JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

J. Agric. Food Chem. **2004,** *52,* 5803–5809 **5803**

Application of Inorganic Element Ratios to Chemometrics for Determination of the Geographic Origin of Welsh Onions

KAORU ARIYAMA,* HIROSHI HORITA, AND AKEMI YASUI

National Food Research Institute, Kannondai 2-1-12, Tsukuba-shi, Ibaraki 305-8642, Japan

The composition of concentration ratios of 19 inorganic elements to Mg (hereinafter referred to as 19-element/Mg composition) was applied to chemometric techniques to determine the geographic origin (Japan or China) of Welsh onions (*Allium fistulosum* L.). Using a composition of element ratios has the advantage of simplified sample preparation, and it was possible to determine the geographic origin of a Welsh onion within 2 days. The classical technique based on 20 element concentrations was also used along with the new simpler one based on 19 elements/Mg in order to validate the new technique. Twenty elements, Na, P, K, Ca, Mg, Mn, Fe, Cu, Zn, Sr, Ba, Co, Ni, Rb, Mo, Cd, Cs, La, Ce, and Tl, in 244 Welsh onion samples were analyzed by flame atomic absorption spectroscopy, inductively coupled plasma atomic emission spectrometry, and inductively coupled plasma mass spectrometry. Linear discriminant analysis (LDA) on 20-element concentrations and 19-element/Mg composition was applied to these analytical data, and soft independent modeling of class analogy (SIMCA) on 19-element/Mg composition was applied to these analytical data. The results showed that techniques based on 19-element/Mg composition were effective. LDA, based on 19-element/Mg composition for classification of samples from Japan and from Shandong, Shanghai, and Fujian in China, classified 101 samples used for modeling 97% correctly and predicted another 119 samples excluding 24 nonauthentic samples 93% correctly. In discriminations by 10 times of SIMCA based on 19-element/Mg composition modeled using 101 samples, 220 samples from known production areas including samples used for modeling and excluding 24 nonauthentic samples were predicted 92% correctly.

KEYWORDS: Chemometrics; linear discriminant analysis; soft independent modeling of class analogy; Welsh onions (*Allium fistulosum* L.); geographic origin; inorganic element

Correct labeling is needed to protect economic benefits of food producers of premium food products and guarantee consumer's food selection by the labeling. In Japan, regulations that require listing the geographic origin on all perishable foods were enacted in 2000. Therefore, the development of techniques to scientifically determine the geographic origin of agricultural products or processed agricultural products has been requested. Numerous approaches for the techniques have already been reported. These techniques have been based on a composition of organic constituents (*1*−*3*) or inorganic elements (*4*−*7*), isotope ratios (*8*, *9*), or their combination (*10*−*12*). This paper reports a technique based on a composition of inorganic element ratios, that is, a 19-element/Mg composition. The technique using inorganic element composition is frequently used for determination of the geographic origin, because inductively coupled plasma atomic emission spectrometry (ICP-AES) and inductively coupled plasma mass spectrometry (ICP-MS) can determine many inorganic elements simultaneously. However, there have been no studies determining the geographic origin of agricultural products or processed agricultural products by a composition of inorganic element ratios excluding isotope ratios as far as we know. Determining the geographic origin of archeological samples such as jasper using this technique has already been reported (*13*). Inorganic elements in archeological samples are measured by X-ray fluorescence (XRF) in many cases, because these samples are not allowed to be destroyed and contain inorganic elements at high concentrations, and the concentration ratios can easily be determined. However, it is difficult to exactly determine elements with XRF, and trace elements in agricultural products cannot be determined due to XRS's high detection limit. Using a composition of inorganic element ratios has the advantage that sample preparation can be simplified. This technique does not require precisely weighing samples and preparing an exact volume of sample solutions. Usually, when inorganic elements in agricultural products are determined, it is necessary to dry samples because concentrations on a dry weight basis are required. However, this technique does not require the drying of samples. Therefore, mistakes in sample preparation can be decreased and the time to prepare a sample can be reduced. In this study, we applied this technique to determine the geographic origin of the Welsh onion (*Allium*

* Corresponding author (telephone +81-48-600-2365; fax +81-48-600-2377; e-mail kaoru_ariyama@cfqlcs.go.jp).

10.1021/jf049333w CCC: $27.50 © 2004 American Chemical Society
Published on Web 08/20/2004

fistulosum L.), a popular vegetable in Japan and mainly produced in East Asia, as quickly as possible.

This study targeted the development of a technique for determining whether Welsh onions originated in Japan or in China as described in a previous paper (*14*). Furthermore, the development of a technique for rapid screening of the geographic origin of Welsh onions sold in Japan was objective, because Welsh onions are a perishable food. Chemometric techniques including linear discriminant analysis (LDA) and soft independent modeling of class analogy (SIMCA) were used for statistical calculations. The previous paper (*14*) established some LDA and SIMCA models by using 20 element concentrations and compared the discrimination results of 216 samples by each model. This paper establishes new LDA models by using 20 element concentrations and 19 elements/Mg in order to validate the new technique and new SIMCA models by using 19 elements/Mg in order to validate the new techniques and determines the origins using the analytical data of 214 samples used in the previous paper (the other two samples were used for rapid analyses) and 30 samples used for rapid analyses.

MATERIALS AND METHODS

Welsh Onion Samples. Four groups of samples consisting of 244 samples were used. The first group were test samples, prepared by pulverizing 25 Welsh onions from a known production area to prepare an average test sample in a production area. These samples, obtained from June 2001 to June 2003, were all from different lots, and all were used for modeling. This group consisted of the two subgroups described below.

