Comprehensive volatile metabolic fingerprinting of bacterial and fungal pathogen groups

Christiaan A Rees¹, Alison Burklund², Pierre-Hugues Stefanuto¹, Joseph D Schwartzman¹,³ and Jane E Hill¹,²

¹ Geisel School of Medicine at Dartmouth, Hanover, NH 03755, United States of America
² Thayer School of Engineering at Dartmouth, Hanover, NH 03755, United States of America
³ Dartmouth-Hitchcock Medical Center, Lebanon, NH 03756, United States of America

E-mail: jane.e.hill@dartmouth.edu

Abstract

The identification of pathogen-specific volatile metabolic ‘fingerprints’ could lead to the rapid identification of disease-causing organisms either directly from ex vivo patient bio-specimens or from in vitro cultures. In the present study, we have evaluated the volatile metabolites produced by 100 clinical isolates belonging to ten distinct pathogen groups that, in aggregate, account for 90% of bloodstream infections, 90% of urinary tract infections, and 80% of infections encountered in the intensive care unit setting. Headspace volatile metabolites produced in vitro were concentrated using headspace solid-phase microextraction and analyzed via two-dimensional gas chromatography time-of-flight mass spectrometry (HS-SPME-GC×GC-TOFMS). A total of 811 volatile metabolites were detected across all samples, of which 203 were: (1) detected in 9 or 10 (of 10) isolates belonging to one or more pathogen groups, and (2) significantly more abundant in cultures relative to sterile media. Network analysis revealed a distinct metabolic fingerprint associated with each pathogen group, and analysis via Random Forest using leave-one-out cross-validation resulted in a 95% accuracy for the differentiation between groups. The present findings support the results of prior studies that have reported on the differential production of volatile metabolites across pathogenic bacteria and fungi, and provide additional insight through the inclusion of pathogen groups that have seldom been studied previously, including Acinetobacter spp., coagulase-negative Staphylococcus, and Proteus mirabilis, as well as the utilization of HS-SPME-GC×GC-TOFMS for improved sensitivity and resolution relative to traditional gas chromatography-based techniques.

Introduction

Bacteria and fungi produce a wide array of metabolic products and byproducts, a subset of which are small molecules capable of existing in the gas phase at ambient temperature (the ‘volatile metabolome’) [1, 2]. A subset of these volatile metabolites possess specific functions, including cell-to-cell signaling, although many arise simply as byproducts of catabolic processes [1]. Due to genetic differences across species, including differences in the genes responsible for the production of these metabolites, it has been hypothesized that different microorganisms produce distinct ‘volatile metabolic fingerprints’. Analysis of these volatile metabolic fingerprints has garnered attention in a range of disciplines, including flavor chemistry (e.g., the aroma profiles of yeasts used for beer making [3]), public safety (e.g., the detection of biothreat agents from culture [4]), and healthcare (e.g., diagnosis of respiratory infections using exhaled breath [5, 6]). Indeed, microbially-derived volatile metabolites, as well as those associated with the host–pathogen interaction, have been proposed as potential diagnostic biomarkers in a variety of clinical settings, including infections of the bloodstream, respiratory tract, and urinary tract [2, 7].
At present, no standardized experimental conditions have been defined for the characterization of volatile metabolites produced by bacteria or fungi. The approximately 160 prior studies in this area have differed from one another in key areas, including: (1) growth parameters (e.g., culture media, incubation time, and oxygen availability), (2) methods for the concentration of volatile metabolites (e.g., solid-phase microextraction, and sorbent traps), and (3) instrumentation employed for metabolite analysis (e.g., ‘electronic noses,’ colorimetric sensors, and gas chromatography-mass spectrometry (GC-MS)). Because of the variability between studies, it is challenging to infer differences between the volatile metabolic profiles of different pathogens unless they have been directly compared in a single experiment. Furthermore, more than half of all prior studies have considered five or fewer species, and approximately half have reported on the volatile metabolites produced by only a single reference strain or clinical isolate for each organism studied. In order to comprehensively assess the volatile metabolic fingerprints of important bacterial and fungal pathogen groups, one must consider both a wide range of groups as well as a diverse collection of isolates in each group.

With the inclusion of 100 isolates across ten pathogen groups, and the use comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry (GC×GC-TOFMS), the present study represents one of the most extensive analyses to-date of the volatile metabolic profiles of pathogenic bacteria and fungi. The pathogen groups included in this study are estimated to account for approximately 90% of bloodstream infections [8, 9], 90% of urinary tract infections [10], and 80% of all infections encountered in the intensive care unit setting [11]. The analytical approach employed in the present study has demonstrated utility in the identification of novel microbially-derived metabolites in vitro [12–15], and has additionally provided insight into underlying microbial processes [16, 17]. As a subset of the volatile metabolites previously detected under in vitro conditions have been translated to ex vivo patient specimens such as breath [5, 18, 19], analysis of in vitro cultures using GC×GC-TOFMS could represent a powerful approach for the identification of novel volatile biomarkers with potential diagnostic utility.

Materials and methods

Isolates collection

The study was approved by the Committee for the Protection of Human Subjects at Dartmouth College (Study#: 28773). Ten pathogen groups (n = 10 isolates per group) were included in this study: Acinetobacter spp., Candida spp., coagulase-negative Staphylococcus, Enterobacter spp., Enterococcus spp., Escherichia coli, Klebsiella spp., Pseudomonas aeruginosa, Proteus mirabilis, and Staphylococcus aureus. Species-level identifications were also available for isolates belonging to Candida spp. (Candida albicans (8) and Candida parapsilosis (2)), Enterobacter spp. (Enterobacter cloacae (7) and Enterobacter aerogenes (3)), Enterococcus spp. (Enterococcus faecalis (6) and Enterococcus faecium (4)), and Klebsiella spp. (Klebsiella pneumoniae (8) and Klebsiella oxytoca (2)), while none were provided for isolates belonging to either Acinetobacter spp. or coagulase-negative Staphylococcus. This isolates collection also encompassed three commonly encountered antimicrobial resistance patterns, namely methicillin-resistant S. aureus (MRSA, five S. aureus isolates), vancomycin-resistant Enterococcus (VRE, three E. faecalis and one E. faecium), and extended-spectrum beta-lactamase producing (ESBL-producing) Klebsiella spp. (two K. pneumoniae and one K. oxytoca). All isolates were obtained from the clinical microbiology laboratory at Dartmouth-Hitchcock Medical Center (Lebanon, NH, United States). Isolates had previously been cultured from either blood or urine. Genus-level identification was confirmed via 16S rDNA sequencing for bacterial pathogens [20, 21], and 5.8S rDNA for Candida spp. [22]. Primers and PCR reaction conditions are presented in supplementary table S1 available online at stacks.iop.org/JBR/12/026001/mmedia. Mannitol salt agar was used to confirm the identities of Staphylococcus strains, and differentiate between S. aureus and coagulase-negative Staphylococcus.

