OXFORD

BIOLOGY Methods & Protocols

Biology Methods and Protocols, 2024, **9(1)**, bpae087

https://doi.org/10.1093/biomethods/bpae087

Advance Access Publication Date: 14 November 2024

Methods Article

Development of a new breath collection method for analyzing volatile organic compounds from intubated mouse models

Alastair Taylor[1], Sylvia Blum[2], Madeleine Ball[1,*], Owen Birch[1], Hsuan Chou[1], Julia Greenwood[1], Shane Swann[1], Lara Pocock[1], Max Allsworth[1], Billy Boyle[1] Kerstin Geillinger-Kaestle[2]

[1]Owlstone Medical, Cambridge, CB4 0GA, United Kingdom
[2]Boehringer Ingelheim, Biberach, 88397, Germany

*Corresponding author. Owlstone Medical, Cambridge, CB4 0GA, United Kingdom. Tel: +01223 428200; E-mail: madeleine.ball@owlstone.co.uk

Abstract

A new pre-clinical method for capturing breath samples from intubated mice is presented. This method significantly reduces background levels, allowing more accurate measurements of VOCs originating from the breath ("on-breath") as opposed to background contamination. The method was developed by integrating industry-standard volatile-capturing sorbent tubes with respiratory mechanics measurement equipment (flexiVent®), resulting in a mouse breath sample that can be transported and analyzed by TD-GC-MS and other central lab technologies. Using the methodology, the discrimination between on-breath VOCs from background compounds provides a cleaner dataset, which can accelerate the validation of VOCs identified from mouse models and their translation to clinical trials. Three metrics were developed to identify on-breath VOCs, with 22 identified using Type 1 (50% of the breath samples exceeding three standard deviations above the mean signal of the system blanks), 34 with Type 2 (P-value $\leq$ .05 between paired breath and blank samples), and 61 with Type 3 (ROC-AUC value $\geq$ 0.8 to differentiate between breath and blank samples). The number of compounds seen at elevated levels on mouse breath was quantified and compared to the levels seen on human breath samples to compare methodologies.

Keywords: breath analysis; VOC analysis; mouse model; mouse breath; volatile organic compounds

Introduction

Exhaled breath contains volatile organic compounds (VOCs) that originate from biological processes in the body, and therefore have the potential to be utilized as biomarkers for clinical use. As an example, breath analysis has demonstrated the potential for good diagnostic performance in respiratory diseases, such as asthma diagnosis and stratification for likely treatment responsiveness [1–6], which are key goals for improved management of asthma. As VOCs can be derived from both the proximal airways as well as the systemic bloodstream [7], exhaled breath contains volatile compounds originating from many distal tissues and organs throughout the body, such as the gastrointestinal tract [8, 9].

As breath biomarkers have been identified across a broad range of disease areas [10–13] they are particularly promising for the broader adoption of next-generation non-invasive diagnostic and monitoring tools. Currently available breath tests in medical practice include fractional exhaled nitric oxide (FE_{NO}) for asthma, ^{13}C-Urea for *Helicobacter pylori* infection in the stomach, and hydrogen–methane for the diagnosis of small intestinal bacterial overgrowth or carbohydrate malabsorption, all of which are non-invasive and highly sensitive [14–16]. However, the translation of breath tests to clinical use remains limited, which is partly due to the lack of consistent methodologies and quality controls across the breath research literature [17, 18]. To address the challenges in clinical translation and expand the range of breath-based tests, a reliable animal model breath analysis method for pre-clinical studies would be advantageous.

Biomarker discovery, validation, and translation can be challenging. One difficulty is establishing the baseline "normal" composition of the sampling matrix, which can be complicated due to diverse factors, such as diet and metabolic variations. This means that clinical trials comparing disease and control groups require a large number of participants to account for confounding variables, raising costs, and trial durations [19]. Exhaled breath is unique in that a large proportion of the total VOCs detectable have been inhaled immediately prior to sampling from the background air, and therefore are unrelated to underlying physiological processes [20]. This represents a significant challenge, and there is a need for standardized methods of background VOC correction [9, 21]. Background VOCs can originate from multiple potential sources such as ambient air or from breath sampling equipment [22, 23]. Establishing a breath analysis method using inbred mice in controlled lab settings as a substitute for human breath can reduce the variability and challenges, allow for a better understanding of breath as a sampling medium using an animal model, and ultimately expediting

Received: 13 September 2024. **Revised:** 24 October 2024. **Editorial decision:** 31 October 2024. **Accepted:** 12 November 2024

the identification and validation of breath biomarkers for clinical use.

Several papers have previously aimed to study the breath of mouse models [24–26]. The methods utilized range from clean mice allowed to freely roam in ventilated cage, to mice in nose-only inhalation tubes. In a preliminary study by Hintzen et al., an acrylic nose cone with sorbent tube mouse breath sampling method tentatively identified 27 volatile compounds in exhaled mouse breath [25]. This included acetone and isoprene, which are two of the most abundant volatile compounds found consistently in human breath [27–29]. These results demonstrate the potential for translating findings from mouse breath research to human breath. However, methods that involve collecting VOCs from unrestrained animals in a cage can include VOCs from the skin, fur, urine, and feces—rather than just the exhaled breath [24]. In order to ensure that VOCs analyzed are truly exclusively from the breath, and to improve the signal-to-noise ratio, collection and analysis of breath from intubated mice can provide complementary data to advance the understanding of the mouse breath matrix.

No studies have directly compared baseline breath VOC profiles between mice and humans using the same analytical method. As the intended benefit of utilizing mice models to help in breath translational studies, assessing the translatability between human and mouse breath is advantageous. Given the numerous volatile compounds detectable in human breath as potential biomarkers, it is likely that there are additional compounds contained in mouse breath that could be detected and characterized in laboratory studies for better inference and translatability to human biology. To enable the detection and characterization of these compounds in mice, the identification of truly breathborne, or "on-breath" VOCs originating from physiological processes in the body and distinguishing these from not on-breath contaminating VOCs inhaled from the environment or collection equipment itself presents a significant challenge. Therefore, it is essential to develop sampling and analytical methods capable of providing a sufficient signal-to-noise ratio in mouse breath analysis.

In this study, we developed a method for accurately characterizing the VOCs in the breath of healthy intubated mice using GC-MS, capable of superior control over the signal-to-noise ratios than is currently possible for human, and mouse breath. We also compared the VOCs contained within mouse breath to system blank samples and incorporated three analytical metrics to determine the number of on-breath compounds in mice, improving the robustness of the method. Finally, as a proof-of-principle and to assess the progress in data analysis quality, we compared the on-breath VOCs identified from intubated mouse models to the on-breath VOCs in human exhaled breath.

Materials and methods

Mouse breath sampling

Animal selection and preparation

Male C57BL/6JRj mice were purchased from Janvier Labs (France) at the age of 8–12 weeks. Mice were housed in groups of 2–5 in individually ventilated cages at 22°C–25°C, at a humidity of 45%–65%, with a 12 hour day/night cycle, and given free access to water and food. Ethical approval was obtained from the regional board for animal care and welfare (Regierungspräsidium Tübingen, Germany, TVV-22-002-G).

