Using an Ion-Trap MS Sensor to Differentiate and Identify Individual Components in Grapefruit Juice Headspace Volatiles

Kevin L. Goodner[†,*] and Russell L. Rouseff[‡]

U.S. Citrus and Subtropical Products Research Laboratory, Agricultural Research Service, U.S. Department of Agriculture, 600 Avenue S, Northwest, Winter Haven, Florida 33881 and Citrus Research and Education Center, University of Florida, 700 Experiment Station Road, Lake Alfred, Florida 33850

An ion-trap mass spectrometer chemical sensor has been utilized to differentiate between grapefruit juices that differ only in the concentration of a single component, and the sensor was able to identify that component. Grapefruit juice was fortified with 40 to 2000 ppm vanillin, a low-level naturally occurring compound in citrus juices. Principal components analysis and discriminant analysis of mass spectral data (*m*/*z* 50−200) provided clear separation of the grapefruit juice samples. Vanillin was observed in the juice headspace at the 40 ppm level, with identification possible at the 100 ppm level using either MS or MS/MS.

Keywords: *Vanillin; mass spectrometry; flavor; electronic nose; statistical analysis; PCA; DFA; multivariate*

INTRODUCTION

The "electronic nose" is a fairly recent technology utilizing an array of chemical sensors and employing chemometrics to differentiate samples. Semiconducting sensors were first used in a commercial device to detect gases by Taguchi (*1*). By 1989, approximately 20 million sensors were being produced commercially each year (*2*). However, shortcomings of early semiconducting sensors included limited reproducibility, stability, sensitivity, and selectivity (*2*). These drawbacks still exist, but are not as pronounced as in the past because of improved manufacturing techniques and new sensor designs. The concept of using an array of sensors for rapid gas analysis was discussed as early as 1983 (*3*). These early sensor arrays involved redundancy, signal averaging, or single compound specificity sensors with a gating circuit for detection. The ability to perform intense statistical calculations, such as pattern recognition and neural networks, has only recently been readily available using personal computers. Since 1993, a number of companies have introduced electronic nose products that combine a sensor array and chemometrics, such as Alpha-MOS (Toulouse, France), Aromascan (Crewe, UK), Neotronics/EEV (Elmsford, NY), Cyrano (Pasadena, CA), and Perkin-Elmer (Norwalk, CT).

Two of the more common statistical analyses used for electronic nose data are principal components analysis (PCA) and discriminant function analysis (DFA). Principal components analysis is a technique that attempts to preserve the inherent structure and variance of the data while reducing the complexity of the representation by reducing the number of dimensions used to represent the data. PCA can be described as projecting an *n*-dimensional data set onto a two- or three-dimensional space. The orthogonal axes for this space are a linear combination of the individual variables such that the first principal component retains as much variation as possible and the second principal component retains as much of the remaining variance as possible. The two main benefits of PCA are that it provides a two-dimensional representation of *n*-dimensional data, and that the magnitude of the coefficients of the variables that comprise the principal components give an indication of their significance for determining data structure. Discriminant function analysis (DFA) operates similarly to PCA in that it uses a linear combination of the variables to obtain the axes. However, DFA creates axes that have the largest separation between predefined classes instead of the inherent structure. Thus, DFA will have the highest coefficients on the variables that most discriminate between the data classes versus those that have the highest variation (*4*).

The applications of electronic noses are many and varied. They have been used for determining the spoilage of beef (*5*), predicting the shelf lives of edible oils (*6*), differentiating grapefruit juice varieties (*7*) and fragrances (*8*), flavor analyses (*9*), and discriminating between good and rancid biscuits (*10*). Common limitations of electronic noses are the drifting of sensor signal over time, limited sensor lifetime, limited types of sensors, and the general lack of specificity of these sensors. One method for avoiding these limitations is to use a mass spectrometer (MS) as a multi-sensor array. In this case, each mass-to-charge ratio (*m*/*z*) is a "sensor" that detects any molecule or fragment with that particular *m*/*z*. This means that a mass-spectrometer-based electronic nose has potentially hundreds of sensors. A mass-spectrometer-based chemical sensor also would not be subject to many of the limitations of electronic noses using semiconductor technology. The reproducibility, stability, and sensitivity of mass spectrometers have been well established. However, the use

* To whom correspondence should be addressed. Tel: (863) 293-4133. Fax: (863) 299-8678. E-mail: goodner@citrus.usda.gov.