Culture conditions and sample preparation

Culture conditions and sample preparation were as described previously [23]. Isolates were pre-cultured overnight (5 ml, 37°C, 200 rpm shaking) in Difco grieving Mueller-Hinton Broth (MHB) (Becton Dickinson, Franklin Lakes, NJ, United States), and inoculated 1:1000 into 20 ml of fresh MHB, which was incubated under identical conditions (37°C, 200 rpm shaking) in a 250 ml Erlenmeyer flask for 12 h. After 12 h, cultures were transferred to 50 ml conical flasks, submerged in ice to quench metabolism, and centrifuged (12 100 × g, 4°C, 5 min). Bacterial and fungal growth was approximated via optical density at 600 nm (OD600). 4 ml of culture supernatant was transferred to a 20 ml air-tight glass headspace vial and sealed with a Silicone/PTFE screw cap (both from Sigma-Aldrich, St. Louis, MO, United States). Samples were immediately stored at −20°C, and analyzed within three months of preparation. One replicate was prepared per isolate.

Concentration and analysis of volatile compounds

Volatile metabolites were concentrated using headspace solid-phase microextraction (HS-SPME) and separated and analyzed via comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry (GC×GC-TOFMS). The GC×GC-
TOFMS instrument was a LECO® Pegasus 4D (LECO Corp., St. Joseph, MI, United States) equipped with a rail autosampler (MPS, Gerstel Inc., Linthicum Heights, MD, United States). SPME fibers were produced by Supelco® (Bellefonte, PA, United States), and gas chromatography (GC) columns by Restek® (Bellefonte, PA, United States). A comprehensive list of instrumental parameters is presented in Table 1.

Data Processing and chromatographic alignment
Chromatographic data were aligned using the Statistical Compare feature of the ChromaTOF® software, v4.50 (LECO Corp.). The baseline was drawn through the middle of the noise, and peaks with a minimum signal-to-noise (S/N) threshold of 50:1 were identified. Subpeaks within a chromatogram were combined if their second dimension retention time shift was ≤0.1 s between subsequent modulation periods, and their mass spectral match score was ≥600/1000. For the alignment of peaks across chromatograms, maximum allowable first and second dimension retention time deviations were set at 6 s and 0.15 s, respectively, and the minimum mass spectral similarity match set at ≥600/1000. For peaks detected at a S/N threshold of 50:1 in at least one chromatogram, all remaining chromatograms were searched at a reduced 5:1 threshold for the detection of low abundance analytes. Suspected chromatographic artifacts and contaminants, as defined previously [24], were removed from the data prior to any statistical analyses, as were known atmospheric gases (e.g., carbon dioxide and argon), and peaks eluting prior to 358 s [13].

Table 1. Instrumental parameters for HS-SPME-GC × GC-TOFMS analysis.

| Parameter                                           | Value                                                                 |
|-----------------------------------------------------|-----------------------------------------------------------------------|
| Fiber length, film thickness, and fiber composition | 2 cm, 50/50 μm, triphasic (divinylbenzene/carboxen/polydimethylsiloxane) |
| Headspace solid-phase microextraction (HS-SPME)     |                                                                        |
| Headspace exposure conditions                       | 60 min, 37 °C, 200 rpm agitation                                       |
| Two-dimensional gas chromatography (GC × GC)       |                                                                        |
| Inlet parameters                                    | 270 °C, splitless injection, 180 s desorption time                    |
| Carrier gas and flow rate                          | Helium, 2 ml min⁻¹ (constant flow)                                    |
| Column composition and dimensions                   | Column 1: Rxi®-624Sil MS, 60 m × 0.25 mm × 1.4 μm²; Column 2: Stabilwax, 1 m × 0.25 mm × 0.5 μm² |
| GC temperature program                              | 35 °C hold for 30 s + 3.5°C min⁻¹ ramp to 230 °C                       |
| Temperature offsets                                 | +5 °C (secondary oven), +25 °C (modulator)                             |
| Modulator parameters                                | 2 s modulation period, alternating 0.5 s hot/0.5 s cold pulses          |
| Transfer line temperature                           | 250 °C                                                                |
| Time-of-flight mass spectrometry (TOFMS)            |                                                                        |
| Acquisition delay                                   | 170 s                                                                 |
| Mass range                                          | 30–500 amu                                                             |
| Acquisition rate                                    | 200 spectra s⁻¹                                                        |
| Detector voltage offset                             | +50 V                                                                 |
| Ion source temperature                              | 200 °C                                                                |

* Dimensions correspond to length × internal diameter × film thickness.

Compound reporting and calculation of retention indices
Chromatographic peaks were assigned putative identifications based on mass spectral matching and retention time data. Specifically, putative compound identifications were assigned to peaks with a forward match score of ≥800/1000 relative to the National Institute of Standards and Technology (NIST) 2011 mass spectral library, and whose experimentally-determined retention index was consistent with that of a midpolar Rxi®-624Sil MS stationary phase, as described previously [13]. Retention indices were calculated using external alkane standards (C₉–C₁₆). The SPME fiber was exposed to a vial containing a pure retention index mixture (Sigma-Aldrich) for 10 min at 50 °C and desorbed at a 40:1 split. The gas chromatographic and mass spectrometric method were as defined previously. Reported RIs are between the literature values for polar and nonpolar column sets, due to the midpolarity of the Rxi®-624Sil MS stationary phase. Retention indices less than 600 (corresponding to C₉) or greater than 1600 (corresponding to C₁₆) were not extrapolated.

Statistical analyses
All statistical analyses were performed using R v3.2.2 (R Foundation for Statistical Computing, Vienna, Austria). Chromatographic data was first normalized using probabilistic quotient normalization [25] and log₁₀-transformed. The Mann–Whitney U test with Benjamini–Hochberg correction [26, 27] was used to identify metabolites that were significantly different between pathogen groups and sterile media, with a corrected p-value of 0.05 established as the threshold.
for statistical significance. Network modeling was used to visualize relationships between metabolites [28]. Random Forest (RF) was used to identify volatile metabolites that were discriminatory between groups and predict pathogen identifications for leave-one-out cross-validation (LOOCV) test samples [29]. For each RF comparison, a total of 100 iterations were performed, with 500 decision trees generated per iteration. For assessing variable importance, all samples were included in the RF analysis, and volatile metabolites were ranked according to their mean decrease in accuracy (MDA), a measure of variable importance. For predicting pathogen identifications, one sample was withheld from the data, and RF was performed using the remaining 99 samples. Class probabilities were calculated for the withheld sample, with simple majority voting used to assign the withheld sample to a given pathogen group. This process was repeated using a different withheld sample each time, until all 100 samples had been withheld once. Principal component analysis (PCA) was used to reduce dimensionality and visualize variance in the data. Hierarchical clustering analysis (HCA) was employed to assess relatedness between samples based on their volatile metabolic profiles using Euclidean distance as the distance metric. For both PCA and HCA, variables were mean-centered and unit-scaled. Centering and scaling was applied to each variable independently. The agraphe package was used to generate network plots, ggplot2 to generate matrices, rgl to generate PC scores plots, and gplots to generate heat maps.

