

## Article

# Combined Experimental and Multivariate Model Approaches for Glycoalkaloid Quantification in Tomatoes

Gabriella Tamasi <sup>1,2,\*</sup> , Alessio Pardini <sup>1,2</sup> , Riccardo Croce <sup>1</sup>, Marco Consumi <sup>1,3,\*</sup> , Gemma Leone <sup>1,3</sup> ,  
Claudia Bonechi <sup>1,2</sup> , Claudio Rossi <sup>1,2</sup> and Agnese Magnani <sup>1,3</sup> 

<sup>1</sup> Department of Biotechnology Chemistry and Pharmacy, University of Siena, via A. Moro 2, 53100 Siena, Italy; alessio.pardini@unisi.it (A.P.); riccardo.croce@student.unisi.it (R.C.); gemma.leone@unisi.it (G.L.); claudia.bonechi@unisi.it (C.B.); claudio.rossi@unisi.it (C.R.); agnese.magnani@unisi.it (A.M.)

<sup>2</sup> Center for Colloid and Surface Science (CSGI), via della Lastruccia 3, 50019 Sesto Fiorentino, Italy

<sup>3</sup> National Interuniversity Consortium of Materials Science and Technology (INSTM), via G. Giusti 9, 50121 Firenze, Italy

\* Correspondence: gabriella.tamasi@unisi.it (G.T.); marco.consumi@unisi.it (M.C.)

**Abstract:** The intake of tomato glycoalkaloids can exert beneficial effects on human health. For this reason, methods for a rapid quantification of these compounds are required. Most of the methods for  $\alpha$ -tomatine and dehydrotomatine quantification are based on chromatographic techniques. However, these techniques require complex and time-consuming sample pre-treatments. In this work, HPLC-ESI-QqQ-MS/MS was used as reference method. Subsequently, multiple linear regression (MLR) and partial least squares regression (PLSR) were employed to create two calibration models for the prediction of the tomatine content from thermogravimetric (TGA) and attenuated total reflectance (ATR) infrared spectroscopy (IR) analyses. These two fast techniques were proven to be suitable and effective in alkaloid quantification ( $R^2 = 0.998$  and  $0.840$ , respectively), achieving low errors (0.11 and 0.27%, respectively) with the reference technique.

**Keywords:** tomatine; chromatography; thermogravimetric analysis; mid-infrared spectroscopy; multivariate analysis



**Citation:** Tamasi, G.; Pardini, A.; Croce, R.; Consumi, M.; Leone, G.; Bonechi, C.; Rossi, C.; Magnani, A. Combined Experimental and Multivariate Model Approaches for Glycoalkaloid Quantification in Tomatoes. *Molecules* **2021**, *26*, 3068. <https://doi.org/10.3390/molecules26113068>

Academic Editor: Giorgio S. Senesi

Received: 29 April 2021

Accepted: 18 May 2021

Published: 21 May 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Fruit ripening is a complex process that causes considerable changes in color, texture and flavor. Additionally, the chemical composition of fruits is affected during this process, including conversion of starch to sugars, biosynthesis and accumulation of pigments and aromatic volatiles, as well as modification of cell wall ultrastructure [1].

In tomatoes, a significant decrease in steroidal glycoalkaloid content (i.e.,  $\alpha$ -tomatine and dehydrotomatine) occurs as a function of the ripening process. These components are not useful for plant growth but play a significant role in defense mechanisms against pathogens. According to their specific function, the glycoalkaloid concentration is highest in stems and leaves, during the first stages of plant growth. In the fruits, the glycoalkaloid content decreases as a function of the ripening process and at the same time, the color changes from green to red [2–5]. It is well known that the tomato glycoalkaloids are biosynthesized and then degraded during fruit ripening. In particular, in tomato fruits, the tomatine content decreases as the fruits grow and it is completely degraded when the fruits turn red [4,6,7].

It is now well known that glycoalkaloids also possess beneficial effects for human health [6,8–10], as well as many other secondary metabolites [11–16]. Thus, rapid methods for the quantification of  $\alpha$ -tomatine and dehydrotomatine in tomato fruits are potentially useful.

Most of the studies on natural glycoalkaloid quantification are based on chromatographic techniques [17], including the quantification of tomatine [6,7,10]. Although high-performance liquid chromatography (HPLC) still remains the gold standard technique for

quantifying organic substances [18], other methods are under development to overcome some disadvantages of this analytical technique. Liquid chromatography is, indeed, very powerful, but it is time-consuming and requires considerable manual work for sample pre-treatments (e.g., extraction and purification of the analytes). In many cases, spectroscopic methods combined with chemometric approaches can overcome these problems. Attenuated total reflection–Fourier transform mid-infrared spectroscopy (ATR-FT-MIR) is widely used because it is a reliable, rapid, low-cost and non-destructive technique. Moreover, combined with chemometric methods, it can highlight the spectral differences between similar samples, modeling the systematic variance of the data and presenting it in a simpler way [19–21]. This approach can allow the quantitative determination of bioactive molecules [22]. Although chemometric methods combined with ATR-FT-MIR have been widely applied to classify various food and agricultural products, according to our knowledge, there are no studies about the determination of tomatine in tomato species. Similarly, differential scanning calorimetry (DSC) is a technique that can be applied for quantitative food analysis [23,24]. On the contrary, very few applications of thermogravimetry (TGA) in food analysis are reported, usually related to evaluating thermal stability during processing and the action of environmental conditions more than analyte quantification [25,26]. However, TGA is a very quick and user-friendly technique, and does not require specific sample pre-treatments.

The objective of this study is to evaluate the ability of TGA and ATR-FT-MIR as fast techniques to predict the tomatine content in tomatoes, through multivariate statistical approaches, using HPLC coupled with an electrospray ionization and triple quadrupole mass spectrometer (HPLC-ESI-QqQ-MS/MS) as a comparative reference technique.

## 2. Results and Discussion

### 2.1. HPLC-ESI-QqQ-MS/MS Glycoalkaloid Determination in Different Industrial Tomato Varieties at Different Vine-Ripe Stages

The present study reports the quantification of  $\alpha$ -tomatine and dehydrotomatine, in eight industrial varieties of tomatoes, carried out via an HPLC-ESI-QqQ-MS/MS protocol in hydroalcoholic acidic extracts of lyophilized samples. The results are summarized in Table 1. In agreement with other studies [4,6,7,27] on the same variety, a higher content of glycoalkaloids was found in green tomatoes. The turning stage represents the ripening phase at which the most significant changes occur. Indeed, up to a 77% decrease of the two glycoalkaloids was observed. A further decrease between 24 to 77% was observed in the pink ripening stage. Only the H7204 red ripened tomato variety showed a very low  $\alpha$ -tomatine content, corresponding to 1.3% of the  $\alpha$ -tomatine found in the corresponding green tomatoes, while dehydrotomatine was detected only in trace amounts.

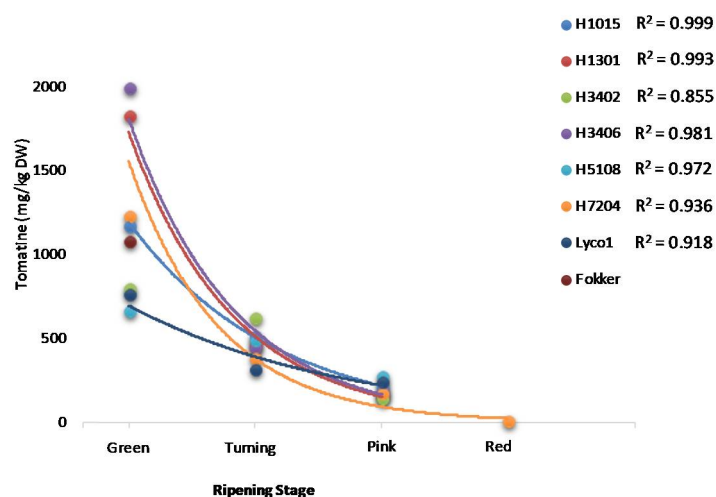
The pseudo-exponential decrease in tomatine (the sum of  $\alpha$ -tomatine and dehydrotomatine) as a function of the ripening stage is strictly related to the variety, and the equations that describe the decreasing pattern are reported in Figure 1.

