JOURNAL OF
AGRICULTURAL AND
FOOD CHEMISTRY

J. Agric. Food Chem. **2005,** *53,* 4041–4045

Determination of the Country of Origin of Garlic (*Allium sativum*) Using Trace Metal Profiling

RALPH G. SMITH

U.S. Customs and Border Protection Laboratory, 214 Bourne Boulevard, Savannah, Georgia 31408

A method for determining the country of origin of garlic by comparing the trace metal profile of the sample to an authentic garlic database is presented. Protocols for sample preparation, high-resolution inductively coupled plasma mass spectrometry, and multivariate statistics are provided. The criteria used for making a country of origin prediction are also presented. Indications are that the method presented here may be used to determine the geographic origin of other agricultural products.

KEYWORDS: Country of origin; agricultural; garlic; *Allium sativum*; HR-ICPMS; multivariate statistics; trace-metal; profiles; prediction

INTRODUCTION

U.S. Customs and Border Protection (CBP) Laboratories have been working on laboratory methods to facilitate scientific identification of the country of origin of various agricultural products. One such method presented here includes the determination of the trace metal profile of the agricultural product for comparison with an established database of trace metal profiles of the product from various countries.

The uptake of trace metals by agricultural products from the soil in which they are grown provides a mechanism for identification of their geographic origin. There are a number of factors such as rainfall, sunshine, temperature, soil characteristics, and plant species that may play an important role in the uptake of trace metals. It is the combination of these factors that influence the uptake of trace metals creating a rough snapshot or historical record of the plant's growth. In most cases, the trace metal profiles of agricultural products from various countries display enough statistical uniqueness to make a definitive country of origin prediction.

The use of trace metal profiling for determining the geographic origin of agricultural products has been cited in the literature as far back as the 1980s. While analytical instrumentation has improved capabilities considerably since that time, the technique was validated with older instrumentation such as inductively coupled plasma (ICP) emission and graphite furnace atomic absorption and quadrupole ICPMS. The introduction of high-resolution ICPMS lowered detection limits and eliminated many of the isobaric interferences prevalent with quadrupole ICPMS instruments.

Schwartz and Hecking (*1*) did a study of the geographic origin of orange juice, macadamia nuts, and pistachio nuts using multivariate analysis of the trace element composition. Nikdel et al. (*2*) also did trace metal work for determining adulteration and country of origin of orange juice. These workers also used multivariate statistics for country of origin prediction. Anderson et al. (*3*) and Anderson and Smith (*4*) have done trace metal profiling along with multivariate statistics to determine the geographic origin of potatoes and coffee, respectively. Petersen et al. (*5*) showed a good correlation of trace metals between soil and crop concentrations.

Bibak et al. (*6*) have studied the effects of three levels of nitrogen on trace metal concentrations of potatoes using high-resolution ICPMS. Their work indicates that while fertilizers do influence trace metal uptake by plants, the relationship is complex and is influenced by a number of factors such as soil type, pH, source of nitrogen, and the presence of micronutrients as impurities in the fertilizers. This study indicates that overall, the use of fertilizers does not have an effect on the ability to predict geographic origin of garlic.

More recently, Anderson and Smith (*7*) demonstrated chemical profiling for determining the geographic origin of pistachio nuts. The authors examined other parameters in addition to trace metals including inorganic anions, organic acids, and carbon and nitrogen isotopic ratios. It is interesting to note that their work illustrates the significance of trace metal profiling as compared to the other parameters examined, as the best predictor of geographic origin.

CBP has an interest in determining the country of origin of various agricultural products in order to enforce existing importation laws and regulations and to collect the appropriate tariffs as provided by law. Import requirements include effective administration of import quotas and special trade agreements and enforcement of trade embargoes and compliance with Department of Commerce antidumping and countervailing duty orders. Often, the import restrictions are quite severe and unscrupulous importers are quick to attempt to circumvent them by transshipping agriculture products from a country with trade restraints through a second country that does not have trade restraints or is a recipient of favorable import requirements. The ability to determine the origin of agricultural products is a powerful tool in the enforcement of import laws and requirements.