Subgroup 1-1 consisted of 65 test samples harvested from 53 areas in Japan; these were obtained directly from the places of production.

Subgroup 1-2 consisted of 36 test samples harvested from 3 areas (Anchiu Shandong, 18; Shanghai, 10; and Amoy Fujian, 8) in China; these were obtained from a trading company and a wholesaler.

The other groups were test samples prepared by pulverizing one Welsh onion; they were used to predict the geographic origin of samples excluding the last SIMCA models in this paper. This is because it is necessary to determine the geographic origin of one Welsh onion for practical application, and the efficiency of a model is evaluated by the accuracy of the prediction. The second group was test samples prepared from one Welsh onion harvested in a known production area from June 2001 to June 2003. This group consisted of the two subgroups described below.

Subgroup 2-1 consisted of 14 test samples harvested from 14 areas in Japan; 11 samples were prepared from Welsh onions of the same lot as samples in subgroup 1-1.

Subgroup 2-2 consisted of 25 test samples harvested from 3 areas (Anchiu Shandong, 8; Shanghai, 9; and Amoy Fujian, 8) in China; 22 samples were prepared from Welsh onions of the same lot as samples in subgroup 1-2.

The third group was test samples prepared by pulverizing one Welsh onion purchased at retailers from September 2002 to July 2003. Some Welsh onion samples labeled as produced in Japan may have been falsely labeled, so the accuracy of predictions cannot be evaluated using these samples. Welsh onions labeled as produced in China can be considered to be authentic because Japanese consumers do not have a good image of imported Welsh onions, which are sold at about half the price of Japanese Welsh onions. This group consisted of the two subgroups described below.

Subgroup 3-1 consisted of 24 test samples labeled as produced in Japan; these were purchased from 22 retailers in 12 cities in Japan and not used to evaluate the accuracy of predictions!

Subgroup 3-2 consisted of 50 test samples labeled as produced in China; these were purchased from 33 retailers in 15 cities and 1 town in Japan.

The fourth group was test samples prepared by pulverizing one Welsh onion harvested in a known production area and one Welsh onion purchased at a retailer. These samples were used for rapid analyses based on 19-element/Mg composition. Analytical results of these samples, which are from a different lot from the samples for modeling, were not all used for modeling. This group consisted of the two subgroups described below.

Subgroup 4-1 consisted of 15 samples from Welsh onions harvested from 15 areas in Japan from November 2003 to July 2003.

Subgroup 4-2 consisted of 15 samples of Welsh onions randomly selected from the same lot as the 50 Welsh onion samples for subgroup 3-2.

Details of all samples are described in the Supporting Information (SI).

Quantitative Aanalyses of Welsh Onion Samples. Twenty elements in samples of groups 1−3 were quantitated. Sodium was determined by FAAS (SOLAAR 969, Nippon Jarrel-Ash). Ten elements (P, K, Ca, Mg, Mn, Fe, Cu, Zn, Sr, and Ba) were determined by ICP-AES (JICP-PS3000UV, Leeman Labs). Nine elements (Co, Ni, Rb, Mo, Cd, Cs, La, Ce, and Tl) were determined by ICP-MS (JMS-PLASMAX2, JEOL). These 20 elements were determined in test samples prepared using the lower 10 cm of Welsh onions. The preparation of a test sample using 25 Welsh onions was described in a previous study (*15*), including a 1% HCl extraction for Na and open-vessel wet acid digestion for the other elements. The preparation of a test sample using one Welsh onion was described in a previous paper (*14*) excluding a 1% HCl extraction for Na.

Rapid Aanalyses of One Welsh Onion. Samples of group 4 were used for rapid analyses. The same part of a Welsh onion washed and cut in the same procedure as in the previous paper (*14*) was pulverized by a mixer with ceramic blades (B400, Nihon Büchi) together with about the same amount of superpure water (>18 MΩcm) purified using the Super Q system (Nihon Millipore); it is not necessary to weigh the sample and measure the volume exactly. About 5 g of samples was taken in Teflon beakers ($n = 3$); it is not necessary to weigh them exactly. These samples were then digested with 5 mL of 68% HNO_3 and 1.25 mL of 70% $HClO_4$ (each acid was ultrapure grade; Tamapure AA-100, Tama Chemical) by heating on a household hotplate in the same way as for the 25 Welsh onions, but HF was not added. Before complete evaporation of the acids, 625 μL of indium of 200 μg/L was added as an internal standard and 2% HNO_3 was added so the volume would be ∼25 mL; it is not necessary to cool the sample to room temperature or give an exact volume. The indium concentration (internal standard) was ∼5 μg/L. The residues were dissolved by heating and shaking and prepared into sample solutions by filtering with membrane filters having pore diameters of 0.45 μm. Three blank solutions were also prepared by performing in the same way without samples. This sample preparation takes <5 h, whereas the method not based on 19-element/Mg composition takes 7 h and needs a moisture measurement. Instrument measurements were performed under the conditions of **Table 1**. Twenty elements in the prepared sample solutions were determined, and 19 elements/Mg were obtained. This series of procedures can be performed within 1.5 days, that is, 12 work hours, on three test samples (each is repeated three times) and three blanks.

