Report for PhD register

Microbial Volatile Organic Compounds - A Path Towards Rapid Non-Invasive Detection of Wound Infections

by
Shane Fitzgerald, BSc

PhD Thesis

Dublin City University

School of Chemical Sciences

Supervisor: Dr. Aoife Morrin
Co-supervisor: Dr. Linda Holland

Submitted: May 2022

1

Declaration of Authorship

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy is entirely my own work, and that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed:
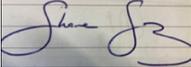
Print: Shane Fitzgerald

Student No.: 14482908

Date: 05/05/22

Pre-examination copy
Submitted 27th May 2022

Thank you to my family, girlfriend, friends and dogs (past and present) for all of their support over the last four years.

Thank you also to my supervisors, Dr. Aoife Morrin and Dr. Linda Holland, and also to my past research colleague Dr. Emer Duffy who provided direction, support and guidance to me in the early stages of this project.

Shane

Table of Contents

Pre-examination copy
Submitted 27th May 2022

Pre-examination copy
Submitted 27th May 2022

6

List of Abbreviations

VOC	Volatile organic compound
mVOC	Microbial volatile organic compound
ATP	Adenosine triphosphate
FA	Fatty acid
SCFA	Short chain fatty acid
CF	Cystic fibrosis
ICU	Intensive care unit
VAP	Ventillator acquired pneumonia
BAL	Bronchoalveolar lavage
TB	Tuberculosis
DFU	Diabetic foot ulcer
SPME	Solid phase microextraction
TD	Thermal desorption
HS	Headspace
SIFT-MS	Selective ion-flow tube - mass spectrometry
SESI-MS	Secondary electrospray - mass spectrometry
PTR-MS	Proton transfer reaction - mass spectrometry
GC-MS	Gas chromatography - mass spectrometry
e-nose	Electronic nose
CF	Cystic fibrosis
PDMS	Polydimethylsilicone
PEG	Poly (ethlene glycol)
PA	Polyacrylate
DVB	Divinylbenzene
Car	Carboxen
OD	Optical density
TSB	Tryptone soy broth
MHB	Mueller-Hinton broth
BHI	Brain heart infusion
LB	Luria-Bertani
PCA	Principal component analysis
HC	Hierarchical clustering

Microbial Volatile Organic Compounds - A Path Towards Rapid Non-Invasive Detection of Wound Infections

Shane Fitzgerald

The detection of microbial volatile organic compounds (VOCs) has previously demonstrated potential as a non-invasive diagnostic tool for a variety of infectious diseases and disorders. The objectives of this thesis were to firstly, illustrate and describe the close association between microbial and clinical volatilomics; secondly, to characterise the volatilomes of key wound-associated pathogens and to explore the factors influencing their VOC emission; and finally, to demonstrate the application of VOC analysis for the detection of wound infections. Chapter 1 of this report is a review of past and current literature surrounding the field of microbial and clinical volatilomics. The most common experimental methods used across microbial and clinical volatilomic studies are discussed with the aim of highlighting the critical need for standardization of these techniques across the field. This chapter finally illustrates the close association between microbial VOCs and disease and describes potential opportunities for clinical applications in the future. Chapter 2 describes the comprehensive *in vitro* volatilomic profiling of prevalent wound-associated bacterial pathogens. In this work, species- and strain-level volatilomic diversity were explored by utilizing a simple experimental workflow coupled with robust multivariate analysis techniques. Temporal stability of microbial VOC emissions was also investigated and measured against to cell growth. Chapter 3 was a further investigation of these pathogens. The aims of this study were to examine the influence of different nutritional media on the VOC output of the bacterial pathogens and to investigate strain-level volatilomic differences in VOC emission kinetics. In Chapter 4, the focus of investigation was shifted to fungal pathogens to characterise the factors surrounding VOC emission in multiple *Candida* species. In this work, volatile metabolites from 10 clinical strains of planktonic *C. parapsilosis* and one strain of planktonic *C. albicans* were profiled. The effect of biofilm formation on the *C. parapsilosis* volatilomes was investigated for the first time by comparing volatilomes of a biofilm-positive strain and a biofilm-negative strain over time using a novel sampling approach. In the final chapter of this thesis, our rapid non-invasive experimental workflow was employed for the analysis of wound swab samples. 23 participants (26 wounds total, 15 infected; 11 non-infected) were included in this work. The volatilomes of infected and non-infected wound samples are characterised and compound-level differences between them are described in this chapter. The results of this ongoing work provide clear insight into the potential of volatilomics for future clinical applications. Overall this collective work demonstrates the close association between microbial and clinical volatilomics and highlights the clear potential for volatilomics to be used for clinical diagnoses of wound infections. This thesis concludes with a short discussion of the future outlook of this work.

Chapter 1: Comprehensive microbial volatilomics as a novel route towards rapid detection of infectious diseases

Abstract

The emergence and spread of novel and old infectious diseases have been a significant challenge faced by humans across history. Although the development of early detection systems and novel therapeutics have significantly reduced the impact of infectious diseases in the modern world, the emergence of the SARS-Cov-2 virus in 2019 highlighted how unprepared it was for such a global emergency. Development of rapid non-invasive disease detection techniques have since been at the forefront of scientific research. The detection of volatile organic compounds (VOCs) emitted from infecting pathogens as metabolic by-products or resulting from inflammation present a unique opportunity for non-invasive rapid detection of disease. This review paper aims illustrate the potential clinical applications of volatilomics in the future. The fundamental metabolic pathways by which microbes produce volatile metabolites will be described to provide context on how these compounds are generated. The most common experimental methods used across microbial and clinical volatilomic studies will then be discussed with the aim of highlighting the critical need for standardisation of these techniques across the field. Past and current microbial volatilomic research on several prominent bacterial and fungal pathogens is then summarised before a discussion of clinical volatilomics. Emerging evidence of applications of volatilomic profiling for the rapid discrimination of disease-associated patients from non-disease associated patients is discussed in this section with respect to previous research. Finally, a brief outlook on the key challenges and potential opportunities in volatilomics is provided.

1.1. Introduction

The growing emergence of infectious diseases across the world has proved to be a significant burden on public health and global economies[1]. On a global level, the spread of disease has been accelerated by rising populations worldwide, increased travel and trade[2], increased human interference with nature and wildlife[3][4] and the overuse of antibiotics in clinics[5] and agriculture[6]. A major factor in the increase in severity and mortality rates of these diseases is the rapid spread of antimicrobial resistance worldwide[5]. Despite the development of molecular diagnostics in the last 25 years, these techniques are highly specialised and are not universally available. Instead, the majority of hospitals rely on classical techniques such as plating, blood tests, and x-rays to identify microbial infections. Although these techniques are highly standardised across the globe, they are limited by time and specific labour requirements, and in the case of blood tests – lack of specificity. Throughout the years, clinicians have used odour as an indicative marker of disease. For example, the sulfide emission in the breath of *Helicobacter pylori*-positive patients with halitosis and gastrointestinal issues[7], the sweaty feet odour of patients with isovaleric acidemia[8], and offensive odours from infected wounds. However, there are many volatile emissions from pathogenic microbes that cannot be detected by the human nose and instead require specific sampling and analysis instruments to be detected. These volatile emissions are collectively termed the "volatilome" of the disease and these volatilomes are composed of many different volatile compounds originating from both the host and the invading pathogens. With the development of robust instrumental volatilomic workflows in the last 20 years, a variety of clinical samples can be collected and rapidly characterised at the compound-level. The identification of infection-specific volatilomes across a wide variety of maladies such as pneumonia, tuberculosis, COVID-19, and chronic wounds may provide a significant opportunity in infection prevention and control.

Volatile compounds are biproducts of the various stages of metabolic pathways utilised by microbes for their survival. The metabolism of sugars, lipids, amino acids, sulfur- and nitrogen-containing compounds, aromatic compounds and the subsequent metabolism of those products potentially give rise to thousands of volatile compounds. The availability of such substrates vary widely across nature which may significantly influence the volatilome of a particular species of microbe across different environments. Other environmental conditions such as oxygen availability[9], phase of cellular growth[10][11], temperature[12] and pH[13] influence the microbial emission of volatile compounds. Many microbes have evolved to be facultative anaerobes – microbes which grow best in the presence of oxygen, but do not require it for survival. As a result, to generate energy, microbes utilise multiple metabolic pathways and electron acceptors - other than oxygen - to survive.

