8244 J. Agric. Food Chem. 2009, 57, 8244–8249
DOI:10.1021/jf901658f

JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY
ARTICLE

Determination of the Country of Origin of Farm-Raised Shrimp (Family Penaeide) Using Trace Metal Profiling and Multivariate Statistics

Ralph G. Smith* and Carson A. Watts

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

The role of Customs and Border Protection on shrimp importations into the United States is discussed. The focus of this work is to present evidence that the comparison of trace metal data from an imported shrimp sample to an authentic reference database using multivariate statistics produces an accurate geographic location of the sample. The authentic reference database includes data from more than 100 sampling sites spread over eight countries along with the accuracy of each site relative to its known geographic origin. Potential sources of bioaccumulation are briefly discussed; however, it is beyond the scope of this work to provide scientific verification of the mechanism for trace metal uptake. Three examples of imported shrimp are provided for demonstrating the exact procedure for determining their true countries of origin.

KEYWORDS: Country of origin; Customs and Border Protection; shrimp; Penaeide; ICPMS; multivariate statistics; trace-metal; profiles

INTRODUCTION

Approximately 90% of all shrimp consumed in the United States are imported farm-raised shrimp. A listing of the major shrimp importations by country over the past nine years is provided in **Table 1**. Early in 2004 the U.S. Department of Commerce (DOC) initiated an investigation into shrimp importations and in July of that year issued a preliminary determination, which was published in the *Federal Register*, implementing antidumping duties on imported shrimp from certain countries. The antidumping determination issued by the DOC requires Customs and Border Protection (CBP) to enforce assessment of antidumping duties on shrimp that are being sold for less in the U. S. market than in the originating country. Should the DOC determine that foreign producers are receiving government subsidies to reduce their costs to produce shrimp, countervailing duties are also assessed at importation. The antidumping duties vary depending on the country and exporter or producer and can be as high as 112% of the value of the shipment depending upon the findings of the DOC. Additional penalties may be assessed against importers attempting to circumvent the antidumping or countervailing duties by claiming an incorrect country of origin or producer. It is interesting to note from **Table 1** that between 2003 and 2004, as a result of the required antidumping duty deposits required on all shrimp importations from certain countries, the exports of shrimp from some countries increased dramatically, whereas other countries significantly decreased their shrimp exports to the United States.

CBP is responsible for enforcing antidumping and countervailing duty rulings and orders issued by the DOC and published in the *Federal Register*. A verifiable and scientifically accepted method and/or investigative technique for determining the true country of origin of shrimp imports is necessary to properly enforce the antidumping requirements of the United States.

Trace metal profiling for determining the geographic origin of agricultural products has been widely reported. Anderson et al. (*1*) and Anderson and Smith (*2*) have done trace metal profiling and multivariate statistics to determine the geographic origin of potatoes and coffee, respectively. Anderson and Smith (*3*) have also demonstrated chemical profiling for determining the geographic origin of pistachio nuts. Smith (*4*) has reported the use of trace metal profiling and multivariate statistics to determine the geographic origin of garlic. Others (*5*, *6*) have reported determining the geographic origin of various agricultural products using trace metal profiling.

The use of trace metals for determining the geographic origin of aquatic and marine organisms is more difficult. Favretto et al. (*7*) reported using trace elements to differentiate between mussels from two different sites. Windom et al. (*8*) compared trace metal data from a bottom-dwelling fish, *Coryphaenoides armatus*, which is found in both the Atlantic and Pacific Oceans. They reported that the higher concentrations of metals in the muscle tissue of the fish from the Pacific Ocean is statistically