THEORETICAL CALCULATIONS

LDA. LDA was performed using Statistica Pro 03J application software (StatSoft Japan). In this program, the coefficients of canonical variates are derived on the basis of the maximization of the correlation between canonical variates in what is called canonical correlation analysis. Actually, these coefficients are equivalent to coefficients of canonical discriminant functions that define a hyperplane that separates each group. At the same time, group classification functions are derived. These are linear discriminant functions that are introduced for each group; a case will be classified into the group for which the score is largest (*16*). In this study, we performed a backward stepwise regression to reduce the number of variables. First, we conducted a regression analysis including all variables. Next, as many variables as possible were removed on the basis of an *F*-to-remove threshold specified in advance. The *F*-to-remove thresholds of LDA on 20-element composition were 3, and those of two kinds of LDA on 19-element/Mg composition were 6, because the best discrimination results were obtained in this condition.

Table 1. Operating Conditions for FAAS, ICP-AES, and ICP-MS Instruments

FAAS	Nippon Jarrel-Ash, SOLAAR 969	ICP-MS	JEOL, JMS-PLASMAX2
flame condition		plasma condition	
flame type	air–C_2H_2	RP power	1.3 kW
fuel flow rate	1.2 L/min	plasma gas	14.0 L/min
burner type	50 mm	auxiliary gas	0.7 L/min
element	wavelength	nebulizer gas	0.96 L/min
Na	589.0 nm	resolution: 500	
		main slit	1400
		α slit	12
		collect slit	294
		element	*m/z*
		Cd	111
ICP-AES Varian, Vista Pro		In	115
plasma condition		Te	125
RP power	1.2 kW	Cs	133
plasma gas	15 L/min	La	139
auxiliary gas	1.5 L/min	Ce	140
nebulizer gas	0.5 L/min	Nd	146
element	wavelength	Sm	149
P	213.618 nm	Eu	153
K	769.897 nm	Tb	159
Mg	280.270 nm	Yb	172
Ca	317.933 nm	W	182
Mn	259.372 nm	Tl	205
Fe	259.940 nm	resolution: 3000	
Cu	324.754 nm	main slit	201
Zn	213.856 nm	α slit	2.5
Sr	421.552 nm	collect slit	50
Ba	455.403 nm	element	*m/z*
		Co	59
		Ni	60
		Rb	85
		Mo	98

SIMCA. SIMCA (*17*) were performed using Pirouette application software (ver. 3.11, Informetrix). This chemometrics technique establishes a principal component model that is called a SIMCA box for each class (it has been called a group in LDA), so that each model has the appropriate number of factors. Next, an unknown case is applied to each model and predicted on the basis of the distance from the models. The ordinary prediction by SIMCA classifies a case into a class or outlier based on 95% levels of significance. In this study, however, we classified a Welsh onion sample into the Japanese class or the Chinese class that it had the higher probability of belonging to that because every sample belongs to the Japanese or Chinese class. The probability was obtained from the *F* value, which is the square of the distance from a class divided by the total residual variance of the class.

Modeling in LDA and SIMCA. Concentrations of 20 elements (Na, P, K, Ca, Mg, Mn, Fe, Cu, Zn, Sr, Ba, Co, Ni, Rb, Mo, Cd, Cs, La, Ce, and Tl) in samples or sample solutions were used for modeling. Three LDA models were established using 20-element concentrations and 19 elements/Mg of samples, and one set of SIMCA model was established using 19 elements/Mg of samples. The first LDA model (LDA 1) was established on 20-element concentrations with group 1, which consisted of samples prepared using 25 Welsh onions from known production areas, and used for two-group discrimination: whether the sample was from Japan or China. The second LDA model (LDA 2) was established on 19 elements/Mg with group 1 for two-group discrimination. The third LDA model (LDA 3) was established on 19 elements/Mg with group 1 for classification into the four groups of Japan and Shandong, Shanghai, and Fujian in China. Ten SIMCA models were established on 19-element/Mg composition with randomly selected 78 samples from group 1 and 23 samples from subgroup 3-2, which consisted of Chinese samples purchased at retailers, and used for two-group discrimination.

RESULTS AND DISCUSSION

Concentration of 20 Elements in Welsh Onions and Selection of a Reference Element for Rapid Analysis. Table 2 summarizes the analytical results of 101 samples for group 1 from known production areas in the same way as the summary of the data on 103 samples in the previous paper (*14*). This table shows that element compositions of Welsh onions differed not only between Japan and China but also within China. These tendencies indicated the possibility of classifying four production areas of Welsh onions: Japan, Anchiu Shandong, Shanghai, and Amoy Fujian.

Inorganic element concentrations in agricultural products are usually expressed on a dry-weight basis because moisture of

Table 2. Analytical Results (Dry-Weight Basis) for 101 Welsh Onion Samples of Group 1