**Results**

**Volatile molecular fingerprints of bacterial and fungal pathogen groups**

We hypothesized that pathogenic microorganisms differ from one another in their production of volatile metabolites. After 12 h of growth in MHB under aerobic conditions, the volatile metabolites produced by ten pathogen groups (ten clinical isolates per group) were concentrated using headspace solid-phase microextraction and analyzed via comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry (HS-SPME-GC×GC-TOFMS). The pathogen groups evaluated in this study consisted of one fungal group (Candida spp.), three Gram-positive groups (Enterococcus spp., coagulase-negative Staphylococcus, and S. aureus), and six Gram-negative groups. Two distinct phylogenetic lineages were represented within the Gram-negatives, namely the Pseudomonadales (Acinetobacter spp., and Pseudomonas aeruginosa) and the Enterobacteriales (Enterobacter spp., Escherichia coli, Klebsiella spp., and P. mirabilis) [30].

After chromatographic alignment and the removal of contaminants (see materials and methods), a total of 811 volatile metabolites were detected. Of these, 203 were: (1) detected in nine or 10 (of 10) isolates for one or more pathogen groups, and (2) significantly more abundant in cultures relative to sterile media controls ($p < 0.05$). This subset of 203 metabolites is used for all subsequent analyses and discussion, unless otherwise specified.

There was notable variability in the number of metabolites that met these two inclusion criteria across different pathogen groups, ranging from 133 for *P. mirabilis*, to only one for *Enterococcus* spp. (figure 1). Twice as many volatile metabolites fulfilled the inclusion criteria, on average, for Gram-negative pathogen groups (range: 47–133) than both Gram-positive pathogen groups (range: 1–24), and fungi (three detected for Candida). Candida, Enterococcus, and coagulase-negative *Staphylococcus* grew slower in MHB relative to the seven other pathogen groups assessed (supplementary figure S1). Indeed, the median culture density for these three groups after 12 h of growth, as approximated by the optical density at 600 nm ($\text{OD}_{600}$), ranged from 0.45 to 1.62, while the median $\text{OD}_{600}$ for the remaining seven pathogens ranged from 2.06 to 2.61. We hypothesize that this slower growth likely contributed to the limited number of metabolites detected for these three pathogen groups relative to the other groups assessed in this study, although we also acknowledge that pathogen-to-pathogen differences in production of volatile metabolites could have also contributed.

Other than the notable difference in the number of compounds detected for Gram-negative bacteria relative to both Gram-positive bacteria and fungi, we observe that the proportions of volatile metabolites shared between pathogen groups do not clearly recapitulate genetic relatedness. For example, among the Gram-negative pathogen groups included in this study, *P. aeruginosa* and *Acinetobacter* are more genetically similar to one another, belonging to the order Pseudomonadales, than either is to the four other Gram-negative pathogens (Enterobacter, *E. coli*, Klebsiella, and *P. mirabilis*), which all belong to the group Enterobacteriales [30]. Despite this, *Acinetobacter* shares more metabolites with both *P. mirabilis* and *Klebsiella* ($n = 63$ and 42, respectively) than it does with *P. aeruginosa* ($n = 35$) (figure 1). We note, however, that when the number of shared metabolites is normalized to the total number of metabolites produced by each pathogen group, a relatively greater proportion of *P. aeruginosa*’s metabolites are shared with *Acinetobacter* compared with either *P. mirabilis* or *Klebsiella* (66% versus 68% or 58%, respectively). However, this trend is not uniformly true across all pathogen groups.

We next performed a network analysis to assess the relationships between the 203 selected volatile metabolites, using Pearson’s correlation coefficient as the measure of relatedness between metabolites. This analysis constrains features (depicted as nodes) within a two-dimensional space as a function of their correlations to all other features, with connecting lines...
representing either significantly positive (red) or negative (blue) correlations [28]. The network plot depicting these metabolites (figure 2) includes two distinct clusters of approximately 50 highly correlated features each (left and upper-left), one relatively diffuse cluster of approximately 60 positively- and negatively-correlated features (bottom-right), and one cluster of seven highly positively correlated features (upper-right).

When these metabolites are displayed as a function of the pathogen groups that produce them, the cluster of highly correlated metabolites on the left side is found to correspond to those produced predominantly by *P. mirabilis* (figure 2(H)), the upper-left cluster to those produced predominantly by *Acinetobacter*, *Klebsiella*, *P. mirabilis*, and *Pseudomonas* (figures 2(G)–(J)), the lower-right cluster to those produced by all six Gram-negative pathogens (figures 2(E)–(J)), and the upper-right cluster to those produced by the *Staphylococci* groups coagulase-negative *Staphylococcus* and *S. aureus* (figures 2(C) and (D)).
respectively). The production of volatile metabolites by these pathogen groups in the context of the network analysis can also be visualized via the calculation of centroids (supplementary figure S2). As we would have predicted based on figure 2, the centroids for P. mirabilis and Acinetobacter are located towards the left and upper-left, respectively, those for Klebsiella, Enterobacter, E. coli, and P. aeruginosa cluster together in the center/upper-right, and those for S. aureus and coagulase-negative Staphylococcus are found in the upper-right. One must be cautious with interpreting the location of the centroids for Candida and Enterococcus, given the relatively few metabolites used to calculate these values (only three and one, respectively).

**Volatile molecular fingerprints predict pathogen-level identification**

Having demonstrated that these pathogen groups differed broadly in their production of volatile metabolites, we hypothesized that a subset of these 203 metabolites could also be useful for classifying samples into different pathogen groups. We obtained a 95% classification accuracy for the discrimination between these ten pathogen groups using the machine learning algorithm RF with LOOCV [29]. Classification accuracies were highest for coagulase-negative Staphylococcus, S. aureus, E. coli, Klebsiella, P. aeruginosa, and P. mirabilis (100% accuracy), followed by Candida, Enterobacter, and Acinetobacter (90%), and finally Enterococcus (80%). Error matrices were generated to visualize occurrences of sample misclassification, as well as sample class probabilities generated from RF, for LOOCV samples (supplementary figure S3). Instances of misclassification were most commonly encountered between Candida and Enterococcus, possibly resulting from slower growth relative to the other pathogen groups assessed, as described previously. In further support of this hypothesis, the one isolate of Acinetobacter that reliably misclassified as Enterococcus also grew poorly in MHB relative to the other isolates from this species. Of note, we report high classification accuracies for the discrimination between closely related groups, including coagulase-negative Staphylococcus and S. aureus (100%), the four Enterobacteriales (98%) (Enterobacter, E. coli, Klebsiella, and P. mirabilis), and the two Pseudomonadales (100%) (P. aeruginosa and Acinetobacter), indicating that while these pathogens are genetically similar to one another, their volatile metabolic profiles are distinct.