[†] USDA, ARS, SAA U.S. Citrus & Subtropical Products Laboratory.

[‡] University of Florida, Citrus Research and Education Center.

10.1021/jf000867g CCC: $20.00
Published on Web 12/15/2000

of mass spectrometers as electronic noses is a fairly recent adaptation and many improvements and different implementations can be made. One example is illustrated by Marsili (*11*) by using a SPME fiber and a 1-m fused silica adapter in order to bypass the GC of a GC−MS system. This setup provided separation of differing milk samples and demonstrates the technique's usefulness. No attempt was made at identification of individual components in the headspace, only differentiation of samples. There are commercially available quadrupole mass-spectrometer-based electronic nose systems (Hewlett-Packard, Palo Alto, CA).

One modification is to use an ion-trap-based mass spectrometer instead of the usual quadrupole-based MS. Quadrupole mass spectrometers, also known as mass selective detectors or mass filters, scan the mass range to produce a mass spectrum. The amount of time required for quadrupole data collection at any particular mass (and hence, the resultant sensitivity), is dependent on the scan rate, d*m*/d*t*. An ion-trap MS traps mass fragments in an electrical field and then ejects the desired masses to obtain a mass spectrum. The ability to trap and isolate individually charged species allows further fragmentation for identification purposes: this is referred to as MS/MS. The purpose of this report is to present evidence that a mass-spectrometer-based electronic nose can be used to differentiate samples, identify constituents, and quantify constituents in both the MS and MS/MS modes.

MATERIALS AND METHODS

A gable-top carton of 100% Florida Ruby Red "not-from-concentrate" grapefruit juice (GJC) was obtained from a local grocery. Vanillin (99.9%+ pure) was obtained from Aldrich (Milwaukee, MI).

All data were collected using a Finnigan GCQ Plus GC−MS system (Thermoquest Corp, San Jose, CA) with 99.999% pure helium in the electron ionization (EI) mode. Carrier gas velocity was 35.0 cm/sec through a 60 m × 0.25 mm i.d., 0.25 m RTX5-MS column (Restek Corp, Bellefonte, PA). Juice samples (20 mL) were housed in 30-mL septa-capped vials that were placed in-line with the GC column which terminated in the vial, and 0.32 mm i.d. deactivated-silica going from the vial to the MS. To place the juice in-line with the GC column, the GC column was removed from the MS interface and inserted into a septa cap. Then a piece of deactivated fused-silica tubing approximately 0.5 m long was also inserted into the septa cap and placed into the MS interface. The helium flow from the column was used to convey the headspace from the sample vials into the MS. The juice samples were spiked using a concentrated solution of vanillin in ethanol. The GC oven was used to adjust the sample temperature. Samples were allowed to equilibrate at 95 °C before mass spectra (*m*/*z* 50−200) were collected once per second for 1 min. The eighteen mass spectra directly preceding the switch from MS to MS/MS were selected as replicates for each concentration. Next, MS/MS spectra were obtained once per second for 1 min by isolating the base peak, *m*/*z* 151, for 12 ms followed by application of an excitation voltage of 0.8V using a quartet (q) value of 0.225 for 30 ms and subsequently scanning from *m*/*z* 50 to 160.

All statistical analyses were conducted using Statistica (Version 5.5, Statsoft, Tulsa, OK). The variables for inclusion in the discriminant function analysis (DFA) studies were chosen by taking the 36 variables with the highest loadings in the principal components analysis (PCA) analysis. For the testing and validation of the DFA, $^1/_2$ of the data points were used to develop a model and the other $^1/_2$ of the data points were used to validate the model. The training data points were chosen as every other data point with the validation points being the remaining points.