Nevertheless, more significant differences were found among varieties. In particular, H1301 and H3406 varieties were the richest in  $\alpha$ -tomatine and dehydrotomatine (Table 1) whereas the lowest values were found in H5108 and Lyco1 varieties. It is possible to notice that the  $\alpha$ -tomatine content at the green ripening stage is strongly influenced by the varietal factor, covering a wide range of concentrations of glycoalkaloids (from  $1772 \pm 33$  to  $552 \pm 45$  mg/kg dry weight, DW). The difference in the  $\alpha$ -tomatine content among the different tomato varieties became less significant proceeding through the ripening stages, from the green to the pink ripening stage. In the latter, the range of the  $\alpha$ -tomatine content was rather narrow, ranging from  $238 \pm 17$  to  $120 \pm 11$  mg/kg DW. Interestingly, the rate of  $\alpha$ -tomatine degradation, due to the ripening process of the fruits, also differed within the different varieties, underlining that this process is mainly influenced by the intrinsic genetic differences linked to the variety of the plants.

**Table 1.**  $\alpha$ -Tomatine and dehydrotomatine contents in vine-ripened industrial tomato varieties. The values are expressed as mg/kg DW and mean  $\pm$  SD (standard deviation;  $n = 9$ ). The content of tomatine (sum of the two glycoalkaloids) is also reported.

Variety	Ripening Stage	$\alpha$ -Tomatine	Dehydrotomatine	Tomatine
H1015	Green	1028 $\pm$ 53 <sup>a</sup>	147 $\pm$ 9 <sup>a</sup>	1176 $\pm$ 54 <sup>a</sup>
	Turning	441 $\pm$ 28 <sup>b</sup>	78 $\pm$ 8 <sup>b,c</sup>	519 $\pm$ 29 <sup>b</sup>
	Pink	180 $\pm$ 19 <sup>c,d,e</sup>	34 $\pm$ 5 <sup>d,e</sup>	213 $\pm$ 20 <sup>c,d</sup>
H1301	Green	1614 $\pm$ 40 <sup>f</sup>	215 $\pm$ 4 <sup>f</sup>	1829 $\pm$ 41 <sup>e</sup>
	Turning	400 $\pm$ 17 <sup>b</sup>	57 $\pm$ 3 <sup>g</sup>	457 $\pm$ 17 <sup>b,f</sup>
	Pink	141 $\pm$ 12 <sup>c</sup>	21 $\pm$ 2 <sup>h</sup>	161 $\pm$ 12 <sup>c,g</sup>
H3402	Green	688 $\pm$ 40 <sup>g</sup>	109 $\pm$ 11 <sup>i</sup>	796 $\pm$ 41 <sup>h</sup>
	Turning	535 $\pm$ 19 <sup>h</sup>	86 $\pm$ 7 <sup>c,j</sup>	621 $\pm$ 20 <sup>i</sup>
	Pink	120 $\pm$ 11 <sup>c</sup>	21 $\pm$ 3 <sup>h</sup>	141 $\pm$ 11 <sup>g</sup>
H3406	Green	1772 $\pm$ 33 <sup>i</sup>	221 $\pm$ 3 <sup>f</sup>	1993 $\pm$ 33 <sup>j</sup>
	Turning	388 $\pm$ 18 <sup>b</sup>	61 $\pm$ 4 <sup>g</sup>	449 $\pm$ 18 <sup>f,k</sup>
	Pink	155 $\pm$ 11 <sup>c,d</sup>	26 $\pm$ 3 <sup>d,h</sup>	181 $\pm$ 12 <sup>c,d,g</sup>
H5108	Green	552 $\pm$ 45 <sup>h</sup>	111 $\pm$ 7 <sup>i</sup>	663 $\pm$ 46 <sup>i</sup>
	Turning	420 $\pm$ 27 <sup>b</sup>	69 $\pm$ 7 <sup>b</sup>	489 $\pm$ 28 <sup>b,f</sup>
	Pink	238 $\pm$ 17 <sup>e,j</sup>	42 $\pm$ 6 <sup>e</sup>	279 $\pm$ 18 <sup>l,m</sup>
H7204	Green	1088 $\pm$ 34 <sup>a</sup>	143 $\pm$ 12 <sup>a</sup>	1231 $\pm$ 36 <sup>a</sup>
	Turning	314 $\pm$ 16 <sup>k</sup>	75 $\pm$ 6 <sup>b,c</sup>	388 $\pm$ 17 <sup>k</sup>
	Pink	148 $\pm$ 17 <sup>c</sup>	28 $\pm$ 4 <sup>d,h</sup>	176 $\pm$ 17 <sup>c,g</sup>
	Red	14 $\pm$ 2 <sup>l</sup>	<LOD	14 $\pm$ 2 <sup>n</sup>
Lyco1	Green	678 $\pm$ 62 <sup>g</sup>	90 $\pm$ 8 <sup>j</sup>	768 $\pm$ 63 <sup>h</sup>
	Turning	294 $\pm$ 22 <sup>j,k</sup>	26 $\pm$ 3 <sup>d,h</sup>	320 $\pm$ 22 <sup>m</sup>
	Pink	217 $\pm$ 32 <sup>d,e</sup>	25 $\pm$ 4 <sup>d,h</sup>	242 $\pm$ 32 <sup>d,l</sup>
Fokker	Green	952 $\pm$ 101 <sup>m</sup>	128 $\pm$ 15 <sup>k</sup>	1080 $\pm$ 102 <sup>o</sup>

Values marked with the same letter within the same column are not statistically different (Tukey's test,  $p > 0.05$ ).



**Figure 1.** Graphical representation of the pseudo-exponential decrease in tomatine content as a function of the ripening stage for the different industrial tomato varieties. The  $R^2$  calculated parameters are also reported.

## 2.2. Thermogravimetric Analysis (TGA)

Thermogravimetry has already been applied to analyze and quantify vegetal compounds in complex matrices [28,29]. The weight loss of each freeze-dried sample in the ranges 120–200 and 200–400 °C are summarized in Table 2.

**Table 2.** Results of TGA analysis: mean  $\pm$  SD ( $n = 3$ ) of the weight loss (in percentage) in the ranges 120–200 °C and 200–400 °C as a function of ripening stages.

Variety	Ripening Stage	Weight Loss	
		120–200 °C	200–400 °C
H1015	Green	17.5 $\pm$ 0.2	38.2 $\pm$ 0.3
	Turning	19.0 $\pm$ 0.2	36.4 $\pm$ 0.2
	Pink	19.3 $\pm$ 0.2	34.4 $\pm$ 0.5
H1301	Green	15.9 $\pm$ 0.1	39.7 $\pm$ 0.8
	Turning	19.0 $\pm$ 0.4	36.3 $\pm$ 0.3
	Pink	20.4 $\pm$ 0.3	33.8 $\pm$ 0.2
H3402	Green	16.9 $\pm$ 0.3	37.8 $\pm$ 0.4
	Turning	17.5 $\pm$ 0.5	38 $\pm$ 0.1
	Pink	20.5 $\pm$ 0.1	33 $\pm$ 0.3
H3406	Green	14.9 $\pm$ 0.1	40.1 $\pm$ 0.2
	Turning	18.5 $\pm$ 0.6	36.7 $\pm$ 0.3
	Pink	20.3 $\pm$ 0.2	33.3 $\pm$ 0.5
H5108	Green	18.4 $\pm$ 0.3	36.7 $\pm$ 0.6
	Turning	18.7 $\pm$ 0.2	36.3 $\pm$ 0.5
	Pink	20.3 $\pm$ 0.4	34.6 $\pm$ 0.2
H7204	Green	14.1 $\pm$ 0.2	40.1 $\pm$ 0.3
	Turning	18.9 $\pm$ 0.3	34.3 $\pm$ 0.4
	Pink	24.2 $\pm$ 0.3	33.2 $\pm$ 0.6
Lyco1	Green	16.9 $\pm$ 0.5	37.7 $\pm$ 0.4
	Turning	19.9 $\pm$ 0.4	33.9 $\pm$ 0.6
	Pink	19.5 $\pm$ 0.2	33.7 $\pm$ 0.4
Fokker	Green	16.8 $\pm$ 0.2	37.4 $\pm$ 0.5

The weight loss in the range 120–200 °C increases from 16  $\pm$  1% for tomatoes at the green stage to 18.7  $\pm$  0.5% and 21  $\pm$  1% for tomatoes at the turning and pink stages, respectively. Concerning the decomposition between 200–400 °C, a clear trend can be observed in the analyzed samples: moving from the green to pink stage, a decrease in weight loss is observed. Indeed, in the green stage, a mean weight loss of about 38  $\pm$  1% is found that decreases to 36  $\pm$  1% in the turning stage and reaches 33.7  $\pm$  0.4% in the pink stage.

