Investigating the Origin of Tomatoes and Triple Concentrated Tomato Pastes through Multielement Determination by Inductively Coupled Plasma Mass Spectrometry and Statistical Analysis

Gabriella Lo Feudo,[†] Attilio Naccarato,[‡] Giovanni Sindona,*[,‡] and Antonio Tagarelli[‡]

[†]INCA Istituto Nazionale delle Conserve Alimentari—Sede di Cosenza, Via N. Parisio 13, I-87100 Cosenza, Italy, and [‡]Dipartimento di Chimica, Università della Calabria, Via P. Bucci Cubo 12/C, I-87030 Arcavacata di Rende (CS), Italy

The concentration of 32 elements (Al, As, Ba, Be, Ca, Cd, Ce, Cu, Dy, Fe, K, La, Lu, Mg, Mn, Na, Nd, Pb, Rb, Sm, Sr, Th, U, V, Zn) was determined in tomatoes harvested in different four Italian regions and in triple concentrated tomato paste samples coming from Italy, China, Greece and California. The resulting multielement profiles were processed using three chemometric techniques to evaluate the possibility of discrimination between different cultivation areas. The closed-vessel microwave digested samples were diluted and analyzed by DRC-ICP-MS with CH_4 as reaction gas. The accuracy of the proposed method was considered acceptable (values in the range 75−120%) for 25 out of the 35 elements of the reference material NCS ZC85006 Tomato. The origin of tomato fruits and the areas of production as "Italy" and "non-Italy" of the triple concentrated pastes were evaluated by three supervised pattern recognition procedures, linear discriminant analysis (LDA), soft independent modeling of class analogy (SIMCA) and *K*-nearest neighbors (KNN).

KEYWORDS: Tomato; trace elements; ICP-MS; authenticity; statistical analysis

INTRODUCTION

Tomatoes (*Lycopersicum esculentum*) are vegetables worldwide used in the diet. Their intake is referred to fresh fruits and also to semifinished products as sauces, purées and pastes (mono, double and triple concentrated). Moreover, tomatoes have important healing effects associated with the presence of antioxidants, such as carotenoids, polyphenols, and vitamins. The availability of these active principles in final products is frequently affected by product manipulation during the several steps of the industrial processing. In this regard, the main adulteration of tomato products consists in the dilution of tomato pastes with water instead of using fresh fruits alone. Another aspect is represented by the importation of triple concentrated pastes from countries other than those of the producers. In 2007, 160 million kilograms of tomato pastes, corresponding to about 25% of the Italian tomato production, have in fact been imported from abroad (*1*). Authentication of food products is of primary importance for consumers, farmers and producers. For consumers, geographical origin assures quality and organoleptic and nutritional characteristics whereas, from an economic point of view, product authentication is fundamental to prevent unfair competition that can eventually affect the regional and even national economy. In order to protect both the consumers and the national production, a decree was issued by Italy in 2006 that

established the obligation from 1 January 2008 to state clearly in the label the origin (region or state) of fresh tomato used in the preparation of sauces (*2*). Similar problems apply to all the tomato production countries. The number of critical points in the production chain of tomato derivatives calls for the introduction of efficient analytical methods to reliably check the origin of the raw materials. A traceability study based on a DNA extraction procedure has been published (*3*), and the metabolomics of tomato fruit has been frequently addressed (*4−8*). Few investigations are known for tomato juice and tomato paste (*9−11*). Two of these studies deal with the determination of geographical origin of cherry tomatoes (*12*) and triple concentrated tomato pastes (*13*) by ^{1}H NMR spectroscopy.

Multielement analysis has been applied to a range of foodstuff to develop methods that allow the identification of their geographical origins (*14*, *15*). Recently, we have used trace element profiling for the characterization and adulteration detection of olive oil and milk (*16*, *17*). Moreover, trace element concentration from tomatoes was subjected to a chemometric approach to distinguish between conventional and organic cultivation (*18*). Atomic absorption spectrometry (AAS) and inductively coupled plasma atomic emission spectroscopy (ICP-AES) are commonly used techniques for the determination of elements in different food matrices (*19−21*). Inductively coupled plasma mass spectrometry (ICP-MS) represents an increasingly popular technique because of its well-known high sensitivity and rapid multielement capability for each single sampling (*22−25*).

*Corresponding author. Phone: +39-0984-492046. Fax: +39-0984-493307. E-mail: sindona@unical.it.

Figure 1. Italy map.

Four Italian production regions, namely, Calabria, Basilicata and Emilia Romagna, and two growing seasons, 2007 and 2008, were considered in this study. 46 elements were initially checked, but only 32 were submitted to statistical analysis to test the feasibility of discriminating between different cultivation areas, since acceptable recovery data were obtained for them. Three supervised pattern recognition procedures, linear discriminant analysis (LDA), K-nearest neighbors (KNN) and soft independent modeling of class analogy (SIMCA), were applied to evaluate the discrimination and classification ability of these multivariate techniques. The same method was used to match a tomato paste sample made in Italy with those produced in other countries such as California, China, and Greece.

