Rapid Identification of Bacteria with a Disposable Colorimetric Sensing Array

James R. Carey,* Kenneth S. Suslick,*‡ Keren I. Hulkower,‡ James A. Imlay,§ Karin R. C. Imlay,§ Crystal K. Ingison,* Jennifer B. Ponder,* Avijit Sen,‡ and Aaron E. Wittrig‡

¹Department of Applied Chemistry, National University of Kaohsiung, 700 Kaohsiung University Road, Kaoshiung 811, Taiwan
²Department of Chemistry, University of Illinois at Urbana—Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801, United States
³Department of Microbiology, University of Illinois at Urbana—Champaign, 601 South Goodwin Avenue, Urbana, Illinois 61801, United States

Supporting Information

ABSTRACT: Rapid identification of both species and even specific strains of human pathogenic bacteria grown on standard agar has been achieved from the volatiles they produce using a disposable colorimetric sensor array in a Petri dish imaged with an inexpensive scanner. All 10 strains of bacteria tested, including Enterococcus faecalis and Staphylococcus aureus and their antibiotic-resistant forms, were identified with 98.8% accuracy within 10 h, a clinically important time frame. Furthermore, the colorimetric sensor arrays also proved useful as a simple research tool for the study of bacterial metabolism and as an easy method for the optimization of bacterial production of fine chemicals or other fermentation processes.

INTRODUCTION

The detection and identification of bacteria are pressing problems in both medicine and industry.¹² A patient may present to the physician symptoms consistent with a bacterial infection, but the physician may be unable to address the infection with the appropriate antibiotic until the identity or antibiotic susceptibility of the bacteria has been determined; as a consequence, sepsis remains one of the leading causes of death even among first-world nations.¹ In industry, many products must be screened after manufacture for bacterial contamination before they may be released, and as a consequence, regulation of the food industry must be particularly stringent.² Existing methods for identification of pathogenic bacteria are severely limited by the necessity of long culturing times, the need for highly trained laboratory personnel, and the requirement of expensive and high-maintenance equipment.³⁴⁻⁵

Bacteria stink: that is, they produce volatile organic compounds (VOCs) to which the mammalian olfactory system is highly responsive. Consequently, an experienced microbiologist can readily identify many bacteria by smell. We have previously developed a simple colorimetric sensing array⁶ for the detection of VOCs⁷ and discrimination among complex mixtures;⁸ we report here that these arrays can readily identify many bacteria by the production of their volatile metabolites. We find that a disposable sensor array placed in a standard Petri dish and imaged with an ordinary flatbed scanner is capable of identifying human pathogenic bacterial strains in less than 10 hours, which represents a substantial improvement over current clinical techniques in terms of speed, ease of use, and cost.

Traditional techniques for both manual and automated bacterial identification are based on the biochemical characteristics of each microorganism as defined by yes/no answers to a series of biochemical tests. In essence, these tests differentiate and identify bacteria by identifying specific bacterial metabolites as a function of available nutrients. A much less explored alternative, however, would be to identify bacteria by monitoring their metabolic output from growth on a single complex nutrient mixture. This strategy arises from the well-established knowledge that different species of bacteria consuming the same nutrients produce different metabolites. Different species, and even different strains of the same species, emit distinct profiles of enzymatic reaction products such as amines, sulfides, and fatty acids.⁹⁻¹¹ Given the high sensitivity of colorimetric sensors to numerous VOCs, we expected that the various volatile metabolites produced by different bacterial strains might provide identifying fingerprints in the response of the sensor arrays.

Toward that end, we have used a cross-responsive colorimetric sensor array to monitor the complex composite of volatile compounds produced by 10 bacterial strains grown in replicate
on solid media in closed Petri dishes. The sensor arrays consist of 36 chemically responsive dyes, including metalloporphyrins, pH indicators, metal salts, and solvatochromic dyes, that change color when exposed to a broad range of volatile analytes. Given the wide range of VOCs produced by bacteria,9–11 the chemical diversity of the sensor elements present in our array is critical to its capability to respond to broad classes of individual analytes and its ability to distinguish among complex mixtures.