Mouse breath collection hardware

The mouse breath sampling system was developed by modifying the commercial flexiVent® FX1 small animal ventilator (SCIREQ, Montreal, Quebec, Canada). We connected an ambient filter and a flexiVent filter to the ventilator with flexible tubing and a sorbent tube for breath collection (Fig. 1). The ambient filter contained 50 g of activated charcoal (Airpel® Desotec Ltd) to reduce the levels of ambient VOCs entering the system, and hence their likelihood of impacting either the mouse breath samples themselves or the equipment blank samples. The filter was fitted with

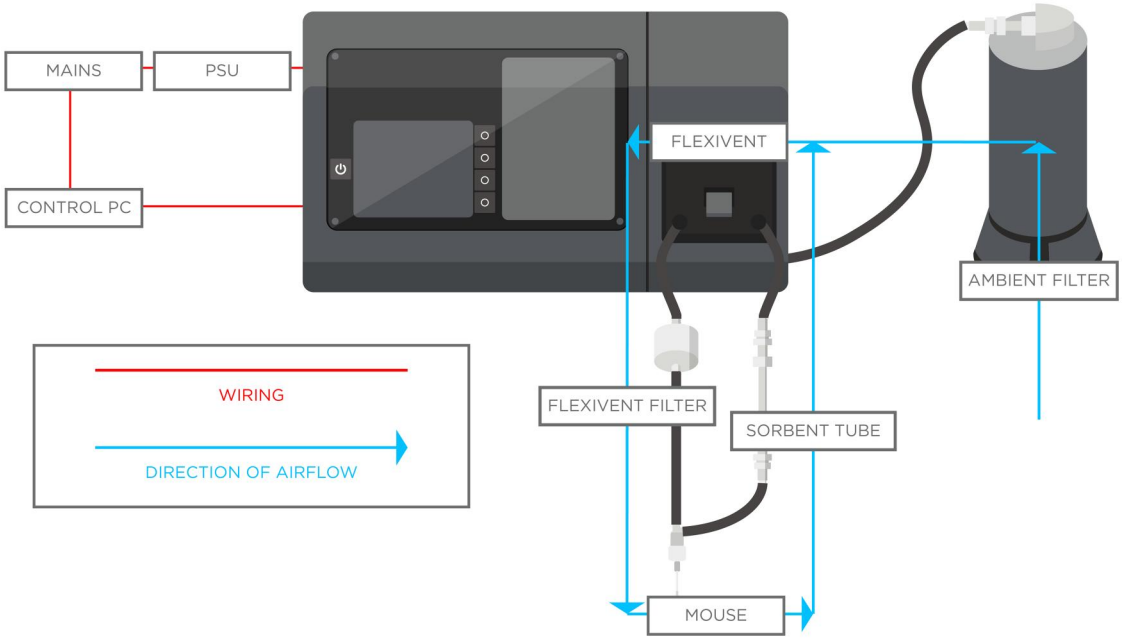

Figure 1. A schematic showing the mouse breath sampling system used for this study. This system consists of a flexiVent small animal ventilator, an ambient filter, and a flexiVent filter both connected to the ventilator by flexible tubing, and a sorbent tube for breath collection. The system blank sample was collected from inside the flexiVent, connected to the ambient filter.

a staged filtration system using ceramic disk and PTFE membranes, filtering down to 1.5 microns to prevent the charcoal from entering the flexiVent and was changed every two weeks. The flexiVent filter was used as a VOC filter fitted on the exit flow of the flexiVent. The filter contained a disc of filter paper (Grade 72 activated charcoal filter paper, Cole-Parmer®) impregnated with activated charcoal. The filter paper was changed at the start of each day during sampling.

A total of 15 breath samples, and 15 system blank samples were collected in a C2-CXXX-5149 bio-monitoring-inert-coated tube with Tenax TA/carbograph 5TD adsorbent material (Markes International, Llantrisant, UK), placed in the flow path between the intubated mouse and the return port on the flexiVent. One tube was used per mouse. After collection, the tubes were sealed with brass caps. The flow path was connected by flexible tubing using Viton rubber ranging from 30 to 125 mm in length. Prior to use, the Viton tubing was baked in an oven at 70°C for 24 hours to remove VOCs that may have been introduced during manufacturing or storage.

Mouse breath collection

A total of 15 breath samples were collected from 15 C57BL/6JRj intubated mice over four days. Prior to anesthetization the mice were removed from their main cages and placed in a holding cage and removed at the required time.

0.5 mg/kg Medetomidin, 5 mg/kg Midazolam and 0.05 mg/kg Fentanyl were applied intraperitoneally prior to tracheotomy. After deep anesthesia was reached, the trachea was prepared and a small incision (~2 mm) was made to insert an 18G metal cannula, which was fixed with a thread. Afterwards, mice were carefully coupled to the flexiVent small animal ventilator, equipped with a custom-made sampling device. Ventilation and breath sampling were controlled using flexiWare 7.5. Mechanical ventilation was started with a tidal volume of 6.5 mL/kg, a frequency of 150 breaths/min, and a positive end-expiratory pressure of 3 cmH_2O. To reduce the contamination levels for the lung in part of the flexiVent system, the volatile compounds generated by the flexiVent (Fig. 1) were attempted to be reduced by adding additional low volume, low pressure drop, and a filter prior to the mouse lung.

The flexiVent uses the forced oscillation technique to experimentally assess the lung function and respiratory system mechanics of mice. A detailed description of how to set up the device for normal mouse intubation has been described previously [30]. After the "mechanical ventilation" section, instead of moving on to "lung measurements," the mice were ventilated for five minutes to allow for more flushing of their lungs with VOC-free air. The existing collection tube was then removed and replaced with a fresh C2-CXXX-5149 bio-monitoring-inert-coated tube with Tenax TA/carbograph 5TD adsorbent material (Markes International, Llantrisant, UK) for sample collection. The mice were ventilated for a further 45 min for VOC breath collection, with a deep lung inflation procedure performed every two minutes during this period, resulting in approximately 1.8 L of breath obtained over 45 min. Further lung function testing could be proceeded either by removing the VOC collection hardware from the flexiVent, or moving the mouse to a second flexiVent system.

Mouse system background collection

For optimized untargeted analysis, a clean background is necessary to maximize the signal-to-noise ratio, so that signals of interest can be identified more easily. The background of a system can be measured through collection of air from various points in the system via a sorbent tube. Since VOCs are ubiquitous in the environment and therefore can be introduced through multiple components and points throughout the analytical process [31], it is important to ensure comparable collection and handling of breath and background samples [22]. A total of 15 system blank samples were collected immediately prior to mouse breath collection. System blank collection was performed using the same procedure used for mouse breath sampling, with the syringe driver removed from inside the flexiVent and the sorbent tube directly connected to the inlet of the ambient filter inside the system. Ambient blanks were collected at the start and end of the day in the room where breath collection was performed, to track environmental contamination (eight ambient blank samples in total).

Human breath sampling

Human participants

All participants provided written informed consent. Healthy adults (≥18 years) who met the inclusion criteria: were free of active respiratory infection symptoms or diagnoses (including COVID-19) and agreed to fast for at least 2 hours prior to providing a breath sample, were enrolled. A total of 13 healthy volunteers were recruited (demographic data in Table 1), and breath was collected over 4 days. Each study participant provided a single breath sample collected using the ReCIVA® Breath Sampler, and a paired system blank sample was collected at the same time (Owlstone Medical, Cambridge, UK).