The analysis of pure cultures of pathogenic microbes under varying environmental conditions has already proved to be a valuable method for elucidating the accessory and core volatilomes of these organisms. For example, in some cases[14][15][16], volatile compounds detected in pure cultures *in vitro* have also been identified *in vivo* and have been used – alongside standardised microbiological validation - to accurately identify pathogens present in particular infections. However, among some of the key challenges in translating microbial volatilomics to clinical volatilomics are: the disproportionate numbers of cells and resulting abundances of volatiles generated in nutrient-rich media relative to real clinical samples; finding appropriate controls for strict background subtraction; and the varying contributions

of the host commensal microbiome to the volatilome of the samples. Despite these challenges, by identifying the volatile compounds that are intrinsically linked to the survival of specific pathogens *in vitro*, we gain insight into potential cellular origins of volatile compounds detected from patients of infectious disease.

Wide variation in experimental and instrumental techniques used across volatilomics have essentially slowed the progress of the field. As a comparison, similar challenges are also experienced in microbiome research. Cross-study validation of results and data have subsequently become difficult. The recent publishing of comprehensive review papers[17][18][19][20] and books[21] as well as the development of the mVOC 2.0 database[22] - a comprehensive microbial volatilomic database – have elevated volatilomics into focus and have highlighted major opportunities. However, it is critical that a wide collaborative move is made towards standardised methods of sample collection, pre-treatment, sampling, and analysis. Such a move will reduce biases and improve reproducibility of results across studies and ultimately lay the foundations for future clinical applications of volatilomics.

Scope of review
The primary aim of this review is to highlight the clinical potential of microbial volatile organic compounds primarily for future diagnostics. An initial discussion of the fundamental pathways from which these metabolites are generated will guide the review towards its key topics. These are: the need for standardisation of sampling and analysis tools in clinical volatilomics; how comprehensive microbial volatilomics will lead to successful clinical volatilomics; the challenges of clinical volatilomics; and the prospect of microbial volatilomics research for novel antibiotic discovery in the future.

1.2.   Discussion
1.2.1.   Metabolic pathways involved in microbial volatilomes
Volatile metabolites are generated as byproducts at each step of multi-step pathways – potentially giving rise to a vast number of compounds. These compounds are typically low molecular weight, with high vapour pressure (boiling points < 250°C), and lipophilic in character (low polarity)[21]. Following emission, these molecules can then be further metabolised into potentially thousands of other volatile metabolites. Primary metabolism, or central metabolism is the fundamental keystone of survival across both eukaryotic and prokaryotic organisms. Primary metabolic products are derived from pathways such as glycolysis, fermentation, the tricarboxylic acid (TCA) cycle, or the various electron transport chains[23][17]. The ultimate objective of primary metabolism is to generate as much biochemical energy – in the form of ATP – as possible.
One of the most important molecules in both prokaryotic and eukaryotic metabolism is acetyl coenzyme A (acetyl coA). Illustrated in Figure 1.1 are some of the key metabolic pathways of which acetyl coA is intrinsically linked. Many of the metabolic reactions illustrated in Figure 1.1 are reversible and proceed based on the current metabolic requirements of the organism. In cellular respiration, acetyl coA is produced by the decarboxylation of pyruvic acid at the end of glycolysis, where it subsequently reacts with oxaloacetate to form citrate to initiate the TCA cycle. This key metabolic pathway generates ATP and other energy carriers (NADH and $FADH_2$) which enter the electron transport chain (ETC) to generate more ATP. The nature of the volatiles produced and the amount of energy that can be generated is dependent on the electron acceptor available to oxidise the electron carriers. For example, in aerobic

respiration of carbohydrates, exogenous O_2 is utilised as the final electron acceptor in the electron transport chain which generates a high abundance of ATP and completely oxidises glucose to CO_2. Many volatiles such as acetoin, acetic acid, and pyruvate are produced from the downstream stages of oxidation of these carbohydrates[21][17]. In anaerobic respiration, exogenous SO_4^{2-}, NO_3^-, or CO_3^{2-} are utilised as electron acceptors to generate relatively less ATP and volatiles such as hydrogen sulfide (H_2S) and ammonia (NH_3). Other volatiles generated through these pathways include sulfur- and nitrogen-containing organic compounds[21]. The lower energy yield in anaerobic respiration is due the lack of a high potential electron acceptor (e.g. oxygen) to oxidise the high energy electron carriers (NADH and $FADH_2$). ATP is generated when the electrons from the low potential electron donor are transferred to the high potential acceptor (e.g. O_2) by redox reactions. Therefore, the amount of ATP generated depends not on the concentration of the electron acceptor and donor but the redox potential difference of all the reactions between the electron donor and the available acceptor[24]. In nature, many bacteria simultaneously utilise multiple electron transport chains – with different electron acceptors - to survive in constantly varying environments[25]. This can allow many microbes to survive in both aerobic and anaerobic conditions – these are labelled as facultative anaerobes. If no electron acceptors are present, some microbes can transfer their electrons to an endogenous acceptor and initiate fermentation. However, bacteria and fungi can excrete fermentation products even in the presence of oxygen and other high potential electron acceptors. This seemingly less efficient and wasteful phenomenon is known as overflow metabolism[26]. It is characterised by the excretion of acetate (the Acetate switch[27]), which can occur anaerobically through mixed-acid fermentation; and aerobically when growth on excess glucose inhibits respiration (i.e. the Crabtree effect)[27]. Ethyl acetate, isoamyl acetate and methyl thioacetate are common volatile acetate molecules frequently reported in both bacterial and fungal profiling studies. Proposals for why microbial cells utilise fermentation over a higher ATP-yielding respiratory pathway for energy have been reported throughout the years[28][29][30][31], however, collective understanding of this complex phenomenon has yet to be achieved.

The majority of work on microbial volatiles has been primarily focused on characterising the primary volatilomes of microbial species, however, high degrees of specificity in volatilomes can only be found by investigating the secondary volatilome of microbes. Secondary metabolites are not essential for growth and development. These compounds typically have unusual structures and are generated by enzyme driven polymerisation reactions of larger molecules (e.g. terpenes and sesquiterpenes)[21].These compounds can also emerge through non-enzymatic reactions, schleiferone A and B are examples of secondary metabolites that are generated from the ex vivo reaction between 2-phenylethylamine and acetoin[32].

The metabolism of amino acids is also a key source of volatile metabolites. Leucine is a short chain amino acid that is readily catabolised by microbes when glucose is in low supply. This catabolism can lead to the generation of common microbial volatiles such as 3-methyl-1-butanol (isoamyl alcohol), 3-methylbutyric acid (isovaleric acid), 2-methylbutyric acid, 3-methyl-1-butanol acetate, and 3-methylbutyrate[33]. The shikimate pathway is a seven-step metabolic pathway that links the metabolism of sugars to the biosynthesis of aromatic compounds[34]. Key amino acids such as tryptophan, tyrosine, and phenylalanine are commonly produced in using this pathway. Phenylethyl alcohol[35], indole[36], and 2-aminoacetophenone[37] are all examples of downstream metabolites that are also produced at various stages of this pathway[17]. In the context of infection, microbes frequently form

biofilms (discussed in section 2.3.1.) *in vivo* to help them survive in challenging environments. A critical factor in the success of this survival is the slowing down of the cell growth rate[38] due to the significant decrease in primary metabolic activity. Conversely, the upregulation of amino acids during biofilm development[39] could shift cell metabolism in favour of amino acids over sugars. This difference in metabolic activity may also translate into a difference in volatilomic activity between planktonic and biofilm cells and potentially lead to the biofilm-specific emergence of more amino acid derived metabolites.

