerine* tomato juice could contribute to the observed enhanced absorption of lycopene from the *tangerine* tomato compared to the red tomato (Cooperstone et al., 2015).

### 3.3. Similarities and differences in polar phytochemicals

As with the non-polar fraction, many phytochemicals belonging to a number of different compound classes were detected in the polar extract using an untargeted approach. The complexity of the data is illustrated by the overlaid extracted ion chromatograms in Fig. 3. Polar phytochemicals, including cinnamic acid derivatives, flavonoids, and glycoalkaloids, were tentatively identified based on LC–MS lists of tomato metabolites in the literature (Adato et al., 2009; Moco et al., 2006, 2007; Vallverdú-Queralt, Jáuregui, Medina-Remón, Andrés-Lacueva, & Lamuela-Raventós, 2010). Many of these compounds were not significantly different or had a fold change less than two between the juices (Table 2). Although not rising above the 2-fold threshold, detection of such a plethora of expected tomato phytochemicals gives us confidence in the compound coverage of the QTOF-MS acquisitions. In analyzing the polar fraction of the tomato juices using ESI-, 346 (73%) of the 474 phytochemicals detected after quality control filtering were found to be significantly different with a corrected  $P < 0.05$  (Benjamini-Hochberg FDR multiple testing correction) between the red and *tangerine* tomato juices. Of those 346 phytochemicals, 77 had a fold change of at least two between the two juices with 55 being higher in the red tomato juice and 22 being higher in the *tangerine* tomato juice.

As with the lipophilic phytochemicals, we performed a hierarchical clustering analysis on the 77 significantly different compounds with a fold change greater than two between the



tomato varieties (Fig. 1B). When comparing the hierarchical clustering results for the lipophilic tomato fraction to the polar fraction, it is apparent from the heat map that the lipophilic phytochemicals vary between the juices to a greater degree than do the polar phytochemicals. While there are a number of carotenoids that are particular to the red or *tangerine* tomato, many of the polar phytochemicals and metabolites appear to be common to both tomato juices and have smaller fold change differences. For example, the abundant tomato flavonoids rutin and naringenin differ in the juices by only 1.78- and 1.48-fold, respectively, while the phenolic acids cinnamic acid and caffeic acid differ by less than 1.3-fold. Chalconaringenin was not detected in the tomato juices as it is largely converted to naringenin during thermal processing (Capanoglu et al., 2008). Our LC-MS method also allowed for the detection of tomato steroidal glycoalkaloids, including  $\alpha$ -tomatine, lycopersides, and esculeosides, which were found to be similar between the juices. Concentrations of representative compounds belonging to the flavonoid, phenolic acid, and glycoalkaloid classes can be found in Table S1.

In this study, data were reduced to focus on those compounds with a fold change of at least two between the red and *tangerine* tomato juices, which may be considered potential drivers of differences in efficacy. Here we combined literature and accurate mass database searching along with in-source and collision induced fragmentation data to identify phytochemicals of interest ( $P < 0.05$  and differing in abundance by a fold change of two or more). MS/MS experiments were performed using the same QTOF-MS system and mobile phase gradient as the untargeted analyses, yielding accurate mass information for the fragment ions as well as the parent compound. With accurate mass MS/MS spectra, the MSC software described in Section 2.8 was used to propose chemical formulas for the parent ions as well as their fragments. Through MSC, the ChemSpider database was queried for candidate chemical structures and those structures were scored by the software based on how well the observed fragments could be chemically rationalized. For example, the compound with an observed  $[M-H]^-$  mass of 353.0876 at retention time 2.99 min with the fragments 191.0565, 173.0455, 179.0352, and 135.0450, was matched with chlorogenic acid (compatibility score of 93 out of 100). Chlorogenic acid is a commonly reported phenolic compound in tomatoes (Gómez-Romero et al., 2010; Vallverdú-Queralt et al., 2010) and following tentative identification by MSC, was confirmed in the tomato juice samples using an authentic standard. This demonstrates the utility of automated MS/MS structure correlation programs for elucidating the structures of undetermined phytochemicals and metabolites.

Of the significantly different polar phytochemicals between the red and *tangerine* tomato juices, 15 were tentatively identified (Table 3). This was particularly interesting, because differences in polar phytochemicals between the red and *tangerine* tomatoes have yet to be reported. Based on the current study, it is not possible to determine if all of these differences can be ascribed to variety and it is likely that other factors, such as maturity at harvest and post-harvest handling, could have had an impact on the chemical composition of the tomato juices. Some of the phenolic compounds differentiating the red and *tangerine* tomato juices (Table 3) have reported bioactivity, including dihydrochalcones and hydroxycinnamic acids. These differences are particularly important as they have the potential to translate into differences in clinical efficacy and nutritional impact between the tomato juices. Of course,

activity *in vivo* will also depend on the individual concentrations of these phytochemicals in the tomato juices and their bioavailability and metabolism in humans.

The dihydrochalcone phloretin-di-*C*-glycoside has been reported to be a significant flavonoid in tomatoes as analyzed by NMR (Slimestad, Fossen, & Verheul, 2008). In fact, it has been found to be present in tomatoes at 5–14% of the total flavonoid content. We have tentatively identified this dihydrochalcone in the tomato juices and detected it at levels greater than 2.5 times higher in the red tomato juice compared to the *tangerine* tomato juice. Gómez-Romero et al. (2010) similarly found significant differences in phloretin-di-*C*-glycoside among different Spanish tomato cultivars with levels being much higher in the red tomato variety compared to the greener colored varieties analyzed. Research has demonstrated a chemoprotective effect of phloretins through suppression of inflammation (Chang, Huang, & Liou, 2012; Lee et al., 2011), inhibition of cell proliferation (Devi & Das, 1993), and induction of apoptosis (Kim, Kwon, Kang, Lee, & Lee, 2009; Kobori, Shinmoto, Tsushida, & Shinohara, 1997).