Human breath collection

Breath collection was performed in a single room for all subjects. Breath samples were collected by adsorption onto the same C2-CXXX-5149 bio-monitoring-inert-coated tubes with Tenax TA/carbograph 5TD adsorbent material (Markes International, Llantrisant, UK) through the ReCIVA® Breath Sampler (software version BSC v3.4.0). Four sorbent tubes were pre-conditioned in a TC-20 (Markes International) by an N2 flow at 20 psi and 320°C for 2 hours. Approximately 1.25 L of breath was sampled over 15 min in each tube at 225 mL/min. The ReCIVA was configured to exclude air from the upper airway to mainly sample breath from the lower lung. Ambient contamination was minimized using the CASPER™ Portable Air Supply, which has been used previously [32–34]. The tubes were purged with a TD-100 (Markes International Ltd Llantrisant, UK), stored at a temperature of 4°C–8°C, and then analyzed.

Table 1. The age and gender of the 13 participants of this study.

Age	Gender
31	Female
24	Male
50	Male
34	Female
27	Female
37	Female
42	Male
25	Male
29	Male
29	Female
41	Male
30	Male
26	Female

Human system background collection

Prior to breath sampling, a system blank sample was collected to allow "on-breath" compounds to be determined, versus those present in the system background. These used the above ReCIVA settings but with the fractionation algorithm modified to continuous to sample a continuous volume from the headspace of the device. New consumables were used for each blank sample (including mouthpiece and cartridge), and the mouthpiece inlet was sealed, to provide a representative sample from the device without a subject present.

Breath analysis

Mice and human breath samples were both analyzed using the Breath Biopsy® OMNI® analysis method [32, 35], within a short period of time after collection (approximately 30 days median) to minimize the impact of storage effects on the analysis. All samples were analyzed using high-resolution thermal desorption gas chromatography and accurate mass spectrometry (TD-GC-MS) using the Q Exactive™ GC Orbitrap™ GC-MS/MS, with specific analytical settings shown in Supplementary Table S1.

Calculation

Data processing and analysis

An untargeted feature table was extracted from the raw chromatograms, and each molecular feature in the feature table was represented by a deconvoluted mass spectrum. The data were generated using an untargeted feature extraction workflow. The steps of the workflow are summarized in the Supplementary Material.

To mitigate sources of analytical drift and variability over the TD-GC-MS run, a normalization algorithm developed by Owlstone Medical was applied. The algorithm utilizes eight isotopically labeled compounds spread throughout the chromatogram based on retention time, spiked at known concentrations as internal standards. Scaling was applied to all molecular features using combinations of these internal standards' peak areas to correct their peak areas. Several quality checks assess normalization performance, the most important being principal component analysis (PCA), inspecting which PC sequence (analytical batch) effects are observed. After normalization, sequence trends are no longer observed in the first few PCs and are pushed into later PCs where they explain negligible amounts of variance in the data.

Three metrics that have been used previously to distinguish on-breath and background VOCs [36] were used in this study to evaluate the performance capabilities of the developed breath sampling method (Fig. 2). These metrics can be used to track performance improvements over time in response to optimization steps. These metrics are designed to give both the breadth of compounds that may be on-breath, as well as impart the ability to rank on-breath VOCs with different confidences due to the ability to cross-reference the classification between different metrics. These metrics were:

- Type 1 (standard deviation): the mean signal for each VOC was calculated across all system blank samples, and a VOC was considered on-breath if the signal in at least 50% of the breath samples exceeded three standard deviations above the mean signal of the system blanks.
- Type 2 (paired t-test): the VOC signal was compared to the signal in the paired system blank sample. A VOC was considered on-breath if it achieved a *P*-value of $\leq$.05 with a mean fold change $\geq$2.
- Type 3 (receiver operating characteristic area under the curve, ROC-AUC): a ROC-AUC was generated for each VOC. A VOC was considered on-breath if the fold difference between breath and blank samples were >1, and the ROC-AUC value was $\geq$0.8.

VOC identification

VOC identification was carried out using several methodologies, each producing an identity with a different level of associated confidence. The highest confidence identification was carried out via the analysis of pure chemical standards alongside the breath samples; VOCs in the breath sample were matched spectrally and via retention time to the chemical standard. VOCs that fell within a tight retention time window and achieved a spectral match score (SI and RSI) >800 were assigned the identity accordingly. These identities are referred to as Tier 1 IDs.

Second in confidence are those VOCs that were matched against Owlstone Medical's Breath Biopsy high-resolution accurate mass (HRAM) library. The library was built over time by analyzing standards on the same analytical method used to analyze the breath samples. VOCs were assigned an identity if they fell

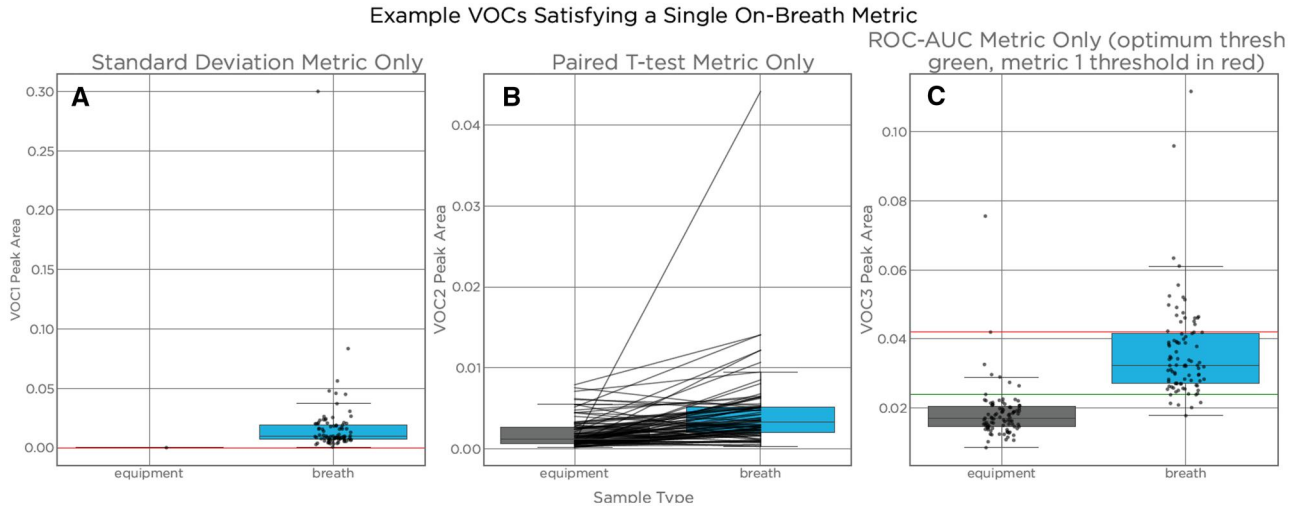

Figure 2. Summary of the on-breath metrics used to discriminate mouse breath volatiles from background features. For metric 3 (ROC-AUC metric), the red line indicates the threshold, and the green line indicates the optimal threshold. The green line being lower demonstrates that the data can still be separable to a high level (detected by metric 3 ROC), even if the VOC is not classified as on-breath by metric 1.

within a narrow retention index window and achieved a spectral match score (SI or RSI) >800 against the HRAM library.

