

# Discrimination of Complex Mixtures by a Colorimetric Sensor Array: Coffee Aromas

Benjamin A. Suslick,<sup>†,‡</sup> Liang Feng,<sup>‡</sup> and Kenneth S. Suslick<sup>\*,‡</sup>

University Laboratory High School, University of Illinois at Urbana–Champaign, 1212 W Springfield Ave., Urbana, Illinois 61801, and Department of Chemistry, University of Illinois at Urbana–Champaign, 600 S. Mathews Ave., Urbana, Illinois 61801

The analysis of complex mixtures presents a difficult challenge even for modern analytical techniques, and the ability to discriminate among closely similar such mixtures often remains problematic. Coffee provides a readily available archetype of such highly multicomponent systems. The use of a low-cost, sensitive colorimetric sensor array for the detection and identification of coffee aromas is reported. The color changes of the sensor array were used as a digital representation of the array response and analyzed with standard statistical methods, including principal component analysis (PCA) and hierarchical clustering analysis (HCA). PCA revealed that the sensor array has exceptionally high dimensionality with 18 dimensions required to define 90% of the total variance. In quintuplicate runs of 10 commercial coffees and controls, no confusions or errors in classification by HCA were observed in 55 trials. In addition, the effects of temperature and time in the roasting of green coffee beans were readily observed and distinguishable with a resolution better than 10 °C and 5 min, respectively. Colorimetric sensor arrays demonstrate excellent potential for complex systems analysis in real-world applications and provide a novel method for discrimination among closely similar complex mixtures.

The evaluation and discrimination of complex mixtures remains an important challenge to chemical analysis. The most common strategy for analysis of mixtures is a complete component-by-component approach, i.e., fractionation of the mixture and characterization of the individual components. This generally implies the use of hyphenated techniques, i.e., the sequential combination of a separation technique (e.g., a chromatography) with single or multiple spectroscopic techniques (e.g., mass spectrometry).<sup>1,2</sup> While gas chromatography/mass spectrometry (GC/MS) is the most popular of all hyphenated techniques, it often proves cumbersome for accurate discrimination among

similar complex mixtures.<sup>2,3</sup> Moreover, even for high-performance separation techniques, the number of compounds that can be differentiated is disappointingly small relative to the extremely large number of components in truly complex mixtures.<sup>3,4</sup>

For complex mixtures with hundreds of components, there are often multiple analytical goals: in addition to the occasional requirement for a full component-by-component analysis, more common needs involve comparisons against a standard, discrimination of subtle differences among similar mixtures, or changes in the mixture as a function of time or conditions. For these other needs, one may also consider a complementary sensor-array approach.<sup>5</sup> An alternative method to discriminate among complex mixtures is to treat the mixture as a single analyte and to collect a combined response simultaneously on an array of sensors, in analogy with our own biological sensors, i.e., an electronic nose or tongue. Compared with traditional analytical technologies, this approach can be less expensive and easier to operate, making it potentially attractive for many industrial applications. Electronic noses<sup>5–16</sup> generally rely on multiple, cross-reactive sensors that yield a response based either on changes in physical properties (e.g., mass, volume, conductivity) or reaction with a surface (i.e., analyte oxidation on heated metal oxides). Specific examples of the sensors used in prior electronic noses include conductive polymers and polymer composites, multiple polymers doped with single fluorescent dye, polymer coated surface acoustic wave (SAW) devices, and metal oxide sensors. These types of sensors are often successful in discriminating among different single unrelated compounds. A common limitation in prior electronic nose technology, however, is a general lack of chemical specificity: this makes differentiation among closely related compounds

\* Corresponding author. Tel: 1-217-333-2794. Fax: 1-217-333-2685. E-mail: ksuslick@uiuc.edu.

<sup>†</sup> University Laboratory High School, University of Illinois at Urbana–Champaign.

<sup>‡</sup> Department of Chemistry, University of Illinois at Urbana–Champaign.

(1) Lobinski, R.; Szpunar, J. *Hyphenated Techniques in Speciation Analysis*; Royal Society of Chemistry: Cambridge, 2003.

(2) Brinkman, U. A. T. *Hyphenation: Hype and Fascination*; Elsevier Science Ltd: Amsterdam, 1999.

(3) Hübschmann, H.-J. *Handbook of GC/MS Fundamentals and Applications*, 2nd ed.; Wiley-VCH: New York, 2009.

(4) Kellner, R.; Mermet, J.-M.; Otto, M.; Widmer, H. M. *Analytical Chemistry*; Wiley-VCH: New York, 1998.

(5) Walt, D. R. *Anal. Chem.* **2005**, *77*, A-45.

(6) Gardner, J. W.; Bartlett, P. N. *Electronic Noses: Principles and Applications*; Oxford University Press: New York, 1999.

(7) Aernecke, M. J.; Walt, D. R. *Sens. Actuators* **2009**, *142*, 464–469.

(8) Anslin, E. V. J. *Org. Chem.* **2007**, *72*, 687–699.

(9) Lewis, N. S. *Acc. Chem. Res.* **2004**, *37*, 663–672.

(10) Röck, F.; Barsan, N.; Weimar, U. *Chem. Rev.* **2008**, *108*, 705–725.

(11) Hierlemann, A.; Gutierrez-Osuna, R. *Chem. Rev.* **2008**, *108*, 563–613.

(12) Hsieh, M.-D.; Zellers, E. T. *Anal. Chem.* **2004**, *76*, 1885–1895.

(13) Grate, J. W. *Chem. Rev.* **2000**, *100*, 2627–2647.

(14) Janata, J.; Josowicz, M. *Nat. Mater.* **2003**, *2*, 19–24.

(15) Wolfbeis, O. S. *J. Mater. Chem.* **2005**, *15*, 2657–2669.

(16) James, D.; Scott, S. M.; Ali, Z.; O'Hare, W. T. *Microchimica Acta* **2005**, *149*, 1–17.

problematic and makes it especially difficult for prior electronic nose technology to discriminate among highly similar complex mixtures.