An alternative production mechanism for acetyl coA is through the ß-oxidation of lipids whereby it serves as the starter unit for the production of fatty acids[40]. The condensation of acetyl coA with malonyl coA results in CO_2 being lost as the new C-C bond between these molecules is formed. Malonyl coA is the thiol ester of CoA and malonic acid and is formed by the acylation of coA with CO_2[41]. This is then followed by several cycles of chain elongation with malonyl coA which terminates at a variety of chain lengths to liberate respective acids. Microbes are capable of utilising a variety of starter units – other than acetyl coA – for this particular reaction and this results in vast diversity in the compounds produced. The decarboxylation of intermediate compounds - produced as biproducts of the chain extension cycle - leads to the generation of various alkanes, 1-alkenes, and methyl ketones[20]. Hydrolysis and reduction reactions of these intermediates also give rise to a variety of compounds such as acids, 1-alcohols and aldehydes. This pathway is utilised by many bacteria, including various clinical pathogens. For example, 1-undecene was detected in the breath of patients with *Acinetobacter baumanii*-positive ventilator-associated pneumonia. As discussed in section 2.4.1.2. of this review, this compound is also a key component of the *P. aeruginosa* volatilome[45][42]. 4-methyl-1-decene is another alkene that has been detected *in vivo*, in the breath of patients following infection with *Mycobacterium tuberculosis*.
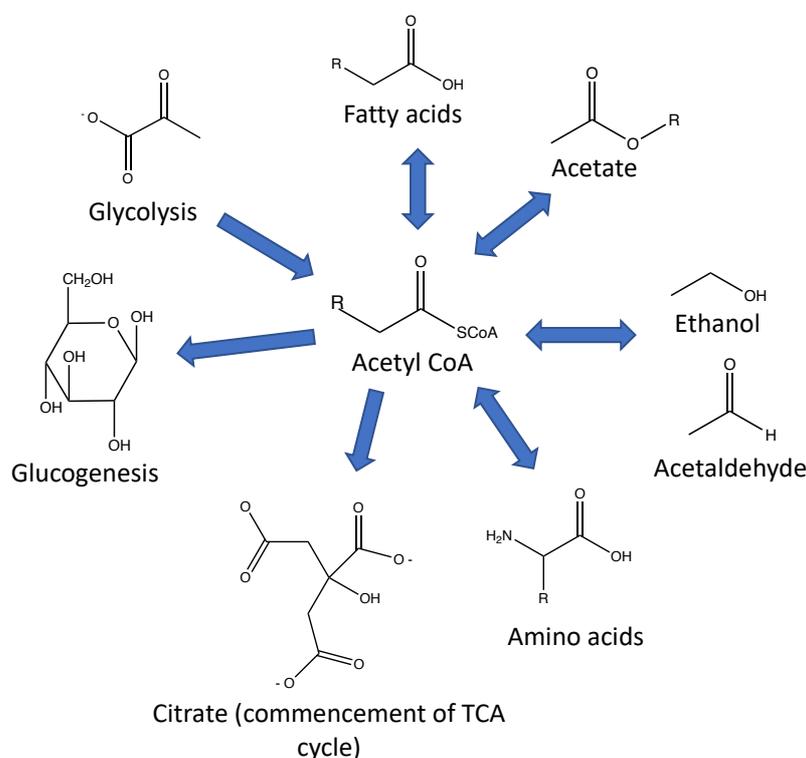
Figure 1.1: Summary of the reversible and irreversible metabolic reactions from which Acetyl CoA is involved.

Sulfur-containing volatile compounds are primarily derived from the oxidation of methanethiol. Methanethiol is a highly common compound that is associated with decaying biomass. It spontaneously dimerises in air oxidatively to form dimethyl disulfide[43] but also can be oxidised metabolically by microbial oxidase enzymes[44] to form sulfides, disulfides, and trisulfides[17]. Dimethyl disulfide and dimethyl trisulfide are volatile compounds that are emitted by a wide range of microbial species and have also been described as indicators of decaying biomass. It should be noted that these compounds are also present in many commercial growth media used in microbial volatilomic studies[15][45]. In living systems, sulfur-containing volatile compounds can also be generated through the metabolism of the sulfur-containing amino acids, cysteine and methionine[46]. These pathways are particularly relevant in the volatilomes of gastrointestinal (GI) - associated pathogens such as *Helicobacter pylori*[47] and *Clostridium difficile*[48] whom both have highly complex volatile sulfurous-compound profiles.

1.2.2. Analytical techniques used in microbial and clinical volatilomics

1.2.2.1. Sample collection

The primary aim of clinical volatilomics is ultimately to deliver a rapid non-invasive means of acquiring useful diagnostic information. Non-invasiveness is achieved by a technique when it does not involve breaking the skin or contact with the mucosa or internal cavities of the body. For example, biopsies would be considered invasive as they require a sample of tissue to be removed from a specific site while swabbing would be considered relatively non-invasive as it only requires a couple of seconds contact with the sample site. There are many approaches to obtaining a human volatilomic sample. Human matrices such as breath, sputum, skin, urine, blood, faeces, sperm, sweat or a wound bed all provide potentially unique diagnostic information about the human condition. However, these matrices also introduce their own interferences (discussed in section 2.4.1.) which must be accounted for via background subtraction. In this section, some frequently used techniques for volatilomic sample collection will be discussed.

Breath volatilomic research has focused on the collection and subsequent analysis of exhaled breath for the detection of disease-associated biomarkers[49]. The collection of breath samples for volatilomic sampling has been comprehensively reviewed [50]. Breath sample collection is typically carried out using gas sampling bags, syringes, evacuated steel containers, and glass bulbs.

Samples from superficial matrices such as skin, wounds and other lesions can also be collected using a variety of techniques. Solvent extraction has also been used for the sampling of skin[51][52], however, the requirement for solvents and for pre-concentration of sample matrix prior to analysis limits its potential role in a clinical workflow. Direct headspace sampling is easily compatible with SPME fibers and has been previously employed for rapid non-invasive sample collection of skin volatiles[52][53]. This technique utilises a small glass housing with a SPME inlet that cups the surface of the skin and is fixed in position using adhesive tape. Polydimethylsiloxane (PDMS)-based membranes and patches have been previously used for the collection of volatiles from skin and wounds. Swabbing is a non-invasive sample collection technique that is universally employed for a wide range of clinical applications and has yet to

15

be comprehensively investigated for volatilomic sampling. Recently, clinical swabs with SPME fibers incorporated into the cotton were employed for direct sampling of human metabolites taken from nasal and oral cavities[54]. Swabbing has also been shown to be sufficient for volatilomic discrimination of infected and non-infected wounds using an e-nose[55], however, broader volatilomic data is required to identify the discriminative compounds. Our group is currently profiling the volatilome of swabs taken from diabetic foot ulcers with varying degrees of infection to broadly screen for and identify infection-specific volatilomic shifts using HS-SPME-GC-MS.

Table 1.1: Cutaneous and superficial sample collection techniques for clinical volatilomics.

Sampling method	Benefits	Limitations
Swab	- Rapid sample collection (6-10s) - Simple and universal - Versatile – can be used for any site on the body (i.e nasal, saliva, axillae) - Non-invasive - Can collect sample at varying depths (wounds)	- Contact - introduces exogenous contamination, - Relatively high background signal intensity - Unsuitable for dry skin VOC analysis
Direct skin	- Minimises exogenous contamination - Rapid turnover of data - Suitable for skin analysis	- Potentially uncomfortable - Limited to specific sites on the body 15 – 20 min *in vivo* sample time
Sorbent membrane/patch	- Simple and universal - Relatively versatile - Non-invasive	- Contact - introduces exogenous contamination - Relatively high background signal intensity - Time consuming - Superficial VOCs only (wounds) - Requires thermal desorption/solvent extraction

1.2.2.2.   Sampling and analysis

The wide variation in sampling and analysis techniques used across clinical volatilomics restricts cross-study comparisons of data and essentially blocks validation of results. As an initial step towards identifying consistent disease-associated volatilomic shifts, sampling and analysis techniques with large reference libraries that support broad untargeted screening of compounds should be employed. Untargeted whole volatilomic screening of clinical samples is a frequently used technique for identifying discriminative patterns in volatilomic data. It allows a wide range of volatile compounds to be identified and assessed on their individual discriminative contribution of the examined samples. Once these volatilomic patterns are

Pre-examination copy
Submitted 27th May 2022

consistently validated for a given disease by other studies, the compounds responsible can then be targeted and quantified by direct mass spectrometric techniques. A good example of this comprehensive workflow was employed by Shestivska et al. [56] to first broadly screen the volatilome of multiple strains of *Strenotrophomonas* using SPME-GC-MS and then quantify target compounds of interest using selective ion flow tube (SIFT)-MS. However, currently many clinical volatilomic studies use these direct mass spectrometric techniques first without knowing what compounds to specifically target and quantify. While these studies highlight the future potential of these direct quantification techniques, they also highlight the critical need for more untargeted clinical volatilomic studies with standardised sampling and analysis protocols.

Thermal desorption tubes / Solid-phase microextraction

Across the literature, thermal desorption sampling tools such as sorption tubes and solid phase micro-extraction (SPME) fibers have been frequently employed for both microbial and clinical volatilomics. Both of these sampling tools are directly compatible with GC-MS, which allows efficient separation of the extracted analytes – which can then be accurately identified using a large reference library. Sorption tubes are packed with pre-determined ratios of sorbent combinations (e.g. porous polymers, carboxen, silica gels) to expand the range of analytes that can be extracted[57]. Volatile analytes travel up the tubes and are adsorbed onto the various sorbents based on their polarity and volatility. For example, highly volatile low-molecular weight compounds that have low polarity[17] and are retained using a strong sorbent such as carboxen in the deeper beds of the tube[58]. As the molecular weight of compounds increases, they become less volatile and more polar, therefore to capture these compounds, porous polymers such as polyethylene glycol (PEG) – that rely on stronger polar interactions - are required. This technique has proven to be highly applicable in clinical breath studies (see Table 1.3) where the breath sample is typically collected in a Tedlar bag and then subsequently transferred into the sorption tubes using a vacuum pump.