Table 1. Major U.S. Importations of Shrimp 2000−2008[a,b]

year	Thailand	Indonesia	Vietnam	Mexico	Ecuador	Malaysia	China	India	total[c]
2000	193.1	16.2	15.5	28.7	18.6	1.01	17.5	28.2	405.0
2001	133.6	15.5	32.7	29.9	25.7	1.45	26.9	32.7	390.4
2002	110.9	16.9	43.9	24.3	29.1	1.46	48.3	45.4	418.8
2003	125.1	21.0	56.3	25.5	33.6	1.22	77.1	44.9	488.4
2004	124.2	46.1	36.1	29.1	38.0	12.6	56.4	40.7	496.8
2005	153.3	51.4	42.8	28.5	48.1	17.2	10.8	35.4	482.2
2006	187.1	58.7	35.8	35.8	58.8	*d*	28.3	27.2	520.5
2007	179.7	57.9	37.6	40.8	58.1	22.8	21.9	20.4	516.5
2008	175.1	80.7	46.7	34.7	55.8	28.4	26.4	14.4	525.8

[a] Source: CBP Trend Analysis and Analytical Selectivity Program (TAP). [b] All values are kilograms $\times 10^6$. [c] Totals for all U.S. imports of shrimp. [d] Missing data.

Published on Web 08/28/2009
This article not subject to U.S. Copyright. Published 2009 by the American Chemical Society

significant with higher concentrations of trace metals found in deep Pacific Ocean waters. Jung and Zauke (*9*) reported on the bioaccumulation of trace metals in brown shrimp using a laboratory-controlled environment.

Potential sources for trace metal accumulation in the edible tissue of farm-raised shrimp may include but are not limited to those listed here. Ponds used for shrimp production vary in size, shape, design, and stocking density. Some of the ponds are completely lined with a thick plastic lining, some are semilined, and others have no lining at all. The reasons for the differences are either financial or proprietary and, therefore, for the purposes of this work were ignored. Some of the ponds are aerated; others are not, depending upon the stocking densities. The stocking densities of the ponds may range from 10 to 160 animals per square meter or more. The water used for the shrimp culturing is based upon availability and generally comes from a river or estuary. The salinity of the water may vary depending upon where along the estuary the water is pumped and whether it is the rainy or dry season. Sometimes the water is filtered, and sometimes it is not. Feeding rates depend upon stocking densities, size of the shrimp, and possibly other parameters. It may be assumed that water quality parameters such as pH, dissolved oxygen, and alkalinity vary widely within a country and perhaps even more so between countries.

Undoubtedly, diet and water used for rearing the shrimp might be expected to be major contributors in the bioaccumulation of trace metals. The difficulty of studying the water quality of up to 10 river systems or sites per country during all growing seasons would be prohibitive. Most of the larger shrimp farmers view the feed used for growing the shrimp as proprietary and are unwilling to provide samples of the feed. In light of a lack of documentation on feeds used, it is likely that in most cases the feed is prepared from locally available inexpensive fish. Therefore, trace metals in the local fish used for shrimp food are assumed to transfer a locality-specific trace metal source. It was not within the scope of this study to provide models for bioaccumulation of metals by the farm-raised shrimp. We merely point out that environmental differences exist and that other workers (*6*−*9*) have reported on bioaccumulation of trace metals in biota. This study is concerned only with the effect of bioaccumulation, specifically whether there is enough diversity of the trace metal profiles between shrimp from the various countries cited to provide a scientifically valid country of origin prediction. The concept of trace metal profiling for determining the geographic origin of agricultural products has been reported (*1*−*5*). This work is unique in that it applies the concept of trace metal profiling to live animals, and we suggest that it is successful because of the diversity of the environments in which the shrimp were grown.