MATERIALS AND METHODS

Chemicals and Instrumentation. The reagents used for mineralization (HNO_3 (65%), H_2O_2 (30%), HCl (30%), HF (40%), $HClO_4$ (70%), H_3BO_3) were Suprapur (Merck, Darmstadt, Germany). All other reagents used for analysis were of analytical reagent grade (Merck, Darmstadt, Germany). Two multielement solutions of Ag, Al, As, Ba, Be, Bi, Ca, Cd, Co, Cr, Cs, Cu, Fe, Ga, In, K, Li, Mg, Mn, Na, Ni, Pb, Rb, Se, Sr, Tl, V, U and Zn (100 mg/L, Merck) and Ce, Dy, Er, Eu, Gd, Ho, La, Lu, Nd, Pr, Sm, Sc, Tb, Th, Tm, Y and Yb (10 mg/L, PerkinElmer) were used to prepare the calibration standards. Aqueous solutions were prepared using ultrapure water, with a resistivity of 18.2 MΩ cm, obtained from a Milli-Q plus system (Millipore, Bedford, MA). All glassware was decontaminated with nitric acid (2%, v/v) overnight, rinsed with ultrapure water and dried.

The accuracy of the method was evaluated by analyzing the certified reference material NCS ZC85006 Tomato (China National Analysis Center for Iron & Steel 2000).

The sample preparation was carried out using the following system for microwave digestion: Anton Paar Multiwave 3000 with programmable power control (maximum power 1400 W) and rotor XF100 (operating pressure up to 120 bar maximum; operating temperature 260 °C maximum; construction material PTFE-TFM for the vessel). The determination of the elements of interest was carried out utilizing an Elan DRC-e ICP-MS instrument (Perkin-Elmer SCIEX, Canada). Samples were introduced by means of a crossflow quartz nebulizer with a Scott-type spray chamber. The ICP torch was a standard torch (Fassel type torch) with platinum injector.

A solution containing Rh, Mg, Pb, Ba and Ce (10 μg/L) was used to optimize the instrument in terms of sensitivity, resolution and mass calibration. The $^{140}Ce^{16}O^+/^{140}Ce^+$ ratio was used to check the level of oxide ions in the plasma that could interfere in the determination of some elements; also, instrumental parameters such as RF power and carrier gas flow were optimized and the level of doubly charged ion monitored by means of the signal $^{137}Ba^{2+}/^{137}Ba^+$. The $^{140}Ce^{16}O^+/^{140}Ce^+$ and $^{137}Ba^{2+}/^{137}Ba^+$ ratios obtained after optimization were 2.8 and 3.2%, respectively.

Sampling. Italian and foreign samples of certain origin were provided by Istituto Nazionale delle Conserve Alimentari (INCA). A total of about 130 tomato fruits of cv. Perfectpeel were hand harvested in August 2007 and 2008 from four different Italian regions (Calabria, Basilicata, Puglia and Emilia Romagna, **Figure 1**) and immediately stored at −20 °C.

One hundred samples (500 g for each samples) of triple concentrated tomato paste, 50 Italian, 10 Chinese, 20 Californian and 20 Greek, were screened. Italian tomato paste samples were obtained from the same producers of the tomato samples (**Table 1**).

Article

J. Agric. Food Chem., Vol. 58, No. 6, 2010 **3803**

Table 1. Number of Tomato and Triple Concentrated Tomato Paste Samples

		tomato		paste	
zone	region	2007[a]	2008[a]	2007[b]	2008[b]
Mesagne	Puglia (Italy)	10	10		
Collecchio	Emilia Romagna (Italy)	10	20	10	10
Matera	Basilicata (Italy)	10	10		
Crotone	Calabria (Italy)	20	20	20	10
unknown	(China)				10
unknown	(California)				20
unknown	(Greece)				20

[a] Harvest year. [b] Production year.

Analytical Procedure. In order to check the effectiveness of the digestion procedure, about 0.5 g of the certified reference material was subjected to digestion treatment using a mixture of 4.5 mL of HNO_3/1 mL of H_2O_2 and 0.5 mL of each of the following third components: HF, HCl, H_3BO_3 and $HClO_4$. The HNO_3/H_2O_2/HF mixture provided the best digestion results in terms of accuracy for the elements certified in reference material NCS ZC85006 Tomato.

An aliquot of tomato sample (10 g) or triple concentrated tomato paste sample (1 g) was weighted directly into the vessel of the microwave system. The digestion was performed by adding 4.5 mL of HNO_3, 1 mL of H_2O_2 and 0.5 mL of HF to each sample. The operating conditions used for the microwave digestion system were 1000 W over ten minutes and hold at this power for eight minutes. After digestion the extracts were quantitatively transferred to a graduated polypropylene test tube, and the volume was adjusted to 50 mL with ultrapure water. The analytical batch consisted of a set of calibration standards, samples, a minimum of three procedural blanks, one procedural blank spiked with a solution containing the elements of interest and the certified reference material. A midrange calibration standard was measured at the end of each analytical run, in order to assess instrumental drift throughout the run.