In recent years, we have developed a rather different approach using a colorimetric sensor array. The design of such an array<sup>17–24</sup> is based on strong dye–analyte interactions, which is quite different from other electronic nose technologies that generally rely on weak, nonspecific intermolecular interactions. Optical arrays have also found other applications for sensing in aqueous solutions of anions, organic compounds, amino acids, beverages, proteins, and even whole cancer cells.<sup>25–35</sup>

Owing to their high specificity and low cost, colorimetric sensor arrays are suitable for both laboratory and industrial applications in the analyses of complex mixtures. To prove this point, we sought a carefully controlled, highly reproducible complex mixture, and for these studies, we have examined commercially available coffees as an archetype of such complex analytes. Coffee is one of the most consumed beverages in the world, and remarkably, the primary industrial method of quality control for coffee remains the use of human smell and taste, in spite of the inherent nonquantitative and often subjective limitations that such “organoleptic” analysis implies.<sup>36</sup> The volatiles that make up the aroma of any coffee, of course, play an important role in sensory analyses and can be considered a “fingerprint” of the product.<sup>37,38</sup> While considerable efforts have been made to chemically characterize the aroma-related substances in coffee,<sup>37–43</sup> reliable discrimination among different coffees remains a difficult

task still under very active investigation.<sup>44–47</sup> Roasted coffee beans contain more than 1000 discrete chemical compounds,<sup>39–43</sup> which makes the identification and discrimination of different coffees extremely difficult by traditional chemical analysis. Furthermore, the roasting of coffee beans is highly dynamic, and the processes that develop the flavor and aroma of coffee are strongly time and temperature dependent.<sup>39–43,48,49</sup>

We report, here, a colorimetric sensor array method for analyzing the aroma of a variety of coffees both in whole bean and ground form. We observe excellent discrimination among 10 different brands of commercial coffee. In addition, we are also able to distinguish the differences among coffee beans roasted at varying temperatures and lengths of time.

## EXPERIMENTAL SECTION

**Materials and Methods.** Ten commercially available roasted coffees (six Columbian coffees, Eight O’Clock Columbian Roast, Folgers Columbian Roast, Folgers Grande Supreme Decaf, Maxwell House Original Roast, Maxwell House Original Roast Decaf, Starbucks Columbian Roast, and four other coffees, Starbucks Espresso Roast, Café Mai Traditional, Starbucks Sumatra Roast, and Eight O’Clock Hazel Nut) were purchased from local supermarkets and stored in the freezer compartment of a conventional kitchen refrigerator. For roasting experiments, Huila Columbian green coffee beans were used (Columbia Street Roastery, Champaign, IL); the Huila Departamento is located in southwest Columbia.

The composition of the colorimetric sensor array used was described previously; the specific nanoporous pigments used in the array are given in the Supporting Information, Table S1. For printing, the formulations were loaded into a 36-hole Teflon ink well. Sensor arrays were printed using an array of 36 floating slotted pins (which delivered approximately 130 nL each) by dipping into the ink well and transferring to the polyethylene terephthalate (PET) film. Once printed, the arrays were aged under a slow nitrogen flow for at least 3 days to ensure removal of solvent vapors. The shelf life of the arrays is excellent with no significant changes in array response after 3 months.<sup>24</sup>

A Masterflex peristaltic pump (#77912–00; pump head #77390–00) was connected with Teflon tubing to a 100 mL vial containing 400 mg of the coffee sample (cf. Supporting Information, Figure S1). Two three-way valves were used to close the system, allowing the volatiles to saturate the air within the tubing. In order to reproducibly expose the array to an air stream fully saturated with coffee volatiles, the pump was

- (17) Suslick, K. S.; Bailey, D. P.; Ingison, C. K.; Janzen, M.; Kosal, M. A.; McNamara, W. B., III; Rakow, N. A.; Sen, A.; Weaver, J. J.; Wilson, J. B.; Zhang, C.; Nakagaki, S. *Quim. Nova* **2007**, *30*, 677–681.
- (18) Rakow, N. A.; Suslick, K. S. *Nature* **2000**, *406*, 710–713.
- (19) Suslick, K. S. *MRS Bull.* **2004**, *29*, 720–725.
- (20) Rakow, N. A.; Sen, A.; Janzen, M. C.; Ponder, J. B.; Suslick, K. S. *Angew. Chem., Int. Ed.* **2005**, *44*, 4528–4532.
- (21) Janzen, M. C.; Ponder, J. B.; Bailey, D. P.; Ingison, C. K.; Suslick, K. S. *Anal. Chem.* **2006**, *78*, 3591–3600.
- (22) Zhang, C.; Suslick, K. S. *J. Am. Chem. Soc.* **2005**, *127*, 11548–11549.
- (23) Musto, C. J.; Lim, S. H.; Suslick, K. S. *Anal. Chem.* **2009**, *81*, 6526–6533.
- (24) Lim, S. H.; Feng, L.; Kemling, J. W.; Musto, C. J.; Suslick, K. S. *Nature Chem.* **2009**, *1*, 562–567.
- (25) Zhang, C.; Bailey, D. P.; Suslick, K. S. *J. Agric. Food Chem.* **2006**, *54*, 4925–4931.
- (26) Zhang, C.; Suslick, K. S. *J. Agric. Food Chem.* **2007**, *55*, 237–242.
- (27) Lim, S. H.; Musto, C. J.; Park, E.; Zhong, W.; Suslick, K. S. *Org. Lett.* **2008**, *10*, 4405–4408.
- (28) Bajaj, A.; Miranda, O. R.; Kim, I. B.; Phillips, R. L.; Jerry, D. J.; Bunz, U. H. F.; Rotello, V. M. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 10912–10916.
- (29) Buryak, A.; Severin, K. J. *Am. Chem. Soc.* **2005**, *127*, 3700–3701.
- (30) Collins, B. E.; Wright, A. T.; Anslyn, E. V. *Top. Curr. Chem.* **2007**, *277*, 181–218.
- (31) Palacios, M. A.; Nishiyabu, R.; Marquez, M.; Anzenbacher, P. *J. Am. Chem. Soc.* **2007**, *129*, 7538–7544.
- (32) Zimmermann-Dimer, L. M.; Machado, V. G. *Quim. Nova* **2008**, *31*, 2134–2146.
- (33) Gonzalez, D. C.; Savariar, E. N.; Thayumanavan, S. *J. Am. Chem. Soc.* **2009**, *131*, 7708–7716.
- (34) Kong, H.; Zhang, S. C.; Na, N.; Liu, D.; Zhang, X. R. *Analyst* **2009**, *134*, 2441–2446.
- (35) Wu, Y. Y.; Na, N.; Zhang, S.; Wang, X.; Liu, D.; Zhang, X. R. *Anal. Chem.* **2009**, *81*, 961–966.
- (36) Meilgaard, M. C.; Civille, G. V.; Carr, B. T. *Sensory Evaluation Techniques*, 4th ed.; CRC Press: Boca Raton, 2006. Stone, H.; Sidel, J. L. *Sensory Evaluation Practices*, 3rd ed.; Elsevier Academic Press: San Diego, 2004.
- (37) Huang, L. F.; Wu, M. J.; Zhong, K. J.; Sun, X. J.; Liang, Y. Z.; Dai, Y. H.; Huang, K. L.; Guo, F. Q. *Anal. Chim. Acta* **2007**, *588*, 216–223.
- (38) Ribeiro, J. S.; Augusto, F.; Salva, T. J. G.; Thomaziello, R. A.; Ferreira, M. M. C. *Anal. Chim. Acta* **2009**, *634*, 172–179.
- (39) Illy, E. *Sci. Am.* **2002**, *286*, 86–91.