This study is devoted entirely to garlic due to the level of interest from a CBP perspective because of the relatively large

10.1021/jf040166+ This article not subject to U.S. Copyright. Published 2005 by the American Chemical Society
Published on Web 04/14/2005

quantity of fresh garlic imported into the United States with the incorrect country of origin stated on the CBP entry form. The primary issue on garlic is determining whether the garlic was grown in a country against which the U.S. Department of Commerce has issued antidumping duties. Because garlic has repeatedly been sold in U.S. markets at prices below market value and due to complaints from the U.S. garlic industry, the Commerce Department has placed significant antidumping duties on garlic imported from some countries. Depending on the country of origin and exporter, antidumping duties may be as high as 376% of the value of the garlic. Exporters are often very resourceful in finding markets for their garlic. To circumvent the high antidumping duties, importers have claimed the country of origin to be other than that where the product was actually grown. It is interesting that in some cases the claimed country of origin is a country that does not export any garlic.

The method described here, while simplistic in nature, requires careful control over several of the experimental parameters. Obviously, the country of origin prediction is only as good as the database that is being used for comparison. The data in the database must be representative of the countries in question and ideally should include samples taken in all major growing regions within the country. Periodic updates of sampling may be required in order to assess possible seasonal variations. Other requirements include careful attention to analytical protocols and quality assurance considerations in order to produce verifiable and defendable data. Also, experience with multivariate statistics is essential in order to successfully build and utilize a suitable database for making country of origin predictions.

MATERIALS AND METHODS

The procurement of authentic reference samples from all of the countries of interest has been an important part of this work. Fortunately, we have found that garlic reference samples harvested as early as 1996 continue to match the profiles of the country from which they were grown when compared to recently harvested garlic. Other considerations included concerns about differences in garlic varieties. The varieties of garlic reference samples reported by our collectors are numerous and usually include common names only. In many cases, the garlic variety reported was regional and may not be found at any other location. Because garlic imported into the United States may or may not specify the variety on any of the paperwork and the difficulties reported above, we decided to simplify the study by only considering the country of origin of the samples without regard to regional differences within a country. Undoubtedly, there are minor regional differences but as long as those differences are smaller than the differences between countries, the differences are beyond the scope of this work. The prediction accuracy of our database that is reported in **Table 1** supports the claim that differences in varieties are less than differences between countries. While our goal was to obtain reference samples from every growing region within a given country, we did not set an objective to obtain any set number of reference samples from each growing region. We plan to continue to collect additional reference samples to expand our database and to monitor the possibility of seasonal changes.

In some instances, reference samples have been obtained through Customs Attaché offices. The most successful method for collecting authentic reference samples has been the use of outside contractors. This method, however, is very expensive and time consuming due to the process time for issuing contracts.

We recognize that from a statistician's point of view, the more reference samples the better. The complexity of collecting reference samples from distant regions of the globe for use by CBP was enormous because in some cases people were uncomfortably suspicious of our intentions. We found that garlic may be grown throughout some of the smaller countries while only in small regions of other larger countries. Information on the major regions sampled within each country is shown in **Table 2**. There may be numerous subregions within these major regions in some cases.

Table 1. Prediction Accuracy of Garlic Database Using Discriminant Analysis and Cross-Validation

from country	as compared to	accuracy (%)
Argentina	suspect country	100
suspect country	Argentina	100
Canada	suspect country	100
suspect country	Canada	100
Chile	suspect country	100
suspect country	Chile	100
China	suspect country	100
suspect country	China	100
South Korea	suspect country	100
suspect country	Korea	98
Mexico	suspect country	100
suspect country	Mexico	100
Peru	suspect country	100
suspect country	Peru	100
Philippines	suspect country	100
suspect country	Philippines	100
Spain	suspect country	100
suspect country	Spain	100
Thailand	suspect country	97
suspect country	Thailand	100
Vietnam	suspect country	100
suspect country	Vietnam	100