We additionally sought to determine whether our stringent compound inclusion criteria may have impacted our ability to differentiate between pathogen groups. To assess this, we repeated RF using all 811 volatile metabolites detected in this study, rather than the selected 203 that we had used previously. The inclusion of all 811 metabolites resulted in a classification accuracy of 96% (supplementary figure S4), a modest improvement from the 95% reported using the 203 selected metabolites. Of note, improvement was reported only in the differentiation between Candida and Enterococcus (from 85% to 90%), while all other classification accuracies remained unchanged.

Finally, we sought to identify the subset of volatile metabolites that were most highly discriminatory between pathogens, and visualize how these metabolites varied across our collection of isolates. From RF, we calculated variable importance (as defined by the MDA) for the 203 selected metabolites in a single ten-class comparison (i.e., all ten pathogen groups treated as individual classes in a single statistical model), as well as ten two-class comparisons consisting of one pathogen group versus the nine others aggregated into a single class. For each model, we extracted the three most discriminatory metabolites (those with the largest MDA) and aggregated these 33 metabolites into a single feature set, for which we generated a principal component (PC) scores plot to visualize differences between pathogens (figure 3(A)). In the case where a metabolite that had previously been selected in one comparison was within the top three most discriminatory metabolites for another comparison, the fourth most discriminatory compound was selected, such that all 11 models were weighted equally.

The first three principal components generated from this set of 33 metabolites reveal clustering of samples by pathogen group, with relatively few exceptions. S. aureus and coagulase-negative Staphylococcus cluster adjacent to one another, as do the two Pseudomonadales (Acinetobacter and P. aeruginosa), and the three closely related Enterobacteriales (Enterobacter, E. coli, and Klebsiella). P. mirabilis, with its unusual metabolite profile (as depicted in figure 2(H)) formed a distinct cluster along PC2, away from other Enterobacteriales, as well as the other Gram-negative, Gram-positive, and fungal pathogens. Of note, Candida, Enterococcus, and one of the Acinetobacter isolates form a relatively indistinct cluster towards the center of the PC scores plot. This observation is consistent with our findings from the previous analysis, which demonstrated that Candida and Enterococcus tended to misclassify as one another, albeit with relatively low frequency (supplementary figure S3). However, it is clear that the first three principal components presented in this plot, which together account for 58% of the total variance in the data, fail to capture a sizeable portion of the variance between Candida and Enterococcus samples, as isolates from these two groups classified with approximately 85% accuracy. Indeed, other principal components not included in these scores plot, such as PC5 (9%) and PC7 (3%) account for much of the variance between these two groups (not shown). For Enterobacter, E. coli, and Klebsiella, the three Enterobacteriales that clustered together on the PC scores plot, a second scores plot was generated using the same suite of 33 discriminatory compounds used for the ten-group comparison, but including only these three pathogen
In this second plot, we observe that Enterobacter, E. coli, and Klebsiella all occupy spatially-distinct regions, demonstrating that even amongst these closely related pathogens, the production of volatile metabolites is differential. Indeed, these pathogens could be distinguished from one another with 100% accuracy using the prior RF model. We generated a heat map as another way of visualizing the relationship between samples using these 33 discriminatory metabolites, with Euclidean distance as the measure of relatedness between samples (figure 4). A subset of the information presented in this heat map was evident via PC scores plot. For example, strains tended to cluster as a function of pathogen group, and more closely related pathogens, such as Klebsiella, E. coli, and Enterobacter, tended to be more similar to one another than they were to other groups. Interestingly, this suggests that while the overall production of volatile metabolites is only modestly associated with genetic relatedness, a subset of discriminatory volatile metabolites can provide a powerful tool for distinguishing closely related pathogens.

Figure 3. Principal component (PC) scores plots for 10 fungal and bacterial pathogen groups, generated using 33 discriminatory volatile compounds identified via Random Forest (RF). 3(A), left: PC scores plot depicting all 10 pathogen groups. Fungi: triangles; Gram-positive cocci: diamonds; Gram-negative rods (Enterobacterales): squares; Gram-negative rods (Pseudomonadales): circles. PC1: 25%, PC2: 19%, PC3: 14%. 3(B), right: PC scores plot depicting only the three closely related Enterobacterales: Enterobacter, E. coli, and Klebsiella, using the same discriminatory volatile compounds for the comprehensive pathogen-level comparison presented in 3(A): PC1: 27%, PC2: 15%, PC3: 11%.

Figure 4. Heat map depicting the relative abundance of the 33 discriminatory volatile metabolites identified using RF (rows) across the 10 pathogen groups assessed in this study (columns). Dendrogram (top) depicts the relatedness between strains, utilizing Euclidean distance as the distance metric. Cell colors correspond to relative compound abundance after mean-centering and unit-scaling, ranging from blue (low relative abundance) to red (high relative abundance). CAN or *: Candida; E + or †: Enterococcus spp.; CNS or #: coagulase-negative Staphylococcus; SA: S. aureus; A or †: Acinetobacter spp.; EC: E. coli; K: Klebsiella spp.; PA: P. aeruginosa; PM: P. mirabilis. Putative compound identifications or class assignments are provided when possible (left). Experimentally-determined retention indices (RIs) are provided for analytes (‘A’ followed by a number), and metabolites for which only compound class assignments could be determined. The statistical model from which a given metabolite was selected is provided (right), with O representing metabolites selected from the overall (ten-group) comparison.

groups (figure 3(B)). In this second plot, we observe that Enterobacter, E. coli, and Klebsiella all occupy spatially-distinct regions, demonstrating that even amongst these closely related pathogens, the production of volatile metabolites is differential. Indeed, these pathogens could be distinguished from one another with 100% accuracy using the prior RF model.

We generated a heat map as another way of visualizing the relationship between samples using these 33 discriminatory metabolites, with Euclidean distance as the measure of relatedness between samples (figure 4). A subset of the information presented in this heat map was evident via PC scores plot. For example, strains tended to cluster as a function of pathogen group, and more closely related pathogens, such as Klebsiella, E. coli, and Enterobacter, tended to be more similar to one another than they were to other groups. Interestingly, this suggests that while the overall production of volatile metabolites is only modestly associated with genetic relatedness, a subset of discriminatory volatile metabolites can provide a powerful tool for distinguishing closely related pathogens.
metabolites may be useful for assessing genetic similarities between pathogen groups. Of interest, the ability of these metabolites to distinguish between Candida and Enterococcus is arguably more apparent when visualized via heat map relative to PC scores plot, with nine of ten Candida isolates forming a single cluster, and seven of ten Enterococcus isolates belonging to one of two distinct clusters. The ability to visualize the relative abundance of the 33 discriminatory compounds across the 100 samples via heat map demonstrates that while a subset of the discriminatory compounds are reliably produced by only a single pathogen group (e.g., 2,4-dithiapentane for P. mirabilis), the majority are produced across groups (e.g., 2-butaneone for Acinetobacter and P. aeruginosa, and 2,4-dimethylheptane for Acinetobacter, Klebsiella, and P. mirabilis). Therefore, although the metabolites included in this analysis represent the most discriminatory features associated with a single pathogen group (excluding the three obtained from the ten-group comparison), a subset of the included metabolites have some discriminatory ability for more than one pathogen group included in this study.