Eight point calibration curves covering the range 0.1−2000 μg/L were used for quantitative analysis. Standard solutions were prepared by diluting the multielement solutions cited in the Chemicals and Instrumentation section.

Statistical Analysis. LDA was performed by Statistica 8.0 (StatSoft 2007 Edition); KNN and SIMCA were executed by V-Parvus 2004 (26).

RESULTS AND DISCUSSION

Analytical Performance. Three aliquots of nearly 0.5 g of the certified reference material in 4.5 mL of HNO_3/1 mL of H_2O_2 were subjected to digestion according to the operating conditions described in the Analytical Procedure section. To set up the best experimental conditions the experiment was repeated in the presence of 0.5 mL of each of the following third components: HF, HCl, H_3BO_3 and $HClO_4$.

The extracts were diluted to 50 mL with ultrapure water and subjected to ICP-MS analysis. Initially, 46 elements were investigated: ^{7}Li, ^{9}Be, ^{23}Na, ^{24}Mg, ^{27}Al, ^{39}K, ^{44}Ca, ^{45}Sc, ^{51}V, ^{52}Cr, ^{53}Cr, ^{55}Mn, ^{57}Fe, ^{58}Ni, ^{60}Ni, ^{59}Co, ^{63}Cu, ^{64}Zn, ^{66}Zn, ^{69}Ga, ^{75}As, ^{82}Se, ^{85}Rb, ^{88}Sr, ^{89}Y, ^{107}Ag, ^{114}Cd, ^{115}In, ^{133}Cs, ^{138}Ba, ^{139}La, ^{140}Ce, ^{141}Pr, ^{142}Nd, ^{152}Sm, ^{153}Eu, ^{158}Gd, ^{159}Tb, ^{164}Dy, ^{165}Ho, ^{166}Er, ^{169}Tm, ^{174}Yb, ^{175}Lu, ^{205}Tl, ^{208}Pb, ^{209}Bi, ^{232}Th and ^{238}U.

The determination of some elements by ICP-MS is known to suffer from polyatomic isobaric interference. For example, regarding the elements of our interest, the scandium signal at m/z 45 is affected by the possible presence of $^{13}C^{16}O_2^+$ and $^{29}Si^{16}O^+$, the chromium signal at m/z 53 by $^{40}Ar^{13}C^+$ and $^{37}Cl^{16}O^+$, the iron signal at m/z 56 by $^{40}Ca^{16}O^+$, the nickel signal at m/z 58 by $^{42}Ca^{16}O^+$, the zinc signal at m/z 64 by $^{48}Ti^{16}O^+$ and $^{32}S_2^+$, the selenium signal at m/z 80 by $^{40}Ar^{40}Ar^+$ and, finally, the europium signal at m/z 153 by $^{137}Ba^{16}O^+$. The dynamic reaction cell (DRC) is proved to be an effective method for relieving such isobaric interferences. Then Sc, Cr, Fe, Ni, Zn, Se and Eu have been monitored in both modes (standard and DRC modes), in the

Table 2. Instrumental Parameters and Operating Conditions for the ICP-MS Instrument

rf power (W)	1130
nebulizer (carrier gas) flow rate (L min^{-1})	0.95
lens voltage (V)	6.25
analog stage voltage (V)	−1900
pulse stage voltage (V)	1050
discriminator threshold (V)	70
quadrupole rod offset (V)	0
resolution (amu)	0.70
detector	Dual
speed of peristaltic pump (rpm)	24
sweeps/reading	20
replicates	3
dwell time	50 ms
scan mode	Peak hopping
DRC parameters	
CH_4 reaction gas flow (mL/min)	0.70 for Cr, Fe Se and Zn
	1.15 for Sc, Ni and Eu
rejection parameter a (RPa)	0
rejection parameter q (RPq)	0.5 for Eu
	0.6 for Sc, Ni, Se and Zn
	0.7 for Cr and Fe

presence of methane as reaction gas. The optimization of the CH_4 flow rate and the RPq value (RPq is an ELAN parameter that is proportional to the Mathieu parameter q) were carried out using matrix blank solution prepared with HCl (2%), HNO_3 (1%), CH_3OH (1%), Ca at 20 mg/L, Si and Ba at 50 μg/L and a matrix blank solution spiked with Sc, Cr, Fe, Ni, Zn and Se at 1 μg/L and Eu at 0.1 μg/L. The best background equivalent concentrations (BEC) were obtained at a flow rate value (mL/min) of 0.7 for Cr, Fe, Se and Zn and at value of 1.15 for Sc, Ni and Eu. The best S/N ratio was obtained with a RPq value of 0.6 for all elements except for Eu (0.5), Cr (0.7) and Fe (0.7). ICP-MS analyses were performed following the operating program and parameters shown in **Table 2**.