In Figure 2, a comparison between thermographs of H7204 variety samples at different ripening stages is depicted, showing the peculiar increase in weight loss between 120 and 200 °C. In this temperature range, the degradation of volatiles occurs [30]. The increase in weight loss observed in this region of the thermogram could be explained by the accumulation of flavor and aromatic compounds as a consequence of the ripening process. Concurrently, a decrease in weight loss between 200 and 400 °C is observed. At these temperatures, macromolecules such as cellulose, hemicellulose, pectin and starch are thermally degraded [30,31]. During the ripening process, many hydrolytic enzymes are involved in cell wall metabolism, causing the softening of fleshy fruits. At the same time, the depolymerization of many polysaccharides to sugars occurs. Consequently, the content of these macromolecules in tomatoes progressively decreases as a function of the ripening process.

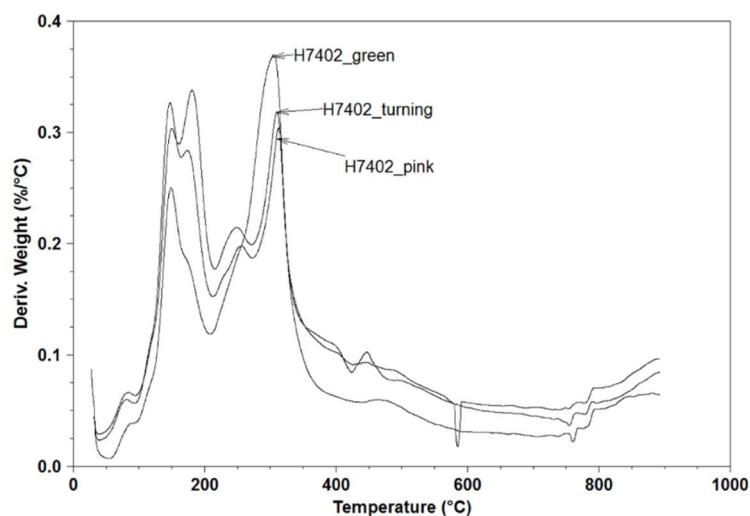


Figure 2. Thermographs of H7204 at different ripening stages.

As shown in Figure 3, strong correlations showing opposite trends were found between the logarithm of the tomatine content detected by HPLC-ESI-QqQ-MS/MS and the weight loss in both of the temperature ranges (Pearson's correlation coefficients,  $r = -0.879$  and  $0.958$ , respectively).

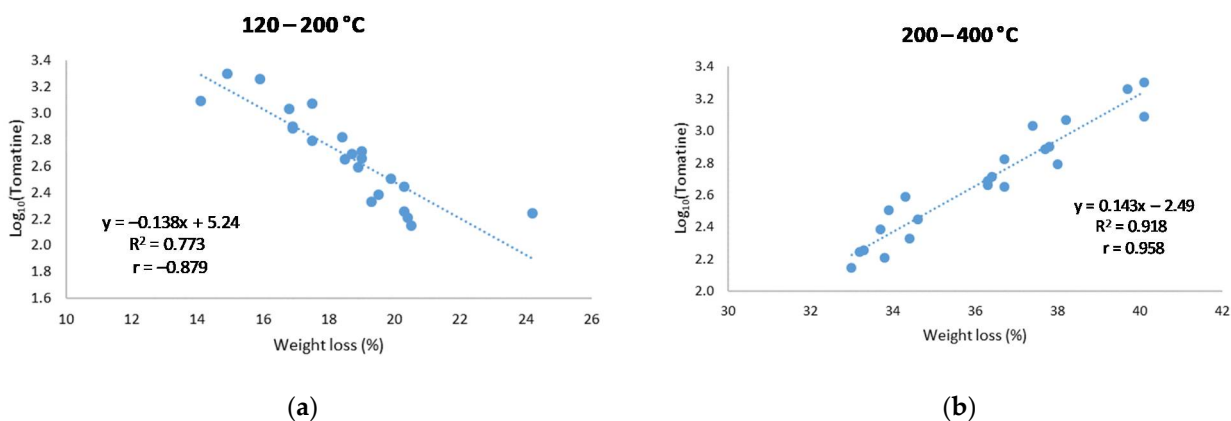
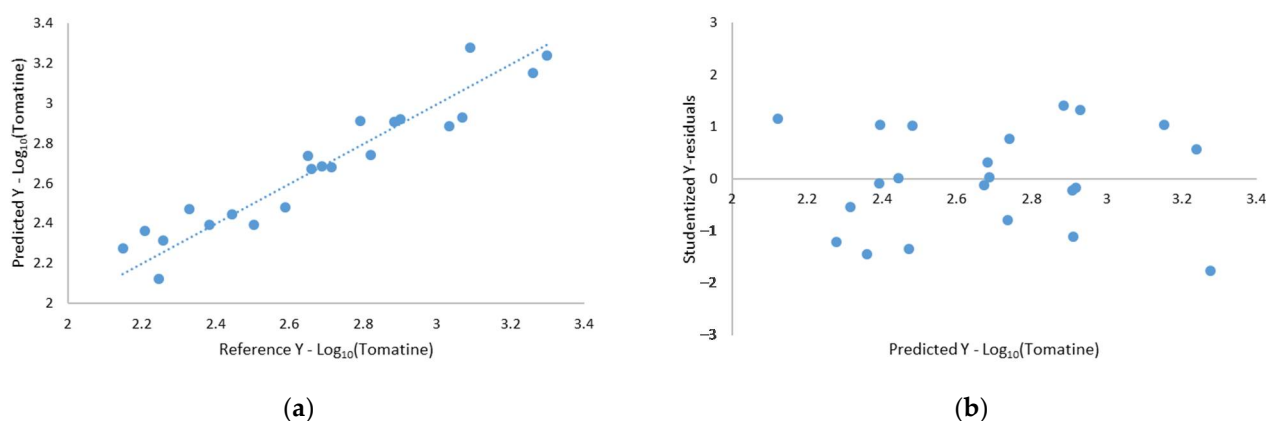


Figure 3. Linear regression model of the  $\text{Log}_{10}(\text{Tomatine})$  vs. the weight loss in the temperature ranges (a) 120–200 °C and (b) 200–400 °C.