Of the 33 metabolites selected via RF, putative compound identifications could be assigned to six, and compound class assignments to an additional 16. The chromatographic characteristics of these metabolites, as well as their relative abundances across different pathogen groups, are presented in supplementary table S2. For the six compounds assigned putative compound identifications, 2,3-butanediol was most highly abundant in cultures of Enterococcus spp., 2,4-dimethylheptane in Acinetobacter, 2,4-dithiapentane in P. mirabilis, 2-butaneone in P. aeruginosa, 2-heptanone in Klebsiella, and toluene in Candida. In considering all 22 compounds assigned either a putative identification or compound class assignment, we note that cultures of S. aureus and coagulase-negative Staphylococcus were enriched in ketones (2-butaneone, 2-heptanone, and four unidentified ketones), while cultures of Acinetobacter were characterized by an abundance of hydrocarbons (2,4-dimethylheptane and two unidentified hydrocarbons). Of note, because this analysis considers only a subset of discriminatory pathogen-derived volatile metabolites, the proportions of different molecular classes reported here are not necessarily representative of the comprehensive volatile metabolic profile associated with each pathogen group. However, the observation that metabolites from a single molecular class are discriminatory between groups is possibly reflective of pathogen-level differences in the utilization of the pathways that result in the generation of these molecules (e.g., fatty acid metabolism for the production of 2-ketones) [1].

Discussion

The present study has considered the volatile metabolic profiles of 100 bacterial and fungal isolates belonging to ten distinct pathogen groups, and represents one of the most extensive analyses of the volatile metabolites produced by pathogens in vitro to date. It additionally represents the single most extensive analyses of bacterial and fungal volatile metabolites performed using GC×GC-TOFMS, a powerful analytical technique well-suited for the characterization of complex mixtures including breath [31], as well as culture headspace metabolites [12–15, 23, 24, 32–34]. In total, we estimate that over 160 prior studies have described qualitative and/or quantitative differences in the composition of headspace volatile metabolites between different pathogen groups under either in vitro or ex vivo conditions. In general, these studies have focused on the volatile metabolic profiling of relatively few organisms, often those most pertinent to a specific disease or disease process, and frequently with only a limited number of clinical isolates or reference strains included per organism. Julak and colleagues provided arguably the broadest analysis with respect to the number of different pathogen groups assessed, with 37 [35], while Coloe reported on the largest number of isolates from a single group, with 61 E. coli clinical isolates [36]. Other studies have sought to compromise between the number of pathogen groups and number of isolates per group, including those by Roine and colleagues, who analyzed a total of 80 isolates across four pathogen groups [37], Lim and colleagues, who analyzed 62 isolates across 18 pathogen groups [38], and Bruins and colleagues, who analyzed 52 isolates across 11 pathogen groups [39]. The present study demonstrates the importance of including both a broad range of pathogen groups, due to the differential metabolite profiles observed across groups, as well as the inclusion of a range of isolates, due to the observation that some rare isolates may produce a volatile molecular profile quite distinct from that of the ‘typical’ isolate.

Our results suggest that both the number and variety of volatile metabolites shared between the pathogen groups included in this study were only modestly associated with the genetic relatedness between the pathogen groups themselves. The selected discriminatory features demonstrated a slightly better association with genetic relatedness relative to all pathogen-derived metabolites, although the relationship was still quite modest, with the fungus Candida clustering adjacent to the Gram-positive Enterococcus, and P. mirabilis exhibiting a very distinctive volatile molecular profile relative to the other Enterobacteriales included in this study. Had our volatile metabolic fingerprints closely reflected genetic relatedness, we would have anticipated that the majority of the variance in our data would be attributed to differences between fungal, Gram-positive, and Gram-negative
pathogen groups, with ‘within-group’ differences (e.g., Enterobacteriales versus Pseudomonadales) accounting for a smaller proportion of the overall variance in the data [40].

Of interest, prior studies conducted using either MHB or Mueller-Hinton agar as a growth medium have reported similar findings. For example, Boots and colleagues noted that the majority of the variance in their volatilomic data resulted from differences in the metabolic profiles of *P. aeruginosa* and *E. coli* versus *S. aureus* and *K. pneumoniae*, despite the substantial genetic similarity between *K. pneumoniae* and *E. coli* [30, 41]. Liang and colleagues observed distinct differences between Gram-positive pathogens, Gram-negative pathogens, and fungi, although the Gram-negatives that they assessed (*E. coli, K. pneumoniae, P. aeruginosa*, and *Acinetobacter baumannii*) clustered together despite belonging to two distinct genetic clades [42]. Therefore, our results are plausibly related to the choice of growth conditions employed in the present study, as it has previously been demonstrated that growth media influences the composition of headspace volatile metabolites detected, even within a single organism [23, 42–44]. Alternative culture conditions might have resulted in metabolite profiles that were more reflective of genetic differences between pathogen groups relative to what was observed in the present study. For example, Lim and colleagues, who utilized BacT/ALERT® aerobic blood culture media for bacterial growth, observed that their bacterial strains tended to cluster in a way that broadly reflected genetic relatedness, with a distinct cluster corresponding to *P. aeruginosa* and *A. baumannii*, another corresponding to a group of Gram-positive pathogens (*Streptococcus* spp. and *Enterococcus* spp.), and a third that encompassed most of the Enterobacteriales (including *E. coli, Klebsiella* spp., *Serratia marcescens*, and *Salmonella enterica*) [38]. Of interest, they also noted that *P. mirabilis* produced a volatile metabolic signature distinct from those of other related Enterobacteriales. At present, there is not sufficient uniformity across studies with respect to experimental design to definitively state that one growth medium is superior to another for assessing volatilomic differences across a wide range of pathogens, although such a study could represent an important future direction for this work.