The hydroxycinnamic acid derivatives coumaric acid-hexose, chlorogenic acid, hydroxyferulic acid, and ferulic acid-hexose were also found to be significantly different between the red and *tangerine* tomato juices. Both chlorogenic acid and ferulic acid have known antioxidant activity *in vitro* and have been reported to inhibit tumor promotion in mice (Huang, Smart, Wong, & Conney, 1988). Chlorogenic acid (3-*O*-caffeoylquinic acid) is the conjugate of coumaric acid with quinic acid and was approximately two times higher in the *tangerine* tomato juice compared to the red tomato juice. Research has suggested that chlorogenic acid may protect against carcinogenesis by modulating phase two detoxifying enzymes (Feng et al., 2005). It has also been shown to significantly attenuate intestinal glucose absorption (Welsch, Lachance, & Wasserman, 1989).

Unlike chlorogenic acid, the ferulic acid derivative hydroxyferulic acid was found to be over two times higher in the red tomato juice. The position of the hydroxyl group could not be determined in this study as the compound had the same accurate mass and retention time as 5-hydroxyferulic acid, but different MS/MS fragmentation. A hexose derivative of ferulic acid was detected at a 4.6-fold higher abundance in the *tangerine* tomato juice compared to the red. Ferulic acid is an effective free radical scavenger, which is believed to be one of the predominant mechanisms through which it exerts a biological effect in humans (Srinivasan, Sudheer, & Menon, 2007). It has been reported to protect against a number of conditions including inflammation, cancer, neurodegeneration, and diabetes.

#### 4. Conclusions

The objective for using an unbiased or untargeted metabolomics approach was to capture a greater number of phytochemical differences between two tomato products being used in human clinical trials than could be accomplished with conventional targeted approaches. Through our analyses we were able to not only detect differences in carotenoid composition, but we were also able to detect and identify other phytochemical similarities and differences that we would have likely missed had we focused on only select groups of compounds. Additionally, an untargeted metabolomics approach allowed us to gain a rich chemical

profile of these foods and avoid the error of assigning a single compound to represent an entire class. Here we demonstrate the effectiveness of using untargeted metabolomics to broadly characterize the chemical composition of foods in effort to enhance nutritional and dietary interventions being conducted with these products.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

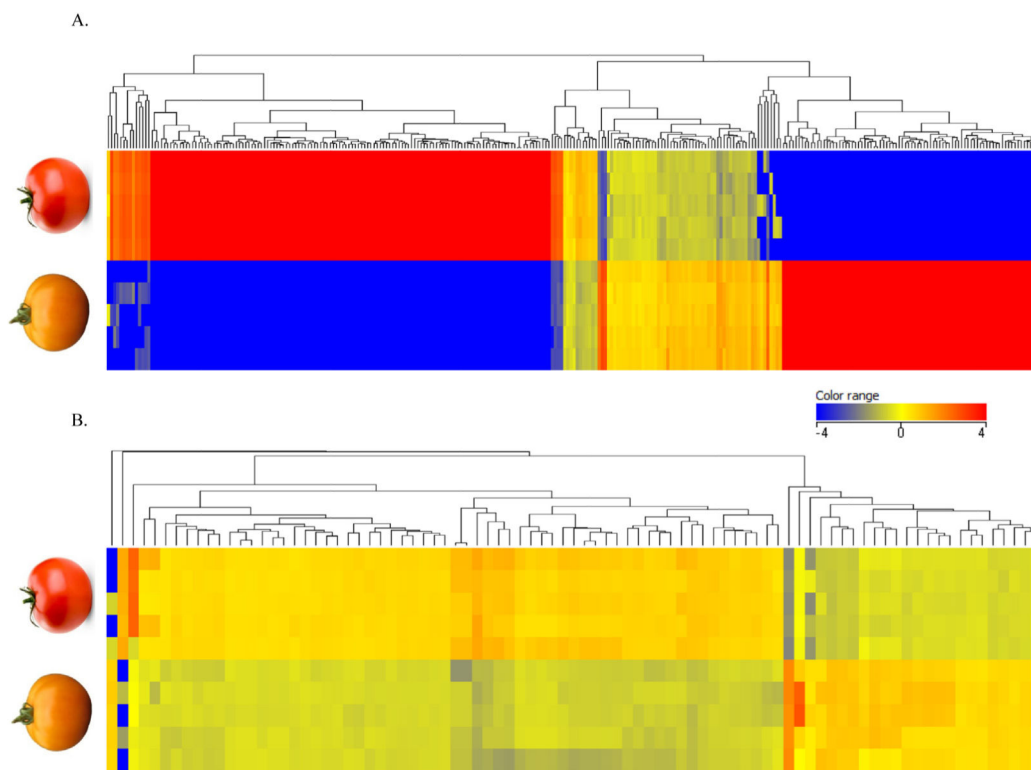
We would like to acknowledge David Francis, Ph.D. (Department of Horticulture and Crop Science, The Ohio State University) for providing the tomatoes used in this study and Dan Cuthbertson, Ph.D. (Agilent Technologies) for his assistance with the metabolomics analysis. This study was performed within the Nutrient and Phytochemical Analytics Shared Resource under The Ohio State University Comprehensive Cancer Center (NIH P30 CA016058).