	Japan ($n = 65$)			China ($n = 36$)								
				Anchiu Shandong ($n = 18$)			Shanghai ($n = 10$)			Amoy Fujian ($n = 8$)		
	mean[a]	SD[b]	RSD[c] (%)	mean	SD	RSD (%)	mean	SD	RSD (%)	mean	SD	RSD (%)
K	17.3	3.3	19	14.3	1.9	13	14.7	1.6	11	20.4	2.8	14
P	2.92	0.62	21	3.46	0.41	12	2.52	0.46	18	3.30	0.65	20
Ca	2.69	0.98	36	3.32	0.96	29	2.24	0.52	23	1.73	0.41	23
Mg	0.983	0.220	22	1.07	0.13	12	0.810	0.150	18	0.810	0.110	13
Na	41	26	62	87	41	47	141	63	44	129	76	59
Zn	29.4	11.2	38	22.3	5.5	24	17.3	3.4	20	31.3	2.5	8.0
Fe	18	5	28	21	6	27	15	6	37	18	3	17
Mn	15.4	14.5	94	14.4	7.1	49	4.4	1.6	37	27.5	10.9	40
Sr	11.4	5.6	50	19.7	6.9	35	13.9	3.9	28	22.8	2.8	12
Rb	6.6	5.4	81	4.7	3.7	78	6.1	1.8	29	11.2	4.5	40
Ba	4.0	4.6	115	3.9	2.0	50	0.6	0.5	93	3.7	3.5	94
Cu	3.1	1.7	55	3.5	1.2	33	5.6	1.4	25	3.4	0.8	23
Ni	1.09	1.01	93	1.39	0.95	69	0.45	0.16	35	1.27	0.32	25
Cd	0.183	0.140	76	0.131	0.071	54	0.055	0.025	46	0.111	0.080	72
Mo	0.16	0.17	112	0.09	0.09	100	0.03	0.02	78	0.12	0.09	74
Co	0.03	0.06	204	0.07	0.05	73	0.00	0.01	230	0.14	0.10	73
Tl	0.033	0.058	178	0.009	0.007	74	0.010	0.013	130	0.056	0.024	44
Ce	0.006	0.009	140	0.020	0.048	240	0.007	0.007	97	0.044	0.036	82
La	0.010	0.022	215	0.010	0.025	260	0.001	0.003	210	0.039	0.026	67
Cs	0.017	0.055	320	0.006	0.005	82	0.007	0.003	42	0.021	0.009	43

[a] For K, P, Ca, and Mg, means and standard deviations are given in mg/g; for all other elements, means and standard deviations are given in μg/g. [b] Standard deviation. [c] Relative standard deviation.

Table 3. Two-Froup Discriminations of 214 Samples by LDA 1 Modeled Using 20 Element Concentrations for Group 1 and Those of 244 Samples by LDA 2 Modeled Using 19 Elements/Mg for Group 1

no.		LDA 1 modeled using 20 element concentrations	LDA 2 modeled using 19 elements/Mg
	producing country	rate classified correctly	
1-1	Japan	98% (64/65)[a]	94% (61/65)[a]
1-2	China	94% (34/36)[a]	92% (33/36)[a]
		mean 97%	mean 93%
	producing country	rate predicted correctly	
2-1	Japan	100% (14/14)[b]	93% (13/14)[b]
2-2	China	96% (24/25)[b]	96% (24/25)[b]
		mean 97%	mean 95%
	country on label	rate of matching the labeling	
3-1	Japan	92% (22/24)[c]	92% (22/24)[c]
3-2	China	92% (46/50)[c]	92% (46/50)[c]
	producing country		rate predicted correctly
4-1	Japan		100% (15/15)[b]
	country on label		rate of matching the labeling
4-2	China		60% (9/15)[c]
			mean 83%

[a] Number in parentheses represents the number of samples classified correctly/the number of analyzed samples. [b] Number in parentheses represents the number of samples predicted correctly/the number of analyzed samples. [c] Number in parentheses represents the number of samples matching the labeling/the number of analyzed samples.

Table 4. Normalized Coefficient of Canonical Functions and *F* Statistics Obtained by LDA for Two-Group Discrimination with 20 Element Concentrations and 19 Elements/Mg for Group 1[a]

	LDA 1 based on 20-element composition		LDA 2 based on 19-element/Mg composition	
	normalized coeff of canonical functions	*F* statistic	normalized coeff of canonical functions	*F* statistic
Na	0.683	36.31	0.767	35.76
P	0.770	25.59	0.881	29.35
K	–0.682	21.61	–1.138	42.55
Ca	–0.354	6.31		
Zn	–0.311	3.54	–0.539	12.26
Sr	0.647	27.94	0.503	15.68
Mo	–0.234	3.11		
Cd	–0.304	5.03		
Ce	0.397	9.83	0.612	21.06
eigenvalue	2.598		2.181	

[a] Elements were chosen by backward stepwise LDA with element concentrations and concentration ratios to Mg; *F*-to-remove thresholds were 3 and 6.

agricultural products varies with cultivation conditions, preservation situations, etc. In this case, element concentrations are greatly influenced by the content of organic constituents because the dry weight mainly reflects the content of organic constituents. In the case of concentrations on a wet-weight basis, element concentrations are greatly influenced by the moisture in a fresh vegetable like a Welsh onion because the moisture is ∼90%. Having focused on inorganic element composition, a concentration based on the total of inorganic element content or on inorganic element contents taken up as cations might be more suitable. However, one element was chosen in this study to achieve the fastest possible analysis. We defined an element set up as a reference in the rapid analysis based on the composition of concentration ratios as a reference element. The selection criteria of the reference element were that the contribution for determining the origin of a Welsh onion was small and could be analyzed precisely and accurately. Magnesium was chosen as the reference element that fulfills these criteria. In a later section, it is shown that the contribution of Mg for determining the origin of a Welsh onion is small.

LDA Modeled on Welsh Onions from Japan and China: LDAs 1 and 2. LDAs on 20-element composition and on 19-element/Mg composition were peformed with group 1 to classify Welsh onions from Japan or China. Two models described as LDA 1 and LDA 2 were established. Nine elements (Na, P, K, Ca, Zn, Sr, Mo, Cd, and Ce) were chosen by backward stepwise regression in LDA 1, and ratios of six elements to Mg (Na, P, K, Zn, Sr, and Ce/Mg) were chosen in LDA 2. **Table 3** summarizes the results. The Supporting Information presents the results for every sample by three LDA models established in this paper. The classifications of group 1 by LDA 1 was 97% correct, and that by LDA 2 on group 1 was 93% correct. The predictions by LDAs 1 and 2 for samples of group 2 and subgroup 3-2, in which the samples are from known production countries, were 94 and 93% correct, respectively. High rates of classification and prediction accuracy were achieved by LDA 2 as well as by LDA 1. Therefore, it was considered that the discrimination by LDA on 19-element/Mg composition as well as that on 20-element composition was effective. The rate of matching the labeling by each model for subgroup 3-1 was the same 92%.