EXPERIMENTAL SECTION

Bacteria. Bacterial strains were purchased from the American Type Culture Collection (Manassas, VA) and were cultured following manufacturer protocols before use. Bacteria were grown in tryptic soy broth (TSB) and plated during log phase growth on tryptic soy agar containing 5% sheep blood (Becton Dickinson, Franklin Lakes, NJ). Bacterial suspensions were prepared by inoculating 5 mL of TSB with a single colony and allowing it to grow overnight. A subculture then was prepared in 5 mL of fresh TSB and shaken at 37 °C for 3 h to achieve the desired inoculum density of 0.5–5 on the McFarland turbidity scale. This equates to \( \sim 3 \times 10^8 \)–1.5 \( \times 10^9 \) CFU/mL. Next, 10–250 \( \mu \)L of the subculture was spread onto a 60 mm TSA/sheep blood Petri dish. A control (10–250 \( \mu \)L of TSB broth without the bacterial inoculum) was conducted in parallel for each experiment.

Colorimetric Sensor Array. The disposable colorimetric sensor array and its construction have been previously described,1 and the specific dyes used for this study are listed in the Supporting Information, Table S1. An array was mounted in an inert platform that was inserted into the lid of the Petri dish. The Petri dish was closed and inverted onto an ordinary flatbed scanner (as shown in Figure 1) housed in an incubator at 35–37 °C; the printed side of the array faced toward the scanner. Data were collected using an Epson Perfection 4390 scanner every 30 min. Color difference maps were generated by averaging the RGB color values for the center of each spot and subtracting from them the RGB averages for the baseline image: \( \Delta R, \Delta G, \Delta B \), i.e., red value after exposure minus red value at 90 min, etc. Substantial sensor array color changes arising from the media occur upon exposure as the media equilibrates to the incubator temperature of 37 °C, so the baseline image was taken at 90 min, which was sufficient for complete equilibration but before any significant bacterial growth had occurred. A complete database is available in the Supporting Information.

RESULTS AND DISCUSSION

Current Clinical Bacterial Detection. Bloodstream infections by bacteria (i.e., bacteremia) are among the most serious medical problems because of their significant morbidity and mortality.1 In the United States, an estimated 750,000 patients annually develop bloodstream infections, with associated mortality rates ranging from 14 to 50%. Sepsis and septic shock are the tenth leading cause of death in the U.S. and the third most common cause of death in Germany; furthermore, sepsis and related complications represent a significant economic burden, with an estimated annual cost beyond $17 billion in the U.S. alone.2,3

There remain major unmet needs to shorten and improve current clinical and laboratory methods for the detection and identification of bloodstream infections, and molecular diagnostic methods (ranging from mass spectral analysis to PCR to peptide–nucleic acid fluorescent probes, etc.) have yet to have a major impact on such diagnoses.4,5 Blood cultures remain the “gold standard” for the diagnosis of bacteremia.1,2 Current standard clinical procedures start with culturing (generally in liquid growth media), for example of blood samples, which generally takes 24–48 h to confirm the presence of bacteria but can take much longer for slow-growing bacteria (e.g., tuberculosis generally requires more than a week). Microorganisms taken from positive blood cultures are then subjected to rapid gram staining (∼10 min but then must be subcultured and analyzed by culture-based systems such as the Analytical Profile Index (API) test or antibiotic susceptibility tests (e.g., bioMérieux’s automated VITEK system), which require another 18–48 h.3 The reader is referred to an excellent recent review of clinical procedures by Riedel and Carroll.14 Nucleic-acid-based identification,4,5 such as fluorescent in situ hybridization (FISH), can be performed directly from positive blood cultures that require only 6 h, although identification by culture-based systems is needed for confirmation.4,5