Table 2. Major Regions of Origin of Garlic Reference Samples

country	region of origin
Argentina	Mendoza Province San Juan Province
Chile	Llay Llay Region El Monte Region Rengo Region
South Korea	Kang Won Province Chung Buk Province Chung Nam Province Jeon Nam Province Jung Buk Province Jeju Province Kyung Province
Peru	Lima Arequipa (Southwest Peru) ICA (Central Coast)
Spain	Andalusia Castilla Puente Genil
Vietnam	Hai Dong Province Thai Binh Province Bac Giang Province Ha Tay Province Luo Cai Province Quang Noai Province Hung Yen Province
Canada	Eastern Ontario Southern Ontario Prince Edward Island
China	Shandong Province Jiangsu Province Henan Province
Mexico	State of Baja California State of Guanajuato Leon Region San Juan
Philippines	Northern Ilocos Sur Southern Ilocas Sur Sinait Ilocos Sur
Thailand	Chang Mai Region Chang Rai Region Phyao Region

Sample Preparation. All sample preparations were carried out in a Class-100 clean room. To obtain a representative sample from whole garlic bulbs, several cloves from multiple bulbs were selected. The cloves were sliced into thin slices and dried in an oven at 85 °C for 48 h. The dried garlic pieces were then ground using a blender (IKA Works, Inc. model A10), creating a garlic powder. Preparation of the garlic powder for trace metal analyses required acid digestion in order to solubilize the sample. Typically, 0.5 g of garlic powder was weighed directly into a Teflon digestion vessel and 10 mL of HNO_3 (EM Merck Omni-Trace) was added. The sample was digested using a microwave digestion system, and after digestion, the digestate was diluted to 50 mL with 18 MΩcm water.

A Thermo Finnigan Element-2 high resolution ICPMS was used in all trace metal analyses reported here. Isobaric interferences common in quadrupole ICPMS instruments were virtually eliminated by selecting the resolution power needed to resolve the isobaric species from the desired element. The high-resolution instrument has a resolution power of approximately 300, 4000, and 10000 at low-, medium-, and high-resolution modes, respectively. The instrument required daily tuning and mass calibration in each of the resolution modes. Typical detector count rates for a 1 ppb indium standard were 1.6×10^6 CPS, 120000 CPS, and 10000 CPS in low-, medium-, and high-resolution modes. Obviously, higher resolving power comes with a substantial sacrifice in sensitivity. Fortunately, most elements can be successfully analyzed in either low- or medium-resolution modes. A CETAC ASX-100 autosampler was used for sample introduction into the ICPMS. The autosampler has a built in Plexiglas cover, which protects the uncovered samples from airborne contaminants that may be present in the instrument room.

To determine which elements to use for trace metal profiling, it was useful to do semiquantitative analyses on a few samples in order to provide information on which elements exist at concentrations above the detection limits of the instrument. This also provides a mechanism for determining how to bracket the sample concentrations with the appropriate calibration standards. Using the information obtained from the semiquantiative analysis, quantitative analysis of the samples may proceed.

Instrument standardization performed in this work included a five-point calibration curve with a correlation coefficient of 0.995 or better. The calibration curve for each element was verified prior to sample analyses by analyzing an initial calibration verification standard (ICV), which is a standard from a source other than the calibration standards. A continuing calibration verification standard (CCV) was analyzed after each 10 samples throughout the sample run in order to verify that the instrument calibration was maintained. The criterion for acceptance for the ICV and CCV was ±10%. A certified reference sample from NIST (SRM1570A Spinach leaf) was included with each sample run with an acceptable quality objective of ±10% for data validation. Also, at least one reagent blank was included with each sample run. All analyses done in medium- or high-resolution included the use of mass offset correction and lock mass to prevent drift of the magnet during the sample run. The instrument software created a text file of the raw data that can be imported into an Excel spreadsheet for the calculation of the trace metal concentrations accounting for sample dilutions, sample mass, and blank contributions.