Finally, lower confidence identification was carried out by spectral match against the National Institute of Standards and Technology (NIST) library of chemicals.

Statistical analysis

Data analysis was performed using the Python programming language (Python Software Foundation, Python Language Reference, version 3.7 https://www.python.org/).

Results

Method background and qualitative assessment

In order to conduct an initial assessment of the relative background of the flexiVent system, 500 mL of air was collected from each of the five different scenarios of blank samples (Control—an unused sorbent tube, A—ambient air in the room, B—filtered ambient air, C—unfiltered air in the flexiVent system, D—filtered air in the flexiVent system) while the system was in operation (Fig. 3). Note that this volume was intended to provide a relative background comparison, but a larger volume was used during actual mouse breath sampling (representing approximately 15 min of operating time vs. 45 min).

The data presented in Fig. 3 demonstrate that contamination originating from the flexiVent can be removed by the presence of a second filter. Subsequent experiments showed that optimal performance can be maintained by replacing the filter paper at the start of each test day.

Raw mouse and system blank features

Mouse breath and paired system blank samples were separately collected using the flexiVent system. A PCA was fit on the mouse feature table, showing a much tighter clustering of blank samples (orange) than breath samples (blue) (Fig. 4). This is to be expected, as biological variance may be greater than the technical repeatability of blank samples. The plot also shows a separation of breath and blank samples on PC1. While there is some overlap, this may again be explained by the variance, and amount of, biological signal.

To reliably distinguish between VOCs that have originated from physiological processes in the body and those that have been inhaled from the environment, a consistently low background from system blank samples is a significant factor. Together with Fig. 2, this demonstrates that the breath collection methodology developed shows a clear ability to generate quality GC-MS data that can be used for downstream analysis.

Quantitative metrics to identify on-breath compounds

A total number of 472 molecular features were identified in mouse breath. However, as previously mentioned, many of these compounds may be from background sources, rather than originating from physiological processes in the body.

To distinguish between compounds that may be originating from the breath (referred as “on-breath”) from those originating from the background air, three different types of metrics were developed to establish quantitative classification thresholds (Fig. 4).

Using all three metrics the number of on-breath VOCs were:

- 21 with Type 1
- 30 with Type 2
- 56 with Type 3

A list of the most confidently identified VOCs can be viewed in Table 2, the full list of on-breath compounds in mice is in Supplementary Table S2.

There is a good degree of overlap between the compounds identified between the three metrics (Fig. 5), leaving a total range of 15–66 on-breath VOCs identified in mice depending on how stringently a definition is set (must be on-breath by one, two, or three metrics).

All the compounds identified using Type 1 were also seen in other types, whereas 31 VOCs were only considered as on-breath in Type 3. Typically, Type 1 (standard deviation) is the most conservative metric, whereas Type 3 (ROC-AUC) is the least

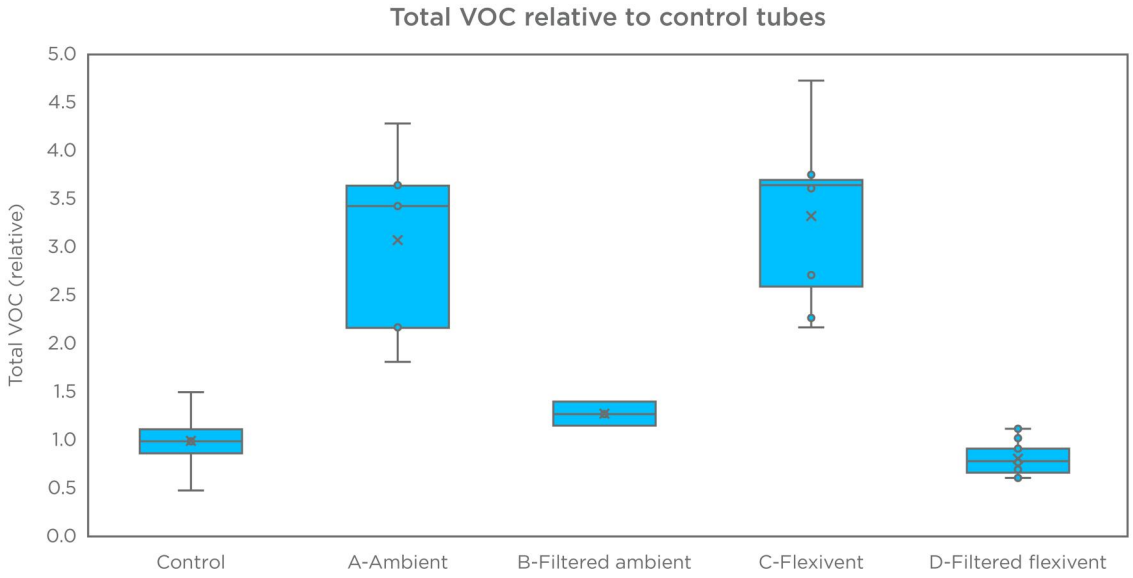

Figure 3. The different levels of total relative VOC concentration in different blank samples to test the background of the flexiVent system. The control was an empty sorbent tube, A was ambient air in the room, B was the filtered ambient air, C was the air in the flexiVent system, and D was the filtered air in the flexiVent system. The X symbol indicates the mean, the line indicates the median.

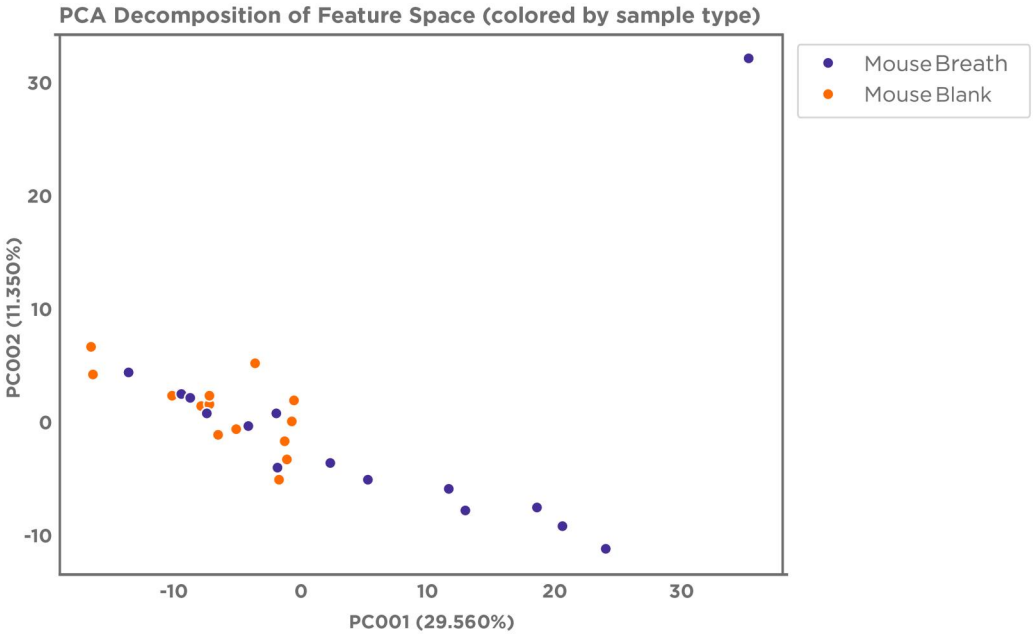

Figure 4. A PCA of mouse breath features vs system blank features.