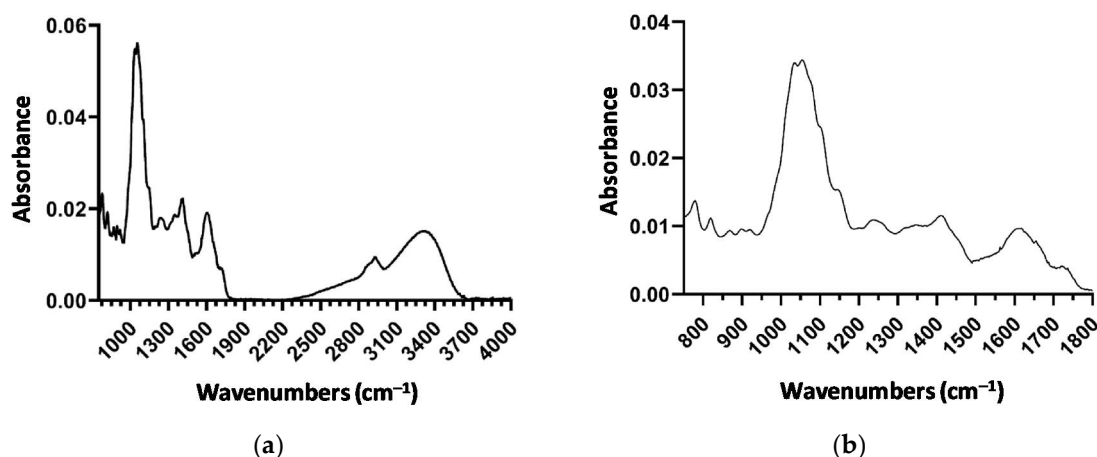
Therefore, a multiple linear regression (MLR) model was proposed to extrapolate the tomatine concentration in lyophilized tomato samples from TGA, taking into account the weight loss in the selected temperature ranges. The MLR analysis gave the equation  $Y = -0.0472X_1 + 0.0983X_2$  ( $R^2 = 0.999$  and adjusted  $R^2 = 0.998$ ). The intercept of the regression model was forced to zero since it was not statistically significant ( $p > 0.05$ ), while both the coefficients of  $X_1$  and  $X_2$  were significantly different from zero ( $p << 0.05$ ). The F-test showed that the overall model is significant ( $p << 0.05$ ), with a root mean square error (RMSE) of 0.11. The relationship between the tomatine concentrations predicted by the MLR model and the measured ones is reported in Figure 4a. The residual plot of the model (Figure 4b), reporting the autoscaled Y-residuals vs. the predicted Y, gave a random distribution between  $\pm 3$  (99% confidence interval) for all the samples, confirming that the MLR model reported here is able to predict tomatine values along the whole range of concentrations employed.



**Figure 4.** (a) MLR output of measured vs. predicted values of tomatine concentration and (b) model residuals.

### 2.3. Attenuated Total Reflection–Fourier Transform Mid-Infrared Spectroscopy (ATR-FT-MIR)

Figure 5a reports the spectrum of a dried tomato sample (as an example). None of the spectra deviate from the example, if not considered in terms of peak intensity. In the spectra displayed, the specific bands of vegetal samples are highlighted and assigned as follows. The band at  $3289\text{ cm}^{-1}$  is characteristic of NH and OH stretching vibrations. The region of  $2923\text{--}2853\text{ cm}^{-1}$  can be assigned to the symmetrical and asymmetric stretching modes of the  $\text{CH}_3$  and  $\text{CH}_2$  groups. The  $1720\text{ cm}^{-1}$  band corresponds to the stretching of the C=O ester carbonyl or carboxylic acid groups, which are characteristic of fatty acids and polysaccharides. The amide I band at  $1652\text{ cm}^{-1}$  results from the C=O stretching in the amides I, II and III, while the amidic band II at  $1520\text{ cm}^{-1}$  originates from the bending vibrations of the N-H groups.



**Figure 5.** (a) Example of ATR-FT-MIR spectrum of lyophilized tomato. (b) Average ATR-FT-MIR spectrum of 22 samples in the range  $750\text{--}1800\text{ cm}^{-1}$ .

Table 3 summarizes bands in the IR region from  $1800\text{ to }900\text{ cm}^{-1}$  that include the “fingerprint” region, which includes bands corresponding to the vibrations of the C-O, C-C, C-H and C-N bonds. This region is, on the one hand, very rich in information, but, on the other hand, difficult to analyze due to its complexity. This area provides important information about organic compounds, such as sugars, alcohols and organic acids, present in the sample by featuring their molecular vibrations (stretching, bending and torsions of the chemical bonds) in specific infrared regions. This region was dominated by a broad band centred at  $1055, 1035\text{ cm}^{-1}$ , with evident shoulders at  $1145$  and  $1100\text{ cm}^{-1}$ , due to strong vibrational modes of various carbohydrates and acids, which are abundant groups in tomatoes. Tomatine shows main bands in the region between  $1100$  and  $950\text{ cm}^{-1}$ , which,



however, are covered by the stronger sugar and polysaccharide absorption and are not useful for quantitative analysis as the absorbance band at  $956\text{ cm}^{-1}$  corresponds to a *trans*  $-\text{CH}=\text{HC}-$  bending out of plane deformation band that is the unique IR marker band specific to lycopene [32].

**Table 3.** Main functional groups assigned to ATR-FTIR spectra of tomato [33].

Wavenumber ( $\text{cm}^{-1}$ )	Proposed Assignment
1720	C=O ester
1650	Amide I, $\beta$ -sheet
1604	C-C aromatic
1551	C-C aromatic
1520	Amide II, $\text{C}\equiv\text{N}$ stretching
1410	$\text{CH}_2$ bending of lipids and fatty acids
1350	$\text{CH}_3$ bending proteins and lipids and $\text{CH}_2$ wagging and twisting
1240	OH bending
1196	C-O-C ester stretching
1145	C-O-C ester stretching
1055	C-O-C glycosidic bond
955	CH( <i>trans</i> OOP)

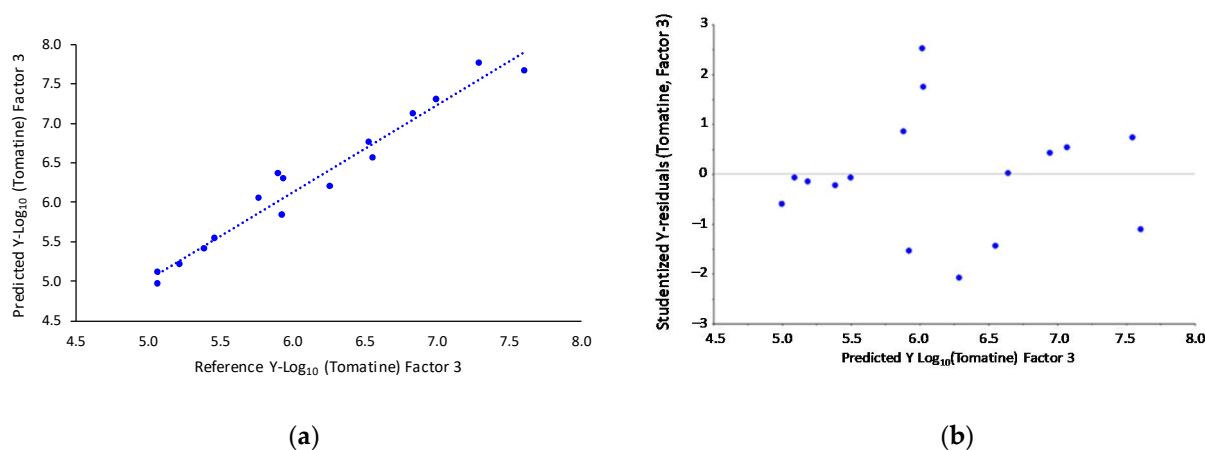
As discussed, the spectra contain a multitude of bands that are characteristic of vegetal samples which do not allow to obtain information on a specific compound without the interference of the matrix as whole. For this reason, we used an ATR-FT-MIR “fingerprint” analytical approach [21,34] for the structural identification of compounds considering that no two chemical structures will have the same ATR-FT-MIR spectrum. ATR-FT-MIR provides a characteristic signature of chemical or biochemical substances present that can be used for chemometric studies.

#### 2.4. Chemometric Approach

The average spectrum of 22 samples, as regards the most informative regions located at low wavenumbers ( $752\text{--}1800\text{ cm}^{-1}$ ), is reported in Figure 5b. This was used to build the regression model. The predictive model for the determination of tomatine in extracted tomato samples was obtained by applying the partial least squares regression (PLSR) model to 16 samples of the calibration set, after data pre-treatment and column mean centering. The uncertainty test identifies 48 significant variables, leading to a decrease in the sample/variable ratio. This also determines a decrease in the risk of finding random correlation and leads to an increase in the reliability of the regression model [35].

The first two components have been identified as significant for tomatine quantification. In more detail, in the final model, the first three latent variables explain 95% of the Y-variance and 73% of the variance in the X-block, with Pearson’s regression coefficient  $R^2 = 0.95$  and root mean square of calibration (RMSEC) = 0.11.

Figure 6a shows the relationship between the measured values and the predicted ones for the calibration set; all the samples are randomly distributed around the regression line with a negligible dispersion for the whole range of variability. In Figure 6b, the relationship between the autoscaled Y-residual and the predicted ones is shown: all the samples are randomly distributed within  $\pm 3$  values that correspond to the 99% confidence interval. Therefore, it is possible to confirm that the regression model presents a comparable predictability along the whole concentration interval with a negligible BIAS value.