- (40) Nijssen, L. M. *Volatile Compounds in Food: Qualitative and Quantitative Data*; TNO-CIVO Food Analysis Institute: Zeist, Netherlands, 1996.
- (41) Grosch, W. *Nahrung-Food* **1998**, *42*, 344–350.
- (42) Clarke, R. J.; Vitzthum, O. G., Eds. *Coffee: Recent Developments*; Blackwell Science: Oxford, 2001.
- (43) Flament, I. *Coffee Flavor Chemistry*; J. Wiley & Sons: Chichester, 2002.
- (44) Lindinger, C.; Labbe, D.; Pollien, P.; Rytz, A.; Juillerat, M. A.; Yeretzian, C.; Blank, I. *Anal. Chem.* **2008**, *80*, 1574–1581.
- (45) Zellner, B. D.; Dugo, P.; Dugo, G.; Mondello, L. *J. Chromatogr., A* **2008**, *1186*, 123–143.
- (46) Rocha, S.; Maetztl, L.; Barros, A.; Cid, C.; Coimbra, M. A. *J. Sci. Food Agric.* **2004**, *84*, 43–51.
- (47) Aishima, T. *J. Agric. Food Chem.* **1991**, *39*, 752–756.
- (48) Czerny, M.; Grosch, W. *J. Agric. Food Chem.* **2000**, *48*, 868–872.
- (49) Franca, A. S.; Oliveira, L. S.; Oliveira, R. C. S.; Agresti, P. C. M.; Augusti, R. J. *Food Eng.* **2009**, *92*, 345–352.

allowed to run for 30 min at a rate of ~22 mL/min. After which, the array was exposed to the volatile compounds for 2 min.

Digital images of the array before and after 2 min exposure to each coffee sample were acquired on an Epson Perfection V200 Photo flatbed scanner. All the analyses of coffee samples were conducted in quintuplet trials. For each trial, a color change profile was obtained by subtracting the "before" image from the "after" image using Photoshop or a customized package, ChemEye (ChemSensing, Inc.). The center two-thirds of each spot was averaged to avoid subtraction artifacts at the edges of the spots.

The color change profiles were compiled into a database library of 108-dimensional vectors (36 red, green, and blue color values) and represent a unique fingerprint for each specific sample. Standard chemometric analyses, including both principal component analysis (PCA) and hierarchical cluster analysis (HCA),<sup>50–54</sup> were performed on the database library using the multi-variance statistical package (MVSP, Kovach Computing Services) software. For all HCA, minimum variance (i.e., Ward's method) was used for classification.

**Roasting of Coffee Beans.** The laboratory roasting process was conducted using the following minor modification of a standard literature method.<sup>46–48</sup> An aliquot of ~15 g of green coffee beans were spread out over a Petri dish, and the green coffee beans were roasted at varying temperatures ranging from 180 ± 1 °C to 240 ± 1 °C for 15 min (Fisher Scientific Isotemp Oven #825F). In separate experiments, green coffee beans were roasted at 220 °C for lengths of time varying from 1 min to 3 h. After each roasting, the beans were then allowed to cool inside a sealed glass container and then stored in the freezer compartment of a standard kitchen refrigerator.

## RESULTS AND DISCUSSION

**Evaluation and Discrimination of Complex Mixture.** The complexity of analytical problems posed by real-world samples remains a challenge for chemists. The number of separate compounds in foods, beverages, or even whole cells is sufficiently large as to overwhelm even the use of hyphenated techniques. GC/MS<sup>3</sup> has emerged as the most versatile and widely used technology for the detection of complex mixtures (e.g., coffee,<sup>37</sup> perfume,<sup>55</sup> fragrance,<sup>56</sup> petroleum,<sup>3</sup> etc.). While the power of GC/MS is undeniable, even it has limitations for the accurate discrimination among similar complex mixtures, especially if there is a wide dynamic range in the concentrations of important components. GC/MS is only sensitive to the low parts per million concentration range; below that, preconcentration is essential, and the use of solid-phase microextraction (SPME) has been invaluable for chemical analysis of low concentration components in simple vapor mixtures.<sup>38,57,58</sup> Unfortunately, SPME inherently

gives highly uneven preconcentration of different components (e.g., due to differences in the solid-phase absorption of polar vs nonpolar compounds) so that the analysis does not necessarily reflect the distribution of components in the original complex mixture.