One of our most interesting observations from the present study is that Gram-negative pathogen groups, on average, produced more than twice as many volatile metabolites as Gram-positive and fungal pathogen groups when grown in MHB. Previous studies have reported conflicting results with regards to differences in the composition of volatile metabolites between Gram-negative and Gram-positive organisms. A subset of studies have described only modest differences in the number of volatile metabolites produced, with Boots and colleagues reporting on four metabolites for *S. aureus* compared with six, seven, and seven for *P. aeruginosa, E. coli*, and *K. pneumoniae*, respectively [41], and Filipiak and colleagues reporting on 37 compounds for *P. aeruginosa* compared with 32 for *S. aureus* [43]. Jünger and colleagues, however, described much more dramatic differences in the production of volatile metabolites between Gram-positive and Gram-negative pathogens. Similar to our own findings, that study reported on approximately three- to four-fold differences in the production of volatile metabolites produced by Gram-negative versus Gram-positive pathogens, with a substantial overabundance of volatile metabolites produced by *P. mirabilis* relative to other species [46]. In contrast, others such as Liang and colleagues [42], have reported on a stronger volatilomic signature from *S. aureus* relative to both Gram-negative pathogen groups and fungi. Again, the lack of uniformity across studies regarding experimental design makes it challenging to draw definitive conclusions about the relative ‘sizes’ of the volatile metabolomes of different pathogen groups. Indeed, prior work from our group has demonstrated that even within a single organism, the size of the volatile metabolome can differ by more than two-fold across different common laboratory growth media [23]. A tempting approach to optimize the number and variety of volatile molecules detected is to utilize different media that have been optimized for the growth of specific organisms, for example: 7H9 or 7H11 media for *Mycobacteria* [47–50], chocolate agar for fastidious organisms [51], or Sabouraud dextrose media for yeast [52–58]. However, while such an approach may optimize growth and the production of volatile metabolites for a given pathogen, one must be cautious when describing volatile metabolic differences between organisms grown on different media. Even with techniques to ‘subtract’ the volatile molecular signature of the media itself, previous studies suggests that a microorganism’s volatile molecular signature is fundamentally media-dependent, and that alteration of growth media could result in a substantially different volatile molecular profile [23, 43, 59].

In addition to qualitatively describing differences in the composition of volatile metabolites between pathogen groups, a key component of this study is the use of a supervised machine learning algorithm to predict the pathogen group to which a given sample belongs. Volatile molecular fingerprints could represent novel biomarkers for a range of infectious pathologies, including pneumonia, sepsis, and urinary tract infections [2, 7]. However, in order for these biomarkers to translate to the clinical setting, they must be able to reliably and accurately differentiate between pathogens. With the exception of studies that have described the development of novel sensors (e.g., ‘electronic noses’ or colorimetric sensors), however, test parameters such as accuracy, sensitivity, or specificity, have seldom been reported [60–63].

The importance of discriminating between pathogen groups, specifically closely related pathogen
groups, is most evident in the comparison of *S. aureus* versus coagulase-negative *Staphylococcus*, which we could discriminate from one another with 100% accuracy when considering both the 203 selected volatile metabolites, as well as all 811 metabolites. In the setting of suspected sepsis, the detection of *S. aureus* in a sample usually indicates an underlying disease process, while the detection of coagulase-negative *Staphylococcus* is often indicative of sample contamination [64], and the ability to discriminate between these two groups is thus critically important. Furthermore, the ability to discriminate between *Klebsiella*, *P. mirabilis*, and *E. coli*, for example, which we were able to achieve with 100% accuracy, is pertinent in the setting of urinary tract infections, as antibiotic susceptibilities can vary across pathogen groups, especially within institutions experiencing outbreaks of multidrug-resistant pathogens [65]. Future studies could assess the translatability of our present findings to more clinically relevant media or real bio-specimens, including urine, blood culture, and sputum, to determine the impact of a heterogeneous sample matrix on the ability to discriminate between pathogen groups.

In considering study strengths, we note that the present work represents one of the most extensive analyses of the volatile metabolites produced by bacteria and fungi to-date, and additionally represents the single most extensive such analysis using GC×GC-TOFMS. Through this approach, we were able to identify specific volatile metabolic ‘fingerprints’ associated with each pathogen group, consisting, in aggregate, of over 200 unique metabolites, many more than have been previously reported using other analytical techniques. In addition, we were able to demonstrate that this collection of volatile metabolites yielded a strong classification accuracy (95%) for discriminating between 10 pathogen groups (with 10 isolates per group), demonstrating the potential utility of volatile biomarkers as diagnostic biomarkers. In considering study limitations, we note that the genomic diversity within each pathogen group that is captured by our collection of isolates is not known, and that additional whole genome sequencing would be required to ensure that we have included isolates representing the greatest extent of genetic diversity for each group. We acknowledge that differences in the number of volatile metabolites to meet our two inclusion criteria for each pathogen group may be due, in part, to this within-group heterogeneity, with more diverse pathogen groups yielding fewer volatile metabolites produced by ≥9 of 10 isolates. However, relaxing this inclusion criterion to potentially capture molecules detected in fewer isolates (e.g., ≥5 of 10 isolates, or ≥7 of 10 isolates) does not dramatically alter the total number of reported compounds associated with *Candida*, *Enterococcus*, or coagulase-negative *Staphylococcus*, as the number of compounds to fulfill our second inclusion criterion is reduced due to larger *p*-values after BH correction.

Admittedly, our choice of culture conditions were not optimized for the specific pathogen groups included in the present study, as we instead opted for the use of a growth medium that is already widely employed in the clinical microbiology laboratory setting, and which is known to support the growth of a wide range of human pathogens. It is possible, however, that alternative culture conditions could have yielded volatile metabolic fingerprints that were even more differential between groups. Furthermore, differences in cell densities observed at the time of harvest may have also contributed to differences in the volatile metabolic profiles observed between pathogen groups. However, it is unlikely that such differences could account for the differences observed between the Gram-negative pathogen groups, as well as *S. aureus*, as these groups attained similar cell densities at 12 h post-inoculation. Finally, although LOOCV was employed, the inclusion of an independent validation set could have served to demonstrate additional robustness of the statistical tools used. Future studies could consider translation of our current findings to clinical specimens, or perhaps address more fundamental questions pertaining to the origins of these volatile metabolites.

**Conclusions**

Volatile metabolic profiling represents a powerful approach for the identification and characterization of different bacterial and fungal pathogen groups. The molecules detected have the potential to aid in the diagnosis of different infectious diseases or provide novel insight into underlying cellular metabolic processes. Translation of volatile metabolic fingerprinting to the clinical setting will require the identification of biomarkers capable of differentiating between key pathogen groups, as well as a careful evaluation of key diagnostic parameters, including accuracy, sensitivity, and specificity.