Bacterial Detection by Electronic Nose. There have been various studies using “electronic nose” technologies12,13 (i.e., electronic sensor arrays made from conductive polymers, metal oxides, etc.) for headspace analysis in attempts to identify bacterial strains.14–16 Most such studies use a single time sampling of the headspace over a mature bacterial culture.15,16 The temporal profile of the gases produced in a closed culture environment, however, provides valuable additional information for bacterial identification. In addition, the VOCs produced by bacteria in a closed environment approximate an integral of growth and may therefore improve the rapidity of analysis. In those few studies in which the headspace was sampled at multiple time points,16 bacterial discrimination was generally still quite limited. In part, this is due to the inherent limitations in low dimensionality of prior array technology,2,3,5 which relies primarily on weak and nonselective analyte–sensor interactions: in general, prior electronic nose technology needs only two dimensions, or at most three, to capture >95% of the total dataset variance. As such, previous bacterial identification efforts with traditional electronic nose technology have typically required complex pattern recognition algorithms in order to achieve modest success, even when attempting to classify small numbers of bacterial species.11–16

Colorimetric Sensor Arrays. The colorimetric sensor arrays are disposable, one-time-use sensors that are simply placed in the headspace of bacterial cultures and imaged with an ordinary flatbed scanner at multiple time points during bacteria growth (Figure 1). The responses of the chemically responsive dyes used in the array are in general reversible. As shown below, the array
can therefore monitor changes in the volatile metabolites evolved by each bacterial strain over time and thereby differentiate one bacterial species from another.

Representative data arising from exposure of the sensor array to Escherichia coli are shown in Figure 2, as both color difference maps and a time response profile that show the change in spot colors as a function of bacterial growth time. Several notable features are readily apparent in Figure 2. First, there is a time lag (~150 min in this case) before appreciable signal is observed above baseline. As will be shown below, the duration of this lag is species and strain dependent. Once signal is observed, it evolves through several stages; different volatile analytes appear to be generated at different points during growth. Note, for example, the steeply negative color changes occurring between 300 and 400 min. In addition, several color channels pass through a maximum or minimum and then reverse direction, returning to their baseline values or crossing the baseline to change sign. This too implies changes in the nature of volatiles emitted over time.

Figure 2. Color difference maps and time response profile resulting from colorimetric sensor array exposure to a growing culture of E. coli, American Type Culture Collection (ATCC) No. 25922. (a) The color difference maps (i.e., $\Delta R$, $\Delta G$, $\Delta B$) at select times were generated by subtracting the average RGB of each spot from a baseline image (taken at 90 min). For the purpose of effective visualization only, the color range shown in the color range maps is expanded from RGB values of 0–31 to 0–255. (b) The color change values versus time plotted for all color channels ($\Delta R$, $\Delta G$, and $\Delta B$ values for each spot, i.e., 108 color channels) at each time point.

Figure 3. Color difference maps for 10 different bacterial strains resulting from colorimetric sensor array exposure to Petri dish growing cultures after 480 min. The color range shown is the same as in Figure 2a.

Figure 4. Time response profiles for 10 different bacterial strains (name and ATCC number given in each panel). For each bacterial strain, the color change values versus time are plotted for all color channels at each time point.

Time response profiles were collected in multiple replicates for nine additional bacterial strains (6–24 replicates per strain, cf. Supporting Information Table S2). Representative color difference maps and time response profiles for each strain are shown in Figures 3 and 4, respectively. All quantitative analyses were based on the full digital data, which are provided in Supporting Information Table S3.

Array Response to Bacterial Growth. As illustrated in Figures 3 and 4, each strain of bacterium has its own individual color change fingerprint. Even by simple visual examination of the color difference map (Figure 3) or of the time response profiles (Figure 4), all 10 of the bacterial strains are readily differentiated; quantitative classification (discussed below) is
highly accurate using standard statistical methods. This includes the differentiation of *Staphylococcus aureus* from a methicillin-resistant *S. aureus* (MRSA), as well as *Enterococcus faecalis* from a vancomycin-resistant *E. faecalis* (VRE). There are, moreover, genera- and species-based resemblances. This is most apparent in the two strains of *E. coli* and *E. faecalis*, respectively, which show the greatest interpair similarity in the time response profiles (Figure 4).