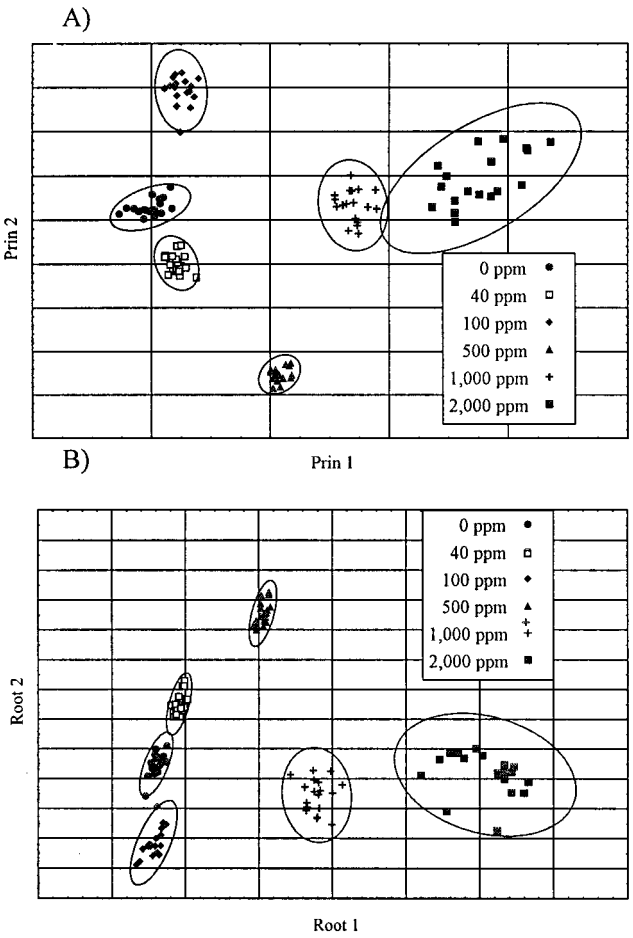

Figure 1. (A) Principal components analysis (PCA) results for the vanillin-fortified samples (PCA computed using correlation matrix 150 variables). (B) Canonical discriminant function analysis (DFA) of 6 variables (*m*/*z* 75, 81, 94, 101, 137, and 152) determined from the 36 initial variables using stepwise discriminant analysis.

RESULTS AND DISCUSSION

Classification. Using principal components analysis (PCA) employing 150 variables (*m*/*z* 50−200), the six concentration levels (0, 40, 100, 500, 1000, and 2000 ppm) formed tight groupings depending on the amount of vanillin present, as can be seen in Figure 1A. The first two principal components (Prin 1 and Prin 2) obtained using correlation matrixes for the PCA analysis were used as the axes in Figure 1A, which accounted for 97.6% of the variance. It is clear from this figure that the mass spectrometer is capable of being used to differentiate samples that differ only in relatively low levels of a single component because there is clear separation between the zero and 40 ppm vanillin-fortified juice. Principal components analysis is used to show the inherent structure of the data. Data from each of the six concentration classses were tightly clustered. There was no overlap in any of the 18 individual data points within each concentration class. The ellipses represent the 95% confidence limit for each group. PCA does not take advantage of any a priori classifications to improve clustering of the data as is done when using discriminant function analysis (DFA).

DFA uses prior classifications of the samples to emphasize those variables which maximize the separation between the different classifications. The variables which do not contribute significantly to separating the

data into groups should not be used. In this study, the variables used for the DFA were chosen from their factor loadings obtained in the PCA analysis. The 36 variables (*m*/*z*) chosen were those that had the highest factor loadings in Prin 1 or Prin 2. Thirty-six variables were chosen, as this is the maximum that is acceptable with 108 data points (18 replicates for 6 concentrations). Generally, at least a 3-to-1 ratio between the number of data points and number of variables is used to model data. Using the 36 variables with DFA, all six concentrations groups are tightly clustered and well separated. However, when attempting to model data, it is best to use as few variables as possible. Therefore, Figure 1B is a canonical DFA using only 6 variables consisting of *m*/*z* 75, 81, 94, 101, 137, and 152. The two most chemically important of these ions are 137 and 152. Terpenes, which are predominant in citrus juice headspaces, have a molecular weight of 136; *m*/*z* 137 is the M+H peak and/or the ^{13}C peak resulting from the natural abundance of ^{13}C in relation to ^{12}C. This ion is also the vanillin fragment M-CH_3 which is the fragment of vanillin that has lost a methyl group. Vanillin has a molecular weight of 152, which explains the importance of *m*/*z* 152. These variables were chosen using backward stepwise discriminant analysis of the original 36 variables with tolerance, 0.01; F to enter, 11; F to remove, 10; and number of steps, 36.