**Acknowledgments**

This work was supported by National Institutes of Health grant R21 AI121076 to JE Hill and Lennart KA Lundblad. Dartmouth College holds an Institutional Program Unifying Population and Laboratory Based Sciences award from the Burroughs Wellcome Fund, and CA Rees was supported by this grant (Grant#1014106). P-H Stefanuto is a Marie-Curie COFUND postdoctoral fellow co-funded by the European Union and the University of Liège. We thank Elissa M Eckhardt and the team at the Clinical Microbiology Laboratory at Dartmouth-Hitchcock Medical Center for generously providing the isolates used in this study.
References

[1] Schulz S and Dickshat J S 2007 Bacterial volatiles: the smell of small organisms Nat. Prod. Rep. 24 818–42

[2] Bos L D, Sterk P J and Schultz M J 2013 Volatile metabolites of pathogens: a systematic review PLoS Pathog. 9 e1003311

[3] Martins C, Brandao T, Almeida A and Rocha S M 2017 Metabolomics strategy for the mapping of volatile exometabolome from Saccharomyces spp. widely used in the food industry based on comprehensive two-dimensional gas chromatography J. Sep. Sci. 40 2228–37

[4] Lonsdale C L, Taba B, Queralt N, Łukaszewski R A, Martinó R A, Rhodes P A and Lim S H 2013 The use of colorimetric sensor arrays to discriminate between pathogenic bacteria PLoS One 8 e62776

[5] Koo S, Thomas H R, Daniels S D, Lynch R C, Fortier S M, Shea M M, Rearden P, Comolli J C, Baden L R and Marty F M 2014 A breath fungal secondary metabolite signature to diagnose invasive aspergillosis Clin. Infect. Dis. 59 1733–40

[6] Schnabel R, Fijten R, Smolinska A, Dallinga J, Boumans M L, Koo S, Thomas H R, Daniels S D, Lynch R C, Fortier S M, Malley M, O'Sullivan F, Robertson D, Skirrow J, Ward C, Williams K P, Gillespie J J, Sobral B W, Lomax E K, Northcote J, Markham F, Vass J, Ammermann A, Weisburg W G, Barns S M, Pelletier D A and Lane D J 1991 16 s ribosomal DNA phylogeny of marine bacteria using advanced analytical instrumentation for the discovery of novel metabolites PLoS One 8 e1003311

[7] Benjamini Y and Hochberg Y 1995 Controlling the false discovery rate—a practical and powerful approach to multiple testing J. R. Stat. Soc. B 57 289–300

[8] Williams K P, Gillespie J, Sobral B W, Nordberg E K, Snyder E E, Shallom J M and Dicksman A W 2010 Phylogeny of gamma-proteobacteria J. Bacteriol. 192 2305–14

[9] Phillips M, Cataneo R N, Chaturvedi A, Kaplan P D, Libardoni M, Mundada M, Patel U and Zhang X 2013 Detection of an extended human volatome with comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry PLoS One 8 e75274

[10] Nizio K D, Perraudeau K A, Troumboukoff A N, Ueland M, Shoma S, Iredell J R, Middleton P G and Forbes S L 2016 In vitro volatile organic compound profiling using GCxGC–TOFMS to differentiate bacterial associated with lung infections: a proof-of-concept study J. Breath Res. 10 026008

[11] Rees C A, Franchina F A, Nordick K V, Kim P J and Hill J E 2016 Expanding the Klebsiella pneumoniae volatile metabolome using advanced analytical instrumentation for the detection of novel metabolites J. Appl. Microbiol. 122 785–95

[12] Weisburg W G, Barns S M, Pelletier D A and Lane D J 1991 16 s ribosomal DNA analysis for phylogenetic study J. Bacteriol. 173 697–703

[13] Frank J A, Reich C L, Sharma S, Weisbaum J S, Wilson B A and Olsen G J 2008 Critical evaluation of two primers commonly used for amplification of bacterial 16 s rRNA genes Appl. Environ. Microbiol. 74 2461–70

[14] Ferrer C, Colom F, Frases S, Mulet E, Abad J I and Allo J L 2001 Detection and identification of fungal pathogens by PCR and by ITS2 and 5.8 s ribosomal DNA typing in ocular infections J. Clin. Microbiol. 39 2873–9

[15] Rees C A, Nordick K V, Smolinska A and Rocha S M 2017 The volatile molecule signature of four mycobacteria species J. Breath Res. 11 031002

[16] Dieterle F, Ross A, Schlotterbeck G and Senn H 2006 Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures, application in 1H NMR metabolomics Anal. Chem. 78 4281–90

[17] Mann H B and Whitney D R 1947 On a test of whether one of 2 random variables is stochastically larger than the other Ann. Math. Stat. 18 50–60

[18] Benjamini Y and Hochberg Y 1995 Controlling the false discovery rate—a practical and powerful approach to multiple testing J. R. Stat. Soc. B 57 289–300

[19] Epalskamp S, Cramer A O J, Waldorp L I, Schmittmann V D and Beesboom D 2012 qgraph: network visualizations of relationships in psychometric data J. Stat. Softw. 48 1–18

[20] Breiman L 2001 Random forests Mach. Learn. 45 5–32

[21] Mann H B and Whitney D R 1947 On a test of whether one of 2 random variables is stochastically larger than the other Ann. Math. Stat. 18 50–60

[22] Phillips M, Cataneo R N, Chaturvedi A, Kaplan P D, Libardoni M, Mundada M, Patel U and Zhang X 2013 Detection of an extended human volatome with comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry PLoS One 8 e75274

[23] Nizio K D, Perraudeau K A, Troumboukoff A N, Ueland M, Shoma S, Iredell J R, Middleton P G and Forbes S L 2016 In vitro volatile organic compound profiling using GCxGC–TOFMS to differentiate bacterial associated with lung infections: a proof-of-concept study J. Breath Res. 10 026008

[24] Rees C A, Franchina F A, Nordick K V, Kim P J and Hill J E 2016 Expanding the Klebsiella pneumoniae volatile metabolome using advanced analytical instrumentation for the detection of novel metabolites J. Appl. Microbiol. 122 785–95

[25] Malley M, O'Sullivan F, Robertson D, Skirrow J, Ward C, Williams K P, Gillespie J J, Sobral B W, Lomax E K, Northcote J, Markham F, Vass J, Ammermann A, Weisburg W G, Barns S M, Pelletier D A and Lane D J 1991 16 s ribosomal DNA phylogeny of marine bacteria using advanced analytical instrumentation for the discovery of novel metabolites PLoS One 8 e1003311

[26] Coloe P J 1978 Head-space gas liquid chromatography for rapid detection of Escherichia coli and Proteus mirabilis in urine J. Clin. Pathol. 31 365–9

[27] Roine A et al 2014 Rapid and accurate detection of urinary pathogens by mobile IMS-based electronic nose: a proof-of-principle study PLoS One 9 e114279

[28] Lim S H et al 2014 Colorimetric sensor array allows fast detection and simultaneous identification of sepsis-causing bacteria in spiked blood culture J. Clin. Microbiol. 52 992–8

[29] Bruins M, Bos A, Petit P L, Eadie K, Rog A, Bos R, van Ramshorst G H and van Belkum A 2009 Device-independent, real-time identification of bacterial pathogens with a metal oxide-based olfactory sensor Eur. J. Clin. Microbiol. Infect. Dis. 28 775–80
Fend R, Kolk A H, Bessant C, Buijtels P, Klatser P R and Junger M, Vautz W, Kuhns M, Hofmann L, Ulbricht S, Filipiak W, Sponring A, Baur M M, Filipiak A, Ager C, Liang J C, Hang Y P, Chingin K, Hu L H and Chen H W 2014 Molecular analysis of volatile metabolites released specifically by Staphylococcus aureus and Pseudomonas aeruginosa cultures: use of selected ion flow tube-mass spectrometry (SIFT-MS) for rapidly diagnosing the causes of diarrhoea J. Breath Res. 016002