There are at least three possible closely related contributions to the changes and reversal of color features in the array, each of which reflect the relationship between the chemical input and output of bacteria. The dominant contributor is likely to be conventional diauxie, in which the bacteria consume one nutrient first and then switch to alternate nutrients when the first is exhausted:17 as the input nutrients change, so too would the evolved volatiles. Consistent with this, the response profiles do depend strongly on the type of solid media (such as TSA, TSA/sheep blood or Luria–Bertani) used as a nutrient source. A second contributor to the temporal changes in the array response (which is generally reversible) may be the bacterial consumption of a previously excreted product. For example, many bacteria ferment rich nutrients (e.g., glucose) and excrete acetate; acetate can then, upon the exhaustion of better carbon sources, be consumed and oxidized to carbon dioxide.18 A third possibility is a transition from aerobic to anaerobic growth, itself a form of diauxie.19 Chemical analysis of the complex mixture of volatiles produced during bacterial growth presents a significant challenge for conventional component-by-component analyses (e.g., GC-MS).9–11 Fortunately, one need not know the identities of the individual volatile products for the identification of the bacteria using a sensor array approach, which gives a composite response to the complex mixture of bacterial metabolites.

The complex structures observed in the bacterial response profiles underscore the importance of incorporating broad chemical diversity in a single sensor array. These structures also emphasize the value in having a reversible, real-time sensor continuously positioned in the headspace of growing bacterial cultures. The additional advantage of having a disposable, inexpensive sensor array allows for the facile monitoring of bacterial cultures individually and continuously.

Quantitative Classification. Classification of specific bacterial strains based on the colorimetric sensor array data was successfully accomplished using a variety of independent techniques. For quantitative analysis, the color change values for each experiment were analyzed as “time-stacked” vectors in order to capture the temporal behavior of the response profiles. The entire time-stacked vector, \( v \), from a single experiment is given by eq 1, using images collected every 30 min for 10 h, with the 90 min image as the baseline for evolution of differences in red, green, and blue values for all 36 spots:

\[
v = \Delta R_{150 \text{ min}}, \Delta G_{120 \text{ min}}, \Delta B_{150 \text{ min}}, \Delta R_{120 \text{ min}}, \Delta G_{120 \text{ min}}, \Delta B_{120 \text{ min}}, \ldots \]

\[
\Delta R_{36 \text{ min}}, \Delta G_{36 \text{ min}}, \Delta B_{36 \text{ min}}, \Delta R_{36 \text{ min}}, \Delta G_{36 \text{ min}}, \Delta B_{36 \text{ min}}, \ldots \]

\[
\Delta R_{1 \text{ min}}, \Delta G_{1 \text{ min}}, \Delta B_{1 \text{ min}}, \ldots \]

\[
\Delta R_{600 \text{ min}}, \Delta G_{600 \text{ min}}, \Delta B_{600 \text{ min}}, \Delta R_{600 \text{ min}}, \Delta G_{600 \text{ min}}, \Delta B_{600 \text{ min}}, \ldots \]

(1)

Surprisingly, accurate identification was possible even using simple Euclidean distances, which compares only the overall total response of the array. Average time-stacked vectors were calculated for each strain of bacteria, and the Euclidean distance between each of the 164 individual experimental vectors and each of the average vectors was determined. Each input experimental vector was classified as the strain whose average vector was nearest. Using the time stack data from 90 to 600 min post inoculation, the correct classification was achieved for 162 of the 164 experiments (i.e., 98.8% accuracy); the only two misclassifications were between *E. faecalis* and its vancomycin resistant mutant (VRE). For the same experiments, the classification accuracy was 95% at 420 min culturing time and 85% at 300 min.