Figure 1 was generated using correlation matrixes; however PCA can also be calculated using covariance matrixes. The differences between using correlation versus covariance matrixes are often slight, are most often seen in the factor loadings, and are useful in understanding the relationships between the variables. The PCA generated from the same 108 data points using covariance matrixes also produced six tightly clustered groups with no overlap. Similar results were also obtained using covariance matrixes and DFA. The only difference was found with canonical DFA where the same backward stepwise discriminant procedure resulted in 8 variables consisting of *m*/*z* 67, 79, 87, 91, 93, 95, 111, and 137. As in the earlier example, *m*/*z* 137 is important because of the changes in terpenes which have a molecular weight of *m*/*z* 136 and as a vanillin fragment. It should be noted that this technique would probably not be possible with a low concentration terpene, as it would be masked by the large quantity of other terpenes found in citrus juices which would have similar fragmentation patterns and would interfere with the compound of choice.

All of the DFA examples in this study were validated by running the DFA using only $^1/_2$ of the data points to train the model, and then applying that model to predict the classification of the validation samples. In all cases, there was 100% correct classification for both the data points used to create the model as well as the unknown data points used for validation.

Quantitation. One feature of mass spectrometers that is commonly utilized is their linear response to concentration changes. This enables calibration curves to be constructed which can be used to determine the concentration of an unknown sample. When all 108 data points using only the response for *m*/*z* 152 are plotted (concentration vs response), the calibration curve has an r^2 of 0.917. Mass/charge 152 was chosen because it is the molecular weight of vanillin; other *m*/*z* have higher r^2, but are fragments of vanillin. The vanillin fragment m-CH_3 at *m*/*z* 137 provides the best calibration

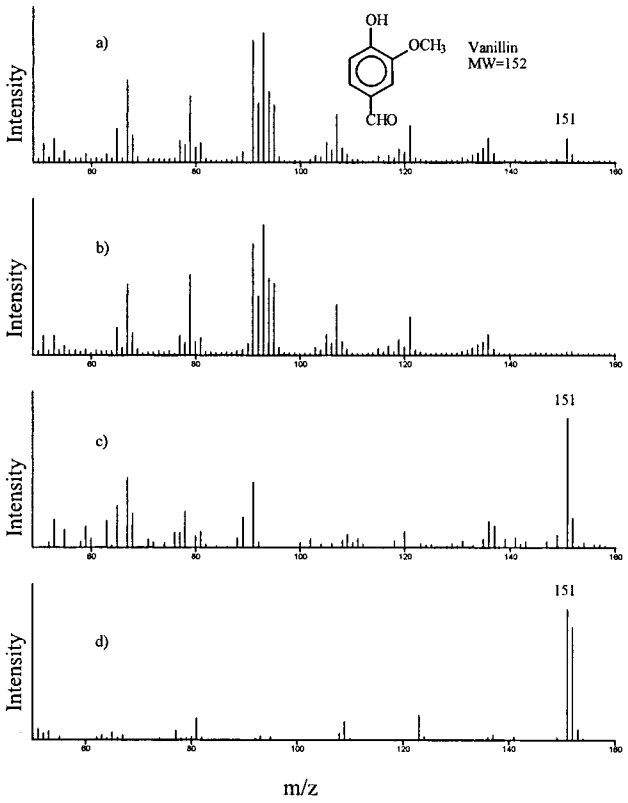

Figure 2. Mass spectra: (a) GFJ fortified with 100 ppm of vanillin; (b) GFJ without fortification; (c) subtraction of the unfortified GFJ from the fortified GFJ; and (d) vanillin standard.

curve (r^2 = 0.962). A partial least-squares regression using Statistica computed an r^2 of 0.999 using only 6 variables which are each linear combinations of the mass/charge intensities.

Identification. One of the most important and useful feature of mass spectrometers is the ability to conduct library-based searches to identify unknown compounds. This technique works well in situations where there are few interferences, but is more difficult with mass spectra of impure compounds. Such a limitation is often circumvented by subtraction of a background spectrum to produce a mass spectrum that is free of chemical or electrical noise. This technique is commonly used in GC−MS to differentiate coeluting compounds or when high background signal is present due to sampling conditions. However, background subtraction can also be used in unseparated headspace applications. Figure 2 shows four mass spectra consisting of (a) GFJ fortified with 100 ppm of vanillin; (b) GFJ without fortification; (c) subtraction of the unfortified GFJ from GFJ fortified with 100 ppm of vanillin; and (d) a vanillin standard. A clear identification can be made at the 100-ppm level by comparing spectra (c) and (d). The 40-ppm level could also be identified, but because of the lower signal it was not as clear a match as in Figure 2. Vanillin naturally occurs in citrus at about the 0.5 ppm level (*12*). The authors feel that it would be possible to detect naturally occurring vanillin in citrus juices with improved sample introduction and optimized conditions. This figure illustrates that when using a mass-spectrometer-based electronic nose it is not only possible to differentiate samples, but also to determine why the samples differ. This is an important and noteworthy demonstration as traditional electronic noses using metal oxide sensors, conducting polymers, or other chemical-based sensors