Under these instrumental conditions, the extracts obtained using various mixtures of digestion reagents were analyzed. The HNO_3/H_2O_2/HF mixture provided the best digestion results in terms of accuracy for the 35 elements certified in reference material NCS ZC85006 Tomato (Be, Na, Mg, Al, K, Ca, Sc, V, Mn, Fe, Ni, Co, Cu, Zn, As, Se, Rb, Sr, Y, Cd, Cs, Ba, La, Ce, Nd, Sm, Eu, Gd, Dy, Er, Yb, Lu, Pb, Th, U). This material was recently used (i) to determine accuracy in a method for the determination of heavy metals by inductively coupled plasma optical emission spectrometry and flame atomic absorption spectrometry after co-preconcentration of the metals (27) and (ii) to validate a method for the determination of Cu, Cr, Pb, Mn and Cd in vegetables by graphite furnace atomic absorption spectrometry (28). Accuracy values were considered acceptable for 25 elements (values in the range 75−120%). Analytical parameters for these elements were shown in **Table 3**.

Although accuracy values for Cu and K were satisfactory, these two elements were not submitted to statistical analysis due to their use in agricultural practice as fertilizers. The chemometric treatments have been applied to the data set containing the concentration of those 23 elements that have shown the best accuracy values, and to the concentration of 9 elements (Ag, Cr, Ga, Ho, Li, Pr, Tb, Tl and Tm), for which results were above the limit of detection (LOD). Among the elements monitored in standard and DRC modes, significant differences have been observed in the assay of ^{56}Fe and ^{52}Cr. Therefore, the statistical analysis has been carried out by means of the data acquired in standard mode except for iron and chromium, whose DRC values were used.

Table 3. Summary of Calibration Parameters, Limits of Detection (LODs), Limits of Quantitation (LOQs) and Mean Accuracies (%) Referred to Certified Reference Material NCS ZC85006 Tomato

	isotope	calibration range (μg/L)	R^2	LOD (μg/kg)[a]	LOQ (μg/kg)[a]	certified value (mg/kg)[b]	found value (mg/kg)[c]	accuracy (%)
Al	27	0.1−2000	0.9998	159	346	0.295 ± 0.043	0.279 ± 0.008	94.5
As	75	0.1−2000	0.9999	0.42	0.55	1.05 ± 0.13	1.03 ± 0.02	98.6
Ba	138	0.1−2000	0.9999	0.68	0.98	55.2 ± 5.2	52.9 ± 0.8	95.9
Be	9	0.1−2000	0.9999	0.031	0.071	(0.084)	0.067 ± 0.009	80.2
Ca	44	0.1−2000	0.9982	1441	1968	5.31 (%) ± 0.19	4.58 ± 0.06	86.3
Cd	114	0.1−2000	0.9999	0.13	0.31	0.82 ± 0.09	0.685 ± 0.004	83.6
Ce	140	0.01−150	0.9999	0.063	0.084	3.08 ± 0.22	2.94 ± 0.26	95.5
Cu	63	0.1−2000	0.9997	0.76	0.95	21.1 ± 2.5	23.5 ± 0.5	111.5
Dy	164	0.01−150	0.9999	0.007	0.014	(0.23)	0.217 ± 0.010	94.3
Fe[d]	56	0.1−2000	0.9999	142	179	0.138 (%) ± 0.015	0.139 ± 0.003	100.9
K	39	0.1−2000	0.9991	92.2	127.2	0.579 (%) ± 0.052	0.564 ± 0.010	97.4
La	139	0.01−150	0.9999	0.031	0.047	1.78 ± 0.17	1.34 ± 0.15	75.4
Lu	175	0.01−150	0.9999	0.001	0.002	(0.019)	0.019 ± 0.002	102.2
Mg	24	0.1−2000	0.9998	50.7	105	0.736 (%) ± 0.057	0.645 ± 0.006	87.6
Mn	55	0.1−2000	0.9997	2.81	4.22	87.1 ± 5.6	104.1 ± 1.1	119.6
Na	23	0.1−2000	0.9997	304	722	(0.13)	0.126 ± 0.003	96.8
Nd	142	0.01−150	0.9999	0.015	0.038	(1.28)	1.36 ± 0.13	106.0
Pb	208	0.1−2000	0.9999	0.24	0.33	4.97 ± 0.54	5.49 ± 0.03	110.5
Rb	85	0.1−2000	0.9999	1.18	2.89	6.66 ± 0.47	6.98 ± 0.01	104.8
Sm	152	0.01−150	0.9999	0.013	0.030	0.270 ± 0.020	0.299 ± 0.019	110.8
Sr	88	0.1−2000	0.9999	2.02	3.15	569 ± 40	610 ± 8	107.2
Th	232	0.01−150	0.9992	0.020	0.042	(0.486)	0.488 ± 0.076	100.5
U	238	0.1−2000	0.9999	0.053	0.086	(0.202)	0.185 ± 0.005	91.4
V	51	0.1−2000	0.9999	1.74	2.65	3.84 ± 0.30	3.83 ± 0.14	99.7
Zn	64	0.1−2000	0.9999	81.8	188	36.2 ± 3.1	33.8 ± 0.56	93.3

[a] LOD and LOQ values are referred to analysis of tomato samples. [b] Certified values without standard deviation are reported in parentheses. [c] ±SD, n = 3. [d] Analyzed in DRC mode.