Multivariate Statistical Analysis. Text files of the trace metal concentrations of the samples were imported into the multivariate statistics software (SAS Institute version 8.2). Subsets of the reference database were created to include only the countries of interest and the samples whose origins were to be determined. A stepwise discriminate analysis was performed to select those elements having sufficient discriminatory power to distinguish one country's metal data from another. The stepwise discriminate analysis was based upon the significance level of the F-test. Using stepwise discriminate analysis, the user may input an SLENTRY value smaller or larger than the default value of 0.15. This value was the probability of the significance of the F-test. Smaller values resulted in the selection of fewer metal variables, and larger values resulted in the selection of more variables. Statistically speaking, the objective was to limit the number of variables selected in order that there were approximately five reference samples from each country for each variable. Therefore, if stepwise discriminate analysis selected five variables, ideally 25 reference samples from each country were needed. It is interesting to note that the elements selected as the best discriminators may vary, depending upon which countries are being compared. Once a prediction model has been established, the model is tested for its accuracy by doing a resubstitution and/or cross-validation analysis. Resubstitution created a calibration data set and a prediction data set and placed all of the reference samples in each data set. The software then produced a prediction accuracy of the model using this criterion. Resubstitution was normally used when the database had a limited number of samples. Cross-validation, which was used in this work, was a bit more rigorous method that removed one reference sample at a time from the calibration data set and placed this reference sample as an unknown sample in the prediction data set for accuracy prediction.

The prediction of the country of origin was accomplished by comparing the trace metal profile of the sample to the profiles of the reference samples from the country claimed by the importer and a country or countries suspected of being the source of transshipped product (e.g., a country against which high antidumping duties have been assessed by the Department of Commerce). Discriminant analysis with cross-validation was performed using only the elements selected from the stepwise discriminant analysis as explained above. The software will then predict the country of origin of the sample and reported the probability of the prediction based upon the current reference samples in the database. This probability did not consider the statistical significance of the number of reference samples in the database. The reference samples in the database should include samples from each growing region of the countries being studied. The results of the discriminant analysis also provided a prediction accuracy of the model based upon the reference samples in the model.

A second statistical method used for reinforcing the country of origin prediction above was canonical discriminant analysis. Canonical discriminant analysis created variables that were linear combinations of the original trace metal variables. The objective was to reduce the number of variables thus simplifying the comparison of trace metal profiles. The method created one less canonical variable than the number of countries in the prediction set or one fewer variables than the number of trace metal variables in the data set, whichever was fewer. Note that the comparison of two countries produced only one canonical variable. By determining the mean and standard deviation of the first canonical variable produced for each country and comparing this to the canonical variable for each unknown sample, a second statistical method for country of origin prediction can be considered. A graphical representation of this, using data from several countries as compared to a suspect country, is presented in **Figure 1**.

Discriminant analysis assumed that the data were normally distributed. However, discriminant analysis can be successfully applied under certain conditions even when this assumption is not valid according to Klecka (*8*). Accordingly, if a particular case has a 90% probability membership in group 1 and only a 10% probability in group 2, then the small inaccuracies created by ignoring the assumptions are unimportant. On the other hand, if there is a 51% probability of membership in group 1 and a 49% probability in group 2, then the assumption of normality cannot be ignored. All of the samples that are reported in this work required a 90% or better probability membership; therefore, we did not consider normality.

RESULTS AND DISCUSSION

Importers are required to state the country of origin of agricultural products on an official entry document when importing them into the United States. If a product is selected by a CBP officer or Agriculture Specialist and sent to the laboratory, it then becomes the responsibility of the laboratory to verify the country of origin claim of the importer or the suspicions of the CBP officer. This is accomplished by comparing the trace metal profile of the sample to a database of authentic reference samples. CBP is interested in knowing whether the trace metal profile of the sample matches the trace