**Figure 6.** PLS output, (a) measured vs. predicted values of tomatine concentration and (b) model residuals.

Finally, the model was validated on the test set. The tomatine contents, predicted by the PLS regression model, were compared with the experimental data to evaluate the actual predictability of the proposed model. The performances were evaluated through the joint analysis of the Pearson's regression coefficient,  $R^2_{\text{pred}}$ , and the root mean square error in prediction, RMSEP:

$$\text{RMSEP} = \sqrt{\frac{\sum_i^N (y_i - \hat{y}_i)^2}{N}} \quad (1)$$

where  $y_i - \hat{y}_i$  is the difference between the measured and predicted value for the  $i$ -th sample in the test set.

The quality parameters calculated for the proposed model confirmed its goodness in predicting tomatine concentration, from spectral features of tomato extract, showing a high  $R^2_{\text{pred}}$  in prediction (0.84) and a low RMSEP (0.27), as well as a negligible BIAS value (0.08).

Table 4 reports the comparison between experimental data obtained via HPLC-ESI-QqQ-MS/MS determination of tomatine and predicted data through the combined experimental/statistic approaches: TGA/MLR and the ATR-FT-MIR/PLS model. The data have been statistically compared via two-way ANOVA followed by Dunnett's test. It is evident that the results are not statistically different considering the 95% of confidence interval (Dunnett's test,  $p \gg 0.05$ ). This result underlines that the two alternative semi-quantitative methods, TGA/MLR and ATR-FT-MIR/PLS, that do not require any pre-treatment extraction of the lyophilized and ground tomato material, are reasonable alternatives to the quantitative HPLC-ESI-QqQ-MS/MS determination (strictly dependent on a solid/liquid extraction).

In conclusion, the complex changes occurring during the ripening process produce significant changes in the chemical composition of tomato fruits. TGA and ATR-FT-MIR analyses revealed significant differences in tomatoes at the different ripening stages, probably due to the variation in the macromolecule content, as a consequence of the enzymatic reactions occurring during fruit growth and maturation. Through chemiometric approaches, these variations were found to correlate with the glycoalkaloid content. The MLR and the PLS regression models constructed on the values detected by HPLC-ESI-QqQ-MS/MS analyses as a reference allowed us to accurately quantify tomatine in the tomato samples via TGA and ATR-FT-MIR analyses. These two techniques may represent a valid alternative in the quantification of tomatine in tomatoes, permitting the omission of the pre-treatments required for chromatographic analyses.



**Table 4.** Comparison between HPLC-ESI-QqQ-MS/MS experimental data and TGA and ATR-FT-MIR predicted data, via MLR and PLS statistic models, respectively. The values within the same row are not statistically different (two-way ANOVA, Dunnett's test,  $p \gg 0.05$ ).

Variety	Ripening Stage	Tomatine		
		HPLC-ESI-QqQ-MS/MS Experimental	TGA/MLR Model Predicted	ATR-FT-MIR/PLS Model Predicted
H1015	Green	1176 ± 54	851	1177
	Turning	519 ± 29	481	482
	Pink	213 ± 20	296	213
H1301	Green	1829 ± 41	1422	1820
	Turning	457 ± 17	470	458
	Pink	161 ± 12	229	160
H3402	Green	796 ± 41	830	797
	Turning	621 ± 20	813	621
	Pink	141 ± 11	189	144
H3406	Green	1993 ± 33	1735	1874
	Turning	449 ± 18	544	454
	Pink	181 ± 12	207	182
H5108	Green	663 ± 46	550	664
	Turning	489 ± 28	486	461
	Pink	279 ± 18	278	280
H7204	Green	1231 ± 36	1893	1182
	Turning	388 ± 17	302	371
	Pink	176 ± 17	133	177
Lyco1	Green	768 ± 63	811	742
	Turning	320 ± 22	248	325
	Pink	242 ± 32	247	239
Fokker	Green	1080 ± 102	766	1039

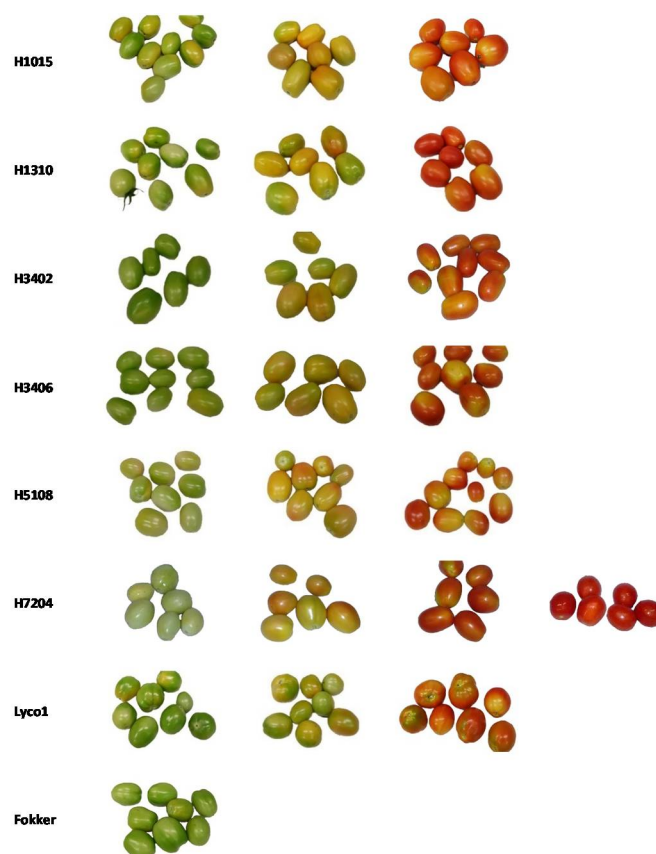
### 3. Materials and Methods

#### 3.1. Chemicals

All the following solvents were purchased from Sigma-Aldrich (Milan, Italy): ethanol (EtOH, gradient grade  $\geq 99.9\%$ ) and methanol (MeOH, gradient grade  $\geq 99.9\%$ ), glacial acetic acid ( $\text{CH}_3\text{COOH}$ , reagent grade  $\geq 99.8\%$ ) and formic acid ( $\text{HCOOH}$ , ACS grade  $\geq 98\%$ ). Tomatidine hydrochloride ( $\geq 95\%$ ) was also purchased from Sigma-Aldrich whereas tomatine ( $>75\%$ ) was purchased from TCI Europe (Belgium). Ultra-pure deionized water was produced using an Acquinity P/7 purifier system (MembraPure GmbH, Berlin, Germany).

#### 3.2. Plant Materials

Eight different vine-ripened industrial varieties of tomato fruits, harvested in summer 2017 at different ripening stages (green, turning, pink and red) were analyzed (Figure 7). The samples were classified according to the definitions provided by the California Tomato Commission and United States Department of Agriculture (USDA, [36]): green (the surface of the tomato is completely green in color, the shade of green may vary from light to dark), turning (more than 10%, but not more than 30%, of the total surface shows a definite change in color from green to tannish-yellow, pink, red or a combination thereof), pink (more than 30%, but not more than 60%, of the total surface is pink or red in color).



**Figure 7.** Industrial tomato varieties at different vine-ripening stages (green, turning, pink and red), classified according to the California Tomato Commission and United States Department of Agriculture (USDA) guidelines.

The samples were washed with deionized water and dried. Six tomatoes of each variety at the different ripening stages were homogenized with a blender (Moulinex, 1000 W) and the homogenates were freeze-dried (VirTis BenchTop Pro lyophilizer,  $-51 \pm 2$  °C,  $1.3 \pm 0.5$  mbar) until a constant weight (5 days, averaged samples water content,  $92 \pm 1$  %). Lyophilized samples were then ground in porcelain mortar, sieved to a 500  $\mu$ m particle size and stored at  $-20 \pm 1$  °C before the subsequent analyses.