The mean concentration and the relative standard deviation in tomato samples and triple concentrated tomato paste samples analyzed by ICP-MS can be seen in the Supporting Information.

Statistical Analysis. Three supervised pattern recognition techniques (LDA, SIMCA and KNN) were chosen in order to obtain classification rules for distinguishing between four Italian cultivation areas of tomatoes (Calabria, Basilicata, Puglia, Emilia Romagna) and between triple concentrated tomato paste samples coming from Italy and from foreign countries (California, China, Greece). These techniques were applied to concentration values (μg/kg) of the selected 32 elements. The reliability of the classification rules needs to be validated, and therefore a cross-validation procedure was performed. The sample set divided into training set including 80% of the components of sample set was used for calculating models, and then the validation set which includes 20% of the components of the sample set was used as an unknown, and classified (five cancellation groups, 5CV). The process was reiterated five times with different random constitutions of both sets to ensure that all the objects were included in the validation set at least once. The goodness of the classification was estimated in terms of prediction ability, which is equal to the percentage of the test set members correctly classified.

One of the most commonly used classification techniques is linear discriminant analysis (LDA). It defines a set of delimiters (depending on the number of considered categories) so that the multivariate space of the objects is divided in as many subspaces as the number of classes. Discriminant functions are obtained as a linear combination of descriptor that maximizes the ratio of variance between categories to variance within categories.

In classification techniques such as LDA, even if the samples that will be predicted by LDA did not belong to any of the categories of the model, each object will be assigned anyway to one of them. Moreover, each object can fall into one and only one category because the multidimensional space is divided into many subspaces as classes. On the other hand, a peculiarity of soft modeling techniques is that an object can belong to one category, to more than one class simultaneously or to none of the categories. Soft independent modeling of class analogy (SIMCA) is a class modeling technique that builds a class model based on the significant principal components (PCs) of the category. The range of the scores of the *N* principal components used to build the model for each category is the edges of a hyper volume, the "normal" SIMCA model. Then, in this technique the models (one for each class) can overlap and/or leave some regions of the multivariate space unassigned. An important consequence of this feature is that SIMCA is able to detect the number of false positives/negatives for each class. Therefore, two parameters to validate the classification can be defined: sensitivity (SENS) and specificity (SPEC) (*29*, *30*). SENS of a class is referred to the percentage of objects belonging to the class correctly accepted by the class model. SPEC of a class corresponds to the percentage of objects not belonging to the class correctly rejected by the class model.

K-nearest neighbors (KNN) is a nonparametric method that categorizes a test object according to proximity to others already classified into classes, by using the Euclidean distance in the multidimensional space as a similarity measurement. The most important feature of the nonparametric method is that it is free from statistical assumption about distribution of variables, while the parametric methods such as LDA and SIMCA are dependent on a given distribution.

Tomato Samples. Chemometric treatments have been applied to the concentration of the 32 elements determined in the samples listed in **Table 1**. It is important to underline that LDA requires the data matrices for each category to have a high ratio between the number of training samples and the number of variables because discriminant analysis using a low ratio of objects to variables could generate an unstable and not very robust model. In our case, the total number of objects (110) is more than three times the number of variables (32) then LDA was performed by
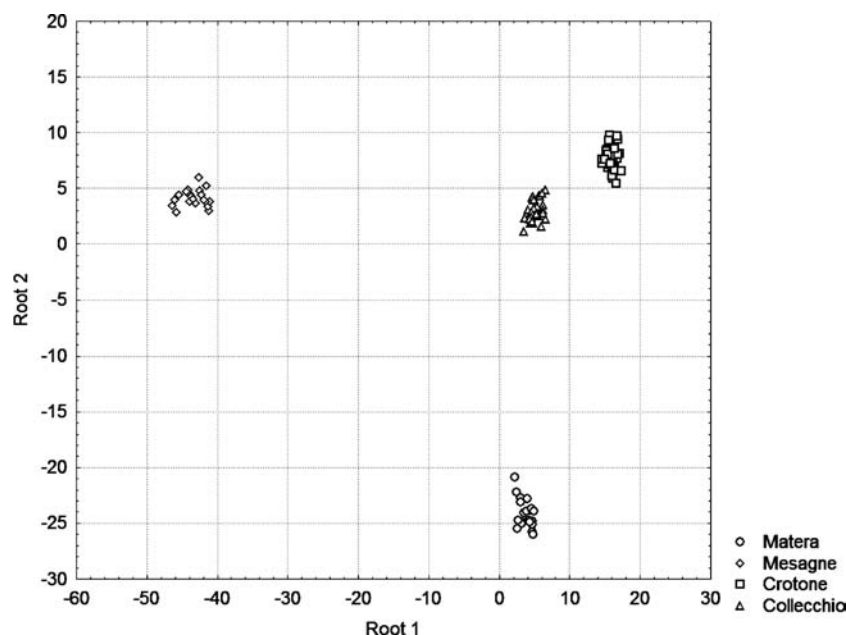
Figure 2. LDA plot for tomato samples.