MATERIALS AND METHODS

It is difficult to obtain current information on the number and location of shrimp farms in each of the countries studied. We do know that most shrimp farms are located near the coast for obvious reasons. However, some countries have over a thousand miles of coast and, in some cases, over very rugged terrain. In some instances, large farms do not allow anyone on the premises, fearing the spread of viruses that can wipe out a shrimp crop. We collected shrimp from at least 10 regions of each country whenever possible and made 3 subsamples from each collection site. A consultant was used to collect the farm-raised shrimp from 10 or more regions of each of the 8 countries listed in **Table 1**. The consultant was required to provide unique sample identifiers, sampling date, GPS coordinates, name of the shrimp farm, variety of shrimp, and location (country, city, and state) for each sample collected. The shrimp were frozen upon collection and packed with dry ice during shipment to the United States and delivery to the laboratory. All of the samples arrived in the laboratory in excellent condition. The most common species of farm-raised shrimp were *Penaeus vannemei* (White leg shrimp) and *Penaeus monodon* (Black Tiger). Approximately 6−12 shrimps, depending upon their size, were selected from each sample, for trace metal analyses. The shrimp were shelled, freeze-dried, and ground into a powder prior to digestion in a microwave digestion system. Duplicate 0.5 g samples were digested in 10 mL of HNO_3 (EM Merck Omni-Trace) using a CEM MarsXpress microwave digestion system and diluted to 50 mL with 18 MΩcm water. A Thermo Finnigan Element-2 high-resolution inductively coupled mass spectrometer (ICPMS) was used for all trace metal analyses. Quality assurance considerations included the following: The ICPMS was calibrated daily using a five-point calibration curve for each element analyzed. The correlation coefficient for the calibration curves was 0.995 or better. Initial calibration verification (ICV) and continuing calibration verification (CCV) standards were analyzed with each data set as described earlier (*4*). A certified reference standard from NIST (SRM1570A) was also digested and analyzed with each data set to ensure that quantitative recoveries for each element analyzed were obtained.

RESULTS AND DISCUSSION

Box and whisker plots of the data for shrimp from each country analyzed are shown in **Figures 1** and **2**. To better clarify box and whisker plots, note that the middle 50% of the data for a given country falls within the box. The whiskers represent the upper and lower 25% on either end of the box. Data points beyond the whiskers are indicative of data outside the 95% confidence level. Therefore, the dispersion of the data for each country may be readily examined. The data demonstrate a considerable overlap of many of the elements, however; boron, rubidium, barium, arsenic, and, to a lesser degree, strontium, selenium, and the Sr/Ba ratio show some disparity between countries. Only through the use of multivariate statistics is it possible to evaluate whether the trace metal data displays enough uniqueness to be useful in classifying the country of origin of shrimp samples.

To provide a level of confidence in determining the true country of origin of imported shrimp, the accuracy of the trace metal data in the reference database for each of the eight countries was examined. In other words, how many of the reference samples from each country are correctly classified when analyzed using multivariate statistics? The accuracy was estimated using linear discriminant analysis and cross-validation. Cross-validation removes one of the reference samples from the database, marks it as "unknown", classifies the sample against the other reference samples, then returns the sample and repeats the process until all of the samples have been classified. A classification accuracy is obtained by determining the percent of correctly classified samples. **Table 2** demonstrates the accuracy of all of the reference samples when one country is compared against another. The classification accuracy was $>90\%$, regardless of which two countries were selected with few exceptions.

Entry documents for imported shrimp must provide a country of origin designation (claimed country) as part of the required documentation for importation. The laboratory was asked to determine whether the imported shrimp matches the claimed country or a country for which antidumping and countervailing duties have been issued (i.e., suspect country).

Multivariate statistics used in this work included three techniques. Stepwise discriminant analysis was used to select the most discriminatory variables (elements). Linear discriminant analysis is used to determine the probability of membership between groups (countries). Canonical discriminant analysis provides canonical variables from the original variables which were used to create a histogram. A description of each of these techniques was described in detail by Klecka (*10*).