SPME relies on a similar mechanism as it utilises a chemically enhanced silica fiber (stationary phase) consisting of various phases to capture a wide range of analytes. Each phase has unique pore sizes and polarity characteristics that allow the retention of both small volatile apolar compounds and larger less volatile polar compounds[59]. Active research is also being carried out on expanding the range of compounds that can be extracted by SPME by developing new fiber coatings[60].

Throughout the literature, in both microbial and clinical volatilomics, the use of these techniques have allowed the detection of a higher number of volatile compounds and have yielded some of the most diverse volatilomic data. For example, in microbial volatilomics, SPME sampling has been widely used to obtain core and accessory volatilomes of highly prevalent clinical bacterial pathogens[61][62][45][42] (Figure 1.2). The use of SPME for *in vivo* volatilomic sampling is constantly evolving for different biological matrices. Although sampling of the skin volatilome using SPME has been demonstrated multiple times, there are yet to be studies investigating cutaneous diseases such as atopic dermatitis, acne, or psoriasis. These diseases have all been shown to have measurable microbiome shifts which - given the close relationship between them - may also indicate potential disease-specific volatilomic shifts. The sampling of breath volatiles has been shown to be more suited to sorption tubes as the breath samples can be readily vacuum transferred from the sample collection apparatus to the tubes. Although these methods enable broad volatilomic screening of a variety of volatilomes, they are limited in that it is highly challenging to obtain accurate

17

quantitative volatilomic data. This can be achieved by incorporating direct mass-spectrometry based methods into the workflow.
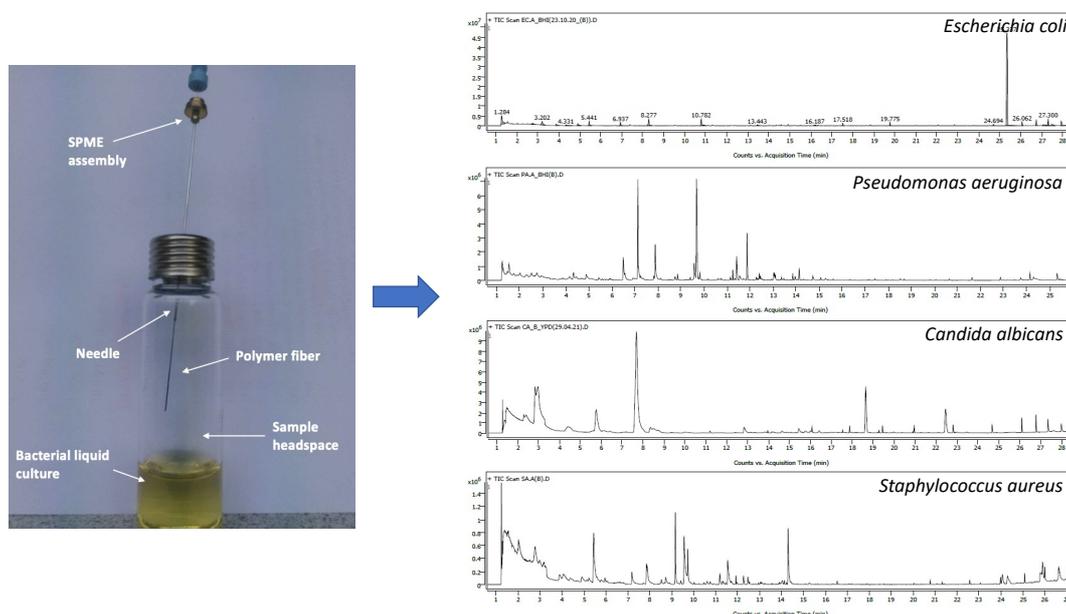
Figure 1.2: Total ion chromatograms of bacterial and fungal volatilomes following HS-SPME-GC-MS analysis after 24 hours of planktonic growth.

Real-time methods of detection are a promising prospect for future volatilomic profiling investigations. The advantages (Table 1.2) of direct detection methods include real-time targeted analysis, absolute quantification, and high sensitivity. Samples can be analysed immediately after being collected, which eliminates potential errors that could arise as a result of sample storage and transportation[63]. Frequently used direct techniques in volatilomics include proton-transfer-reaction (PTR)-MS, selected ion flow tube (SIFT)-MS, ion molecule reaction (IMR)-MS, ion mobility (IMS-MS), and secondary electrospray ionisation (SESI)-MS. Among these techniques, the most frequently employed in volatilomics are SIFT-MS and PTR-MS.

SIFT-MS is an analytical technique used for real-time quantification of volatile analytes. The mechanisms behind this analytical technique are described in great detail in the reviews [64][65][66]. In summary, sample gas molecules are introduced to the ion flow tube and are chemically ionised by precursor ions (H_3O^+, NO^+, and O_2^+) in the helium carrier gas. Precursor ions and sample gas molecules react to form product ions which are then detected and counted by the downstream mass spectrometer.

In a PTR-MS system, primary ions, H_3O^+, react with sample molecules under defined conditions. Water vapour is introduced to a chamber containing a hollow-cathode discharge source that generates hydronium ions (H_3O^+). This process begins at the hollow-cathode discharge source and is facilitated further in the source drift region. Ions in the source drift region are dragged by an electric field into the drift tube. H_3O^+ ions travel through a buffer gas within the drift tube to which sample gaseous molecules are added. The sample molecules are protonated by the H_3O^+ ions. The ion detection system then measures the count rates of remainder H_3O^+ ions and the protonated sample ions, which are proportional to the

18

respective densities of these ions. More detailed accounts of this instrumental technique can be found elsewhere[67].

These techniques have been adapted for both microbial and clinical volatilomics studies (Table 1.2). These studies have demonstrated the clear potential for these methods to be used for accurate quantification of target metabolites in the future. However, due to limited reference libraries, these methods are currently non-applicable for untargeted screening of analytes[68]. As a result, the datasets reported in studies using these methods have relatively low numbers of compounds compared to datasets obtained using GC-MS. It is recommended that volatilomic studies that employ these direct detection techniques also include GC-MS analyses to validate the results[69]. In the future, following the validation of target metabolites, the true value of these quantitative mass spectrometric methods will be realised.

Table 1.2: Overview of frequently used sampling and sampling-analysis methods in microbial and clinical volatilomics