### 3.3. HPLC-ESI-QqQ-MS/MS Quantification of Glycoalkaloids

Glycoalkaloid extraction and analytical quantification were carried out via HPLC-ESI-QqQ-MS/MS as previously validated and reported [6,7], with slight modifications. Briefly, lyophilized samples were extracted by an hydroalcoholic acidic mixture consisting of EtOH/CH<sub>3</sub>COOH 1%, (70:30, *v/v*) and the extraction was ultrasound assisted (two-cycle extraction on the solid phase). The extracts were dried under nitrogen flow with subsequent lyophilization and then reconstituted in MeOH 80%. The analytical determination of  $\alpha$ -tomatine and dehydrotomatine was performed by HPLC-ESI-QqQ-MS/MS analyses. Each sample was extracted in triplicate. The separation was carried out by a reverse-phase column (Phenomenex Luna C18(2)-HST, 2.5  $\mu$ m, 100  $\times$  2.0 mm, 5  $\mu$ m, 100 Å equipped with a Phenomenex SecurityGuard ULTRA pre-column) thermostatted at  $30 \pm 1$  °C and a linear gradient with (A) H<sub>2</sub>O and (B) MeOH both acidified with formic acid 0.1% (*v/v*): linear gradient 40–90% B in 15 min (flow rate of 0.2 mL/min). The injection volume was 5  $\mu$ L. The ESI-MS conditions were optimized through the direct injection of a tomatine standard solution in MeOH in positive ionization mode and  $\alpha$ -tomatine and dehydrotomatine were quantified by multiple reaction monitoring mode (MRM; collision energy –31 V and scan time, 500 ms). The analytical quantification was carried out via an external calibration method: linearity range, 0.5–30 mg/L of tomatine (a mixture of  $\alpha$ -tomatine, 85% and

dehydrotomatine, 13%). The tomatidine was used as an internal standard, taking into account that it does not naturally occur as aglycone in tomatoes [4,6,7]. Calibration curves showing equations with  $R^2 > 0.990$  were used for the quantification. The limit of detection (LOD) and limit of quantification (LOQ) were: 0.20//0.05 and 0.50//0.20 mg/L for  $\alpha$ -tomatine and dehydrotomatine, respectively. The results were expressed as mg/kg of sample dry weight (DW).

All the samples were extracted in triplicate and each extract was analyzed three times ( $n = 9$ ). and results were reported as mean  $\pm$  standard deviation (SD). The significant differences between means were assessed by analysis of variance (ANOVA) followed by Tukey's post hoc test. All the statistical treatments were run on Microsoft Office Excel 365, implemented with the Real Statistic subroutine, setting the level of significance at  $p < 0.05$ .

### 3.4. Thermogravimetric Analysis (TGA)

The analyses were performed using a Q1000 thermogravimeter (TA, Instruments) applying a thermal program from 30 to 900 °C with a heating ramp of 10 °C/min, under constant nitrogen flow [37]. Three lyophilized aliquots for each variety (15 mg in weight) were analyzed.

### 3.5. Attenuated Total Reflection–Fourier Transform Mid-Infrared Spectroscopy (ATR-FT-MIR)

All the samples were analyzed by FTIR using a Nicolet IS50 FTIR spectrophotometer (Thermo Nicolet Corp., Madison, WI, USA), equipped with a single-reflection germanium ATR crystal (Pike 16154, Pike Technologies, Madison, WI, USA) and a deuterated triglycine sulphate (DTGS) detector. The spectra were acquired (32 scans per sample or background) in the range of 4000–800  $\text{cm}^{-1}$  at a nominal resolution of 4  $\text{cm}^{-1}$ . The spectra were corrected using the background spectrum of air. The analysis was carried out at room temperature by spreading a lyophilized sample onto the surface of the ATR crystal. Before each sample was analyzed, the ATR crystal was carefully cleaned with water-wet cellulose tissue and dried using a flow of pure nitrogen gas. The cleaned crystal was checked spectrally to ensure that no residue was retained from the previous sample. The spectrum of every sample was collected 3 times to check the reproducibility and do a statistical analysis. It should be noted that the individual spectra of the same tomato did not vary in band wavelength and varied to some extent in the absorbance. However, the characteristic signatures remained very similar, so averaged spectra are shown in the results.

The frequency scale was internally calibrated with a helium–neon reference laser to an accuracy of 0.01  $\text{cm}^{-1}$ . OMNIC software (OMNIC software system Version 9.8 Thermo Nicolet) was used for spectra acquisition and manipulation [38].

### 3.6. Multiple Linear Regression (MLR)

The multiple linear regression (MLR) analysis was performed in Microsoft Office Excel 365, implemented with the Real Statistic subroutine, plotting tomatine concentration (Y) determined by HPLC-ESI-QqQ-MS/MS analyses vs. the weight loss detected by TGA analyses in the range of temperature, 120–200 °C ( $X_1$ ) and 200–400 °C ( $X_2$ ). Since the concentration of tomatine ranged over one order of magnitude, the values were converted in the logarithmic form to make the data distribution symmetrical. To build the MLR model, the H7204 sample at the red ripening stage was discarded as a leverage outlier, since its tomatine content was too low. In this way, the distribution of tomatine content was fairly symmetrical, allowing us to build a reliable regression model.

### 3.7. Chemometric Approach

#### 3.7.1. Data Processing

The ATR-FT-MIR spectra were imported into UnscramblerX Software (Camo Analytics, Oslo, Norway) for multivariate data analysis. ATR-FT-MIR spectra absorbances at different wavelengths were stored in a data matrix with samples placed in rows and reference values of tomatine placed in the first column. As extensively reported in the literature [39–42],

it is necessary to split the entire database into two subsets, to evaluate the performance of the prediction model. The first subset, representing 75% of the available samples, was used to build the regression model, whilst the second subset, consisting in 25% of total samples, was used to estimate the actual predictability of the model. Therefore, the dataset was divided into a calibration set and a test set using random sample selection in order to avoid a biased model. In addition, an internal cross-validation step was performed using the venetian blind algorithm to find the best modeling settings (i.e., the number of latent variables, [43]). This procedure consisted in building, optimizing and testing models to obtain reliable prediction of tomatine concentration for future extracted tomato samples.

### 3.7.2. Spectral Pre-Processing

The ATR-FT-MIR spectra of the 22 samples were averaged and the most informative region was located at low wavenumbers and the regression model was built using the spectral features from  $752\text{ cm}^{-1}$  to  $1800\text{ cm}^{-1}$ . The H7204 sample at the red ripening stage was also discarded in this case.

Usually, when a regression model is built from spectral data, it is necessary to pre-treat the input variables to remove useless information resulting from unwanted systematic variations. In the present work, a combination of standard normal variate (SNV) and second derivative were used to pre-treat spectral data, removing the baseline offset and differences in global signal intensity [44–46]. SNV transforms the original data according to the following equation:

$$(x_{ij})_{\text{SNV}} = \frac{x_{ij} - \bar{x}_i}{s_i} \quad (2)$$

where  $x_{ij}$  represents the absorbance for the  $j$ -th wavelength and the  $i$ -th sample,  $\bar{x}_i$  and  $s_i$  are the average and the standard deviation of the wavelengths for the  $i$ -th spectrum, respectively. For this reason, SNV transformation is also known as row autoscaling.

A second derivative transformation was also applied to FT-MIR spectra pre-treated with standard normal variate (SNV) transformation. In more detail, a Savitszi–Golay algorithm was applied using a second order polynomial and symmetric kernel option with 7 smoothing points [46].

### 3.7.3. Partial Least Squares Regression (PLSR) and Martens' Uncertainty Test

Partial least squares regression (PLSR) was used on the calibration set to obtain the model that was subsequently used to predict tomatine content in industrial tomato samples. Additionally, in this case, tomatine concentration values were converted into the logarithmic form to make the distribution fairly symmetrical.

A kernel PLS algorithm was used to correlate the logarithm of tomatine concentration and spectral features transformed as mentioned previously [47]. The best number of latent variables that was used in the regression function was estimated for the cross-validation procedure, taking into account the cross-validation root mean square error (RMSECV) trend with respect to the number of retained components. In the present work, only three components were necessary to explain the majority of the total X- and Y-variance. Once the optimal number of latent variables were chosen, Martens' uncertainty test [48] was performed in order to remove unimportant information within spectral data. This powerful tool allowed us to improve the predictability of the model, retaining only the significant variable by giving a more reliable estimate of the prediction error when the model was tested on new samples. Moreover, since a reduced number of spectral variables were used, a simpler model was generated.