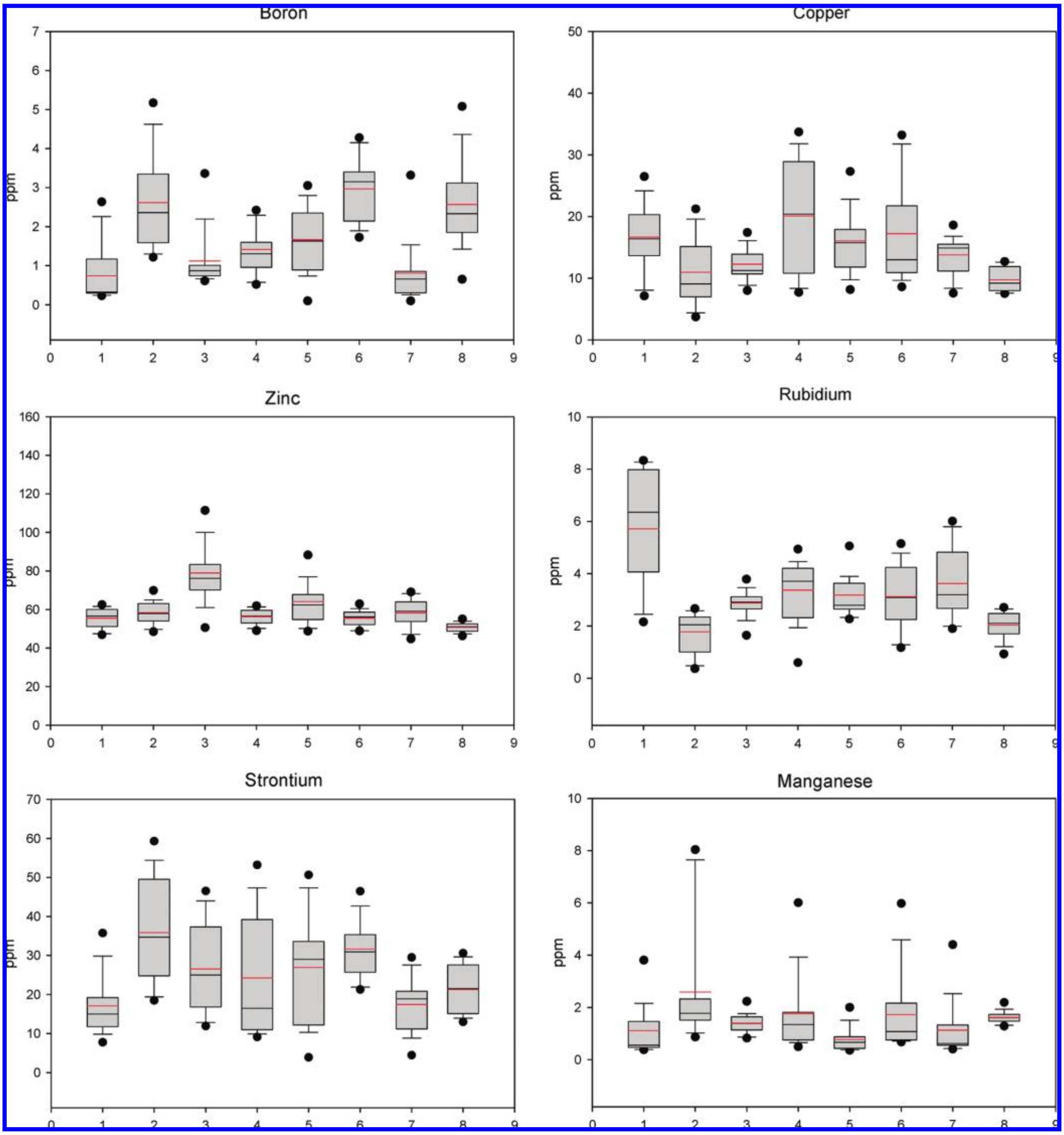

Figure 1. Box and whisker plots: 1, China; 2, Ecuador; 3, India; 4, Indonesia; 5, Malaysia; 6, Mexico; 7, Thailand; 8, Vietnam. The black horizontal line within the box is the median, and the red horizontal line is the mean.

Data files of samples whose origin was to be determined were imported into the multivariate statistics software (SAS version 9.1). Subsets of the reference database were created to include only the data from the claimed country and a suspect country along with sample data for which the origin was to be determined. A procedure known as stepwise discriminant analysis selected the elements that have the greatest discriminatory power for the two countries chosen (claimed vs suspect country). The list of selected elements may be different for a different set of countries, and the identities of these elements are considered to be law enforcement sensitive and are not discussed here. Linear discriminant analysis with cross-validation was performed using the element list determined above. This resulted in a country of origin prediction of the samples with a probability of membership in each of the two countries selected. A second statistical method, canonical discriminant analysis, was used to strengthen and validate the country of origin prediction. Canonical discriminant analysis created variables that are linear combinations of the original trace metal variables. The number of canonical variables created is one less than the number of countries in the prediction set or one less than the number of trace metal variables, whichever is less. In this case, only two countries, claimed and suspect, were examined; therefore, only one canonical variable was provided. To obtain useful information from one canonical variable, a