Biological sensory systems show unmatched ability to discriminate among complex mixtures, as we each know from personal experience. The alternative of electronic nose technologies is, therefore, a potentially attractive technique for analysis of complex mixtures because they treat the mixture as a single analyte and generate a combined response. While electronic nose techniques generally do not give component by component information, they have the potential to provide a fast and easy method to tell one mixture from another, which has particular appeal for industrial QC/QA applications.<sup>5–16</sup> Past electronic nose technology, however, has not always fulfilled this promise in large part because the chemical specificity (and hence the dimensionality of the sensor data) has been often highly limited.<sup>17,20</sup>

**Overview of Coffee and Its Roasting.** The main species of coffee are *Coffea arabica* and *Coffea canephora* (often also called *Coffea robusta*). Robusta beans yield a product that has substantial body, an earthy aroma, and an elevated caffeine content (2.4–5.8 wt %). Arabica beans yield a product that has an intense aroma that can be reminiscent of flowers, fruit, honey, chocolate, caramel, or toasted bread and never have caffeine content higher than 1.5 wt %. Arabica beans account for two-thirds of the world's total coffee production and are generally considered to produce a higher quality cup of coffee.<sup>39</sup> The chemistry of coffee aroma is highly complex and is still not completely understood. The main families of chemical compounds, which decompose into the volatiles during roasting, are alkaloids (e.g., trigonelline), chlorogenic acids, carbohydrates, free sugars (e.g., sucrose), lipids, and proteins. There are more than 300 volatile compounds identified in unroasted, green coffee alone.<sup>43</sup> During the roasting process, the composition of coffee beans is drastically changed and there are more than 1000 volatile compounds that make up the aroma of roasted coffee.<sup>39–43</sup> A coffee's aroma will vary as a function of changes in soil, microclimate, altitude, types and species of bean used, the roasting process, and the preparation of the coffee. These various conditions affect the concentration and composition of the various aroma volatiles, which include<sup>37–43</sup> carboxylic acids, alcohols, aldehydes, alkanes, alkenes, aromatics, esters, furans, ketones, lactones, oxazoles, phenols, pyridines, pyrazines, pyrroles, thiazoles, and thiophenes.

A variety of processes occur during the roasting of green coffee beans: there is loss of moisture (~8 wt %), loss of solid mass (mostly CO<sub>2</sub>, roughly 6–10 wt %), changes in physical morphology (swelling, puffing, and introduction of significant internal porosity), and there are a wide variety of chemical reactions, including Strecker degradations (i.e., the reaction of amino acids with carbonyls to yield ketones and aldehydes and release of CO<sub>2</sub>), Maillard reactions (i.e., nonenzymatic browning reactions of peptides with sugars to form nitrogen heterocycles), fragmentation of carbohydrates (yielding volatile acids), and caramelization (i.e., saccharide polymerization through dehydration). As coffee beans are heating, water vapor and gases evolve, and around 190 to 210 °C, the beans undergo

(50) Hair, J. F.; Black, B.; Babin, B.; Anderson, R. E.; Tatham, R. L. *Multivariate Data Analysis*, 6th ed.; Prentice Hall: Upper Saddle River, NJ, 2005.

(51) Hasswell, S. *Practical Guide to Chemometrics*; Dekker: New York, 1992.

(52) Scott, S. M.; James, D.; Ali, Z. *Microchim. Acta* **2007**, *156*, 183–207.

(53) Johnson, R. A.; Wichern, D. W. *Applied Multivariate Statistical Analysis*, 6th ed.; Prentice Hall: Upper Saddle River, NJ, 2007.

(54) Ciosek, P.; Wróblewski, W. *Sens. Actuators, B* **2006**, *114*, 85–93.

(55) Mondello, L.; Casilli, A.; Tranchida, P. Q.; Dugo, G.; Dugo, P. *J. Chromatogr., A* **2005**, *1067*, 235–243.

(56) Leijs, H.; Broekmans, J.; van Pelt, L.; Mussinan, C. *J. Agric. Food Chem.* **2005**, *53*, 5487–5491.

(57) Kolb, B.; Ettre, L. S. *Static Headspace-Gas Chromatography: Theory and Practice*, 2nd Ed.; J. Wiley & Sons: Hoboken, NJ, 2006.

(58) Ouyang, G.; Pawliszyn, J. *TrAC Trends Anal. Chem.* **2006**, *25*, 692–703.

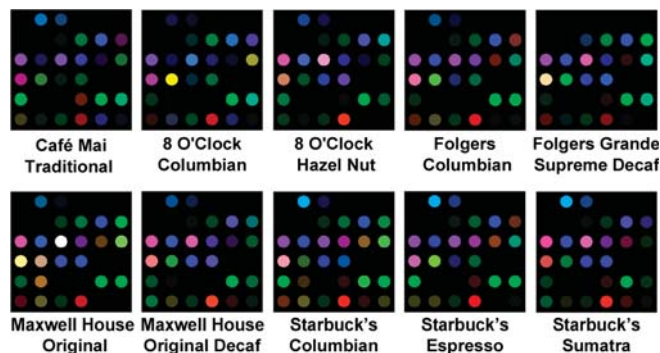


the “First Crack”, a fracturing that is accompanied by a loud sound similar to the popping of popcorn; at this point, the coffee beans have a brown color. Upon further heating to about 220 to 240 °C, a more shallow snapping noise is heard, which is referred to as the “Second Crack” at which point the bean structure and its woody cellulose matrix begin to fracture and the bean color is dark brown to black.

**Design of a Colorimetric Sensor Array.** The design of our colorimetric sensor arrays<sup>17–24</sup> is based on the strong dye–analyte interactions, which is quite different from other electronic nose technologies that generally rely on weak, nonspecific intermolecular interactions, primarily van der Waals and physical adsorption interactions. More specifically, we have chosen chemically responsive dyes in four classes (as illustrated in the Supporting Information, Figure S2): (1) metal ion containing dyes (e.g., metalloporphyrins) that respond to Lewis basicity (i.e., electron pair donation, metal ion ligation), (2) pH indicators that respond to Brønsted acidity/basicity (i.e., proton acidity and hydrogen bonding), (3) dyes with large permanent dipoles (e.g., vapochromic or solvatochromic dyes) that respond to local polarity, and (4) metal salts that respond to redox reactions.