In LDA, the *F* statistic and the absolute value of the normalized coefficient of a canonical function indicate how much selected elements contributed to the classification; a large value indicates a high contribution to classification. **Table 4** shows that Na, P, K, and Sr contributed significantly to the classification by LDA 1, and Na, P, K, and Ce/Mg contributed significantly to that by LDA 2. The elements selected for ratios to Mg in LDA 2 were all selected in LDA 1, too, and LDA 2 was effectively modeled with a fewer number of elements. It was indicated that Mg was not selected as an effective element in LDA 1 and was suitable as a reference element in LDA 2.

LDA Modeled on Welsh Onions from Four Production Areas: LDA 3. Most Welsh onions imported from China are produced in Shandong, Shanghai, and Fujian. These areas are far apart, so modeling to classify samples from four areas may be more useful in determining by LDA whether a Welsh onion originates in Japan or China. The LDA model on 19-element/Mg composition was established to classify Welsh onions from Japan and Anchiu Shandong, Shanghai, and Amoy Fujian in China with group 1, because LDA 2 modeled using 19 elements/Mg as well as LDA 1 modeled using 20-element concentrations was an effective model. This model is described as LDA 3. Ratios of 11 elements to Mg (Na, P, K, Ca, Co, Cu, Sr, Cd, Cs, Ce, and Tl/Mg) were chosen by backward stepwise regression. **Table 5** summarizes the results of these discriminations. The four-group discrimination obtained by LDA 3 on group 1 was 97% correct, and the prediction by this model for group 2 was 95% correct. These results suggested that Welsh onion samples could be classified into samples from Japan and three areas in China. The two-group discrimination by this model was also performed between Welsh onions from Japan and China. **Table 5** also summarizes the results of this discrimination. The classification by LDA 3 for group 1 was 97% correct. The prediction by this model for group 2 and subgroup 3-2 in which samples are from known production countries was 96% correct. This discrimination accuracy was slightly higher than that by LDA 2. The rate of label matching for subgroup 3-1 was 92%.

Table 5. Four-Group Discriminations of 140 Samples and Two-Group Discriminations of 244 Samples by LDA Modeled Using 19 Elements/Mg for Group 1

			LDA 3 modeled using 19-element/Mg compsition	
no.			four-group discrimination	two-group discrimination
	producing area		rate classified correctly	
1-1	Japan		97% (63/65)[a]	97%
1-2	China			97% (35/36)[a]
		Shandong	94% (17/18)[a]	mean 97%
		Shanghai	100% (10/10)[a]	
		Fujian	100% (8/8)[a]	
			mean 98%	
	producing area		rate predicted correctly	
			93% (13/14)[b]	93%
2-1	Japan			100% (25/25)[b]
2-2	China	Shandong	100% (8/8)[b]	mean 97%
		Shanghai	89% (8/9)[b]	
		Fujian	100% (8/8)[b]	
			mean 95%	
	country on label			rate of matching the labeling
3-1	Japan			92% (22/24)[c]
3-2	China			94% (47/50)[c]
	producing area			rate predicted correctly
4-1	Japan			100% (15/15)[b]
	country on label			rate of matching the labeling
4-2	China			73% (11/15)[c]
				mean 87%

[a] Number in parentheses represents the number of samples classified correctly/the number of analyzed samples. [b] Number in parentheses represents the number of samples predicted correctly/the number of analyzed samples. [c] Number in parentheses represents the number of samples that matching the labeling/the number of analyzed samples.

Figure 1 visually classifies group 1 into four groups by plots of scores 1−3 for the canonical functions of LDA 3. Values of score 1 are the main factors for classifying whether Welsh onions were produced in Japan or China. In addition, these values are also the factors for classifying whether Welsh onions were produced in Amoy Fujian or the other areas. **Table 6** indicates that P, Sr, Cd, Cs, and Tl/Mg contributed to the classification by LDA 3. Values of score 2 were the main factors for classifying whether Welsh onions were produced in Shanghai or the other areas. Na, Cu, Cs, and Tl/Mg contributed to this classification. Values of score 3 were the main factors for classifying whether Welsh onions were produced in Anchiu Shandong or the other areas. P, K, Cs, and Tl/Mg contributed greatly to this classification.