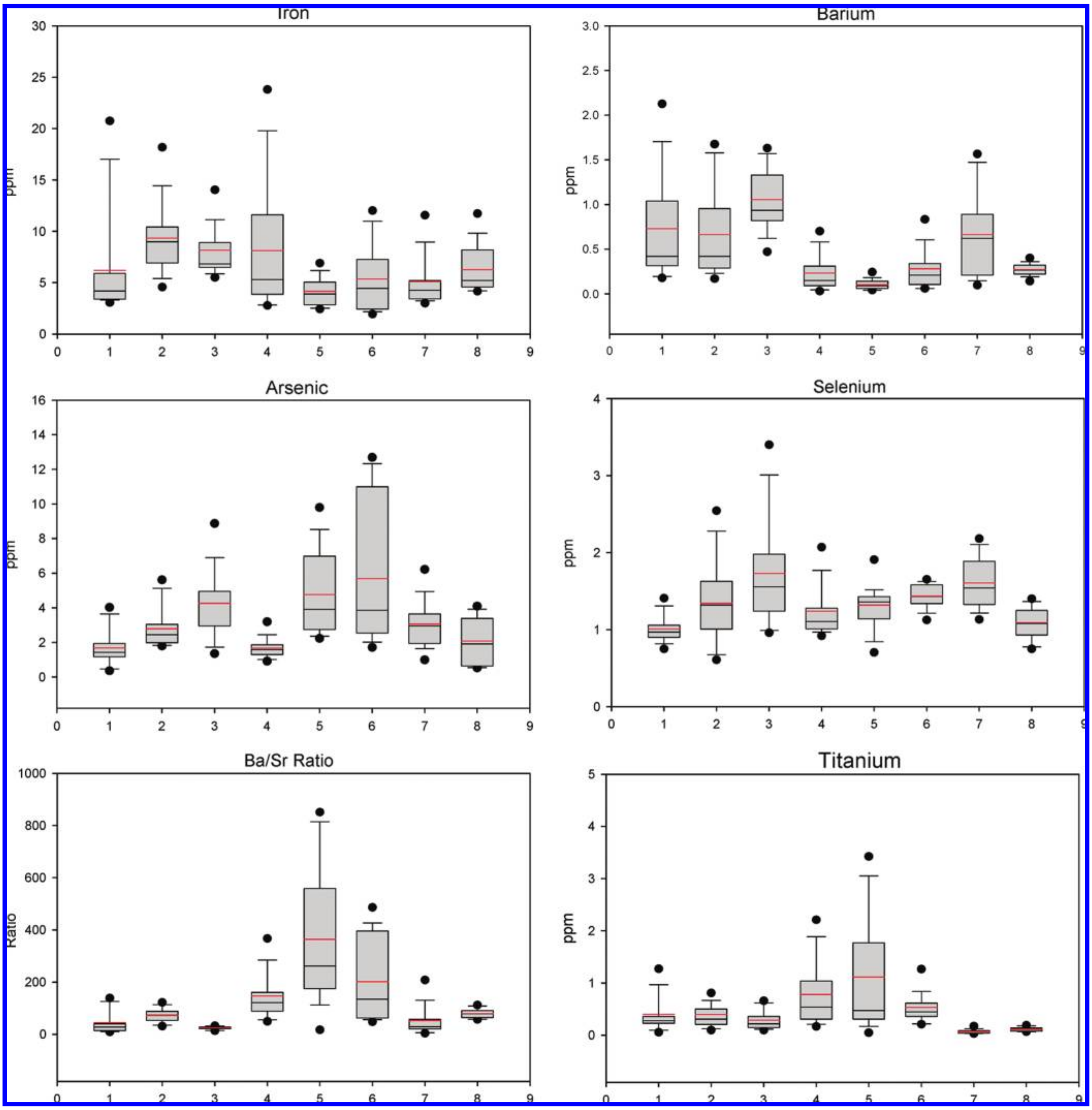

Figure 2. Box and whisker plots: 1, China; 2, Ecuador; 3, India; 4, Indonesia; 5, Malaysia; 6, Mexico; 7, Thailand; 8, Vietnam. The black horizontal line within the box is the median, and the red horizontal line is the mean.