Importantly, our most recent sensing array methodology used in this work is based on nanoporous pigments<sup>23,24</sup> created by the immobilization of chemically responsive dyes in organically modified siloxanes (ormosils<sup>59,60</sup>). Porous sol–gel ormosils provide an excellent matrix for colorants due to high surface area, good stability over a wide range of pH, relative inertness in many environments, and transparency in the UV–visible spectrum. In addition, the physical and chemical properties of the matrix (e.g., hydrophobicity, porosity) can be modified by simply changing the sol–gel constituents. The nanoporous pigments produced by immobilizing colorants in ormosils significantly improves the array’s stability and shelf life.<sup>23,24,59,60</sup> Furthermore, these sol–gel formulations can be modified to make an ink capable of printing onto ordinary polymer flat surfaces.<sup>24</sup> We have also found that the porous matrix serves as a preconcentrator and consequently improves the sensitivity of the sensor.<sup>24</sup>

**Application of a Colorimetric Sensor Array to Coffee Aroma.** While there have been limited attempts to use prior electronic nose technology for analysis of coffee,<sup>61–67</sup> the ability to discriminate among a large number of similar coffees and to monitor the effects of roasting conditions has not been previously reported. In order to explore the ability of our colorimetric sensor arrays to discriminate among highly similar complex mixtures, we have examined the response of the array to a diverse set of



**Figure 1.** Color change profiles after 2 min of array exposure to the saturated vapors from 10 representative commercial coffees. While the overall patterns are all similar, a close examination shows highly reproducible differences between brands. A full digital database is provided in the Supporting Information, Table S2. For display purposes, the color range of these difference maps are expanded from 4 to 8 bits per color (RGB range of 4–19 expanded to 0–255).

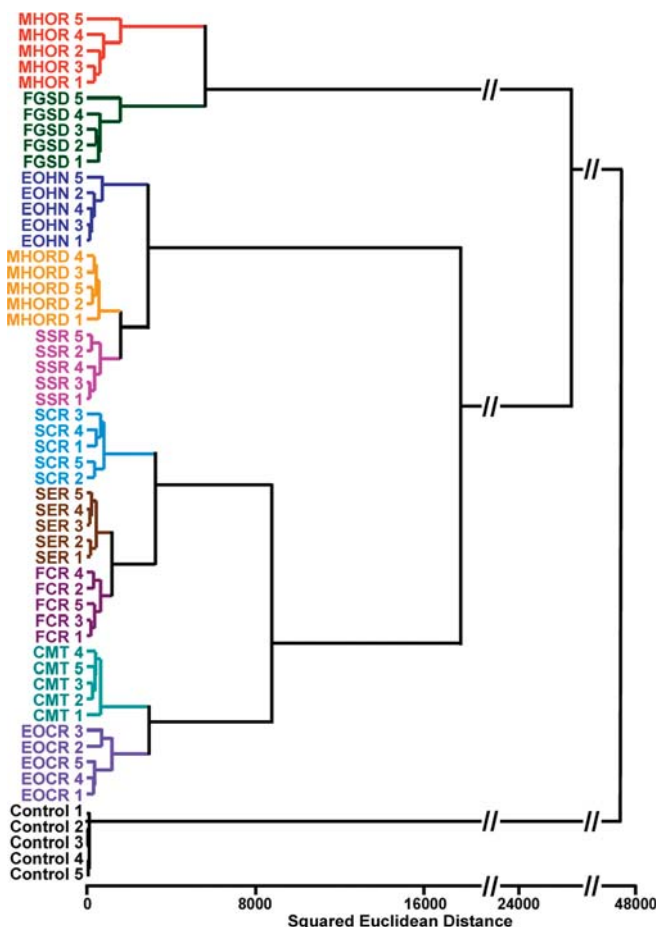
10 commercial coffees. To explore a wide variety of coffee characteristics, decaffeinated coffees, espressos, and various blends were included. In addition, we have used the colorimetric sensor array to examine the effects of roasting of green coffee beans, both as a function of roasting temperature for a fixed time and as a function of time at a fixed temperature.

The printed sensor arrays were digitally imaged with an ordinary flatbed scanner prior to and shortly after exposure to air saturated with the volatiles from ground coffee. A difference map (red minus red, green minus green, blue minus blue) was generated for each analysis, and the resulting color difference profiles (i.e., 108-dimensional vectors made up of the changes in red, green, and blue values of the 36 nanoporous pigments) represent a unique “molecular fingerprint” for each coffee aroma. As shown in Figures 1 and 2, these color difference profiles are able to provide robust discrimination among 10 different brands of commercial coffee. While the overall patterns in the difference maps are all similar (Figure 1), close examination shows highly reproducible differences between brands. Based on the responses of the individual nanoporous pigment components of our colorimetric sensor array, the response of the array to coffee aroma is due at least in part to acidic components (e.g., formic, acetic, propionic acid, etc.).

For quantitative comparison of the difference maps, we can define a 108-dimensional vector (i.e., 36 changes in red, green, and blue values for the after-exposure image compared to the before-exposure image for our 6 × 6 array of nanoporous pigments). Each experimental trial is represented by its 108-dimensional color change profile, and these vectors may be compared by standard chemometric techniques. The complete database is available in the Supporting Information and consists of the 108-dimensional color change profiles for each of the quintuplet trials for each of the 10 coffee samples plus a control.