Set of SIMCA Modeled on Welsh Onions from Japan and China. Modeling in SIMCA requires much more information than in LDA because a SIMCA model is more complicated mathematically than the linear discriminant function of LDA (*14*). Welsh onions from Shangdon and Fujian in China used for modeling of LDAs 1−3 were actually from the smaller areas, Anchiu Shangdon and Amoy in Fujian, although Welsh onions from China are imported from various areas in Shangdon and Fujian. Therefore, SIMCA models on 19-element/Mg composition were established by incorporating the data of samples in subgroup 3-2 into some data in group 1. One hundred and one samples were selected randomly: 60 samples from 65 Japanese samples of subgroup 1-1, 9 samples from 18 samples of Shandong, 5 samples from 10 samples of Shanghai, 4 samples

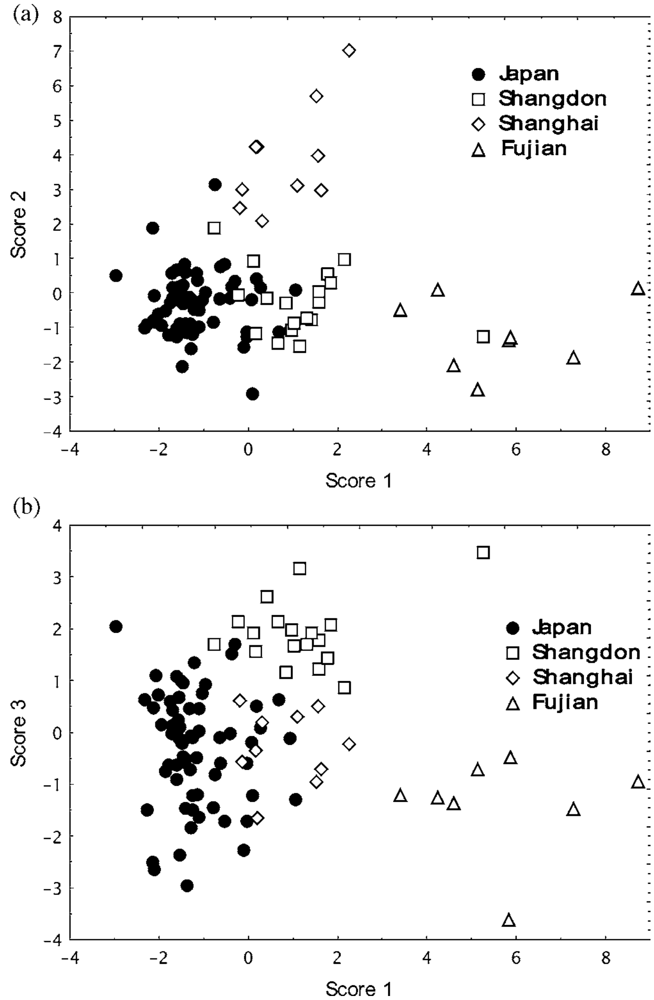

Figure 1. Plots of score 1 versus score 2 (**a**) and score 1 versus score 3 (**b**) of group 1, consisting of 101 samples from known production areas, by LDA 3 modeled using 19 elements/Mg for four-group discrimination.

Table 6. Normalized Coefficient of Canonical Functions and *F* Statistics Obtained by LDA for Four-Group Discrimination with 19 Elements/Mg on Group 1[a]

	normalized coefficient of canonical functions			
	1	2	3	*F* statistics
Na	0.368	0.670	0.100	10.87
P	0.741	−0.269	0.619	13.73
K	−0.492	0.097	−1.276	20.42
Ca	−0.513	0.113	0.251	6.22
Co	0.350	−0.408	−0.129	6.23
Cu	−0.003	0.835	−0.145	14.25
Sr	0.812	−0.363	0.225	17.40
Cd	−0.786	−0.046	0.427	10.50
Cs	−0.907	0.741	0.868	6.08
Ce	0.644	−0.123	0.256	10.25
Tl	0.973	−1.050	−1.102	7.12
eigenvalue	3.845	1.775	0.91452	

[a] Elements were chosen by backward stepwise LDA (*F*-to-remove threshold is 6) with 19 elements/Mg.

from 8 samples of Fujian in subgroup 1-2, and 23 samples from 50 Chinese samples of subgroup 3-2. These random selections were carried out 10 times, and 10 SIMCA models were established. Every Japanese class in these models had six factors. The Chinese class in these models had six or four factors. **Table 7** summarizes the average of discrimination results by 10

Table 7. Two-Group Discrimination of 244 Samples by SIMCA Modeled Using 19 Elements/Mg in 101 Samples[a] Randomly Selected from Subgroups 1-1, 1-2, and 3-2

no.		mean[b] of prediction by SIMCA
	producing country	rate predicted correctly ± SD[c] (%)
1-1	Japan	95 ± 2
1-2	China	94 ± 4
		mean 95 ± 1
2-1	Japan	92 ± 2
2-2	China	90 ± 6
		mean 91 ± 3
	country on label	rate of matching the labeling ± SD (%)
3-1	Japan	78 ± 11
3-2	China	90 ± 5
	producing country	rate predicted correctly ± SD (%)
4-1	Japan	83 ± 5
	country on label	rate of matching the labeling ± SD (%)
4-2	China	89 ± 2
		mean 86 ± 4

[a] Random selections of 60 samples from subgroup 1-1, 18 samples from subgroup 1-2, and 23 samples from subgroup 3-2 were carried out 10 times. [b] Mean of 10 predictions by 10 SIMCA models. [c] Standard deviation.