Technique	Benefits	Limitations	Clinical volatilomics	Microbial volatilomics
SPME	• Diverse range • Cost efficient • Used widely • Sensitive and selective • Easily standardized • Directly compatible with GC-MS. • Superficial human volatilomics (skin, wounds)	• Absolute quantification challenging • Storage of sample	• Skin[53] • Direct breath[70][71] • Wounds[150][151] • GI disorders (Faeces[72])	• HS vials [73][16][15][14][42][45] • Culture bottle [11]
Sorption tubes	• Diverse range • Sensitive and selective • Storage of samples • Compatible with GC-MS • Highly suitable for breath volatilomics	• Special attachment needed for GC-MS analysis • Relatively expensive – requires pump	• COVID-19 (breath) [74][75] • VAP (breath) [14][15] • Tuberculosis (breath) [16][76] • GI disorders (Faeces[77]) • Gastric cancer[78]	• HS vials [10][16] • Culture flasks [79][80] •
Direct syringe	• Simple operation • Compatible with GC-MS • Relatively inexpensive	• Lack of selectivity • Limited number of compounds • Pre-concentration required	• Influenza (breath) [81] • COVID-19[82]	• HS vials [83][84]
Direct quantification (SIFT-, PTR-, SESI-MS)	• Quantification • Real-time analysis • Highly sensitive • Suitable for targeted analysis • Versatile – can be adapted for many biological matrices	• Challenging for screening of unknowns • Relatively expensive • Limited VOC profiles • Highly specialised	• Tuberculosis (breath) [16] • Skin [85] • Healthy breath profiling[86] • Gastric cancer (breath)[87][88][89] • COVID-19 [90]	• HS vials [91][92][93][94][56] • Biofilm assay [95] • Culture bottle[96]
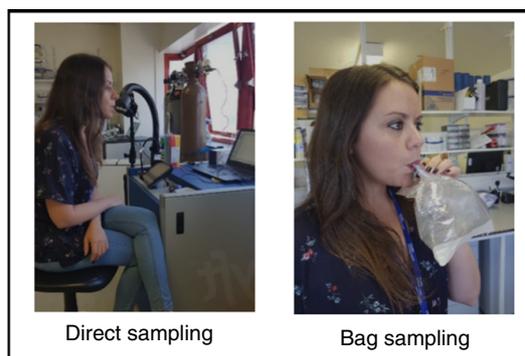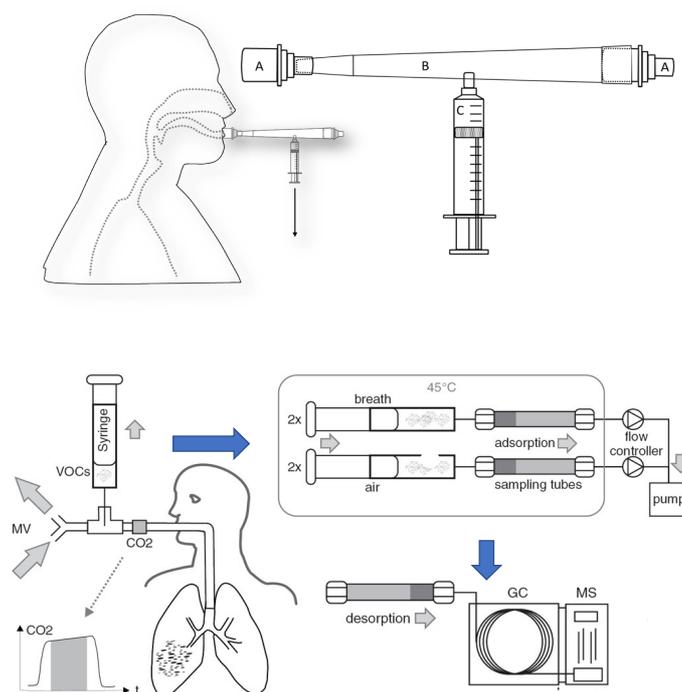
Fig. 9 | Breath analysis by SIFT-MS. Top) Schematic representation of the proposed workflow for breath analysis using online and offline sampling apparatus employed by Ruszkiewicz et al.[82] for the breath profiling of COVID-19 positive patients. Middle) Breath profiling of VAP patients using a syringe-sorption tube sampling workflow followed by GC-MS analysis[...]. Bottom) Direct breath sampling and bag sampling of breath for direct quantification of target analytes by SIFT-MS[64].

1.2.3. *In vitro* microbial volatilomics and chemical volatilomics

1.2.3.1. Microbial volatilomics

Comprehensive characterisation of microbial volatiles involves compiling volatilomic data from microbes across varying environments, growth phases and varying experimental systems. Reported volatilomic results are highly dependent on the analytical methods employed and therefore to obtain the whole volatilome of a particular organism, a variety of sampling and analysis techniques must be employed. Subsequent validation of results must always be carried out using GC-MS as it has the largest reference library of compounds available. *In vitro* analyses of microbial pure cultures allows production mechanisms of specific VOCs.

specific metabolites to be determined. This then aids identifying the cellular origin of volatile compounds that are detected in clinical *in vivo* systems. In this section, the volatilomes of some highly prevalent bacterial and fungal pathogens are discussed.

Staphylococcus aureus

S. aureus is a highly prevalent pathogen that acts as the driving force behind many cutaneous, wound and respiratory infections. The emergence of penicillin-resistant *S. aureus* in the 1950s followed by a methicillin-resistant *S. aureus* (MRSA) epidemic in the 1960s placed this pathogen into international focus. The MRSA epidemic in particular caused significant rates of sepsis in many European countries and its re-emergence and prominence over the last 60 years have led to large global efforts to develop early detection systems[97]. *Staphylococcal* species have been shown to emit a diverse range of volatile metabolites, however, many share similar volatilomes and work is needed to deduce the species- and strain-level diversity of the volatilome of this phyla. In this section, the volatilome of *S. aureus* will be discussed.

S. aureus emits a diverse range of acids, ketones, fatty acid esters, alcohols, and aldehydes. Early stages of *S. aureus* growth has been marked by the relatively high abundance of aldehydes 2-methylbutanal[92][98] and 3-methylbutanal[99][98] which are then consumed and metabolised into other compounds in the latter stages of growth[45]. It is a facultative anaerobe that primarily relies on oxygen to support its metabolism of sugars and amino acids to generate core primary volatile alcohol and acid metabolites such as ethanol, 3-methyl-1-butasnol, isovaleric acid, propanoic acid, 2-methyl propanoic acid, and acetic acid. The abundance of acids and ester compounds produced by *S. aureus* has been shown to be dependent on the availability of sugars[42][61]. Acetoin and 2,3-butanediol are two uncharged molecules that are generated from the metabolism of pyruvate[17]. These molecules are produced in the exponential phase of growth in response to the accumulation of acidic metabolites. The neutrality of the molecules reduce over-acidification of the intracellular environment by neutralising the extracellular environment[100][101]. Other compounds emitted by *S. aureus* include dimethyl disulfide and 1-butanol[14][10] – all of which are highly common across bacterial and fungal volatilomes. Although it should be noted that dimethyl disulfide and 3-methylbutanal are frequently found to be present in blank growth media[98][42][45], the mechanisms by which they are generated in microbes are illustrated in the review by Weisskoff et al.[17]

As highlighted in section 2.4.6. of this review, *S. aureus,* as well as *P. aeruginosa,* and *E. coli* are highly prevalent wound pathogens and are responsible for particularly severe infections in diabetic foot ulcers (DFUs)[102][103]. Malodour has been used as a clinical indicator of wound infection for many years, however, characterisation on a molecular-level of this malodour has never been comprehensively demonstrated. Detection of *S. aureus* volatiles *in vivo* has however been demonstrated by Filipiak et al. [14] whom detected 3-methylbutanal, ethyl acetate, acetic acid, and ethanol in the breath of patients with *S. aureus*-positive ventilator-associated pneumonia. In terms of strain-level volatilomic diversity, A.W Boots et al observed significant differences in the abundance of discriminatory volatiles emitted from MRSA (Methicillin resistant) and MSSA (Methicillin sensitive) cultures. These results highlight the potential influence of antibiotic sensitivity[79] on bacterial volatilomes.

Pseudomonas aeruginosa

P. aeruginosa is a common gram-negative bacillus that is frequently associated with severe infections of burns[104] and diabetic foot ulcers[105], and it has been labelled as the most common cause of persistent, fatal respiratory infections in cystic fibrosis patients[106][107]. Eradication of these infections can be highly difficult due to the exceptional resistance to many antibiotic therapies that *P. aeruginosa* has developed[108].

Clinical *P. aeruginosa* isolates (n=24) were grown *in vitro*[62] analysed to identify 65 unique compounds across these isolates. Among them were first-time reports of 2-decanone and several 3- and 4- ketones. By analysing a high number of strains, this study comprehensively reports many of the core and accessory volatiles associated with the *P. aeruginosa* volatilome. The *P. aeruginosa* volatilome is characterised by a series of compounds that are derived primarily from the fatty acid pathway, these include 1-undecene[10][109][110][111], 2-nonanone[45][93], 2-undecanone[42][45][106], 2-nonanol[42][62], and 2-undecanol[42][45]. Figure 1.4[42] illustrates how these compounds collectively contribute to the discriminative nature of the *P. aeruginosa* volatilome when compared to other pathogens. However, the most characteristic compound linked to the *Pseudomonas* volatilome is 2-aminoacetophenone - produced in the amino acid degeneration (shikimate) pathway via the loss of a hydroxyl group on anthranilic acid[17]. 2-aminoacetophenone is a quorum sensing molecule that has been shown to have a significant dose-dependent promotion of antibiotic resistance in *P. aeruginosa* cells[112]. Interestingly in our own recent study, we consistently detected significant differences in the emission of 2-aminoacetophenone - across multiple growth media - between two *P. aeruginosa* strains[42]. This highlighted potential strain-level differences in antibiotic-sensitivity across the two examined strains while also demonstrating how volatilomics and metabolomics could be employed to identify these potential differences.

1-undecene has been detected *in vivo* in patients with *Acinetobacter baumanii* infections in the ventilator-associated-pneumonia (VAP) study by Gao et al.[15] As discussed this compound is produced in high abundances by *P. aeruginosa* and it has also been detected from a ex vivo sputum model colonised with *P. aeruginosa*[113]. Ex vivo models have been used to simulate wounds infected by *P. aeruginosa*. Using this perfusion wound meat model, 2-aminoacetophenone and 2-nonanone were labelled as potential biomarkers of *Pseudomonas* wound infection as they were detected *ex vivo*[112].