Table 2. A list of the Tier 1 (highest confidence VOC identification via spectral and RT match against pure chemical standard) on-breath VOCs that were confidently identified in mice.

Compound name	Tiers	On-breath metrics (mice)	On-breath metrics (human)
O-cresol	Tier 1	Type 1, 2, 3	Type 1, 2, 3
3-Carene	Tier 1	Type 1, 2, 3	Type 1, 2, 3
2-Butanol	Tier 1	Type 1, 2, 3	N/A
Beta-pinene	Tier 1	Type 1, 2, 3	Type 1, 2, 3
Alpha-pinene	Tier 1	Type 1, 2, 3	Type 1, 2, 3
Nonane, 3-methyl-	Tier 1	Type 1, 3	N/A
Ethanol, 2-(2-ethoxyethoxy)-	Tier 1	Type 2	Type 2, 3
Cyclopentene	Tier 1	Type 2, 3	Type 1, 2, 3
Tetrachloroethylene	Tier 1	Type 3	Type 1, 2, 3
Pyridine	Tier 1	Type 3	Type 1, 2, 3
Pentane	Tier 1	Type 3	N/A
Nonane, 2-methyl-	Tier 1	Type 3	N/A
Heptasiloxane, hexadecamethyl-	Tier 1	Type 3	N/A
Decane	Tier 1	Type 3	N/A
Benzene	Tier 1	Type 3	Type 3

Compounds found common in both humans and mice are highlighted in bold.

conservative, which is reflected in the results of the relative numbers of on-breath VOCs classified. The combined three metrics maximize the opportunity for discovering potentially informative VOCs that can be considered as on-breath, as well as provide higher confidence in the results. For example, a VOC that is classified as on-breath in multiple metrics can lower the probability that it is a false positive hit, while still allowing visibility of the wider range of potential lower-confidence VOCs that are captured by at least one metric. This can give a broader understanding of the composition of the breath matrix and provide adjustable on-breath stringency throughout different stages of pre-clinical studies.

Proof-of-principle: the commonality with human breath data

In order to assess the transferability of the mouse breath VOCs to human VOCs and highlight data analysis improvements, the same metrics were used to quantify the number of on-breath VOCs in human breath compared to paired system blanks sampled. The common high-confidence VOCs between mice and humans are bolded in Table 2, a list of the total 49 on-breath VOCs common to both human and mouse breath can be viewed in Supplementary Table S3. The table also shows how the combination of different stringency of on-breath metrics can help capture a wider list of translatable compounds between mouse and human.

A PCA was plotted for human breath and blank samples (Fig. 6), which combined with the mouse breath and blank sample PCA presented in Fig. 4, showed a clear separability in PC1 by species.

PC2 shows a separability between breath and blank for both species, showing the presence of biological signals in the breath samples. The average fold change in signal intensity compared to blank for all VOCs for both mouse and human breath using all metrics were compared (Supplementary Fig. S1). The majority of compounds identified as in common between mouse and human breath with the greatest fold change in breath to blank were associated with the gut microbiome or were plant-derived. Examples of these include trimethylamine (TMA) and dimethyl sulfone (Fig. 7). These were identified via match against

Owlstone's HRAM library (Supplementary Table S3). There are certain compounds that appear to be specific to mouse or human breath. For example, methyl nitrate (NIST match—Supplementary Table S2) and 2-butanol (Tier 1) appeared to be exclusively on-breath in mice, appearing with a high signal intensity (Fig. 7).

Discussion

Using animal models for breath biomarker discovery work allows a more controlled study, reducing the variability from inter-individuals, diet, and other environmental factors. The reduced risk of false biomarker discovery accelerates the process of identifying potential biomarkers for validation in clinical trials. Here, we present a newly developed method that captures mouse breath with superior control over background contamination, providing more confidence that the VOCs measured are genuinely breath-borne. Analysis of chemical standards has provided high confidence VOC identification, allowing genuine comparison between mouse and human breath. Both an increased signal-to-noise ratio and preservation of VOC integrity are likely to increase the number of VOCs detectable from breath, providing a higher chance for identifying potential biomarkers from discovery work. We also developed three different metrics (Fig. 3, Type 1, 2, and 3) for VOC "on-breath" classification between breath and system blank samples. The use between one or multiple metrics allow different stringency for distinguishing on-breath VOCs from background air, accommodating different study designs and purposes.

One of the challenges across pre-clinical and clinical breath research studies is the control of signal-to-noise. In this study, we addressed this challenge by employing low-background breath collection from intubated mice. We modified one major step of the flexiVent breath collection, which is replacing the lung function measurements with additional ventilation, namely, the flexiVent filter, for breath collection. Although this method requires anesthetics, breath is collected strictly from the lower tract of the lungs. The additional ventilation minimizes contamination from the upper respiratory tract during breath collection and enhances the representation of blood-circulated metabolic processes exchanged at the alveoli in the lungs. Additionally, the sampling method substantially reduces the amount of water collected in breath samples, which may contribute to unexpected signal peaks during sample processing, thereby affecting data accuracy. Reducing water content can also preserve the integrity of the sampled VOCs, as water aerosols can react with VOCs or alter their chemical properties, leading to artifact formation during sample collection and storage. Given that VOC concentrations depend on the breath volume collected, preventing loss during collection is crucial. The intubation

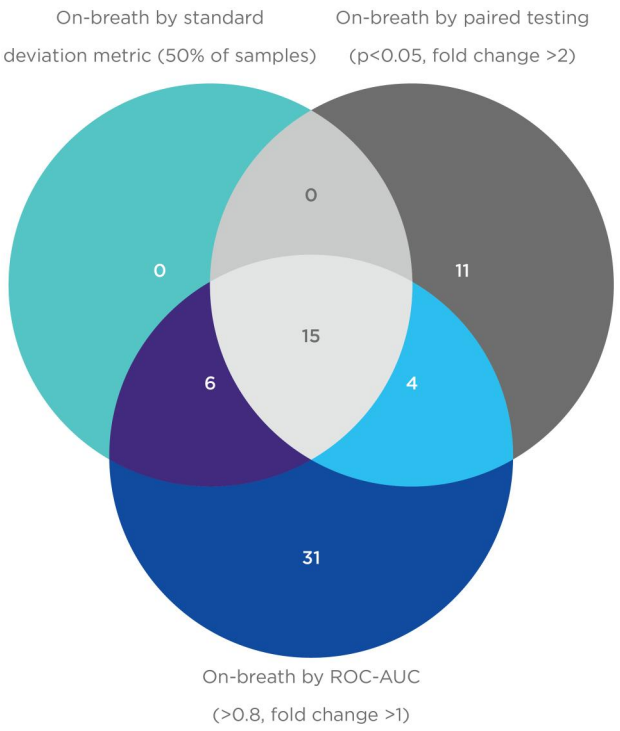

Figure 5. The number of overlapping identified "on-breath" VOCs by each metric type.