do not have the ability to determine specific compositional differences between samples. This suggests that the MS-based electronic noses of the future may have headspace specific libraries. This is important as the mass spectra obtained using a total headspace sampling technique are different from those obtained using other sampling procedures.

The MS/MS technique allows for isolation of an ion and then further fragmentation for identification. The first MS stage was used to filter all but ions *m*/*z* 151−152 and these were then further fragmented. Spectra obtained using 100-ppm vanillin-fortified GFJ were similar to those shown in Figure 2 for the single MS case. The subtracted spectrum was not as clear a match with the vanillin MS/MS spectrum as in the single MS.

An ion-trap mass-spectrometric-based electronic nose can perform functions similar to those performed by a sensor-based electronic nose, with the ion-trap MS-based system being able to identify and quantify individual components in the sample in addition to classifying the sample. The MS-based system has the ability to obtain signal from compounds that are present in low concentrations which means that the instrument can more often measure the aroma-active compound instead of a compound in higher abundance that correlates with the aroma-active compound.

LITERATURE CITED

(1) Taguchi, N. Gas Detecting Device. U.S. Patent 3,631,436, 1971.
(2) Madou, M. J.; Morrison, S. R. *Chemical Sensing with Solid State Devices*; Academic Press: San Diego, CA, 1989.
(3) Clifford, P. K.; Tuma, D. Characteristics of semiconductor gas sensors. I. Steady-state gas response. *Sens. Actuators* **1983**, *3*, 233−254.
(4) Massart, D. L.; Vandeginste, B. G. M.; Deming, S. N.; Michotte, Y.; Kaufman, L. *Chemometrics: A textbook*; Elsevier: Amsterdam, 1988; Vol. 2.
(5) Blixt, Y. *Electronic Noses in the Food Industry*; Institutet För Livsmedel Och Bioteknik, Matforsk: Stockholm, Sweden, 1998; p 108.
(6) Shiers, V.; Adechy, M. Use of multi-sensor array devices to attempt to predict shelf lives of edible oils. *Semin. Food Anal.* **1998**, *3*, 43−52.
(7) Bazemore, R.; Rouseff, R.; Nordby, H.; Goodner, K.; Jella, P.; Bowman, K. Discrimination of grapefruit juice varietal differences by an electronic nose equipped with metal oxide sensors. *Semin. Food Anal.* **1997**, *2*, 239−246.
(8) Hodgins, D.; Conover, D. Evaluating the electronic NOSE. *Perfum. Flavor.* **1995**, *20*, 1−2.
(9) Strassburger, K. J. Electronic nose evaluation in the flavor industry: it really works. *Food Test Anal.* **1996**, *2*, 22−26, 37.
(10) Persaud, K. C.; Travers, P. J. Arrays of broad specificity films for sensing volatile chemicals. In *Handbook of Biosensors and Electronic Noses*; Kress-Rogers, E., Ed.; CRC Press: Boca Raton, FL, 1997; pp 565−592.
(11) Marsili, R. T. SPME-MS-MVA as an electronic nose for the study of off-flavors in milk. *J. Agric. Food Chem.* **1999**, *47*, 648−654.
(12) Goodner, K. L.; Jella, P.; Rouseff, R. L. Determination of Vanillin in Grapefruit and Other Citrus Juices using GC-olfactometry and GC/MS/MS *J. Agric. Food Chem.* **2000**, *48*, 2882−2886.

Received for review July 12, 2000. Revised manuscript received October 27, 2000. Accepted October 30, 2000. Mention of a trademark or proprietary product is for identification only and does not imply a guarantee or warranty of the product by the U.S. Department of Agriculture. The U.S. Department of Agriculture prohibits discrimination in all its programs and activities on the basis of race, color, national origin, gender, religion, age, disability, political beliefs, sexual orientation, and marital or family status.

JF000867G