considering all the elements without applying any method for reduction of the variables. Standard LDA has been applied using four groups corresponding to the four cultivation zones, as input a priori. The bidimensional plot of the first two roots shows a very clear separation of four clusters corresponding to the four cultivation zones (**Figure 2**).

The differentiation between groups is very significant since the very low Wilks λ value (0.0000011) shows that the model is highly discriminating, whereas the high value (325.1) of the F (96.22) parameter indicates a significant difference between the means of the groups. Finally, the probability of a correct classification is very high, because the *p*-level is very low (< 0.00001). To verify the goodness of the method in terms of prediction ability, cross validation was performed. The proposed model has correctly predicted all samples for each zone and this suggests that the presented method may be a potential choice for checking origin of tomatoes. The analyses of the discriminating functions show that Cd, alkaline metals and alkaline earth metals are the most important variables in the distinction of geographical origin. The importance of alkaline earth metals in differentiation of cultivation zones stood out also in a previous study about the identification of the origin of virgin olive oil (*16*).

SIMCA was applied to the same data matrix subjected to LDA. The models obtained were based on normal range, 5% as the significance level for critical distance and a number of components for each category so that explains about 90% of class variance. To evaluate the predictive capability of SIMCA, the same cross-validation procedure was applied. Also the SIMCA model has correctly predicted all samples for each zone (**Table 4A**). SENS and SPEC values obtained were very satisfactory as SIMCA model presented a mean sensitivity of 84.5% and a mean specificity of 100% (**Table 4B**). Finally, also in this case, the discriminant powers of variables show that alkaline metals and alkaline earth metals are among the compounds that mostly contribute to discrimination of cultivation zones.

KNN was applied to the same data set subjected to LDA and SIMCA using the square inverse of the Euclidean distance and the decision criterion of majority vote. The number of neighbors was chosen after evaluation of the success in classification with K values between 1 and 12. The best results (**Table 4C**) were

Table 4. Prediction Matrices for Supervised Pattern Recognition Techniques in the Cross-Validation Procedure for Tomato Samples (Rows Represent the True Class, Columns Report the Assigned Class)

(A) LDA

	Matera	Mesagne	Crotone	Collecchio
Matera	20	0	0	0
Mesagne	0	20	0	0
Crotone	0	0	40	0
Collecchio	0	0	0	30

(B) SIMCA

	Matera	Mesagne	Crotone	Collecchio	SENS (%)	SPEC (%)
Matera	20	0	0	0	80	100
Mesagne	0	20	0	0	80	100
Crotone	0	0	40	0	85	100
Collecchio	0	0	0	30	90	100

(C) KNN

	Matera	Mesagne	Crotone	Collecchio
Matera	17	0	3	0
Mesagne	0	20	0	0
Crotone	1	0	35	4
Collecchio	0	0	5	25

obtained with $K = 3$. The differentiation between geographical zones can be considered acceptable for all classes (mean prediction ability of 88.2%), but better results were obtained by LDA and SIMCA techniques.

Triple Concentrated Tomato Paste Samples. The three multivariate techniques have been applied to the concentration of the same 32 elements considered in the tomato samples. Also in this case, the total number of objects (100, see **Table 1**) is more than

3806 *J. Agric. Food Chem.*, Vol. 58, No. 6, 2010

Lo Feudo et al.

Table 5. Prediction Matrices for Supervised Pattern Recognition Techniques in the Cross-Validation Procedure for Triple Concentrated Tomato Paste Samples (Rows Represent the True Class, Columns Report the Assigned Class)

(A) LDA		
	Italy	non-Italy
Italy	49	1
non-Italy	1	49

(B) SIMCA			
	Italy	non-Italy	SENS (%)	SPEC (%)
Italy	48	2	90	100
non-Italy	0	50	80	100

(C) KNN		
	Italy	non-Italy
Italy	50	0
non-Italy	0	50

three times the number of variables (32) so LDA has been carried out without applying any method for reduction of the number of variables. Standard LDA has been performed using two categories associated to samples coming from Italy (Italy group) and samples from foreign countries (non-Italy group), as input a priori. Also for triple concentrated tomato paste samples statistical parameters are very satisfactory (Wilks $\lambda = 0.00975$, $F(32.67) = 212.6$ and p-level < 0.00001). The cross-validation procedure was performed as for tomato samples and has shown a prediction ability of 98% for each class (**Table 5A**). The analyses of the discriminating functions have again underlined that two alkaline metals (Li and Rb) are the most important variables in the distinction of geographical origin. For the samples considered, the contribution of alkaline earth metals to LDA model seems not important for differentiation of triple concentrated paste tomato samples.