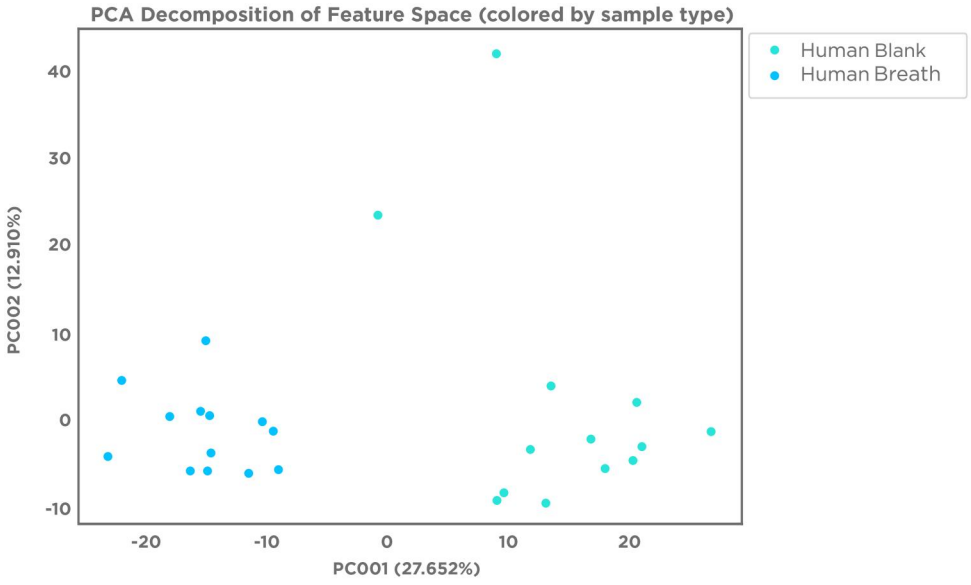

Figure 6. A PCA of human breath features, and system blank features, to compare with the mouse breath and system blank features presented in Figure 3

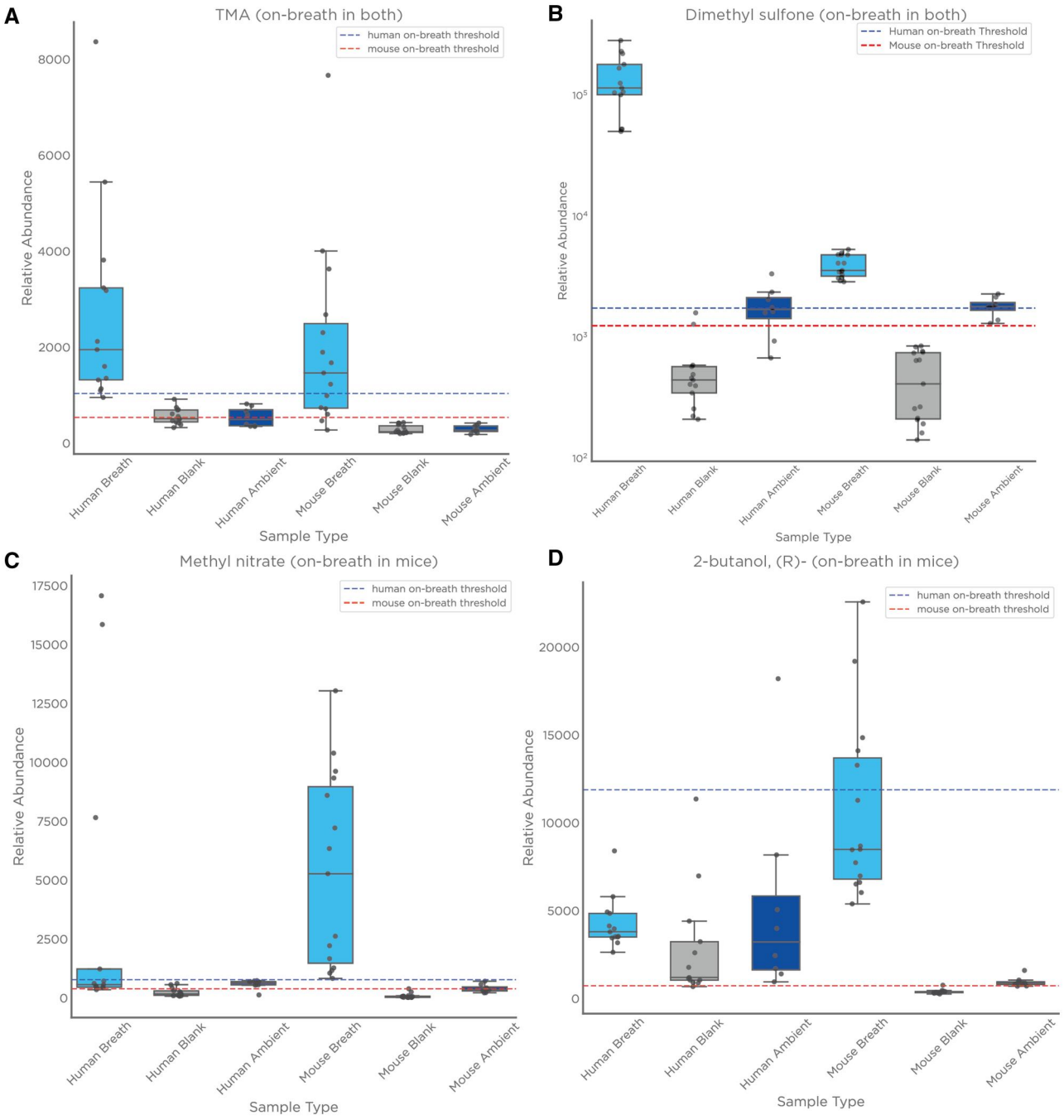

Figure 7. (**A**) Trimethylamine (TMA) is an example of an identified on-breath compound in both humans and mice that is significantly different from background signal. (**B**) Dimethyl sulfone is on-breath in both humans and mice (Note log axis). (**C**) Methyl nitrate and (**D**) 2-butanol are on-breath compounds in mice, but not in humans. The significantly cleaner background for mouse blank samples can be seen in this example. The thresholds are defined as the mean ±3 standard deviations.

method in this study has precise volume control, eliminating potential losses. Furthermore, periodic 'deep lung' inflation every two minutes potentially improves the detection of VOCs from the lower lung. These features of the novel breath collection method presented here are especially beneficial for biomarker discovery studies, where capturing relevant compounds in a specific disease area and at a wider range is vital.

Existing literature has demonstrated the feasibility of murine models for breath VOC research, though the number of studies is limited [24, 26]. These studies all utilized a non-anesthetic approach with breath collection methods. Some require additional preparation work to ensure the mice were cleaned with no contamination from fur due to headspace collection using a respiratory chamber [24], while other studies collect breath through a more closed or semi-closed environment [24–26]. We compared one of these studies, which placed mice in glass restrainers and had them breathe through a nose cone with filtered air using CASPER [25]. The study targeted 27 VOCs and we found 18 of them present within our untargeted analysis dataset. These compounds are mainly alkanes and aldehydes, which are indicators of inflammation and oxidative stress, with confirmed IDs. However, these compounds were not considered on-breath using the three metrics we

developed. As the study by Hintzen *et al.* did not describe clearly how breath compounds differed from system blanks, the differences in the findings of healthy mice used in both studies could be due to the analytical approach we used to distinguish "on-breath" compounds. With the three metrics utilized to distinguish on-breath compounds in our dataset, only 3-carene and alpha-pinene, two compounds originating from a plant-based diet, were found in both studies. Interestingly, our findings also indicate that three exogenous VOCs often associated with environmental sources-namely, xylene, toluene, and the tentatively identified styrene from NIST- were not identified as on-breath compounds within our dataset. This suggests that the intubation method and analytical approach enabled the exclusion of these environmental compounds, allowing discrimination of the on-breath compounds in our analysis. As on-breath compounds are determined by the abundance differences between breath and blank samples, it is important to note that upon utilizing this mouse breath system in a disease model, more compounds reflecting the physiological state as opposed to healthy, homeostasis mouse, may be detectable on-breath—such as patterns of lipid peroxidation products in inflammatory processes [37].