A standard chemometric analysis, hierarchical cluster analysis (HCA),<sup>50–54</sup> was utilized to analyze this database. Colorimetric sensor array data are highly dispersed across a large number of dimensions, and therefore, their analysis requires a classification algorithm that uses the full dimensionality of the data. HCA is the simplest statistical approach and bases its classification on the Euclidean distance between data points (i.e., color change

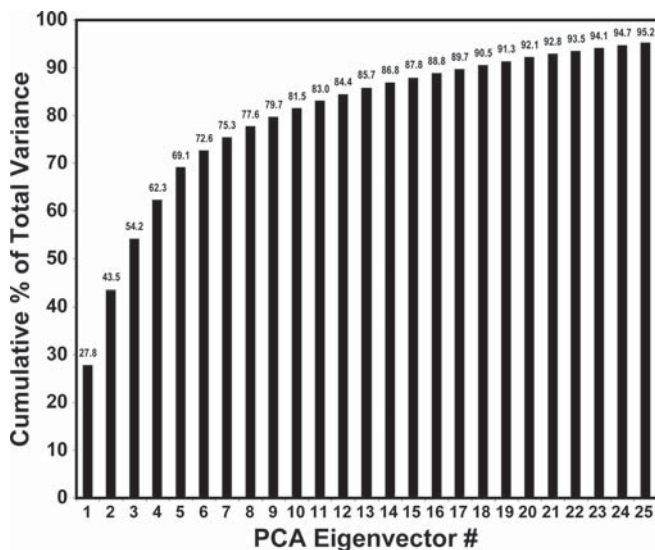
(59) Podbielski, H.; Ulatowska-Jarza, A.; Muller, G.; Eichler, H. J. *Optical Chemical Sensors*; Springer: Erice, Italy, 2006.  
 (60) Dunbar, R. A.; Jordan, J. D.; Bright, F. V. *Anal. Chem.* **1996**, *68*, 604–610.  
 (61) Gardner, J. W.; Shurmer, H. V.; Tan, T. T. *Sens. Actuators, B* **1992**, *6*, 71–75.  
 (62) Pardo, M.; Niederjaufner, G.; Benussi, G.; Comini, E.; Faglia, G.; Sberveglieri, G.; Holmberg, M.; Lundstrom, I. *Sens. Actuators, B* **2000**, *69*, 397–403.  
 (63) Albert, K. J.; Walt, D. R.; Gill, D. S.; Pearce, T. C. *Anal. Chem.* **2001**, *73*, 2501–2508.  
 (64) Pardo, M.; Sberveglieri, G. *IEEE Trans. Instrum. Meas.* **2002**, *51*, 1134–1339.  
 (65) Shilbayeh, N. F.; Iskandarani, M. Z. *Am. J. Appl. Sci.* **2004**, *1*, 129–135.  
 (66) Falasconi, M.; Pardo, M.; Sberveglieri, G.; Ricco, I.; Bresciani, A. *Sens. Actuators, B* **2005**, *110*, 73–80.  
 (67) Wang, X. D.; Ye, M. Y.; Duanmu, C. J. *Sens. Actuators, B* **2009**, *140*, 143–148.



**Figure 2.** Hierarchical cluster analysis (HCA) for 10 commercial coffees and a control. All experiments were run in quintuplet trials; no confusions or errors in classification were observed in 55 trials, as shown. The HCA used minimum variance (i.e., “Ward’s Method”) for clustering. Abbreviations: Maxwell House Original Roast, MHOR; Folgers Grande Supreme Decaf, FGSD; Eight O’Clock Hazel Nut, EOHN; Maxwell House Original Roast Decaf, MHORD; Starbucks Sumatra Roast, SSR; Starbucks Columbian Roast, SCR; Starbucks Espresso Roast, SER; Folgers Columbian Roast, FCR; Café Mai Traditional, CMT; Eight O’Clock Columbian Roast, EOCCR; the number indicates  $n^{\text{th}}$  trial. Control = no coffee present.

vectors) in their full dimensionality. The advantage of HCA compared to other model-dependent statistical analysis (e.g., linear discriminant analysis) is that it makes no assumptions about the classification of results that one is trying to establish. Hierarchical clustering generates a dendrogram that quantitatively compares the Euclidean distances among all the experimental trials, as shown in Figure 2. Remarkably, in quintuplet trials, all 10 coffee samples and a control were accurately identified against one another with no error or misclassifications out of 55 cases. An important lesson from such a clustering analysis is that once a library of array responses is created, the similarity of a new analyte (e.g., a new coffee or other complex mixture) can be quantitatively compared to the existing library entries: e.g., the colorimetric sensor array can tell us what the unknown is “like”.

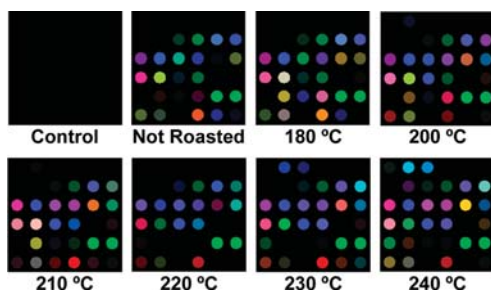
Another standard chemometric technique, principal component analysis (PCA), can be used to probe the dimensionality of sensor array data; PCA creates linear combinations of the array’s responses (i.e., 108 changes in RGB values) so as to maximize the total variance among the data into as few dimensions as



**Figure 3.** Scree plot of the principal components from PCA of 55 trials using 10 coffees and a control. The colorimetric sensor array has an extremely high level of dispersion: for centered, standardized color difference vectors, 18 dimensions are required to define 90% of the total variance and 25 dimensions are required to define 95%.

possible. PCA for most other electronic nose technology is dominated by only two or three independent dimensions: in fact, there is often a single dominant dimension that accounts for >90% of the total discrimination and roughly corresponds to sensor hydrophobicity. This limited dimensionality (or “dispersion”) means that very little of the total diversity of chemical properties is being probed in traditional electronic nose technology, which is the inherent result of relying primarily on van der Waals interactions (e.g., physisorption onto metal oxide surfaces or into polymer films) for molecular recognition. Because of this highly limited dispersion, the data obtained with traditional electronic nose technologies is typically plotted against the two most important PCA dimensions. Such a two-dimensional plot can discriminate among substantially different analytes, but only because the dimensionality of data from most electronic nose technologies is extremely limited. Two- or even three-dimensional PCA plots are not suitable, however, for colorimetric sensor array data because the discrimination among analytes is spread over many dimensions. The advantage of high dimensional data, however, is that it greatly broadens one’s ability to discriminate among closely related analytes or among very similar complex mixtures.