SIMCA was also applied to the same data matrix, and models obtained were based on five components for category "Italy" and four components for category "non-Italy" (explaining about 90% of variance for both classes). Again the cross-validation procedure was applied and, also in this case, the SIMCA model has provided very good results in terms of prediction ability (96% for Italy, 100% for non-Italy). SENS and SPEC values obtained were very satisfactory as the SIMCA model presented a mean sensitivity of 85% and a mean specificity of 100% (**Table 5B**).

KNN was applied using the same criteria chosen for tomato samples. The number of neighbors was evaluated, and the best results were obtained with $K = 5$ (**Table 5C**). In this case, the KNN method is capable to classify correctly all samples for both categories, and this technique has provided better results than those obtained by the LDA and SIMCA methods.

In conclusion, this study shows that the discrimination between tomato samples cultivated in different areas and between tomato paste samples coming from different countries might be achieved by a simple and rapid method, such as ICP-MS, followed by a convenient statistical processing of multielement profile. The ICP-MS step was preceded by digestion treatment using various inorganic acids and hydrogen peroxide in a closed-vessel microwave system, and the overall analytical procedure was evaluated in terms of accuracy compared to the values of the certified elements in reference material NCS ZC85006 Tomato. Accura-

cies were considered acceptable for 25 elements (values in the range 75−120%).

Three supervised pattern recognition procedures (LDA, SIMCA and KNN) were applied to concentration of the selected elements, and prediction ability of these multivariate techniques was tested using cross-validation procedure. For tomato samples, even if KNN produced a model capable of achieving good percentage for prediction (88.2%), LDA and SIMCA were preferred since these techniques are capable of classifying correctly all samples for each category. Moreover, a sensitivity (SENS) mean value of 84.5% and specificity (SPEC) mean value of 100% were achieved by the SIMCA model. The three chemometric approaches were applied also to triple concentrated tomato paste samples coming from Italy and from foreign countries. In this case, excellent results were achieved by all models and, in particular, the KNN method has correctly classified all samples for both categories "Italy" and "non-Italy".

Another significant achievement of the proposed method is the observation that the classification is unaffected by the production year for the concentrated paste and the harvesting years for fresh tomatoes.

Supporting Information Available: The mean concentration of the elements in tomato samples and triple concentrated tomato paste samples analyzed by ICP-MS. This material is available free of charge via the Internet at http://pubs.acs.org.

LITERATURE CITED

(1) http://www.coldiretti.it.

(2) Decree of the Italian ministry of agriculture (MiPAF). *Gazz. Ufficiale* **2006**, *57*, 49.

(3) Turci, M.; Savo Sardaro, M. L.; Visioli, G.; Maestri, E.; Marmiroli, M.; Marmiroli, N. Evaluation of DNA extraction procedures for traceability of various tomato products. *Food Control* **2010**, *21*, 143–149.

(4) Arvanitoyannis, J. S.; Vaitsi, O. B. A review on tomato authenticity: quality control methods in conjunction with multivariate analysis (chemometrics). *Crit. Rev. Food Sci.* **2007**, *47*, 675–699.

(5) Clement, A.; Dorais, M.; Vernon, M. Multivariate approach to the measurement of tomato maturity and gustatory attributes and their rapid assessment by Vis-NIR spectroscopy. *J. Agric. Food Chem.* **2008**, *56*, 1538–1544.

(6) Raffo, A.; Leonardi, C.; Fogliano, V.; Ambrosino, P.; Salucci, M.; Gennaro, L.; Bugianesi, R.; Giuffrida, F.; Quaglia, G. Nutritional value of cherry tomatoes (Lycopersicon esculentum cv. Naomi F1) harvested at different ripening stage. *J. Agric. Food Chem.* **2002**, *50*, 6550–6556.

(7) Iijima, Y.; Suda, K.; Suzuki, T.; Aoki, K.; Shibata, D. Metabolite profiling of chalcones and flavanones in tomato fruit. *J. Jpn. Hortic. Sci.* **2008**, *77*, 94–102.

(8) Mattoo, A. K.; Sobolev, A. P.; Neelam, A.; Goyal, R. K.; Handa, A. K.; Segre, A. Nuclear magnetic resonance spectroscopy-based metabolite profiling of transgenic tomato fruit engineered to accumulate spermidine and spermine reveals enhanced anabolic and nitrogen-carbon interactions. *Plant Physiol.* **2006**, *142*, 1759–1770.

(9) Sobolev, A. P.; Segre, A.; Lamanna, R. Proton high field NMR study of tomato juice. *Magn. Reson. Chem.* **2003**, *41*, 237–245.