We compared the potential translatability of the on-breath VOCs identified in mice using this new method with those found in human breath. The human breath samples in this study were collected while subjects breathed filtered air using CASPER, which removes most contaminating volatile compounds from ambient air and has been utilized in other studies [32–34], Following the same on-breath metrics, we identified more VOCs classified as on-breath in humans across all metric types compared to mice (Supplementary Tables S2 and S3). The PCA (Fig. 6), also indicated clear separability based on species. This is not unexpected due to the relative body size difference, the difference in lung volume, and the differences in environmental exposure between mice and humans. Some VOCs, such as methyl nitrate and 2-butanol, were exclusively on-breath in mice (Fig. 7), which may represent a species-specific difference in breath composition or reflect a difference in environmental or dietary exposure. Despite the differences, 49 on-breath VOCs were found common between human and mouse (Supplementary Table S3). A substantial portion of these compounds with confirmed identities originated from the microbiome, including TMA [38, 39], 2,3-butanoediol [40], dimethyl sulfone [41], and o-cresol [42, 43]. This is important as these volatile compounds are targets for biomedical research into the intricate interplay between the microbiota and disease. TMA, predominately produced through microbial metabolism of the dietary-derived compounds like choline, carnitine, and betaine in the gut, has demonstrated associations with kidney disease and colorectal cancer in human studies. These strong signals of microbial-produced VOCs measured in breath also suggest that, unlike isoflurane- which is known to affect gut microbiome diversity- the use of Medetomidin/Midazolam/Fentanyl for anesthesia did not significantly impact metabolic changes [44]. Other volatile compounds common between human and mice breath include those originating from plant-based dietary sources, such as alpha-pinene and beta-pinene. The notable presence of certain common on-breath VOCs in mouse and human breath implies the potential to investigate disease-relevant volatile compounds for human clinical studies in mouse models. Together, this method supports a promising avenue for translating the findings from pre-clinical mouse models to clinical human studies.

After developing a system to understand the VOC composition of breath from healthy mice, future studies could use a disease mouse model in comparison to healthy mice. This could amplify the detection of significant signals as a characteristic overabundance of certain compounds related to the disease pathology. For example, the specific VOCs in the breath of a respiratory disease mouse model that differ from healthy mice could serve as candidate biomarkers for the underlying pathophysiology. Using the method developing in this study, it is also possible to simultaneously collect lung function data via the flexiVent, and correlate this to VOC abundance. Given that the lung is often modelled as a resistor and a capacitor for lung function assessment, which is how the flexiVent system was originally designed, it is essential that the additional flexiVent filter introduces negligible additional resistance and volume. Further assessment to allow for simultaneous VOC and lung function data collection will offer researchers the opportunity to explore the relationship between specific VOCs and lung function metrics. For drug discovery efforts against respiratory disease, this system could test the success of drug treatments on the disease mouse models, observe whether the VOC composition of breath altered back towards the healthy breath state, and whether this correlates to lung function improvements.

Conclusion

In this work, we present a pre-clinical method for the capture and analysis of the volatile compounds contained in mouse breath, offering advantages over existing methods in the field. The results of this study present a reliable mouse breath sampling and analysis platform that can be used to compare the composition of mouse breath with human breath and establish mice as viable animal model for the pre-clinical study of breath biomarkers.

- We saw 472 compounds in total in mouse breath.
- We identified 15.47% (73) of these compounds as 'on-breath', meaning we reliably distinguished them from background contaminating signals based on three quantitative metrics.
- There were 49 common VOCs identified between mouse and human breath.
- The compounds identified as on-breath and shared between humans and mice were linked to suspected biological functions.

Supplementary data

Supplementary data is available at *Biology Methods and Protocols* online.

Conflict of interest statement. B.B. is the CEO and Co-Founder of Owlstone Medical Ltd. A.T., M.B., O.B., H.C., J.G., S.S., L.P., and M. A. are employees of Owlstone Medical Ltd. S.B. and K.G.-K. are employees of Boehringer Ingelheim.

Funding

None declared.

Data availability

The data underlying this article are available in the article and in its online supplementary material.

Ethical statement

This study was performed in accordance with the Nurenberg Code. The human study was conducted in accordance with the principles embodied in the Declaration of Helsinki and local statutory requirements. All adult participants provided written informed consent to participate in this study. Ethical approval for the animal study was obtained from the regional board for animal care and welfare Regierungspräsidium Tübingen, Germany—approval: TVV-22-002-G.