Beyond using just the Euclidean distance of the array response, there is much greater information available in the variance of the specific spots of the array. The ability of the colorimetric sensor array to discriminate among different bacteria is due, in part, to the high dimensionality of the data. Principal component analysis (PCA) uses the variance in the array response to evaluate the relative contributions of independent dimensions and generates optimized linear combinations of the original 108 dimensions so as to maximize the amount of variance in as few dimensions as possible (46–49). Using standard PCA, the data from 164 experiments were analyzed. As we have observed with colorimetric sensor arrays applied to many other systems,5–8 the data have exceptionally high dispersion, requiring 7 and 23 dimensions to capture 90% and 95% of the total variance, respectively (SI Figure S1 provides the PCA score plot). The first three dimensions in PCA space account for only 79% of the total variance. Nonetheless, a three-dimensional PCA score plot (Figure 5) shows very good clustering of 164 experimental trials on the 10 bacteria.

The extremely high dispersion of our colorimetric sensor array data reflects the wide range of chemical-property space being probed by the choice of 36 chemically responsive dyes. Consequently, chemically diverse mixtures of volatiles produced by bacteria are easily recognizable, and even closely related bacteria can be distinguished. In contrast, data from most prior electronic nose technologies are dominated by only two or three independent

![Figure 5. PCA score plot using the three most important principal components based on all 164 trials of 10 bacterial strains and controls. The resolution between bacterial classes is in fact much better than can be shown by any three-dimensional PCA plot because the first three principal components account for only 79% of the total variance. Color code: black, *S. aureus* red, MRSA; dark green, *S. epidermidis* purple, *S. sciuri* orange, *P. aeruginosa* white, *E. faecium* light blue, *E. faecalis* gray, *E. faecalis* VRE; yellow, *E. coli* 25922; light green, *E. coli* S3502; dark blue, control.](image-url)
dimensions (one of which, analyte hydrophobicity, generally accounts for >90% of total variance); this is the inherent result of relying on van der Waals and other weak interactions for molecular recognition.

Given the high dimensionality of data from the colorimetric sensor array, the usual two- or three-dimensional PCA plots (e.g., Figure 5) cannot adequately represent the discrimination among experimental trials. Instead, we prefer the use of another quite standard chemometric approach, hierarchical cluster analysis (HCA), which is based on the grouping of the analyte vectors according to their spatial distances in their full vector space. HCA has the advantages of being model-free (unlike, for example, linear discriminant analysis or neural nets) and of according to their spatial distances in their full vector space. As shown in Figure 6, HCA generates dendrograms based on clustering of the array using the full dimensionality of the data. The changes in the array response as a function of time for each bacterial strain imply that each strain of bacterium produces different volatile metabolic products at different rates. This suggests in turn that such arrays may also prove useful as an inexpensive research tool for the study of bacterial metabolism.

The biomedical applications of this colorimetric sensing array technology are also potentially significant. The sensor arrays successfully identified all 10 strains of bacteria tested, including E. faecalis and S. aureus together with their antibiotic resistant forms, by monitoring gases evolved during a 10 h growth of bacterial cultures. This time frame is clinically important: because we are able to quickly monitor growth rates (even after just 3 h as shown in Figure 4), parallel monitoring of bacterial growth in various antibiotic-doped media should provide physicians with valuable and timely information to guide treatment. Because the technology developed here is inexpensive and builds easily on conventional culturing, it may find widespread application in less economically developed regions.

The initial studies reported herein are being used to guide development of arrays with greater sensitivity and classification capabilities for bacteria. These arrays are currently being applied to blood culture systems using liquid growth media with very low initial inoculum concentrations. In addition, this colorimetric sensor array has also proven highly effective in point-of-care diagnosis of bacterial sinusitis and in lung cancer screening, and the possibility of other applications of colorimetric sensor arrays for medical diagnosis by breath are likely to emerge.

Figure 6. HCA dendrogram of 164 trials of 10 human pathogenic bacteria and control, using minimum variance (Ward’s method). The dendrogram is based on the 23 dimensions from PCA that capture 95% of the total variance. The ATCC number is given below the name of each bacterial strain.