Although the volatilomic detection of *P. aeruginosa in vivo* has not been demonstrated directly, there are some indications that potentially highlight the clinical persistence of these volatiles. Firstly, the detection of methyl thiocyanate in the breath of *P. aeruginosa*-positive cystic fibrosis (CF) patients was attempted by Shestivska et al.[94] In this study, relatively higher concentration ranges of methyl thiocyanate were detected in *P. aeruginosa*-positive CF patients than *P. aeruginosa*-negative CF patients; however, statistically significant differences in the mean concentrations of this compound across the two groups were not observed. Despite this, this study was limited by low numbers of patients in the *P. aeruginosa*-positive group and did demonstrate the occurrence of another volatile metabolite that has been detected across both *in vitro* and *in vivo* systems.

Escherichia coli

E. coli is a gram-negative facultative anaerobe that is highly abundant in the human intestinal tract. It has been reported to be one of the most predominant pathogens in hospital-

associated infections[114] and has been associated with multiple disorders such as pneumonia[115] and neonatal meningitis[116]. *E. coli* is also a frequently isolated pathogen from chronic wounds.

The volatilome of *E. coli* is characterised by a high abundance of indole[20]. Indole is commonly found in human faeces as a product of *E. coli* activity in the human gut[117], and its high abundance is likely responsible for the characteristic foul odour of the culture. Indole has been documented as an intercellular signal molecule amongst diverse bacteria. It is produced by over eighty species of bacteria[118], though very few produce comparable abundances to *E. coli*. It is produced in the downstream metabolism of aromatic amino acids (shikimate pathway). *E. coli* is highly diverse metabolically and utilises a variety of pathways for its survival. For example, in Figure 1.4 , the detection of even and odd carbon numbered 1- and 2-alcohols ($C_6 - C_{16}$) and odd carbon numbered ketones (2-nonanone, 2-undecanone, 2-tridecanone and 2-pentadecanone) indicates a highly active fatty acid pathway[42][45]. The emission of some of these compounds has been shown to correlate with the respective growth phase of the cells[42]. *E. coli* produces a high abundance of a number of fatty acid esters through the combination of alcohols and acetyl coA enzymes[119]. As a result of this metabolic ability, *E. coli* has been studied for its ability to produce petrochemically useful compounds such as styrene[120], fatty acids[121], and fatty acid methyl esters[122][123]. Some of the alcohols and their resulting fatty acid esters can be seen in the heatmap in Figure 1.4[42]. The detection of *E. coli* infections *in vivo* has yet to be thoroughly investigated. Stool samples have been previously employed for the volatilomic detection GI-associated pathogens such as *Clostridium difficile*[77] and *Giardia spp.*[72] . These studies highlight the potential value of stool volatilomic sampling for the detection of *E. coli*-associated GI-infections.

Other common pathogenic bacteria

Klebsiella pneumonia and *Streptococcus pneumoniae* are members of two different genera, but as their names would suggest, they both contribute to similar disorders, namely pneumonia[124][125] and also meningitis[126][127]. The VOC emissions of both these species have been investigated, both emit alcohols and various 2-ketones such as 2-butanone, 2-pentanone and 2-nonanone[80][128][129] – all of which are common to other bacterial VOC signatures. The two pneumonia-causing species also emit characteristic VOCs; *S. pneumoniae* emits low-molecular weight hydrocarbons (E)-2-butene and (Z)-2-butene(TSB media), while *K. pneumonia* emitted various 3- , 4- and methyl ketones (LB media). 3-methyl-1-butanol is a very common VOC emitted from a many bacteria including *Staphylococcus epidermidis (S. epidermidis)*[45], *S. pneumoniae*[130], *and K. pneumoniae*[129]. Lemfack et al. analysed VOCs produced by commensals *Corynebacterium* and *Staphylococcaceae* species[32]. When cultured in brain heart infusion (BHI) media, *Cornebacterium* emitted compounds such as 2-nonanone, 2-phenylethanol, and 2-undecanone – all of which are also emitted by *E. coli*[45]. GI-associated pathogens such as *H. pylori*[47] and *C. difficile*[48] have been reported to have complex sulfurous volatilomes comprised of various sulfides, thiols, thioesters, and sulfur-containing heterocyclic rings.

Candida

The investigation of fungal volatilomes in a clinical context is relatively understudied, though there has been previous attention given to the emission of volatile antibiotics from specific fungal species. *Candida albicans* is one of the most prominent clinically prevalent fungal

pathogens. It is a polymorphic fungus that can grow either as ovoid-shaped budding yeast; as an elongated ellipsoid cell with pseudohyphae; or a parallel-walled true hyphal form. Physiological and environmental cues such as temperature, pH, O_2 and CO_2 content, and quorum sensing interactions can signal the transition of the Candida cell from one morphology to another. As the fungus transitions from the yeast cell to the ellipsoid cell to the true hyphal form, its virulence increases – this is also reversible. As a result, severe *Candida*-associated infections typically involve cells that have formed true hyphae. For example, severe co-infections can arise between *C. albicans* and *S. aureus* as the bacteria can utilise the hyphae of the fungi as an adhesive surface to generate persistent biofilms[131].

Investigating potential volatilomic shifts that may occur as the cells transition across these morphologies *in vitro* could be a relatively simple route towards characterising the virulent volatilome of *C. albicans*. Recent volatilomic studies have investigated highly common pathogenic *Candida* species: *C. albicans*[132], *C. tropicalis*[132], and *C. glabrata* [132][133]. However, volatilomic data on *C. parapsilosis* is currently very limited - the only study available[134] identified only 3 compounds. Among the discriminative compounds produced by *Candida*, farnesol is a biproduct along the ergosterol synthesis pathway[135] and is a quorum sensing molecule critically used in *C. albicans* biofilm development. It is primarily involved in the control of morphogenic transitions in *Candida*, mediating the transitions from a hypha-to-yeast[136] and inhibiting yeast-to-hypha transitions[137] – which is critical to *Candida* biofilm formation and virulence. Farnesol has also demonstrated strong antimicrobial properties showing anti-biofilm activity against several *Staphylococcal spp*[138][139]. Another key bioactive metabolite emitted by *Candida* is phenylethyl alcohol. This molecule has been linked to the stimulation of filamentous growth[140] and promotes biofilm formation in yeasts[141][142]. Interestingly, despite this filamentous growth this molecule stimulates in yeasts, it has demonstrated inhibition of hyphal formation in *C. albicans* cells[143][140] - which is a core step in *C. albicans* biofilm development[144] . However, hyphae formation is promoted by *C. albicans* cells through the production of tyrosol[140] – a quorum sensing molecule derived from phenylethyl alcohol.

Figure 1.4: Heatmap hierarchical clustering representation demonstrating the clear volatilomic dissimilarities observed across S. aureus, P. aeruginosa and E.coli samples cultured in brain heart infusion (BHI) broth (Fitzgerald et al[42]).

1.2.3.2. Clinical volatilomics

Similar to the way in which many diseases cause dysbiotic shifts in the human microbiome, many diseases cause measurable shifts in the human volatilome. As a result, there are opportunities to develop innovative novel techniques and technologies to accurately diagnose diseases by characterising their volatilome. The major benefit of using volatilomics for potential clinical diagnoses is that volatile compounds - by their nature – can be non-invasively collected and rapidly analysed. Disease-associated volatilomic shifts have been reported for a variety of diseases and maladies, however, there is currently wide variation in the individual compounds that are responsible for these shifts across the literature. As discussed in section 2.2., a major contributor to this variation is the lack of standardised protocols for sample collection, sampling and analysis. The collective move towards standardisation of these protocols in the near future will give rise to major leaps towards uncovering the realistic clinical applicability of these methods. Low sample sizes in clinical volatilomic studies also majorly contribute to the cross-study variation observed across the literature. The aim of this section is to draw comparisons and contrast the results taken from the past and most recent clinical volatilomic studies. This section closes with a table summary (Table 1.3) of reported discriminatory compounds collected from a variety of clinical volatilomic studies and the analysis methods used in each study. While this table clearly illustrates the variation in reported compounds across the literature, it highlights potential cross overs in specific discriminatory compounds across diseases. As well as a mVOC database, building a clinical volatilome database in the future would have substantial benefits for researchers in this developing field.