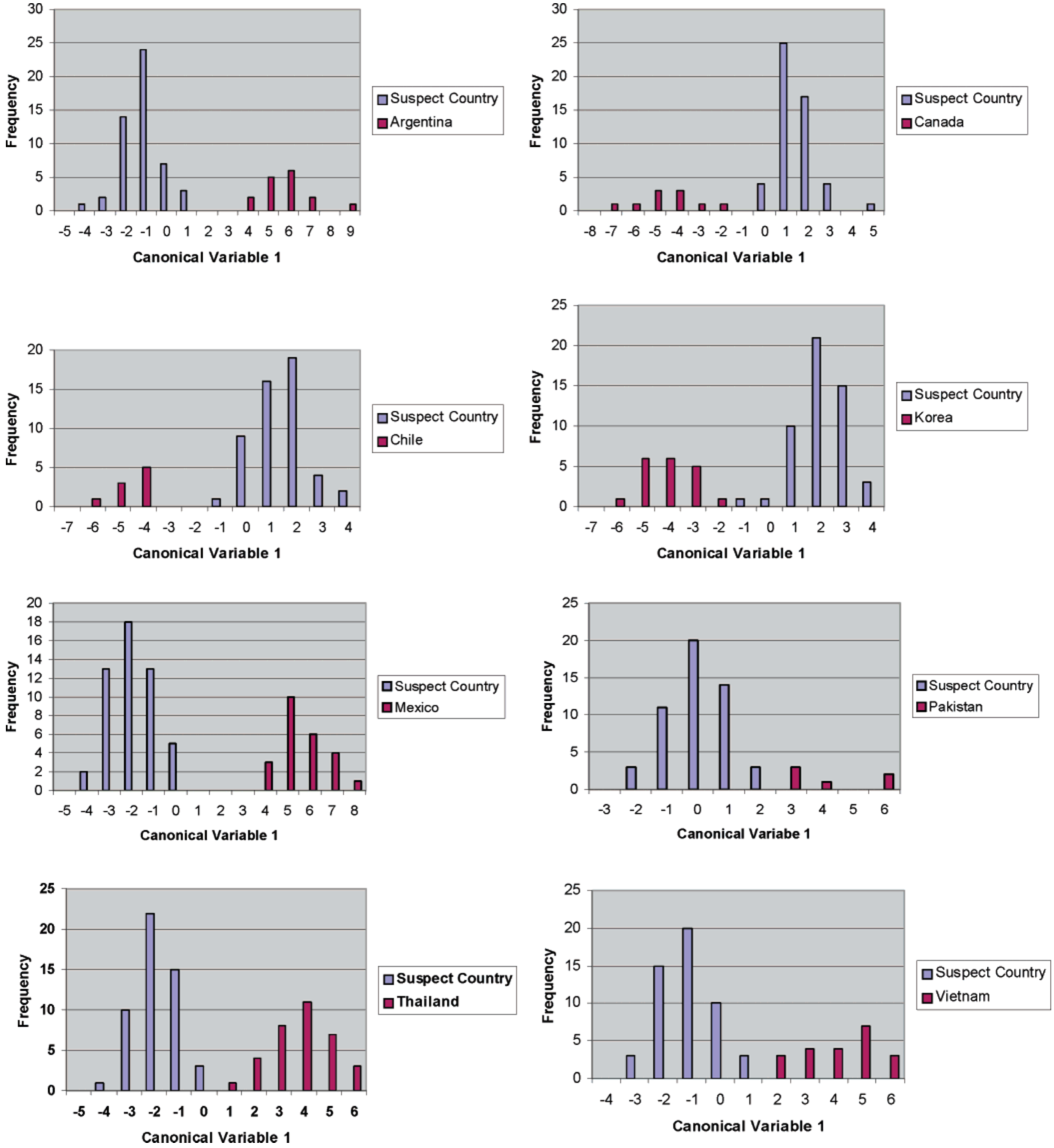

Figure 1. Histogram of canonical variable: suspect country vs comparison country (source: reference database).

metal profile of the country stated on the entry document or whether the profile matches that of a country subject to trade sanctions or increased tariffs. Therefore, comparing sample profiles to every country in our database is unnecessary for our purposes. If, however, an agricultural product were subject to a tariff or sanctions from more than one country, the trace metal profile of the sample would be compared with all of the countries of concern.

Garlic samples were analyzed for 18 elements by high-resolution ICPMS including Li, B, Na, Mg, P, S, Ca, Ti, Mn, Fe, Cu, Ni, Zn, Rb, Sr, Mo, Cd, and Ba. Potassium was analyzed by flame atomic absorption merely for convenience.

The criteria used for reporting a country of origin different from that claimed on the CBP entry document included the following: There must be a greater than 90% probability match as determined by discriminant analysis with cross-validation, and there must be a match with the canonical discriminant variable using two standard deviations around the mean for each country. This conservative approach provides a mechanism for the laboratory to report that the sample either (i) matches the profile of the country claimed, (ii) matches the profile for a suspect country, or (iii) declare the results to be inconclusive.