Table 8. Means[a] of Discrimination Powers and Modeling Powers in SIMCA Modeled Using 19 Elements/Mg on 101 Samples Selected from Subgroups 1-1, 1-2, and 3-2

element	discrimination power ± SD[b]	modeling power ± SD
Na	15.8 ± 4.5	0.51 ± 0.04
P	3.5 ± 0.8	0.60 ± 0.04
K	6.4 ± 1.6	0.58 ± 0.04
Ca	3.8 ± 1.4	0.58 ± 0.05
Mn	7.2 ± 0.9	0.59 ± 0.02
Fe	3.4 ± 2.1	0.49 ± 0.09
Co	7.0 ± 2.2	0.52 ± 0.04
Ni	2.5 ± 0.3	0.53 ± 0.03
Cu	4.7 ± 1.8	0.52 ± 0.05
Zn	4.3 ± 1.1	0.60 ± 0.03
Rb	5.0 ± 1.2	0.56 ± 0.03
Sr	7.3 ± 2.5	0.58 ± 0.05
Mo	6.0 ± 1.8	0.46 ± 0.06
Cd	13.7 ± 2.6	0.54 ± 0.03
Cs	28.0 ± 9.2	0.66 ± 0.03
Ba	6.0 ± 1.6	0.53 ± 0.05
La	2.9 ± 0.6	0.53 ± 0.02
Ce	17.2 ± 7.1	0.60 ± 0.03
Tl	3.8 ± 1.1	0.62 ± 0.20

[a] Mean of 10 times of SIMCA. [b] Standard deviation.

SIMCA models. The average of predictions for group 1 was 95% correct, and that for group 2 and subgroup 3-2 in which samples are from known production countries was 91% correct. The average of label matching rates for subgroup 3-1 was 78%. The average of predictions for 190 samples of groups 1 and 2 and subgroup 3-2 including 101 samples used for modeling was 93% correct, whereas in SIMCA on 20-element composition in the previous paper, the prediction accuracy for 192 samples including 103 samples used for modeling was 96% correct (*14*). Like these, the prediction accuracy on 19-element/Mg composition was high, but slightly lower than that on 20-element composition, although both models were established using different data sets for modeling. We are sure that this SIMCA model on 19-element/Mg composition is effective to determine the geographic origin of a Welsh onion.

In SIMCA, a discrimination power indicates how much each element/Mg contributes to classifying a case; a large value indicates a high contribution to classification (*17*). A modeling power indicates how much each element contributes to making a principal component model of each class. A value close to 1 (0) indicates a high (low) contribution to the model (*17*). **Table 8** shows that Cs/Mg greatly contributed to the classification by this SIMCA model, and Na, Cd, and Ce/Mg contributed, too. This tendency in SIMCA differed from those in LDA, which indicated how much both models differ. The effective elements in SIMCA on 19-element/Mg composition were the same as those on 20-element composition in the previous paper (*14*), excluding Cd, and exhibited a similar tendency about the degree of contribution for each element. The contribution of Mg for determining the origin of a Welsh onion was small in SIMCA on 20-element composition, and Mg was suitable as a reference element in SIMCA on 19-element/Mg composition. As mentioned before, Mg was not selected in LDA 1 and was suitable as a reference element in LDAs 2 and 3. These indicated that Mg concentrations in Welsh onions were not distinctive among different cultivation areas, which may be a point of interest in crop science.

Discrimination by Rapid Analysis. Rapid analyses were performed on 15 Japanese and 15 Chinese samples prepared by using one Welsh onion that was not of the same lot as those used for modeling. The prediction accuracies were 83% for LDA 2, 87% for LDA 3, and 86% for SIMCA. Although high prediction accuracies were achieved, the prediction accuracies were lower than those for group 2 or subgroup 3-2. The reason for this was not clarified. We believe prediction accuracies will be improved by performing much more rapid analysis, because we were unable to do this on many samples.

Discrimination by Combined LDA and SIMCA. The discrimination results can be cross-checked by judging the combination of LDA and SIMCA because LDA and SIMCA establish mathematically different models based on different criteria. LDA and SIMCA based on 19-element/Mg composition exhibited high discrimination accuracies. Hence, combined discriminations of LDA and SIMCA were performed using discrimination results by LDA 3, which showed a higher discrimination accuracy than LDA 2, and SIMCA. The SIMCA model was randomly chosen from 10 SIMCA models. The discrimination accuracies for groups 1, 2, 3-1, 3-2, and 4 were 97, 92, 88, 88, and 93%, respectively. The Supporting Information presents the results for every sample by this model. Three samples in subgroups 3-2 and 4-2 were mistakenly judged by both LDA 3 and SIMCA, probably due to incomplete modeling or analytical mistakes. In addition, one sample in subgroup 3-1 did not match the labeling by both models, and this could be due to incomplete modeling, analytical mistakes, or false labeling. The criterion for judging a Welsh onion of Chinese origin by combined LDA and SIMCA was framed. When the criterion to be a Chinese Welsh onion is that both models determine it as a Chinese one, the unacceptable error that a Japanese Welsh onion was judged as a Chinese one was 0 in 94 samples of subgroups 1-1, 2-1, and 4-1 (Supporting Information). Seventeen Chinese samples were judged not to be Chinese Welsh onions in 126 samples of subgroups 1-2, 2-2, 3-2, and 4-2 (Supporting Information). In other words, the first type of error was 0, and the second type of error was 14%.

Conclusions. The contribution of Mg for determining the origin of a Welsh onion was small, and Mg was suitable as a reference element for rapid analysis. It was suggested that Mg concentrations in Welsh onions were not distinctive among different cultivation areas. LDA and SIMCA based on 19-

element/Mg composition were effective in determining the geographic origin of a Welsh onion. In discriminations by LDA 3 using 101 samples of group 1 from known production areas, the classification for group 1 was 97% correct, and the prediction for 119 samples of group 2, subgroup 3-2, and group 4 was 93% correct. In discriminations by 10 times SIMCA modeled using 101 samples in group 1 and subgroup 3-2, 220 samples from known production areas including 101 samples used for modeling were predicted 92% correctly. Therefore, it is possible to screen whether a Welsh onion sold in Japan is Japanese or Chinese, within 2 days, by using these models. In the combined discrimination of these models, the unacceptable error that a Japanese Welsh onion was judged to be a Chinese one was 0 in 94 samples, and Chinese samples judged not to be Chinese Welsh onions were 17 samples in 126 samples. Hence, it was possible to perform a cross-check of a discrimination result by combining these models.