(10) Tiziani, S.; Schwartz, S. J.; Vodovotz, Y. Profiling of carotenoids in tomato juice by one- and two-dimensional NMR. *J. Agric. Food Chem.* **2006**, *54*, 6094–6100.

(11) Capanoglu, E.; Beekwilder, J.; Boyacioglu, D.; Hall, R.; De Vos, R. Changes in antioxidant and metabolite profiles during production of tomato paste. *J. Agric. Food Chem.* **2008**, *56*, 964–973.

(12) Sequi, P.; Dell'Abate, M. T.; Valentini, M. Identification of cherry tomatoes growth origin by means of magnetic resonance imaging. *J. Sci. Food Agric.* **2007**, *87*, 127–132.

(13) Consonni, R.; Cagliani, L. R.; Stocchero, M.; Porretta, S. Triple Concentrated Tomato Paste: Discrimination between Italian and Chinese Products. *J. Agric. Food Chem.* **2009**, *57*, 4506–4513.

Article

J. Agric. Food Chem., Vol. 58, No. 6, 2010   **3807**

(14) Kelly, S.; Heaton, K.; Hoogewerff, J. Tracing the geographical origin of food: The application of multi-element and multi-isotope analysis. *Trends Food Sci. Technol.* **2005**, *16*, 555–567.

(15) Arvanitoyannis, I. S.; Tzouros, N. E. Agricultural produces: Synopsis of employed quality control methods for the authentication of foods and application of chemometrics for the classification of foods according to their variety or geographical origin. *Crit. Rev. Food Sci. Nutr.* **2001**, *41*, 287–319.

(16) Benincasa, C.; Lewis, J.; Perri, E.; Sindona, G.; Tagarelli, A. Determination of trace element in Italian virgin olive oils and their characterization according to geographical origin by statistical analysis. *Anal. Chim. Acta* **2007**, *585*, 366–370.

(17) Benincasa, C.; Lewis, J.; Sindona, G.; Tagarelli, A. The use of multi element profiling to differentiate between cow and buffalo milk. *Food Chem.* **2008**, *110*, 257–262.

(18) Kelly, S. D.; Bateman, A. S. Comparison of mineral concentrations in commercially grown organic and conventional crops − Tomatoes (*Lycopersicon esculentum*) and lettuces (*Lactuca sativa*). *Food Chem.* **2010**, *119*, 738–745.

(19) Capar, S. G.; Mindak, W. R.; Cheng, J. Analysis of food for toxic elements. *Anal. Bioanal. Chem.* **2007**, *389*, 159–169.

(20) Harrington, C. F.; Clough, R.; Hansen, H. R.; Hill, S. J.; Pergantise, S. A.; Tyson, J. F. Atomic Spectrometry Update. Elemental speciation. *J. Anal. At. Spectrom.* **2009**, *24*, 999–1025.

(21) Gonzalvez, A.; Armenta, S.; de la Guardia, M. Trace-element composition and stable-isotope ratio for discrimination of foods with Protected Designation of Origin. *Trends Anal. Chem.* **2009**, *28*, 1295–1311.

(22) Montaser, A. *Inductively Coupled Plasma Mass Spectrometry*; Wiley-VCH: New York, 1998.

(23) Beauchemin, D. Inductively coupled plasma mass spectrometry. *Anal. Chem.* **2006**, *78*, 4111–4136.

(24) Cubadda, F. Inductively Coupled Plasma-Mass Spectrometry for the Determination of Elements and Elemental Species in Food: A Review. *J. AOAC Int.* **2004**, *87*, 173–204.

(25) Caroli, S. *The Determination of Chemical Elements in Food − Applications for Atomic and Mass Spectrometry*; Wiley: Hoboken, NJ, 2007.

(26) Forina, M.; Lanteri, S.; Armanino, C.; Cerrato Oliveros, C.; Casolino, C. *V-Parvus* (2004) Department of Chimica e Tecnologie Farmaceutiche e Alimentari, University of Genova: Genova, Italy. (Free download at www.parvus.unige.it).

(27) Feist, B.; Mikula, B.; Pytlakowska, K.; Puzio, B.; Buhl, F. Determination of heavy metals by ICP−OES and F-AAS after preconcentration with 2,2'-bipyridyl and erythrosine. *J. Hazard. Mater.* **2008**, *152*, 1122–1129.

(28) Bakkali, K.; Ramos Martos, N.; Souhail, B.; Ballesteros, E. Characterization of trace metals in vegetables by graphite furnace atomic absorption spectrometry after closed vessel microwave digestion. *Food Chem.* **2009**, *116*, 590–594.

(29) González-Arjona, D.; López-Pérez, G.; González, A. G. Holmes, a program for performing Procrustes Transformations. *Chemom. Intell. Lab. Syst.* **2001**, *57*, 133–137.

(30) Forina, M.; Armanino, C.; Leardi, R.; Drava, G. A class-modelling technique based on potential functions. *J. Chemom.* **1991**, *5*, 435–453.