Hertel M, Hartwig S, Schütte E, Raiissi B, Preissner R, Schmidt-Weisthausen A M, Paris S, Kastner I and Preissner S 2016 Identiﬁcation of signature volatiles to discriminate Candida albicans, glabrata, krusei and tropicalis using gas chromatography and mass spectrometry Mycoses 59 117–26

Kuntzel A, Fischer S, Bergmann A, Oertel P, Steffens M, Sahgal N, Monk B, Wasil M and Magan N 2008 Fungal volatile ﬁngerprints: differentiation between dermatophyte species and strains by means of an electronic nose Sensors Actuators B 131 117–20

Aksebzeci B H, Ayasli M H, Kahraman Y, Er O, Kaya E, Ozbilge H and Kara S 2010 Classiﬁcation of root canal microorganisms using electronic-nose and discriminant analysis Biomed. Eng. Online 9 77

Bazemore R A, Feng J, Caie L and Podilla G K 2012 Biomedically important pathogenic fungi detection with volatile biomarkers J. Breath Res. 6 016002

Probert C S, Jones P R and Ratcliffe N M 2004 A novel method for rapidly diagnosing the causes of diarrhoea Gut 53 58–61

Scotter J M, Allardycse R A, Langford V S, Hill A and Murdoch D 2006 The rapid evaluation of bacterial growth in blood cultures by selected ion ﬂow tube-mass spectrometry (SIFT-MS) and comparison with the BacT ALERT automated blood-culture system J. Microbiol. Methods 65 628–31

Ballabio C, Cristoni S, Puccio G, Kohler M, Sala M R, Brambilla P and Martinez-Lozano Sinues P 2014 Rapid identiﬁcation of bacteria in blood cultures by mass-spectrometric analysis of volatiles J. Clin. Pathol. 67 743–6

Dolch M E, Janitza S, Boulesteix A L, Grassmann-Lichtenauer C, Prun S, Denzer W, Schelling G and Schubert S 2016 Gram-negative and -positive bacteria differentiation in blood culture samples by headspace volatile compound analysis J. Biol. Res. 23 5

Hall K K and Lyman J A 2006 Updated review of blood culture contamination Clin. Microbiol. Rev. 19 778–802

Mehrad B, Clark N M, Zhan G G and Lynch J P 3rd 2015 Antimicrobial resistance in hospital-acquired gram-negative bacterial infections Chest 147 1413–21

Hug I A et al 2016 A new view of the tree of life Nat. Microbiol. 1 16048

Boots A W, Smolinska A, van Berkel J J, Fijten R R, Stobberingh E E, Boumans M L, Moenen E J, Wouters E F, Dallinga J W and Van Schooten F J 2014 Identiﬁcation of microorganisms based on headspace analysis of volatile organic compounds by gas chromatography–mass spectrometry J. Breath Res. 8 027106

Liang J C, Hang Y P, Chingin K, Hu L H and Chen H W 2014 Rapid differentiation of microbial cultures based on the analysis of headspace volatiles by atmospheric pressure chemical ionization mass spectrometry Res. Adv. 4 25532–9

O’Hara M and Mayhew C A 2009 A preliminary comparison of volatile organic compounds in the headspace of cultures of Staphylococcus aureus grown in nutrient, dextrose and brain heart bovine broths measured using a proton transfer reaction mass spectrometer J. Breath Res. 3 027001

Dryahina K, Sovova K, Nemec A and Spanel P 2016 Differentiation of pulmonary bacterial pathogens in cystic ﬁbrosis by volatile metabolites emitted by their in vitro cultures: Pseudomonas aeruginosa, Staphylococcus aureus, Stenotrophomonas maltophilia and the Burkholderia cepacia complex J. Breath Res. 10 037102

Filipiak W, Sponning A, Baur M M, Filipiak A, Ager C, Wiesenhofer H, Nagl M, Trompjan J and Amann A 2012 Molecular analysis of volatile metabolites released speciﬁcally by Staphylococcus aureus and Pseudomonas aeruginosa BMC Microbiol. 12 113

Jung M, Vautz W, Kuhms N, Hofmann L, Ulbricht S, Baumback J J, Quintel M and Perl T 2012 Ion mobility spectrometry for microbial volatile organic compounds: a new identiﬁcation tool for human pathogenic bacteria Appl. Microbiol. Biotechnol. 93 2603–14

Fend R, Kolk A H, Bessant C, Buitjels P, Klaster P R and Woodman A C 2006 Prospects for clinical application of electronic-nose technology to early detection of Mycobacterium tuberculosis in culture and sputum J. Clin. Microbiol. 44 2059–45

Crespo E, De Ronde H, Kuiper S, Pol A, Kolk A H, Cristescu S M, Anthony R M and Harren F J 2012 Potential biomarkers for identiﬁcation of mycobacterial cultures by proton transfer reaction mass spectrometry analysis Rapid Commun. Mass Spectrom. 26 679–85

McNerney R, Mallard K, Okolo P J and Turner C 2012 Production of volatile organic compounds by mycobacteria FEMS Microbiol. Lett. 328 150–6

Navarthi T, Mgoode G F, Weetjens B, Kaufmann S H and Schlur S 2012 The volatiles of pathogenic and nonpathogenic mycobacteria and related bacteria Beilstein J. Org. Chem. 8 290–9

Preti G, Thaler E, Hanson C W, Troy M, Eades J and Gelperin A 2009 Volatile compounds characteristic of sinus-related bacteria and infected sinus mucus: analysis by solid-phase microextraction and gas chromatography–mass spectrometry J. Chromatogr. B 877 2011–8

Scotter J M, Langford V S, Wilson P F, McEwan M J and Chambers S T 2005 Real-time detection of common microbial volatile organic compounds from medically important fungi by selected ion ﬂow tube-mass spectrometry (SIFT-MS) J. Microbiol. Methods 63 127–34

Casalinuovo I A, Di Pierro D, Bruno E, Di Francesco P and Coletta M 2006 Experimental use of a new surface acoustic wave sensor for the rapid identiﬁcation of bacteria and yeasts Lett. Appl. Microbiol. 42 24–9

Sahgal N, Monk B, Wasil M and Magan N 2006 Trichophyton species: use of volatile ﬁngerprints for rapid identiﬁcation and discrimination Br. J. Dermatol. 155 1209–16

Sahgal N and Magan N 2008 Fungal volatile ﬁngerprints: differentiation between dermatophyte species and strains by means of an electronic nose Sensors Actuators B 131 117–20

Bazemore R A, Feng J, Caie L and Podilla G K 2012 Biomedically important pathogenic fungi detection with volatile biomarkers J. Breath Res. 6 016002