ACKNOWLEDGMENT

We appreciate Hiroaki Ito and Naoki Usami, Ministry of Agriculture, Forestry and Fisheries of Japan, and persons in charge of Center for Food Quality, Labeling and Consumer Services for help in collecting samples. We are grateful for the assistance provided in statistical analysis by Dr. Shigehiro Naito, National Food Research Institute.

Supporting Information Available: Classification and prediction results of 244 samples by LDA and SIMCA. This material is available free of charge via the Internet at http://pubs.acs.org.

LITERATURE CITED

(1) Etiévant, P.; Schlich, P.; Bouvier, J. C.; Symonds, P.; Bertrand, A. Varietal and Geographic Classification of French Red Wines in Terms of Elements, Amino Acids and Aromatic Alcohols. *J. Sci. Food Agric.* **1988**, *45*, 25−41.

(2) Sanz, S.; Perez, C.; Herrera, A.; Sanz, M.; Juan, T. Application of a Statistical Approach to the Classification of Honey by Geographic Origin. *J. Sci. Food Agric.* **1995**, *69*, 135−140.

(3) Armanino, C.; Acutis, R. D.; Festa, M. R. Wheat lipids to discriminate species, varieties, geo geographical origins and crop years. *Anal. Chim. Acta* **2002**, *454*, 315−326.

(4) Latorre, M. J.; Garía-Jares, C.; Mèdina, B.; Herrero, C. Pattern Recognition Analysis Applied to Classification of Wines from Galicia (Northwestern Spain) with Certified Brand of Origin. *J. Agric. Food Chem.* **1994**, *42*, 1451−1455.

(5) Anderson, K. A.; Magnuson, B. A.; Tschirgi, M. L.; Smith, B. Determining the Geographic Origin of Potatoes with Trace Metal Analysis Using Statistical and Neural Network Classifiers. *J. Agric. Food Chem.* **1999**, *47*, 1568−1575.

(6) Yasui, A.; Shindoh, K. Determination of the geographic origin of brown-rice with trace-element composition. *Bunseki Kagaku* **2000**, *49*, 405−410.

(7) Frías, S.; Conde, J. E.; Rodríguez-Bencomo, J. J.; García-Montelongo, F.; Pérez-Trujillo, J. P. Classification of commercial wines from the Canary Islands (Spain) by chemometric techniques using metallic contents. *Talanta* **2003**, *59*, 335−344.

(8) Martin, G. J.; Guillou, C.; Martin, M. L.; Cabanis, M. T.; Tep, Y.; Aerny, J. Natural Factors of Isotope Fractionation and the Characterization of Wines. *J. Agric. Food Chem.* **1988**, *36*, 316−322.

(9) Oda, H.; Kawasaki, A.; Hirata, T. Determination of the Geographic Origin of Brown-Rice with Isotope Ratios of $^{11}B/^{10}B$ and $^{87}Sr/^{86}Sr$. *Anal. Sci.* **2001**, *17*, i1627−i1630.

(10) Day, M. P.; Zhang, B.; Martin, G. J. Determination of the Geographical Origin of Wine Using Joint Analysis of Elemental and Isotopic Composition. II—Differentiation of the Principal Production Zones in France for the 1990 Vintage. *J. Sci. Food Agric.* **1995**, *67*, 113−123.

(11) Kallithraka, S.; Arvanitoyannis, I. S.; Kefalas, P.; El-Zajouli, A.; Soufleros, E.; Psarra, E. Instrumental and sensory analysis of Greek wines; implementation of principal component analysis (PCA) for classification according to geographical origin. *Food Chem.* **2001**, *73*, 501−514.

(12) Brescia, M. A.; Caldarola, V.; Giglio, A. D.; Benedetti, D.; Fanizzi, F. P.; Sacco, A. Characterization of the geographical origin of Italian red wines based on traditional and nuclear magnetic resonance spectrometric determinations. *Anal. Chim. Acta* **2002**, *458*, 177−186.

(13) Warashina, T. Allocation of Jasper Archaeological Implements by Means of ESR and XRF. *J. Archaeol. Sci.* **1992**, *19*, 357−373.

(14) Ariyama, K.; Horita, H.; Yasui, A. Chemometric Techniques on Inorganic Elements Composition for the Determination of the Geographic Origin of Welsh Onions. *Anal. Sci.* **2004**, *20*, 871−877.

(15) Ariyama, K.; Horita, H.; Yasui, A. Establishment of an inorganic elements measuring method for determining the geographic origin of Welsh onion and preliminary examination. *Bunseki Kagaku* **2003**, *52*, 969−978.

(16) Jennrich, R. I. Stepwise discriminant analysis. In *Statistical Methods for Digital Conputers*; Enslein, K., Ralston, A., Wilf, H. S., Eds.; Wiley: New York, 1977; pp 76−95.

(17) Wold, S.; Sjöström, M. SIMCA: method for analyzing chemical data in terms of similarity and analogy. In *Chemometrics: Theory and Application*; Kowalski, B. R., Ed.; American Chemical Society: Washington, DC, 1977; pp 243−282.

Received for review April 26, 2004. Revised manuscript received July 2, 2004. Accepted July 7, 2004. This study was supported by a Special Research Grant from the Ministry of Agriculture, Forestry and Fisheries of Japan.

JF049333W