Table 2. Relative Accuracies of Database Using Discriminant Analyses and Cross-Validation[a]

	Thailand	Indonesia	Vietnam	Mexico	Ecuador	Malaysia	China	India
Thailand		98.6	96.7	100	89.7	100	92.6	96.7
Indonesia	90.9		100	93.1	93.1	84.6	85.2	93.3
Vietnam	93.9	91.4		96.6	72.4	96.2	100	100
Mexico	93.9	97.1	100		100	100	100	96.7
Ecuador	90.9	87.1	100	100		94.6	100	100
Malaysia	90.9	95.7	100	100	96.6		86.9	100
China	100	100	100	100	100	88.5		100
India	96.9	98.6	100	100	100	100	100	

[a] All values represent the percent of reference samples that match the database.

histogram of the canonical variable produced for each sample in the reference database and for the samples under investigation was plotted. This provided a visual insight to the data results as well as corroborating evidence for the linear discriminant prob-

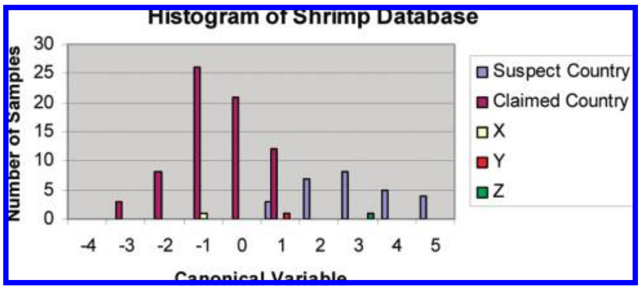

Figure 3. Histogram produced from canonical discriminant analysis.

ability predictions. Each bar on the graph represents the number of samples that have a range of canonical variables between plus and minus 0.5 of the variable shown between the bars. For example, referring to **Figure 3**, it may be seen that the claimed country has 26 reference samples with a canonical variable between −1.5 and −0.5.

Linear discriminant and canonical discriminant analyses are based upon the Euclidian distance between the group centroids (centermost value) of the two countries being evaluated. The proximity of the sample data to either centroid was used by SAS statistical software to calculate the probability of membership. Issues concerning normally distributed data were discussed in an earlier publication (*4*).

From a CBP perspective, a conservative approach is taken in regard to reporting that a sample's country of origin is different from that claimed on the entry documentation. Only after SAS predicts a 90% or greater probability match with a suspect country and canonical discriminant analysis also reveals a match with the suspect country is a report issued stating that the claimed country of origin is incorrect. Therefore, the reported results will either indicate a match with the claimed country of origin or a match with a suspect country whenever SAS statistical software returns a ≥90% with either claimed or suspect country, respectively. If SAS returns <90% match with either country, the results are reported as inconclusive. This conservative approach requires a preponderance of evidence to generate a report that disagrees with the entry documents.

Sample data from three imported shrimp samples labeled X, Y, and Z claiming the same country of origin are shown in **Table 3**. Clearly the data in **Table 3** demonstrate significant differences between the three samples for specific elements, although the entry documents of each of the samples claimed the same country of origin. **Table 4** provides the means and standard deviations of the claimed and suspect countries to which the sample data were compared. The results of linear discriminant analysis shown in **Table 5** indicate that sample X has a >99% probability match with the claimed country, sample Y has inconclusive results because it has a <90% probability match with either country, and sample Z has a >99% probability match with the suspect country. The linear discriminant results compare well with the histogram results shown in **Figure 3**. Note that each of the three sample plots precisely as predicted by linear discriminant analysis. Also note that although sample Y has a >77% probability match with the claimed country, it plots (**Figure 3**) in a region of uncertainty between the claimed and suspect countries.