Tuberculosis (TB)

TB is an infectious disease caused by the bacillus *Mycobacterium tuberculosis* which can be spread from one person to another via breath. The disease primarily affects the lungs (pulmonary TB) but the TB bacteria can also enter the bloodstream and affect other sites of the body (extrapulmonary TB). Throughout the years TB has been diagnosed using microscopy, culturing, and recently molecular methods[145]. However, despite the advances in these diagnostic techniques, millions of TB cases are missed worldwide[146]. This is primarily due to the insensitivity and the required duration of these tests. For example, a mycobacterium culture takes between 4-8 weeks to complete[147].

Syhre et al. [148] investigated whether volatile metabolites produced *in vitro* by various *Mycobacterium* could be used as a rapid and selective diagnostic method. They identified two potential disease-associated VOCs, methyl p-anisate and methyl nicotinate, that were emitted exclusively by *M. tuberculosis* and *M. bovis* – another causative pathogen of TB. However, these compounds were not detected in pure *Mycobacterium* cultures carried out by McNerney et al[149], whom instead reported compounds such as 2-phenylethanol, 2-methyl-1-butanol as being discriminatory markers. McNerney also specifically addressed that the targeted approach previously proposed by Syhre et al. is challenging, and that "a more robust approach is likely to be achieved by obtaining the whole spectra of samples for TB analysis and subjecting these to multivariate analysis". An example of this more robust approach was demonstrated earlier by Phillips et al.[76] whom firstly analysed the volatilome of *M. tuberculosis* cultures before detecting TB-specific patterns in breath samples. In this work, 134 compounds were detected in the breath of TB patients (n=42) and healthy controls (n=59) using sorbent traps coupled with GC-MS. These compounds were then subsequently

used to discriminate the groups using principal component analysis which demonstrated a high degree of sensitivity and specificity.

Untargeted studies have investigated the breath profile of patients with TB, by employing healthy volunteers with no symptoms as the control group, and smear positive cases as the study group[150]. In this study, the potential of VOC profiling was demonstrated for therapeutic treatment monitoring by using a PCA biplot to effectively visualise the therapy timeline of the study group against a control group[150]. A more targeted approach was employed by Mellors et al.[16] whom investigated potential biomarkers of TB found in the breath of animals. In this work, *in vitro* culture based volatilomic analysis was carried out on a single strain of *M. tuberculosis,* and an *in vivo* breath analysis of animals infected with the same strain of *M. tuberculosis.* They detected 4-(1,1-dimethylpropyl)phenol and 4-ethyl-2,2,6,6-tetramethylheptane in both the *in vitro* pure cultures and in the *in vivo* breath analysis of the infected animals. These results indicate that some of the substrates found in the growth media used for the *in vitro* VOC profiling were also present in the lungs of the animals, and that the pathogen was potentially able to exploit this availability of nutrients *in vivo*.

Recently, an interesting novel approach of TB diagnosis was reported by Vishinkin et al.[151] whom employed both headspace GC-MS and nano-based sensor arrays to analyse volatile biomarkers emitted from the skin. In this two-site study in which a total of 636 subjects were analysed, the authors reported that patients with confirmed pulmonary active TB had elevated abundances of acetic acid, 2-ethyl-1-hexanol, toluene, and ethyl cyclopropane. However, it must be noted that despite the highly impressive sample size of the study, compounds such as 2-ethyl-1-hexanol and toluene are present in many environments[52][152] and background concentrations of these compounds vary widely across these environments.

SARS–CoV–2 and COVID-19
In 2020, the rapid spread of the *SARS–CoV–2* virus across the world significantly highlighted the immediate need for non-expensive rapid detection techniques for infectious diseases such as COVID-19. Wide-spread testing was universally accepted as being one of the most powerful tools countries could use at slowing the spread of the virus. Though polymerase chain reaction (PCR) testing is the gold standard in terms of detecting the virus, the shortage of required reagents and kits in the early stages of the pandemic clearly highlighted the need for new rapid methods of detection. Volatilomic screening of the breath of COVID-19 patients has emerged as a potential approach for the identification of novel biomarkers. Although antigen testing has emerged as a rapid alternative approach to PCR testing, many countries have been slow to accept this technique due to accuracy concerns. Respiratory viral infection was previously shown to cause discriminatory volatilomic shifts in human cells *in vitro*[153], highlighting a potential novel route for the detection of viruses *in vivo.* Breath profiling has subsequently been investigated as a route towards non-invasive testing of COVID-19. Though none of the studies described demonstrate high degrees of accuracy in their discriminatory patterns between COVID-positive and COVID-negative patients, they do highlight the need for further investigation. Detecting volatiles specific to a virus is challenging as they do not produce their own metabolites and instead rely on the metabolism of the host[154]. Therefore detection of viruses using volatilomics would most likely rely on the detection of inflammatory shifts in the whole human volatilome rather than detection of a specific indicative compound.

Ruszkiewicz et al.[82] carried out a volatilomic study across two different hospitals with 98 participants, 31 of which had COVID-19. The group used a simple sample collection system which is illustrated in Figure 1.3. Multivariate analysis of the breath volatilomes demonstrated a relatively good discrimination between the breath samples of the COVID-19 positive patients and the COVID-19 negative patients. The compounds responsible for the observed trend were ethanal, heptanal, octanal, acetone, 2-butanone and methanol. In another COVID-19 breath study by Grassin-Delyle[90], was used to analyse the breath of 40 patients, 28 of whom were COVID-positive. Four compounds presented significantly higher abundances in the breath of PCR-positive patients versus PCR-negative patients, these were methylpent-2-enal, 2,4-octadiene, 1-chloroheptane, and nonanal. The discriminatory compounds identified in both these studies were not discriminatory markers in a recent study[74] published in the *European Respiratory Journal*. In this study, the breath of 81 patients (52 PCR-positive cases) was collected using sorption tubes and analysed using GC-MS. Following the analyses, 68% sensitivity and 85% specificity was observed for the discrimination of the breath samples of the two cohorts, of which benzaldehyde, 1-propanol, 3,6-methylundecane, camphene, beta-cubebene, iodobenzene were the discriminatory markers. Most recently, the breath of different groups of children with and without COVID-19 was analysed using sorption tubes coupled with GC-MS to broadly screen for metabolic shifts in their volatilome[75]. The results showed interesting discriminatory patterns between the groups based on shifts in abundance of regular aldehydes heptanal, octanal, and nonanal. Though these aldehydic compounds are common components of the human volatilome, this study is the third COVID-19 breath analysis to report increases in their abundance in COVID-positive patients. These results also broadly agree with recent correlations made between oxidative stress in the breath of cancer patients and elevated aldehyde abundances[155][156][157].

Although, these four COVID-19 studies demonstrate minor similarities, the patterns already emerging highlight the potential application of breath volatilomics as a clinical aid for COVID-19 detection. However, the differences in results across these studies highlights the need for standardised sampling and analysis platforms for such investigations. As discussed in section 2.2. of this review, one of the major challenges in clinical volatilomics lies in validating results due to high variation in experimental methods used across the field. Even with clinical volatilomic studies that use the same techniques, cross-study comparisons are challenging due to factors associated with the sample size, variation of disease severity across participants, lifestyle factors across participants and participant comorbidities. Therefore, in order to develop and elevate clinical volatilomics towards future practical applications, a move towards standardised methodology must be made - this will ultimately require large collaboration across the field.

Ventilator associated pneumonia

Ventilator-associated pneumonia (VAP) is a hospital-acquired infection that commonly occurs in patients in intensive care units (ICUs). The mortality rate is estimated to be approximately 9%, however there are some estimations ranging from 24% to 50%, and upwards to as high as 76% depending on what pathogen(s) are driving the infection[158]. The pathogens associated with VAP include methicillin resistant *S. aureus, P. aeruginosa, Acinebacter spp. Haemophilus spp.,* and *Streptococcus pneumoniae (S. pneumonia)*[159][160]. In severe cases, VAP can lead to bacteremia – a highly deadly systemic infection circulating in the blood[161].

Volatilomic studies have aimed at characterising both VAP-specific VOCs and whole breath volatilomes of VAP patients. Schnabel et al.[162] demonstrated that VAP positive ICU patients and VAP negative ICU patients could be differentiated from each other based on their respective breath volatilomes. 104 critically ill patients with clinical suspicions of VAP were diagnosed either positive or negative for VAP using a bronchoalveolar lavage (BAL). 12 VOCs were subsequently identified that differentiated the two groups PCA results. Among these compounds were 2-methylbutane, heptane, dodecane, tetradecane, and ethanol. Though none of these compounds are individually unique with respect to microbial volatilomes, their collective emission pattern was unique enough to allow a moderate discrimination (sensitivity: 75.8% and specificity: 73.0 %) of the breath profiles of patients with VAP and non-VAP patients. Filipiak et al. analysed breath samples of 28 mechanically ventilated patients, 22 of whom had VAP. *S. aureus* was microbiologically confirmed in 5 VAP patients, which correlated with the identification of compounds such as 3-methylbutyraldehyde, ethyl acetate, acetic acid, and ethanol – all of which had been confirmed as metabolites for *S. aureus* in previous *in vitro* studies[14]. These compounds are all primary metabolites and originate from the breakdown of glucose and amino acids.