## References

- Adato A, Mandel T, Mintz-Oron S, Venger I, Levy D, Yativ M, ... Aharoni A (2009). Fruit-surface flavonoid accumulation in tomato is controlled by a SLMYB12-regulated transcriptional network. *PLoS Genetics*, 5(12), e1000777. 10.1371/journal.pgen.1000777. [PubMed: 20019811]
- Böhm V, Puspitasari-Nienaber NL, Ferruzzi MG, & Schwartz SJ (2002). Trolox equivalent antioxidant capacity of different geometrical isomers of  $\alpha$ -carotene,  $\beta$ -carotene, lycopene, and zeaxanthin. *Journal of Agricultural and Food Chemistry*, 50(1), 221–226. 10.1021/jf010888q. [PubMed: 11754571]
- Boileau TW-M, Liao Z, Kim S, Lemeshow S, Erdman JW, & Clinton SK (2003). Prostate carcinogenesis in N-methyl-N-nitrosourea (NMU)-testosterone-treated rats fed tomato powder, lycopene, or energy-restricted diets. *Journal of the National Cancer Institute*, 95(21), 1578–1586. 10.1093/jnci/djg081. [PubMed: 14600090]
- Canene-Adams K, Lindshield BL, Wang S, Jeffery EH, Clinton SK, & Erdman JW (2007). Combinations of tomato and broccoli enhance antitumor activity in Dunning r3327-h prostate adenocarcinomas. *Cancer Research*, 67(2), 836–843. 10.1158/0008-5472.CAN-06-3462. [PubMed: 17213256]
- Capanoglu E, Beekwilder J, Boyacioglu D, Hall R, & de Vos R (2008). Changes in antioxidant and metabolite profiles during production of tomato paste. *Journal of Agricultural and Food Chemistry*, 56(3), 964–973. 10.1021/jf072990e. [PubMed: 18205308]
- Chang WT, Huang WC, & Liou CJ (2012). Evaluation of the anti-inflammatory effects of phloretin and phlorizin in lipopolysaccharide-stimulated mouse macrophages. *Food Chemistry*, 134(2), 972–979. 10.1016/j.foodchem.2012.03.002. [PubMed: 23107715]
- Clinton SK, Emehiser C, Schwartz SJ, Bostwick DG, Williams AW, Moore BJ, & Erdman JW (1996). Cis-trans lycopene isomers, carotenoids, and retinol in the human prostate. *Cancer Epidemiology, Biomarkers & Prevention*, 5(10), 823–833. Retrieved from <http://cebp.aacrjournals.org/content/5/10/823.short>.
- Clough JM, & Pattenden G (1979). Naturally occurring poly-cis carotenoids. Stereochemistry of poly-cis lycopene and its congeners in “Tangerine” tomato fruits. *Journal of the Chemical Society, Chemical Communications*, 14, 616–619. 10.1039/c39790000616.
- Cooperstone JL, Ralston RA, Riedl KM, Haufe TC, Schweiggert RM, King SA, ... Schwartz SJ (2015). Enhanced bioavailability of lycopene when consumed as cis-isomers from tangerine compared to red tomato juice, a randomized, cross-over clinical trial. *Molecular Nutrition and Food Research*, 59 (4), 658–669. 10.1002/mnfr.201400658. [PubMed: 25620547]
- Devi MA, & Das NP (1993). In vitro effects of natural plant polyphenols on the proliferation of normal and abnormal human lymphocytes and their secretions of interleukin-2. *Cancer Letters*, 69(3), 191–196. 10.1016/0304-3835(93)90174-8. [PubMed: 8513446]