Statistical comparisons of sample trace metal data to only claimed and suspect countries provides a far less complicated approach than comparison to all countries in the database in determining the country of the origin of imported farm-raised shrimp. Statistically much better separation of the data between two countries occurs than when all of the samples in our database are compared. Also, CBP is interested only in knowing whether the country of origin of the sample matches the paperwork from the importer or a suspect country. Obviously, we are unable to verify the country of origin of shrimp samples from countries not in our database. Evidence that marine organisms including shrimp bioaccumulate trace metals was presented. Varied environmental conditions, thought to be instrumental in the mechanisms of trace metal uptake into the muscle tissue of shrimp and other marine organisms, were discussed. Estimated values for the accuracy of our shrimp database were shown to be >90% in most cases. Trace metal data from three imported shrimp samples were presented along with their multivariate statistical evaluation. Given the relative accuracy of the database as determined by cross-validation shown in **Table 2** and the agreement of two statistical analyses (discriminant and canonical discriminant), the results reported in **Table 5** and **Figure 3** provide persuasive evidence that the proposed method is valid for determining the geographic origin of farm-raised shrimp.

Table 3. Trace Metal Data from Actual Samples[a]

sample	B	Cu	Zn	Rb	Sr	Mn	Fe	Ba	Ti	As	Se	Sr/Ba
X	7.05	7.80	43.0	1.26	39.8	1.00	9.70	0.19	0.43	1.52	1.22	205
Y	1.65	8.60	44.8	1.35	15.3	1.06	4.64	0.15	0.11	0.97	0.89	102
Z	2.19	12.5	50.2	1.80	13.7	1.15	6.79	1.04	0.20	1.18	0.90	13.1

[a] All concentrations are in ppm.

Table 4. Mean and Standard Deviations (SD) Claimed versus Suspect Countries[a]

country	statistic	B	Cu	Zn	Rb	Sr	Mn	Fe	Ba	Ti	As	Se	Sr/Ba
claimed	mean	1.42	20.1	56.3	3.38	24.3	1.76	8.13	0.23	0.78	1.67	1.24	146
	SD	0.82	9.58	4.74	1.26	1.26	1.53	6.37	0.20	0.66	0.60	0.37	96
suspect	mean	0.74	16.6	55.5	5.72	17.1	1.12	6.18	0.73	0.40	1.69	1.08	45
	SD	0.78	5.87	5.40	2.13	1.13	1.24	5.45	0.67	0.44	1.02	0.18	45

[a] All concentrations are in ppm.

Table 5. Results of SAS Prediction[a]

sample	from country	classified into country	suspect %	claimed %
X		claimed	0.02	99.80
Y		claimed	22.76	77.24
Z		suspect	99.85	0.15

[a] The DISCRIM Procedure Classification Results for Calibration Data: Work. Shrimp_Test Cross-Validation Results Using Linear Discriminant Function.

LITERATURE CITED

(1) 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.

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

(3) 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.

(4) Smith, R. G. Determination of the country of origin of garlic (*Allium sativum*) using trace metal profiling. *J. Agric. Food Chem.* **2005**, *53*, 4041–4045.

(5) 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.

(6) Samsoe-Peteresen, L.; Larsen, E. H.; Larsen, P. B.; Preben, B. Uptake of trace elements and PAH's by fruit and vegetables

from contaminated soils. *Environ. Sci. Technol.* **2002**, *36*, 3057–3063.

(7) Favretto, L.; Favretto, L. G.; Marletta, G. P.; Saitta, M. Principal component analysis: a chemometric aid for classification of polluted and unpolluted mussels. *Anal. Chim. Acta* **1989**, *220*, 135–144.

(8) Windom, H.; Stein, D.; Sheldon, R.; Smith, R., Jr. Comparison of trace metal concentrations in the muscle tissue of a benthopelagic fish (*Coryphanoides armatus*). *Deep-Sea Res.* **1987**, *34*, 213–220.

(9) Jung, K.; Zauke, G. P. Bioaccumulation of trace metals in the brown shrimp *Cragnon cragnon* (Linnaeus, 1758) from the German Wadden Sea. *Aquat. Toxicol* **2008**, *88*, 243–249.

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

Received May 18, 2009. Revised manuscript received July 29, 2009. Accepted August 22, 2009.