In Martens' uncertainty test, the regression coefficients  $B_i$  for each cross-validation sub-model, chosen with the venetian blind option, were calculated and the differences from the regression coefficient of the total model,  $B_{\text{tot}}$ , were computed. The sum of the squares of the differences in all sub-models was finally evaluated in order to obtain an expression of the variance of the  $B_i$  estimate for a specific wavelength. With a  $t$ -test, the significance (confidence level of 95%) of the estimate of  $B_i$  was calculated and the resulting regression

coefficients were presented with uncertainty limits. Variables with uncertainty limits that did not contain the zero were significant variables. This procedure was iteratively repeated until the difference between RMSEC and RMSECV reached a minimum value.

**Author Contributions:** Conceptualization G.T., M.C. and C.R.; methodology G.T. and M.C.; software A.P. and C.B.; validation A.P., R.C., G.T. and A.M.; formal analysis A.P., M.C. and G.L.; investigation G.T., M.C. and A.M.; data curation A.P., G.T. and R.C.; funding acquisition C.R.; supervision G.T., and C.R.; roles/writing—original draft G.T., M.C. and A.P.; writing—review and editing G.T., M.C., A.P., R.C., C.B., G.L., A.M. and C.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was partially funded by the Italian Ministry for Economic Development (MISE) financing the project “New industrial biotechnology processes for the recovery and use of bioactive tomatine from tomato by-products” in which Società Agricola Cooperativa Consorzio Casalasco del Pomodoro was the Coordinator and requested the University of Siena cooperation, as third part.

**Data Availability Statement:** Data sharing not applicable.

**Acknowledgments:** Toscana Life Sciences Foundation (TLS) was acknowledged just for the access to HPLC-ESI-QqQ-MS/MS instrumentation.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

**Sample Availability:** Samples are not available.

## References

1. Giovannoni, J. Molecular biology of fruit maturation and ripening. *Annu. Rev. Plant. Biol.* **2001**, *52*, 725–749. [[CrossRef](#)] [[PubMed](#)]
2. Friedman, M. Tomato glycoalkaloids: Role in the plant and in the diet. *J. Agric. Food Chem.* **2002**, *50*, 5751–5780. [[CrossRef](#)]
3. Friedman, M.; Levin, C.E.; Lee, S.U.; Kim, H.J.; Lee, I.S.; Byun, J.O.; Kozukue, N. Tomatine-containing green tomato extracts inhibit growth of human breast, colon, liver, and stomach cancer cells. *J. Agric. Food Chem.* **2009**, *57*, 5727–5733. [[CrossRef](#)] [[PubMed](#)]
4. Friedman, M. Anticarcinogenic, cardioprotective, and other health benefits of tomato compounds lycopene,  $\alpha$ -tomatine, and tomatidine in pure form and in fresh and processed tomatoes. *J. Agric. Food Chem.* **2013**, *61*, 9534–9550. [[CrossRef](#)]
5. Liu, J.; Kanetake, S.; Wu, Y.H.; Tam, C.; Cheng, L.W.; Land, K.M.; Friedman, M. Antiprotozoal effects of the tomato tetrasaccharide glycoalkaloid tomatine and the aglycone tomatidine on mucosal trichomonads. *J. Agric. Food Chem.* **2016**, *64*, 8806–8810. [[CrossRef](#)]
6. Tamasi, G.; Pardini, A.; Bonechi, C.; Donati, A.; Pessina, F.; Marcolongo, P.; Gamberucci, A.; Leone, G.; Consumi, M.; Magnani, A.; et al. Characterization of nutraceutical components in tomato pulp, skin and locular gel. *Eur. Food Res. Technol.* **2019**, *245*, 907–918. [[CrossRef](#)]
7. Pardini, A.; Consumi, M.; Leone, G.; Bonechi, C.; Tamasi, G.; Sangiorgio, P.; Verardi, A.; Rossi, C.; Magnani, A. Effect of different post-harvest storage conditions and heat treatment on tomatine content in commercial varieties of green tomatoes. *J. Food Comp. Anal.* **2021**, *96*, 103735. [[CrossRef](#)]
8. Serrati, S.; Porcelli, L.; Guida, S.; Ferretta, A.; Iacobazzi, R.M.; Cocco, T.; Maida, I.; Tamasi, G.; Rossi, C.; Manganello, M.; et al. Tomatine displays antitumor potential in in vitro models of metastatic melanoma. *Int. J. Mol. Sci.* **2020**, *21*, 5243. [[CrossRef](#)]
9. Marcolongo, P.; Gamberucci, A.; Tamasi, G.; Pardini, A.; Bonechi, C.; Rossi, C.; Giunti, R.; Barone, V.; Borghini, A.; Fiorenzani, P.; et al. Chemical characterisation and antihypertensive effects of locular gel and serum of *Lycopersicon esculentum* L. var. “Camone” tomato in Spontaneously Hypertensive Rats. *Molecules* **2020**, *25*, 3758. [[CrossRef](#)]
10. Tamasi, G.; Baratto, M.C.; Bonechi, C.; Byelyakova, A.; Pardini, A.; Donati, A.; Leone, G.; Consumi, M.; Lamponi, S.; Magnani, A.; et al. Chemical characterization and antioxidant properties of products and by-products from *Olea europaea* L. *Food Sci. Nutr.* **2019**, *7*, 2907–2920. [[CrossRef](#)] [[PubMed](#)]
11. Schreiner, M.; Mewis, I.; Huyskens-Keil, S.; Jansen, M.A.K.; Zrenner, R.; Winkler, J.B.; O’Brien, N.; Krumbein, A. UV-B-induced secondary plant metabolites-potential benefits for plant and human health. *Crit. Rev. Plant. Sci.* **2012**, *31*, 229–240. [[CrossRef](#)]
12. Postler, T.S.; Ghosh, S. Understanding the holobiont: How microbial metabolites affect human health and shape the immune system. *Cell Metab.* **2017**, *26*, 110–130. [[CrossRef](#)] [[PubMed](#)]
13. Bonechi, C.; Lamponi, S.; Donati, A.; Tamasi, G.; Consumi, M.; Leone, G.; Rossi, C.; Magnani, A. Effect of resveratrol on platelet aggregation by fibrinogen protection. *Biophys. Chem.* **2017**, *222*, 41–48. [[CrossRef](#)] [[PubMed](#)]
14. Leicach, S.R.; Chludil, H.D. Plant secondary metabolites: Structure–activity relationships in human health prevention and treatment of common diseases. *Stud. Nat. Prod. Chem.* **2014**, *42*, 267–304.