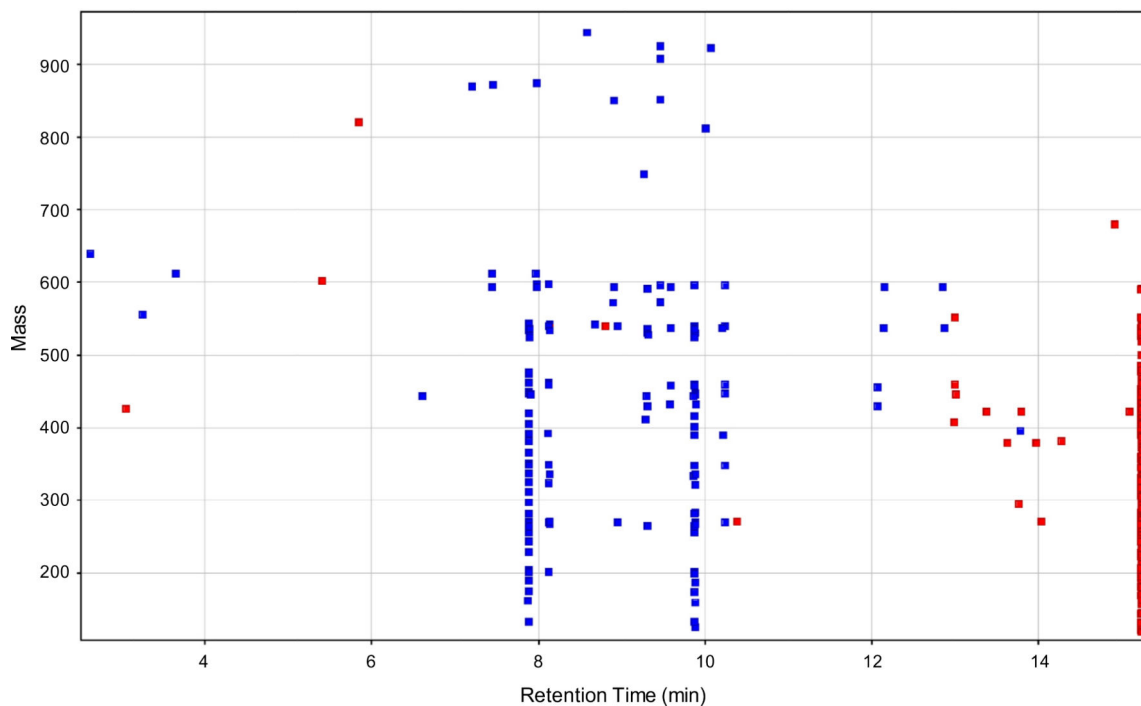
- Di Mascio P, Kaiser S, & Sies H (1989). Lycopene as the most efficient biological carotenoid singlet oxygen quencher. *Archives of Biochemistry and Biophysics*, 274 (2), 532–538. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/2802626>. [PubMed: 2802626]
- Dobson G, Shepherd T, Verrall SR, Conner S, McNicol JW, Ramsay G, ... Stewart D (2008). Phytochemical diversity in tubers of potato cultivars and landraces using a GC-MS metabolomics approach. *Journal of Agricultural and Food Chemistry*, 56(21), 10280–10291. 10.1021/jf801370b. [PubMed: 18937493]
- Erdman JW, Bierer TL, & Gugger ET (1993). Absorption and transport of carotenoids. *Annals of the New York Academy of Sciences*, 691, 76–85. 10.1111/j.1749-6632.1993.tb26159.x. [PubMed: 8129321]
- Erlund I (2004). Review of the flavonoids quercetin, hesperetin, and naringenin. Dietary sources, bioactivities, bioavailability, and epidemiology. *Nutrition Research*, 24(10), 851–874. 10.1016/j.nutres.2004.07.005.
- Feng R, Lu Y, Bowman LL, Qian Y, Castranova V, & Ding M (2005). Inhibition of activator protein-1, NF- $\kappa$ B, and MAPKs and induction of phase 2 detoxifying enzyme activity by chlorogenic acid. *Journal of Biological Chemistry*, 280(30), 27888–27895. 10.1074/jbc.M503347200. [PubMed: 15944151]
- Giovannucci E (1999). Tomatoes, tomato-based products, lycopene, and cancer: review of the epidemiologic literature. *Journal of the National Cancer Institute*, 91 (4), 317–331. [PubMed: 10050865]
- 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. 10.1016/j.phytochem.2010.08.002. [PubMed: 20810136]
- Hadley CW, Miller EC, Schwartz SJ, & Clinton SK (2002). Tomatoes, lycopene, and prostate cancer: Progress and promise. *Experimental Biology and Medicine*, 227(10), 869–880. [PubMed: 12424328]
- Heim KE, Tagliaferro AR, & Bobilya DJ (2002). Flavonoid antioxidants: Chemistry, metabolism and structure-activity relationships. *Journal of Nutritional Biochemistry*, 13(10), 572–584. 10.1016/S0955-2863(02)00208-5. [PubMed: 12550068]
- Holcapek M, Jandera P, Zderadicka P, & Hrubá L (2003). Characterization of triacylglycerol and diacylglycerol composition of plant oils using high-performance liquid chromatography-atmospheric pressure chemical ionization mass spectrometry. *Journal of Chromatography A*, 1010(2), 195–215. [PubMed: 12974290]
- Huang M-T, Smart RC, Wong C-Q, & Conney AH (1988). Inhibitory effect of curcumin, chlorogenic acid, caffeic acid, and ferulic acid on tumor promotion in mouse skin by 12-O-tetradecanoylphorbol-13-acetate. *Cancer Research*, 48(21), 5941–5946. [PubMed: 3139287]
- Jandri Z, Roberts D, Rathor MN, Abraham A, Islam M, & Cannavan A (2014). Assessment of fruit juice authenticity using UPLC-QToF MS: A metabolomics approach. *Food Chemistry*, 148, 7–17. 10.1016/j.foodchem.2013.10.014. [PubMed: 24262519]
- Johanningsmeier SD, & McFeeters RF (2011). Detection of volatile spoilage metabolites in fermented cucumbers using nontargeted, comprehensive 2-dimensional gas chromatography-time-of-flight mass spectrometry (GCxGC-TOFMS). *Journal of Food Science*, 76(1), C168–C177. 10.1111/j.1750-3841.2010.01918.x. [PubMed: 21535646]
- Kim M-S, Kwon JY, Kang NJ, Lee KW, & Lee HJ (2009). Phloretin induces apoptosis in H-Ras MCF10A human breast tumor cells through the activation of p53 via JNK and p38 mitogen-activated protein kinase signaling. *Annals of the New York Academy of Sciences*, 1171, 479–483. 10.1111/j.1749-6632.2009.04692.x. [PubMed: 19723092]
- Kobori M, Shinmoto H, Tsushida T, & Shinohara K (1997). Phloretin-induced apoptosis in B16 melanoma 4A5 cells by inhibition of glucose transmembrane transport. *Cancer Letters*, 119(2), 207–212. 10.1016/S0304-3835(97)00271-1. [PubMed: 9570373]
- Kopec RE, Cooperstone JL, Cichon MJ, & Schwartz SJ (2012). Analysis methods of carotenoids In Xu Z & Howard LR (Eds.), *Analysis of antioxidant rich phytochemicals* (pp. 105–148). Blackwell, Oxford, UK: Wiley-Blackwell 10.1002/9781118229378.ch4.