References

1. Schleich FN, Zanella D, Stefanuto P-H *et al.* Exhaled volatile organic compounds are able to discriminate between neutrophilic and eosinophilic asthma. *Am J Respir Crit Care Med* 2019;**200**:444–53.
2. Bush A. Pathophysiological mechanisms of asthma. *Front Pediatr* 2019;**7**:68.
3. Ibrahim B, Basanta M, Cadden P *et al.* Non-invasive phenotyping using exhaled volatile organic compounds in asthma. *Thorax* 2011;**66**:804–9.
4. van der Kamp MR, Driessen JMM, van der Schee MP *et al.* Volatile organic breath components and exercise induced bronchoconstriction in asthmatic children. *Allergy Asthma Clin Immunol* 2021;**17**:121.
5. Djukanović R, Brinkman P, Kolmert J, SoMOSA study team and the U-BIOPRED study team *et al.* Biomarker predictors of clinical efficacy of the anti-IgE biologic, omalizumab, in severe asthma in adults: results of the SoMOSA Study. *Am J Respir Crit Care Med* 2024;**210**:288–97.
6. Peltrini R, Cordell RL, Wilde M *et al.* Discovery and validation of a volatile signature of eosinophilic airway inflammation in asthma. *Am J Respir Crit Care Med* 2024;**210**:1101–12.
7. Mochalski P, King J, Mayhew CA, Unterkofler K. Modelling of breath and various blood volatilomic profiles-implications for breath volatile analysis. *Molecules* 2022;**27**:2381.
8. Read NW, Al-Janabi MN, Bates TE *et al.* Interpretation of the breath hydrogen profile obtained after ingesting a solid meal containing unabsorbable carbohydrate. *Gut* 1985;**26**:834–42.
9. Chou H, Godbeer L, Allsworth M *et al.* Progress and challenges of developing volatile metabolites from exhaled breath as a biomarker platform. *Metabolomics* 2024;**20**:72.
10. Ratiu IA, Ligor T, Bocos-Bintintan V *et al.* Volatile organic compounds in exhaled breath as fingerprints of lung cancer, asthma and COPD. *J Clin Med* 2020;**10**:32.
11. Van Malderen K, De Winter BY, De Man JG *et al.* Volatomics in inflammatory bowel disease and irritable bowel syndrome. *EBioMedicine* 2020;**54**:102725.
12. Fernández Del Río R, O'Hara ME, Holt A *et al.* Volatile biomarkers in breath associated with liver cirrhosis–comparisons of pre- and post-liver transplant breath samples. *EBioMedicine* 2015;**2**:1243–50.
13. Hanna GB, Boshier PR, Markar SR, Romano A. Accuracy and methodologic challenges of volatile organic compound–based exhaled breath tests for cancer diagnosis: a systematic review and meta-analysis. *JAMA Oncol* 2019;**5**:e182815.
14. Patel SK, Pratap CB, Jain AK *et al.* Diagnosis of Helicobacter pylori: what should be the gold standard? *World J Gastroenterol* 2014;**20**:12847–59.
15. Heffler E, Carpagnano GE, Favero E *et al.* Fractional Exhaled Nitric Oxide (FENO) in the management of asthma: a position paper of the Italian Respiratory Society (SIP/IRS) and Italian Society of Allergy, Asthma and Clinical Immunology (SIAAIC). *Multidiscip Respir Med* 2020;**15**:36.
16. Rezaie A, Buresi M, Lembo A *et al.* Hydrogen and methane-based breath testing in gastrointestinal disorders: the North American Consensus. *Am J Gastroenterol* 2017;**112**:775–84.
17. Issitt T, Wiggins L, Veysey M *et al.* Volatile compounds in human breath: critical review and meta-analysis. *J Breath Res* 2022;**16**:024001.
18. Jia Z, Patra A, Kutty VK, Venkatesan T. Critical review of volatile organic compound analysis in breath and in vitro cell culture for detection of lung cancer. *Metabolites* 2019;**9**:52.
19. Martin L, Hutchens M, Hawkins C, Radnov A. How much do clinical trials cost? *Nat Rev Drug Discov* 2017;**16**:381–2.
20. Drabińska N, Flynn C, Ratcliffe N *et al.* A literature survey of all volatiles from healthy human breath and bodily fluids: the human volatilome. *J Breath Res* 2021;**15**:034001.
21. Herbig J, Beauchamp J. Towards standardization in the analysis of breath gas volatiles. *J Breath Res* 2014;**8**:037101.
22. Di Gilio A, Palmisani J, Ventrella G *et al.* Breath analysis: comparison among methodological approaches for breath sampling. *Molecules* 2020;**25**:5823.
23. Westphal K, Dudzik D, Waszczuk-Jankowska M *et al.* Common Strategies and Factors Affecting Off-Line Breath Sampling and Volatile Organic Compounds Analysis Using Thermal Desorption-Gas Chromatography-Mass Spectrometry (TD-GC-MS). *Metabolites* 2022;**13**:8.
24. Kistler M, Szymczak W, Fedrigo M *et al.* Effects of diet-matrix on volatile organic compounds in breath in diet-induced obese mice. *J Breath Res* 2014;**8**:016004.
25. Hintzen KFH, Smolinska A, Mommers AGR *et al.* Non-invasive breath collection in murine models using a newly developed sampling device. *J Breath Res* 2022;**16**:027102.
26. Jonasson S, Magnusson R, Wingfors H *et al.* Potential exhaled breath biomarkers identified in chlorine-exposed mice. *J Anal Toxicol* 2024;**48**:171–9.
27. Mochalski P, King J, Mayhew CA, Unterkofler K. A review on isoprene in human breath. *J Breath Res* 2023;**17**:037101.
28. Saasa V, Beukes M, Lemmer Y, Mwakikunga B. Blood ketone bodies and breath acetone analysis and their correlations in Type 2 diabetes mellitus. *Diagnostics (Basel)* 2019;**9**:224.
29. Amann A, Costello B. D L, Miekisch W *et al.* The human volatilome: volatile organic compounds (VOCs) in exhaled breath, skin emanations, urine, feces and saliva. *J Breath Res* 2014;**8**:034001.
30. McGovern TK, Robichaud A, Fereydoonzad L *et al.* Evaluation of respiratory system mechanics in mice using the forced oscillation technique. *J Vis Exp* 2013;**75**:50172.
31. Pham YL, Holz O, Beauchamp J. Emissions and uptake of volatiles by sampling components in breath analysis. *J Breath Res* 2023;**17**:037102.
32. Ferrandino G, Ricciardi F, Murgia A *et al.* Exogenous Volatile Organic Compound (EVOC®) breath testing maximizes classification performance for subjects with cirrhosis and reveals signs of portal hypertension. *Biomedicines* 2023;**11**:2957.
33. Chou H, Arthur K, Shaw E *et al.* Metabolic insights at the finish line: deciphering physiological changes in ultramarathon runners through breath VOC analysis. *J Breath Res* 2024;**18**:026008.
34. Doran SLF, Romano A, Hanna GB. Optimisation of sampling parameters for standardised exhaled breath sampling. *J Breath Res* 2017;**12**:016007.
35. OMNI® - The breath biomarker discovery service [Internet]. https://www.owlstonemedical.com/products/breath-biopsy-omni/ (13 November 2023, date last accessed).
36. Arulvasan W, Chou H, Greenwood J *et al.* High-quality identification of volatile organic compounds (VOCs) originating from breath. *Metabolomics* 2024;6**20**:102.

37. Ratcliffe N, Wieczorek T, Drabińska N *et al.* A mechanistic study and review of volatile products from peroxidation of unsaturated fatty acids: an aid to understanding the origins of volatile organic compounds from the human body. *J Breath Res* 2020; **14**:034001.
38. Grabowska-Polanowska B, Faber J, Skowron M *et al.* Detection of potential chronic kidney disease markers in breath using gas chromatography with mass-spectral detection coupled with thermal desorption method. *J Chromatogr A* 2013; **1301**:179–89.
39. Bain MA, Fornasini G, Evans AM. Trimethylamine: metabolic, pharmacokinetic and safety aspects. *Curr Drug Metab* 2005; **6**:227–40.
40. Ji XJ, Huang H, Ouyang PK. Microbial 2,3-butanediol production: a state-of-the-art review. *Biotechnol Adv* 2011;**29**:351–64.
41. He X, Slupsky CM. Metabolic fingerprint of dimethyl sulfone ($DMSO_2$) in microbial–mammalian co-metabolism. *J Proteome Res* 2014;**13**:5281–92.
42. Bermudez-Martin P, Becker JAJ, Caramello N *et al.* The microbial metabolite p-Cresol induces autistic-like behaviors in mice by remodeling the gut microbiota. *Microbiome* 2021;**9**:157.
43. Persico AM, Napolioni V. Urinary p-cresol in autism spectrum disorder. *Neurotoxicol Teratol* 2013;**36**:82–90.
44. Serbanescu MA, Mathena RP, Xu J *et al.* General anesthesia alters the diversity and composition of the intestinal microbiota in mice. *Anesth Analg* 2019;**129**:e126–9–e129.

Biology Methods and Protocols, 2024, 9, 1–11
https://doi.org/10.1093/biomethods/bpae087
Methods Article