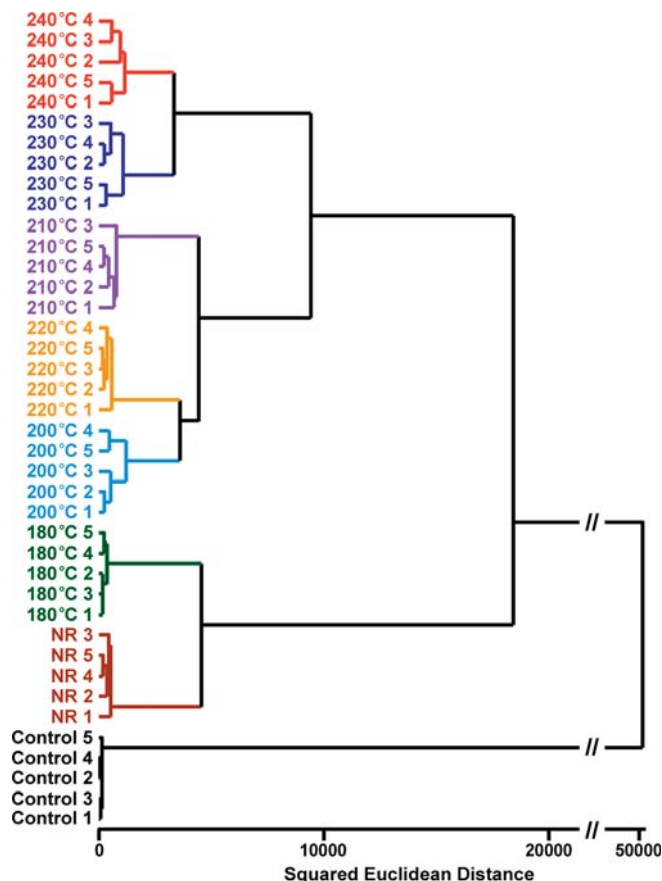
In contrast to other electronic nose technology, the colorimetric sensor array is not limited to van der Waals interactions but rather employs a diverse range of chemical interactions that probes a broad volume of chemical-properties space. Our  $6 \times 6$  array has 108 total possible dimensions (i.e., red, green, and blue color changes for 36 dyes), but there is of course significant redundancy. Nonetheless, the PCA of the colorimetric sensor array data on 10 different roasted coffees plus controls reveals an extremely high level of dispersion. As shown in Figure 3, using 55 trials with centered, standardized color difference vectors, 18 dimensions are required to define 90% of the total variance and 25 dimensions are required to define for 95%. Standardization of the channels (i.e.,  $\Delta\text{RGB}$  values) implies that all of the dyes are



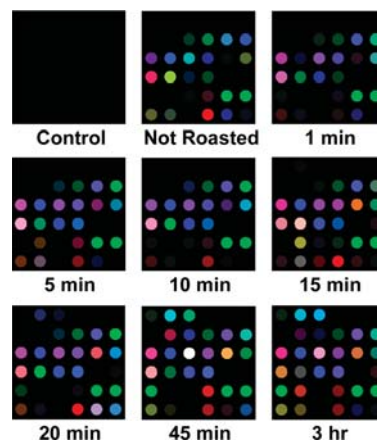
**Figure 4.** Color change profiles after 2 min of array exposure to Columbian Huila green coffee beans roasted for 15 min at 180, 200, 210, 220, 230, and 240 °C. A full digital database is provided in the Supporting Information, Table S2. For display purposes, the color range of these difference maps are expanded from 4 to 8 bits per color (RGB range of 4–19 expanded to 0–255).

of equal importance in contributing to the analysis. Alternatively, without standardization, one implies that the dyes with the largest inherent color changes are the most discriminatory (which is generally a less true assumption); without standardization, the PCA still requires 7 dimensions for 90% of total variance and 13 dimensions for 95%. Regardless, the colorimetric sensor arrays demonstrate extremely high dimensionality in the analysis of coffee aroma, and it is that high dimensionality that permits facile discrimination among such similar complex mixtures as the 10 brands of coffee.

**Colorimetric Sensor Array Response vs Coffee Bean Roasting.** The colorimetric sensor array is also able to identify the effects of various roasting conditions. Most of the aroma of coffee is developed during roasting, and the composition of the coffee aroma will vary with the processing conditions. Moisture will be released from the interior of the bean, as well as carbon dioxide causing the beans to expand rapidly. The external color of the beans will change from light to dark brown as roasting temperature or time increases. Numerous chemical components will decompose or undergo complex chemical reactions. To probe the ability of the colorimetric sensor array to discriminate among complex mixtures directly relevant to quality control processes, we examined the array response to coffee beans before and after being roasted at six different temperatures (180, 200, 210, 220, 230, and 240 °C) for 15 min each, as seen in Figure 4. The colorimetric sensor array was successfully able to identify and discriminate among these seven complex analytes; in quintuplicate trials, there were no errors or misclassifications in the HCA of the six various roasting temperatures and the unroasted (green) coffee beans, as shown in Figure 5. One may conclude that these sensor arrays have a resolution of roasting temperature that is better than 10 °C for the roasting of coffee beans. The clustering within the HCA illustrates the trend seen in the concentration of the volatiles. There is relatively little difference between unroasted beans and 15 min of roasting at 180 °C. Most of the roasting process occurs between 200 and 240 °C, which contains the so-called first and second “crack” of the roasting process. The HCA clustering shows that the samples roasted over that temperature range are easily distinguishable but still relatively similar. After the second crack (i.e., above 220 °C), further decomposition occurs, and larger differences in the composition of the volatiles are observed.