- Kopec RE, Riedl KM, Harrison EH, Curley RW, Hruszkewycz DP, Clinton SK, & Schwartz SJ (2010). Identification and quantification of apo-lycopenals in fruits, vegetables, and human plasma. *Journal of Agricultural and Food Chemistry*, 58(6), 3290–3296. 10.1021/jf100415z. [PubMed: 20178389]
- Kozukue N, & Friedman M (2003). Tomatine, chlorophyll, b-carotene and lycopene content in tomatoes during growth and maturation. *Journal of the Science of Food and Agriculture*, 83(3), 195–200. 10.1002/jsfa.1292.
- Lee JH, Regmi SC, Kim JA, Cho MH, Yun H, Lee CS, & Lee J (2011). Apple flavonoid phloretin inhibits *Escherichia coli* O157:H7 biofilm formation and ameliorates colon inflammation in rats. *Infection and Immunity*, 79(12), 4819–4827. 10.1128/IAI.05580-11. [PubMed: 21930760]
- Lenucci MS, Serrone L, De Caroli M, Fraser PD, Bramley PM, Piro G, & Dalessandro G (2012). Isoprenoid, lipid, and protein contents in intact plastids isolated from mesocarp cells of traditional and high-pigment tomato cultivars at different ripening stages. *Journal of Agricultural and Food Chemistry*, 60(7), 1764–1775. 10.1021/jf204189z. [PubMed: 22264157]
- Meyer AS, Donovan JL, Pearson DA, Waterhouse AL, & Frankel EN (1998). Fruit hydroxycinnamic acids inhibit human low-density lipoprotein oxidation in vitro. *Journal of Agricultural and Food Chemistry*, 46(5), 1783–1787. 10.1021/jf9708960.
- Moco S, Bino RJ, Vorst O, Verhoeven HA, de Groot J, van Beek TA, ... de Vos CHR (2006). A liquid chromatography-mass spectrometry-based metabolome database for tomato. *Plant Physiology*, 141(4), 1205–1218. 10.1104/pp.106.078428. [PubMed: 16896233]
- Moco S, Capanoglu E, Tikunov Y, Bino RJ, Boyacioglu D, Hall RD, ... De Vos RCH. (2007). Tissue specialization at the metabolite level is perceived during the development of tomato fruit. *Journal of Experimental Botany*, 58(15–16), 4131–4146. 10.1093/jxb/erm271. [PubMed: 18065765]
- Pumilia G, Cichon MJ, Cooperstone JL, Giuffrida D, Dugo G, & Schwartz SJ (2014). Changes in chlorophylls, chlorophyll degradation products and lutein in pistachio kernels (*Pistacia vera* L) during roasting. *Food Research International*, 65 (Part B), 193–198. 10.1016/j.foodres.2014.05.047.
- Schwartz SJ, Woo SL, & Von Elbe JH (1981). High-performance liquid chromatography of chlorophylls and their derivatives in fresh and processed spinach. *Journal of Agricultural and Food Chemistry*, 29(3), 533–535. 10.1021/jf00105a025.
- Slimestad R, Fossen T, & Verheul MJ (2008). The flavonoids of tomatoes. *Journal of Agricultural and Food Chemistry*, 56(7), 2436–2441. 10.1021/jf073434n. [PubMed: 18318499]
- Srinivasan M, Sudheer AR, & Menon VP (2007). Ferulic acid: Therapeutic potential through its antioxidant property. *Journal of Clinical Biochemistry and Nutrition*, 40(2), 92–100. 10.3164/jcbn.40.92. [PubMed: 18188410]
- Vaclavik L, Schreiber A, Lacina O, Cajka T, & Hajslova J (2012). Liquid chromatography-mass spectrometry-based metabolomics for authenticity assessment of fruit juices. *Metabolomics*, 8(5), 793–803. 10.1007/s11306-011-0371-7.
- Vallverdú-Queralt A, Jáuregui O, Medina-Remón A, Andrés-Lacueva C, & Lamuela-Raventós RM (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. 10.1002/rcm.4731. [PubMed: 20872631]
- Welsch CA, Lachance PA, & Wasserman BP (1989). Dietary phenolic compounds: Inhibition of Na<sup>+</sup>-dependent D-glucose uptake in rat intestinal brush border membrane vesicles. *The Journal of Nutrition*, 119(11), 1698–1704. [PubMed: 2600675]