15. Bonechi, C.; Donati, A.; Tamasi, G.; Leone, G.; Consumi, M.; Rossi, C.; Lamponi, S.; Magnani, A. Protective effect of quercetin and rutin encapsulated liposomes on induced oxidative stress. *Biophys. Chem.* **2018**, *233*, 55–63. [[CrossRef](#)] [[PubMed](#)]
16. Tiwari, R.; Rana, C.S. Plant secondary metabolites: A review. *Int. J. Eng. Res. Gen. Sci.* **2015**, *3*, 661–670.
17. Liao, B.; Chen, X.; Han, J.; Dan, Y.; Wang, L.; Jiao, W.; Song, J.; Chen, S. Identification of commercial Ganoderma (Lingzhi) species by ITS2 sequences. *Chin. Med.* **2015**, *10*, 22–30. [[CrossRef](#)]
18. Bonechi, C.; Donati, A.; Tamasi, G.; Pardini, A.; Rostom, H.; Leone, G.; Lamponi, S.; Consumi, M.; Magnani, A.; Rossi, C. Chemical characterization of liposomes containing nutraceutical compounds: Tyrosol, hydroxytyrosol and oleuropein. *Biophys. Chem.* **2019**, *246*, 25–34. [[CrossRef](#)]
19. Pei, Y.; Wu, L.; Zhang, Q.; Wang, Y. Geographical traceability of cultivated Paris polyphylla var. yunnanensis using ATR-FTMIR spectroscopy with three mathematical algorithms. *Anal. Methods* **2019**, *11*, 113–122. [[CrossRef](#)]
20. Carballo, T.; Gil, M.V.; Gomez, X.; Gonzalez-Andres, F.; Moran, A. Characterization of different compost extracts using fourier-transform infrared spectroscopy (FTIR) and thermal analysis. *Biodegradation* **2008**, *19*, 815–830. [[CrossRef](#)]
21. Durazzo, A.; Kiefer, J.; Lucarini, M.; Camilli, E.; Marconi, S.; Gabrielli, P.; Aguzzi, A.; Gambelli, L.; Lisciani, S.; Marletta, L. Qualitative analysis of traditional italian dishes: FTIR approach. *Sustainability* **2018**, *10*, 4112. [[CrossRef](#)]
22. Bunghez, I.R.; Raduly, M.; Doncea, S.; Aksahin, I.; Ion, R.M. Lycopene determination in tomatoes by different spectral techniques (UV-VIS, FTIR and HPLC). *Digest J. Nanomater. Biostruct.* **2011**, *6*, 1349–1356.
23. Tan, C.P.; Che Man, Y.B. Comparative differential scanning calorimetric analysis of vegetable oils: I. Effects of heating rate variation. *Phytochem. Anal.* **2002**, *13*, 129–141. [[CrossRef](#)]
24. Kotti, F.; Chiavaro, E.; Cerretani, L.; Barnaba, C.; Gargouri, M.; Bendini, A. Chemical and thermal characterization of Tunisian extra virgin olive oil from Chetoui and Chemlali cultivars and different geographical origin. *Eur. Food Res. Technol.* **2009**, *228*, 735–742. [[CrossRef](#)]
25. Tian, Y.; Li, Y.; Xu, X.; Jin, Z. Starch retrogradation studied by thermogravimetric analysis (TGA). *Carbohydr. Polym.* **2011**, *84*, 1165–1168. [[CrossRef](#)]
26. Manara, P.; Vamvuka, D.; Sfakiotakis, S.; Vanderghem, C.; Richel, A.; Zabaniotou, A. Mediterranean agri-food processing wastes pyrolysis after pre-treatment and recovery of precursor materials: A TGA-based kinetic modeling study. *Food Res. Int.* **2015**, *73*, 44–51. [[CrossRef](#)]
27. Kozukue, N.; Friedman, M. Tomatine, chlorophyll,  $\beta$ -carotene and lycopene content in tomatoes during growth and maturation. *J. Sci. Food Agric.* **2003**, *83*, 195–200. [[CrossRef](#)]
28. Leone, G.; Consumi, M.; Franzi, C.; Tamasi, G.; Lamponi, S.; Donati, A.; Magnani, A.; Rossi, C.; Bonechi, C. Development of liposomal formulations to potentiate natural lovastatin inhibitory activity towards 3-hydroxy-3-methyl-glutaryl coenzyme A (HMG-CoA) reductase. *J. Drug Deliv. Sci. Technol.* **2018**, *43*, 107–112. [[CrossRef](#)]
29. Leone, G.; Consumi, M.; Pepi, S.; Lamponi, S.; Bonechi, C.; Tamasi, G.; Donati, A.; Rossi, C.; Magnani, A. Alginate-gelatin formulation to modify lovastatin release profile from red yeast rice for hypercholesterolemia therapy. *Ther. Deliv.* **2017**, *8*, 843–854. [[CrossRef](#)]
30. Khiari, B.; Moussaoui, M.; Jeguirim, M. Tomato-processing by-product combustion: Thermal and kinetic analyses. *Materials* **2019**, *12*, 553. [[CrossRef](#)]
31. Vega, D.; Villar, M.A.; Failla, M.D.; Vallés, E.M. Thermogravimetric analysis of starch-based biodegradable blends. *Polym. Bull.* **1996**, *37*, 229–235. [[CrossRef](#)]
32. De Nardo, T.; Shiroma-Kian, C.; Halim, Y.; Francis, D.; Rodriguez-Saona, L.E. Rapid and simultaneous determination of lycopene and  $\beta$ -carotene contents in tomato juice by infrared spectroscopy. *J. Agric. Food Chem.* **2009**, *57*, 1105–1112. [[CrossRef](#)]
33. Heredia-Guerrero, J.A.; Benítez, J.J.; Domínguez, E.; Bayer, I.S.; Cingolani, R.; Athanassiou, A.; Heredia, A. Infrared and Raman spectroscopic features of plant cuticles: A review. *Front. Plant. Sci.* **2014**, *5*, 305–316. [[CrossRef](#)]
34. Yap, K.Y.L.; Chan, S.Y.; Lim, C.S. Infrared-based protocol for the identification and categorization of ginseng and its products. *Food Res. Int.* **2007**, *40*, 643–652. [[CrossRef](#)]
35. Osborne, J.W.; Costello, A.B. Sample size and subject to item ratio in principal components analysis. *Pract. Assess. Res. Eval.* **2004**, *9*, 11–20.
36. USDA. *United States Standards for Grades of Fresh Tomatoes*; United States Department of Agriculture: Washington, DC, USA, 1997.
37. Leone, G.; Consumi, M.; Lamponi, S.; Bonechi, C.; Tamasi, G.; Donati, A.; Rossi, C.; Magnani, A. Thixotropic PVA hydrogel enclosing a hydrophilic PVP core as nucleus pulposus substitute. *Mater. Sci. Eng. C* **2019**, *98*, 696–704. [[CrossRef](#)] [[PubMed](#)]
38. Leone, G.; Consumi, M.; Pepi, S.; Pardini, A.; Bonechi, C.; Tamasi, G.; Donati, A.; Lamponi, S.; Rossi, C.; Magnani, A. Enriched gellan gum hydrogel as visco-supplement. *Carbohydr. Polym.* **2020**, *227*, 115347. [[CrossRef](#)] [[PubMed](#)]
39. Garcia-Jares, C.M.; Médina, B. Application of multivariate calibration to the simultaneous routine determination of ethanol, glycerol, fructose, glucose and total residual sugars in botrytized-grape sweet wines by means of near-infrared reflectance spectroscopy. *Fresenius J. Anal. Chem.* **1997**, *357*, 86–91. [[CrossRef](#)]
40. Oliveri, P. Class-modelling in food analytical chemistry: Development, sampling, optimisation and validation issues. A tutorial. *Anal. Chim. Acta* **2017**, *982*, 9–19. [[CrossRef](#)]
41. Shah, N.; Cynkar, W.; Smith, P.; Cozzolino, D. Use of attenuated total reflectance midinfrared for rapid and real-time analysis of compositional parameters in commercial white grape juice. *J. Agric. Food Chem.* **2010**, *58*, 3279–3283. [[CrossRef](#)]



42. Li, B.; Martin, E.; Morris, J. Latent variable selection in partial least squares modelling. *IFAC Proc. Vol.* **2001**, *34*, 463–468. [[CrossRef](#)]
43. Croce, R.; Malegori, C.; Oliveri, P.; Medici, I.; Cavaglioni, A.; Rossi, C. Prediction of quality parameters in straw wine by means of FT-IR spectroscopy combined with multivariate data processing. *Food Chem.* **2020**, *305*, 125512. [[CrossRef](#)]
44. Silalahi, D.D.; Midi, H.; Arasan, J.; Mustafa, M.S.; Caliman, J.P. Robust generalized multiplicative scatter correction algorithm on pretreatment of near infrared spectral data. *Vib. Spectrosc.* **2018**, *97*, 55–65. [[CrossRef](#)]
45. Barnes, R.J.; Dhanoa, M.S.; Lister, S.J. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Appl. Spectrosc.* **1989**, *43*, 772–777. [[CrossRef](#)]
46. Savitzky, A.; Golay, M.J.E. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **1964**, *36*, 1627–1639. [[CrossRef](#)]
47. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130. [[CrossRef](#)]
48. Forina, M.; Latenri, S.; Cerrato Oliveros, M.C.; Pizarro Millan, C. Selection of useful predictors in multivariate calibration. *Anal. Bioanal. Chem.* **2004**, *380*, 397–418. [[CrossRef](#)]