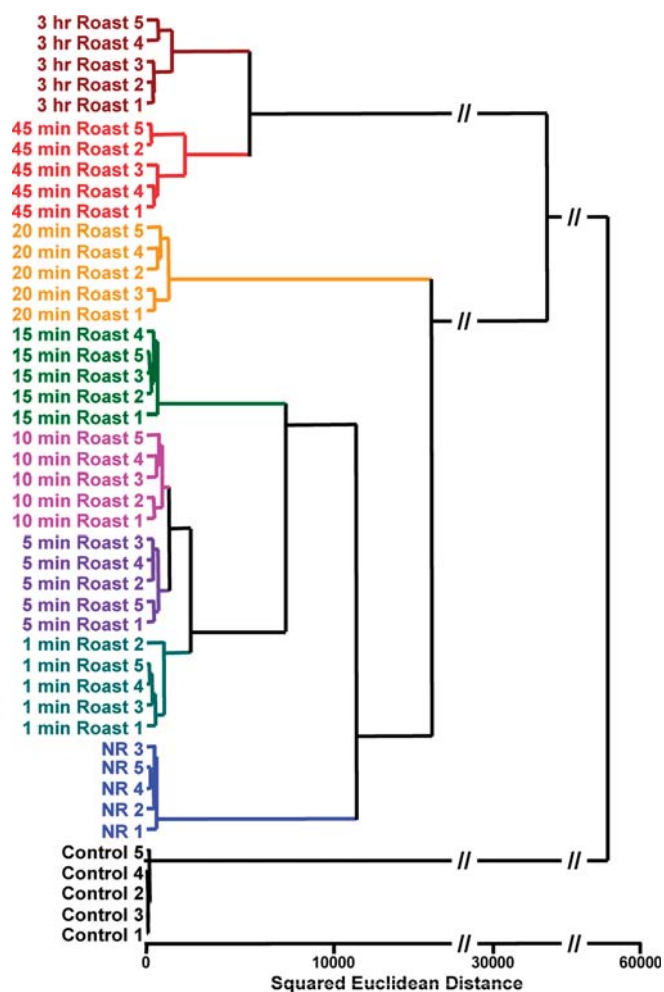
**Figure 5.** Hierarchical cluster analysis (HCA) for Columbian Huila green coffee beans roasted for 15 min at temperatures ranging from 180 to 240 °C. All experiments were run in quintuplet trials; no confusions or errors in classification were observed in 40 trials, as shown. NR = green coffee, not roasted; the number indicates  $n^{\text{th}}$  trial. Control = no coffee present.



**Figure 6.** Color change profiles after 2 min of array exposure to Columbian Huila green coffee beans roasted at 220 °C for times ranging from 1 min to 3 h. A full digital database is provided in the Supporting Information, Table S2. For display purposes, the color range of these difference maps are expanded from 4 to 8 bits per color (RGB range of 4–19 expanded to 0–255).

In addition to various roasting temperatures for a constant time, the effects of various roasting times at a constant temperature (220 °C) were also investigated, as shown in Figure 6. As expected, the response of the array to the coffee aroma increases with increasing roasting time. The increase in response over time is





**Figure 7.** Hierarchical cluster analysis (HCA) for Colombian Huila green coffee beans roasted at 220 °C for times ranging from 1 min to 3 h. All experiments were run in quintuplet trials; no confusions or errors in classification were observed in 45 trials, as shown. NR = green coffee, not roasted; the number indicates  $n^{\text{th}}$  trial. Control = no coffee present.

due in large part to the decomposition of volatiles to carboxylic acids. After very long roasting times (>3 h), the array response decreases, presumably due to increasing loss of volatiles and to further decomposition of initially formed volatiles. For quintuplicate trials, there were again no errors or misclassifications in the HCA for discrimination of coffee aroma at different roasting times among 45 trials (Figure 7). One may conclude that these sensor arrays have a resolution of roasting time that is better than 5 min. In addition, the clustering within the HCA showed an interesting finer trend: the unroasted beans were separate but closest to the lightly roasted beans (from 1 to 10 min), which were separate from the medium roasting (15 min) and well separated from the heavily roasted beans (45 min to 3 h).

## CONCLUSION

In summary, we have successfully created a disposable colorimetric sensor array of nanoporous pigments that is capable of discriminating among extremely similar complex mixtures, in this case specifically, different coffee aromas. This sensor array technique does not, of course, give information about individual

components and so this approach is complementary to, rather than competitive with, more traditional chemical analysis. If one requires specific knowledge about the concentrations of specific analytes within a complex mixture, then no sensor array technique will be appropriate.

For other analytical goals with complex mixtures, however, sensor arrays are a complementary approach. The colorimetric sensor array has proved extremely effective for comparisons against a standard, for discrimination of subtle differences among similar mixtures and for monitoring changes in the mixture as a function of time or conditions. Unique “odor fingerprints” are easily obtainable for any coffee sample from the color changes of an array of 36 chemically responsive colorants. Hierarchical cluster analysis (which is free of any predetermined statistical model) demonstrated flawless discrimination among 10 different brands of commercial coffees. In addition, we were also successful in facile discrimination of coffee aromas as a function of the roasting of green coffee beans under varying roasting times and under varying roasting temperatures. Principal component analysis reveals that the colorimetric sensor array has an extremely high dimensionality, which contributes to the ability of the array to discriminate among highly similar complex mixtures. This high dimensionality provides the promise of future applications to the correlation of objective colorimetric sensor array responses to human sensory evaluation: it is likely that appropriate linear combinations of the 108 channels of our sensor array data can be found to match sensory descriptors of flavor or aroma, which may prove useful for the objective evaluation of issues of quality and consumer preference.

In addition, we have recently developed a functional prototype hand-held device that makes use of an inexpensive white LED (light-emitting diode) and an ordinary CMOS (complementary metal-oxide semiconductor) camera, as shown in the Supporting Information, Figure S3. Combined with a low dead-volume cartridge (Supporting Information, Figure S4), this hand-held device can provide a rapid and highly sensitive method for portable monitoring.

## ACKNOWLEDGMENT

B.A.S. and L.F. contributed equally to this work. Substantial portions of this work were done in the University Laboratory High School Science Department under the supervision of Mr. David R. Bergandine. In addition, we thank Dr. Sung H. Lim and Jonathan W. Kemling for their assistance in the formulation of the arrays. Finally, we acknowledge the early related efforts in our group by Dr. Jennifer B. Ponder and Professor Shirley Nakagaki. This work was supported through the NIH Genes, Environment and Health Initiative through Award U01ES016011.

## SUPPORTING INFORMATION AVAILABLE

Additional information as noted in text. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Received for review December 11, 2009. Accepted January 26, 2010.

AC902823W