Because of the law enforcement nature of this work, we are unable to provide the raw trace metal profile data. However,

Figure 1 demonstrates the discriminatory nature of the trace metal data. Clearly, the canonical variable of a sample produced from the trace metal data is easily discernible between a suspect country and the comparison country as demonstrated in **Figure 1**.

In conclusion, U.S. Customs and Border Protection (formerly known as U.S. Customs Service) implemented a special 60 day "Garlic Intervention" in 2002, targeting a portion of fresh garlic importations. This laboratory found that approximately 40% of 134 samples tested during the intervention were from a country other than that claimed on entry documents. These falsifications resulted in the collection of millions of dollars in antidumping duties.

Because of the large amount of money involved when importers are charged antidumping duties up to 376% of the value of the imported product, they are quick to challenge the scientific method. One of the most prevalent claims is that the imported garlic found by the laboratory to be from a country other than that stated on the entry documents was grown from seed garlic of another country. Importers further claim that this might explain a false positive match with a country having trade restrictions or increased tariffs. To address this issue, we have twice obtained garlic grown from seed garlic for our country of origin determination. In one case, the Chinese seed garlic was grown in the United States and the trace metal profile matched the profile for U.S. garlic. A second study involved very immature garlic bulbs grown in Mexico. Despite the fact that the garlic was Chinese seed garlic that had only been in the ground for approximately 30 days, the trace metal profile matched the profile for Mexican garlic.

Other concerns have included the application of fertilizer and its influence on the trace metal profile of the sample. It is important to keep in mind that all of the reference garlic samples in our database were grown for the purpose of sale without regard to issues such as country of origin. Therefore, garlic farmers applied fertilizers as needed or in the case of poor countries, as they were able to. There probably is a great deal of variability in the fertilizer application from field to field as well as country to country. Despite this, the prediction accuracy of our database is quite good as can be seen in **Table 1**. Given the complexity of the relationship of fertilizer usage and other factors that influence trace metal uptake by plants, it is unlikely that fertilizers could be used to manipulate the trace metal profiles of agricultural products in such a way as to change the country of origin prediction from one country to another.

ACKNOWLEDGMENT

I thank Charles Rawlinson for his assistance in the preparation of hundreds of samples used in this study. I also thank Carson Watts and Cecil Clements for their careful review of the manuscript and Dr. Robert Schwartz for giving me my initial training in multivariate statistics.

LITERATURE CITED

(1) Schwartz, R. S.; Hecking, L. T. Determination of geographic origin of agricultural products by multivariate analysis of trace element composition. *J. Anal. At. Spectrosc.* **1991**, *6*, 637−642.

(2) Nikdel, S.; Nagy, S.; Attaway, J. A. Trace metals: Defining geographical origin and detecting adulteration of orange juice. In *Adulteration of Fruit Beverages*; Nagy, S., Attaway, J. A., Rhodes, M., Eds.; Marcel Dekker: New York, 1988; pp 81−105.

(3) 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* (4), 1568−1575.

(4) Anderson, K. A.; Smith, B. W. Chemical profiling to differentiate geographic growing origins of coffee. *J. Agric. Food Chem.* **2002**, *50* (7), 2068−2075.

(5) Samsoe-Peteresen, L.; Larsen, E. H.; Larsen, P. B.; Bruun, P. Uptake of trace elements and PAH's by fruit and vegetables from contaminated soils. *Environ. Sci. Technol.* **2002**, *36* (14), 3057−3063.

(6) Bibak, A.; Sturup, S.; Haahr, V.; Gundersen, P.; Gundersen, V. Concentrations of 50 major and trace elements in Danish agricultural crops measured by inductively coupled plasma mass spectrometry. 3. Potato (*Solanum tuberosum Folva*). *J. Agric. Food Chem.* **1999**, *47*, 2678−2684.

(7) Anderson, K. A.; Smith, B. W. Use of chemical profiling to differentiate geographic growing origin of raw pistachios. *J. Agric. Food Chem.* **2005**, *53*, 410−418.

(8) Klecka, W. R. *Discriminant Analysis: Series: Quantitative Applications in the Social Sciences*; Lewis-Beck, M. S., Ed.; Sage Publications: Beverly Hills, 1980; pp 1−71.

Received for review March 31, 2004. Revised manuscript received February 28, 2005. Accepted March 3, 2005.

JF040166+