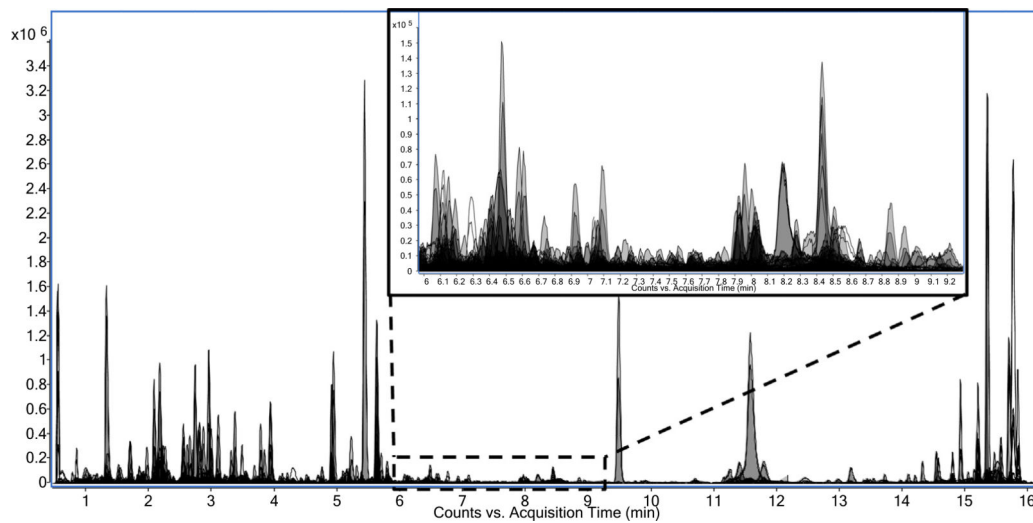


**Fig. 1.** Heat map from hierarchical clustering analysis (Euclidean distance and Ward's linkage rule) performed on significantly different compounds in the red and *tangerine* tomato juices with a fold change  $>2$  in the lipophilic fraction (A) and polar/semi-polar fraction (B). The heat map is colored by normalized relative ion intensities, with red representing relatively high abundance compounds and blue representing relatively low abundance or absent compounds.





**Fig. 2.** Retention time versus mass plot of significantly different non-polar compounds detected in the tomato juices ( $P < 0.05$ ; fold change  $> 2$ ) showing the compound clusters (“metabolite streaks”) in the data. Red colored compounds were more abundant in the red tomato juice, while blue colored compounds were more abundant in the *tangerine* tomato juice.



**Fig. 3.** Overlaid extracted ion chromatograms of polar compounds detected in the tomato juices illustrating the complexity of the data. The inset is zoomed to the region between 6 and 9.3 min to show the many compounds detected within the dynamic range of the instrument.

Table 1

Identified carotenoids and other lipophilic phytochemicals differentiating red and tangerine tomato juices.

Retention Time (min)	Observed [M+H] <sup>+</sup>	Tentative ID	Chemical Formula	Mass Error ( ppm)	Fold-Change (Red vs. Tangerine)
<i>Carotenoids</i>					
7.92	545.5085	Phytoene	C <sub>40</sub> H <sub>64</sub>	0.73	-4.46
8.17	543.4926	Phytofluene	C <sub>40</sub> H <sub>62</sub>	0.37	-3.60
8.92	541.4762	ζ-Carotene isomer	C <sub>40</sub> H <sub>60</sub>	-1.11	<i>b</i>
9.33	537.4458	Tetra- <i>cis</i> lycopene	C <sub>40</sub> H <sub>56</sub>	0.56	<i>b</i>
9.62	539.4608	Neurosporene isomer	C <sub>40</sub> H <sub>58</sub>	-0.56	<i>b</i>
9.91	541.4769	ζ-Carotene isomer	C <sub>40</sub> H <sub>60</sub>	0.18	<i>b</i>
10.28	541.4760	ζ-Carotene isomer	C <sub>40</sub> H <sub>60</sub>	-1.48	<i>b</i>
12.18	539.4603	Neurosporene isomer	C <sub>40</sub> H <sub>58</sub>	-1.48	<i>b</i>
12.89	539.4609	Neurosporene isomer	C <sub>40</sub> H <sub>58</sub>	-0.37	<i>b</i>
13.03	553.4404	Lycopene 1,2-epoxide	C <sub>40</sub> H <sub>56</sub> O	0.00	<i>c</i>
15.27	537.4461	All- <i>trans</i> lycopene (S)	C <sub>40</sub> H <sub>56</sub>	1.12	49.6
<i>Chlorophyll derivatives</i>					
7.24	871.5726	Pheophytin <i>a</i> (S)	C <sub>55</sub> H <sub>74</sub> N <sub>4</sub> O <sub>5</sub>	-0.69	-5.06
10.04	813.5681	Pyropheophytin <i>a</i>	C <sub>53</sub> H <sub>72</sub> N <sub>4</sub> O <sub>3</sub>	0.49	-4.15
<i>Neutral lipids</i>					
7.48	613.4816	DG (18:3/18:3/0:0)	C <sub>39</sub> H <sub>64</sub> O <sub>5</sub>	-1.79	-2.67
7.49	873.6969	TG (18:3/18:3/18:3)	C <sub>57</sub> H <sub>92</sub> O <sub>6</sub>	0.23	-5.48
8.01	875.7122	TG (18:3/18:3/18:2)[iso3] <sup>a</sup>	C <sub>57</sub> H <sub>94</sub> O <sub>6</sub>	-0.11	-3.21
8.94	851.7121	TG (18:3/18:3/16:0)[iso3] <sup>a</sup>	C <sub>55</sub> H <sub>94</sub> O <sub>6</sub>	-0.23	-2.99
9.50	853.7278	TG (18:3/18:2/16:0)[iso6] <sup>a</sup>	C <sub>55</sub> H <sub>96</sub> O <sub>6</sub>	-0.23	-2.18

(S) = standard confirmed.

TG: Triglyceride; DG: Diglyceride.

<sup>a</sup>Fatty acid positions on the glycerol backbone were not confirmed. Number of positional isomers possible is listed in brackets.

Compound was not detected in the red tomato juice.  
Compound was not detected in the tangerine tomato juice.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Tomato phytochemicals identified in the juices with P &gt; 0.05 and/or fold change &lt;2.

Retention Time (min)	Observed [M-H] <sup>-</sup>	Tentative ID	Chemical Formula	Mass Error ( ppm)	P-value (Corr)	Fold-Change (red vs. tangerine) <sup>a</sup>
<i>Cinnamic acid derivatives</i>						
1.37	147.0450	Cinnamic acid	C <sub>9</sub> H <sub>8</sub> O <sub>2</sub>	-1.4	6.73E-07	1.26
2.13	353.0877	Caffeoylquinic acid (I)	C <sub>16</sub> H <sub>18</sub> O <sub>9</sub>	-0.3	2.50E-02	-1.02
2.22	341.0864	Caffeic acid-hexose (I)	C <sub>15</sub> H <sub>18</sub> O <sub>9</sub>	-4.1	3.83E-09	1.36
2.28	325.0920	Coumaric acid-hexose (I)	C <sub>15</sub> H <sub>18</sub> O <sub>8</sub>	-2.8	1.64E-08	1.34
2.36	341.0868	Caffeic acid-hexose (II)	C <sub>15</sub> H <sub>18</sub> O <sub>9</sub>	-2.9	6.74E-01	-
2.60	341.0876	Caffeic acid-hexose (III)	C <sub>15</sub> H <sub>18</sub> O <sub>9</sub>	-0.6	2.50E-07	1.44
2.66	327.1083	Hydroxycinnamic acid-hexose	C <sub>15</sub> H <sub>20</sub> O <sub>8</sub>	-0.6	3.40E-10	1.82
2.77	355.1032	Ferulic acid-hexose (I)	C <sub>16</sub> H <sub>20</sub> O <sub>9</sub>	-0.8	9.67E-01	-
2.84	341.0877	Caffeic acid-hexose (IV)	C <sub>15</sub> H <sub>18</sub> O <sub>9</sub>	-0.3	2.28E-06	1.25
2.91	341.0877	Caffeic acid-hexose (V)	C <sub>15</sub> H <sub>18</sub> O <sub>9</sub>	-0.3	2.87E-07	1.49
3.14	353.0872	Caffeoylquinic acid (III)	C <sub>16</sub> H <sub>18</sub> O <sub>9</sub>	-1.7	7.57E-08	-1.33
3.36	179.0354	Caffeic acid	C <sub>9</sub> H <sub>8</sub> O <sub>4</sub>	2.2	7.13E-08	1.28
6.11	515.1191	Dicafeoylquinic acid (I)	C <sub>25</sub> H <sub>24</sub> O <sub>12</sub>	-0.8	1.91E-03	-1.21
6.33	515.1190	Dicafeoylquinic acid (II)	C <sub>25</sub> H <sub>24</sub> O <sub>12</sub>	-1.0	1.70E-04	-1.40
6.80	515.1191	Dicafeoylquinic acid (III)	C <sub>25</sub> H <sub>24</sub> O <sub>12</sub>	-0.8	1.27E-04	-1.45
8.05	312.1240	Feruloyltyramine	C <sub>18</sub> H <sub>19</sub> NO <sub>4</sub>	-0.3	4.93E-05	-1.61
8.89	677.1504	Tricafeoylquinic acid	C <sub>34</sub> H <sub>30</sub> O <sub>15</sub>	-1.2	1.49E-01	-
<i>Flavonoids</i>						
3.15	903.2398	Quercetin-dihexose-pentose-deoxyhexose	C <sub>38</sub> H <sub>48</sub> O <sub>25</sub>	-1.5	3.82E-03	1.26
3.45	771.1986	Quercetin-dihexose-deoxyhexose	C <sub>33</sub> H <sub>40</sub> O <sub>21</sub>	-0.4	3.94E-02	1.16
3.50	595.1662	Naringenin-dihexose	C <sub>27</sub> H <sub>32</sub> O <sub>15</sub>	-1.0	6.29E-01	-
4.98	741.1879	Quercetin-3-O-trisaccharide	C <sub>32</sub> H <sub>38</sub> O <sub>20</sub>	-0.7	3.69E-05	1.66
5.46	449.1086	Dihydrokaempferol-hexoside	C <sub>21</sub> H <sub>22</sub> O <sub>11</sub>	-0.7	3.07E-05	1.64
5.48	609.1457	Rutin (S)	C <sub>27</sub> H <sub>30</sub> O <sub>16</sub>	-0.7	2.52E-05	1.78

Retention Time (min)	Observed [M-H] <sup>-</sup>	Tentative ID	Chemical Formula	Mass Error ( ppm)	P-value (Corr)	Fold-Change (red vs. tangerine) <sup>a</sup>
6.19	593.1505	Kaempferol 3-O-rutinoside	C <sub>27</sub> H <sub>30</sub> O <sub>15</sub>	-1.2	2.00E-04	1.52
6.45	433.1137	Naringenin-hexose	C <sub>21</sub> H <sub>22</sub> O <sub>10</sub>	-0.7	8.38E-01	-
7.13	917.2340	Quercetin-dihexose-deoxyhexose- <i>p</i> -coumaric acid	C <sub>42</sub> H <sub>46</sub> O <sub>23</sub>	-1.9	1.84E-04	-1.46
9.52	271.0602	Naringenin	C <sub>15</sub> H <sub>12</sub> O <sub>5</sub>	-3.7	2.35E-03	-1.48
<i>Glycoalkaloids</i>						
4.96	1226.5790	Esculeoside B (I)	C <sub>56</sub> H <sub>93</sub> NO <sub>28</sub>	-1.7	5.49E-02	-
5.67	1226.5785	Esculeoside B (II)	C <sub>56</sub> H <sub>93</sub> NO <sub>28</sub>	-1.7	1.08E-03	-1.17
5.83	1268.5872	Lycoperside G or Lycoperside F or Esculeoside A	C <sub>58</sub> H <sub>95</sub> NO <sub>29</sub>	-3.5	2.71E-04	1.29
6.79	1048.5310	Lycoperside H or Hydroxy- $\alpha$ -tomatine (I)	C <sub>50</sub> H <sub>83</sub> NO <sub>22</sub>	-2.3	1.28E-05	-1.49
6.58	1048.5310	Lycoperside H or Hydroxy- $\alpha$ -tomatine (II)	C <sub>50</sub> H <sub>83</sub> NO <sub>22</sub>	-2.3	5.08E-06	-1.53
8.70	1032.5360	$\alpha$ -Tomatine (S)	C <sub>50</sub> H <sub>83</sub> NO <sub>21</sub>	-2.4	2.58E-04	-1.70
<i>Others</i>						
1.30	299.0771	Hydroxybenzoic acid-hexose	C <sub>13</sub> H <sub>16</sub> O <sub>8</sub>	-0.3	8.48E-08	1.40
1.37	164.0715	Phenylalanine	C <sub>9</sub> H <sub>11</sub> NO <sub>2</sub>	-1.2	2.24E-07	1.26
1.75	380.1563	Pantothenic acid-hexose	C <sub>15</sub> H <sub>27</sub> NO <sub>10</sub>	0.3	6.13E-03	1.04
1.86	153.0197	Protocatechuic acid	C <sub>7</sub> H <sub>6</sub> O <sub>4</sub>	2.6	4.21E-05	1.20
2.13	203.0826	Tryptophan	C <sub>11</sub> H <sub>12</sub> N <sub>2</sub> O <sub>2</sub>	0.0	3.06E-05	1.59
3.52	401.1449	Benzyl alcohol-hexose-pentose	C <sub>18</sub> H <sub>26</sub> O <sub>10</sub>	-1.0	1.98E-08	1.32

(S) = standard confirmed.

<sup>a</sup>Fold changes are reported for compounds with a P < 0.05.



Table 3

Identified polar and semi-polar phytochemicals differentiating the red and *tangerine* tomato juices.

Retention Time (min)	Observed [M-H] <sup>-</sup>	Tentative ID	Chemical Formula	Mass Error ( ppm)	MS/MS fragments	Fold-Change (red vs. <i>tangerine</i> )
<i>Cinnamic acid derivatives</i>						
2.99	353.0876	Chlorogenic acid (S)	C <sub>16</sub> H <sub>18</sub> O <sub>9</sub>	-0.6	191.0565, 173.0455, 179.0352, 135.0450	-2.06
3.04	325.0927	Coumaric acid-hexose (II)	C <sub>15</sub> H <sub>18</sub> O <sub>8</sub>	-0.6	163.0398, 119.0500	2.38
3.59	355.1032	Ferulic acid-hexose (II)	C <sub>16</sub> H <sub>20</sub> O <sub>9</sub>	-0.7	175.0398, 193.0506, 160.0167	-4.39
3.73	209.0458	Hydroxyferulic acid	C <sub>10</sub> H <sub>10</sub> O <sub>5</sub>	1.2	137.0606, 165.0554	2.37
<i>Flavonoids</i>						
5.74	597.1821	Phloretin-di-C-glycoside	C <sub>27</sub> H <sub>34</sub> O <sub>15</sub>	-0.7	357.0963, 387.1066, 477.1381, 417.1171	2.54
<i>Others</i>						
1.42	110.0247	3,4-Dihydroxypyridine	C <sub>3</sub> H <sub>5</sub> NO <sub>2</sub>	-0.9	-	2.40
1.76	249.1247	Caffeoylputrescine	C <sub>13</sub> H <sub>18</sub> N <sub>2</sub> O <sub>3</sub>	0.8	-	-2.00
2.30	285.0609	Dihydroxybenzoic acid-pentose	C <sub>12</sub> H <sub>14</sub> O <sub>8</sub>	-2.4	152.0111, 108.0212	-2.86
2.70	175.0609	2-Isopropylmalic acid (S)	C <sub>7</sub> H <sub>12</sub> O <sub>5</sub>	-1.7	115.0339, 85.0654, 59.0136, 113.0606	2.45
2.85	343.1033	Homovanillic acid-hexose	C <sub>15</sub> H <sub>20</sub> O <sub>9</sub>	-0.5	181.0505, 137.0605, 59.0136	2.13
2.89	431.1557	Benzyl alcohol-dihexose	C <sub>19</sub> H <sub>28</sub> O <sub>11</sub>	-0.5	-	2.36
3.32	411.1869	(iso)pentyl dihexose	C <sub>17</sub> H <sub>32</sub> O <sub>11</sub>	-0.7	-	2.04
5.27	245.0929	Acetyl tryptophan	C <sub>13</sub> H <sub>14</sub> N <sub>2</sub> O <sub>3</sub>	-1.1	116.0500, 98.0243	2.11
11.29	329.2330	Hydroxyoctadecanedioic acid	C <sub>18</sub> H <sub>34</sub> O <sub>5</sub>	-1.1	171.1023, 139.1124, 211.1339	3.29
14.96	474.2621	LysoPE(18:3/0:0)	C <sub>23</sub> H <sub>42</sub> NO <sub>7</sub> P	-1.1	277.2161, 152.9952, 78.9587	-2.64

(S) = standard confirmed.