Detection of *in vitro* biomarkers in the breath of VAP patients was also observed by Gao et al.[15] who demonstrated a clear discrimination between the breath profiles of the *Acinetobacter baumanii* - free control group, the colonisation group, and the infection group, with the largest difference being between the control and infection groups. Before analysing the breath of the A. buamanii-positive VAP patients, the volatilome of *A. baumanii* pure cultures was analysed to identify potential growth-associated volatile biomarkers. Among the eight compounds that were present in both the pure cultures and the breath of the VAP patients were 1-undecene, longifolene, tetradecane, nonanal, and decanal. Particular precaution must be applied to the inclusion of compounds such as nonanal and decanal as possible biomarkers of human disease as both of these compounds are present in abundance across many sites of the human body[52]. Despite this, the detection of 1-undecene in the breath of the VAP patients is very interesting as this compound is also a metabolite that is emitted by *P. aeruginosa*[42][110] – another highly dangerous respiratory pathogen[163]. Interestingly, *P. aeruginosa* was shown to emit this 1-undecene in an ex vivo sputum model[113] discussed in the next paragraph. The detection of short-chain alkanes such as tetradecane was also observed in VAP volatilomic study carried out by Schnabel et al.[162]. In another *in vivo* investigation, Szikszay et al.[164] artificially induced *P. aeruginosa*-associated and *E. coli*-associated VAP on two individual groups of rabbits and analysed the resulting breath profiles using GC-IMS. The group detected compounds that were previously detected *in vitro*[165] that allowed the discrimination of the breath of infected rabbits and non-infected control group of rabbits. However, indole was the only compound of these five that was successfully identified, again highlighting the need for broad screening via GC-MS for such untargeted analyses.

Ex vivo sputum models have also been explored by Lawal et al.[166][50] as a potential route towards VAP biomarker detection. Similar to the idea of the ex vivo wound models discussed in section 2.4.4., the sputum models are artificial mixtures of various components of sputum and used as a source of growth for the bacteria. In these studies, following the analysis of various VAP-associated pathogens in both standard nutrient broth and the sputum model, multiple compounds such as indole (emitted by *E. coli*) , 1-undecene (emitted by *P. aeruginosa*), cyclopentanone (emitted by *E. coli* and *P. aeruginosa*), 1-hexanol (emitted by *E.*

29

coli), and thiocyanic acid, methyl ester (emitted by *P. aeruginosa*) were marked as potential indicators of bacterial infection in the lungs as they were detected in either both media or only the sputum model.

During the COVID-19 pandemic the international demand for ventilators significantly increased as the numbers of hospitalisations increased. Though there are currently few studies describing the incidence of VAP in ventilated COVID-19 patients, a recent observational study[167] determined that they are at a higher risk of developing VAP than ventilated non-COVID 19 patients. The authors propose that this is potentially due to the prolonged period of ventilation caused by COVID-19. If this holds true universally, investment in non-invasive rapid detection methods as a tool for detecting early stage VAP will be a critical step in preparing for the next pandemic.

Cystic fibrosis

Cystic fibrosis (CF) is a lethal genetic disease that primarily affects the lungs. Persistent inflammation of the lungs leads to progressive lung injury and destruction and ultimately causes a gradual in lung function[168]. As the lungs decline they become increasingly susceptible to infection by pathogens such as *P. aeruginosa* and *S. aureus*[169].

Microbial volatilomic *in vitro* studies have reported potential diagnostic markers of CF infections by *P. aeruginosa*[170] and *S. aureus*[171][172]. The potential diagnostic markers reported in these studies were absent from *in vivo* studies, instead, discrimination of infected and non-infected CF patients was achieved by broad untargeted analyses of whole volatilomes[173][174] . Neerincx et al.[173] reported that the compounds responsible for this discrimination between *S. aureus* infected and non-infected CF patients were 1,4-pentadiene, ethanol, acetone, 2-butanone, acetoin, hexanal, undecane, 2-methyl-naphthalene and isopropyl myristate. Although the CF group in the study carried out by Robroeks et al.[174] included patients a variety of microbial infections (60% *S. aureus* positive), they reported ten different discriminatory markers between the CF group and healthy controls; these included 3,3-dimethylhex-1-ene, 2-buten-1-ol, N-methyl-2-methylpropylamine and various poly-unsaturated and saturated hydrocarbons. The contrasting results of these two studies highlight the challenges of determining individual target compounds as volatile biomarkers of disease. These results are summarised in Table 1.3.

Gastric cancer

The human gastrointestinal tract has a critical role in health and disease and is one of the most diverse and abundant microbiomes on Earth. Many diseases have been characterised by the shifts in the gut microbiome they induce and vice versa. Gastric cancer is an example of this; as dysbiosis in the gut leads to *Helicobacter-pylori* infection which can then develop to tumorigenesis[175].

In a recent study[87] investigating the breath profile of patients with gastric cancer, significant increases in the abundance of fatty acids were detected in the breath of patients with gastric cancer compared to healthy controls. These results were validated with direct measurement of fatty acids from ex vivo tissue biopsies from the stomachs of cancer patients and healthy controls. The specific fatty acids that were detected in higher abundance in the gastric cancer patients are listed in Table 1.3. A proposed mechanism by which these fatty acids are released through exhaled breath has been linked to the potential partitioning of microbial metabolites across the digestive tract and into the airways[176]. Higher abundances of fatty acids in the breath of patients with gastric cancer have also been detected by Kumar et al.[88][89]. In these

studies, abundances of phenols (phenol, methyl phenol and ethyl phenol) and aldehydes were also found to be higher in the breath of gastric cancer patients compared to healthy controls.

Another potential route for detecting volatile biomarkers of gastric cancer could be through the analysis of faecal samples. Volatilomic analyses of human faecal samples has been previously demonstrated in a number of studies[77][72] and could potentially be highly valuable for the diagnosis of a variety of diseases.

Chronic wounds

It is estimated that around one in four people with diabetes will develop a diabetic foot ulcer (DFU) in their lifetime[177]. The feet of diabetic patients can become ischaemic due to macrovascular disease; this pressure results in the skin becoming very thin and susceptible to breaking. Once the skin is broken, many processes contribute to delayed healing and ulceration can occur[178]. Infections of DFUs are related with poor outcomes, a 12-month observational study reported that out of 299 participants, ulcers only healed in 136 participants (45.5%) but recurred in 13; lower extremity amputations were recorded in 52 (17.4%) participants and revascularization surgery in 18 (6.0%) participants; 45 (15.1%) of the participants died[179]. Ulcer duration has been positively correlated with bacterial diversity while the severity (depth) of ulcers has been positively correlated with the abundance of anaerobic bacteria[180]. The strain-level diversity of *S. aureus* has also been positively correlated with poor outcomes[181]. Endogenous blood biomarkers such as white blood cell count, erythrocyte sedimentation rate, C-reactive protein and procalcitonin are regularly used as indicators of infection and a relative measure of infection severity[182][183]. However, as these are only inflammatory biomarkers, they are relatively unreliable for the discrimination of infection-associated inflammation and non-infection-associated inflammation.

Early volatilomic profiling of chronic wounds and health skin controls highlighted discriminatory patterns from the wound samples[184]. Healthy skin is not an appropriate control for chronic wounds as there are significant differences in the microbial composition[185] and chemical environment[186] between wounds and healthy skin. The instability of the skin volatilome has been demonstrated by measurable shifts following minor barrier disruption via several rounds of tape stripping. Open wound beds are also regions of high oxidative stress[187], which can generate volatiles through the molecular breakdown of the cellular components of the surrounding tissue[188]. From a clinical stand point, non-infected wounds may serve as more suitable controls than healthy skin for volatile profiling. Non-infected wounds are wounds that show no visual or microbiological indications of infection. M. Haalboom et al. demonstrated relative differentiation of the volatile signatures of infected wounds (n=37) and non-infected wounds (n=40) using an e-nose[55]. Although this study highlights the volatilomic discrimination between infected and non-infected wound samples, the compounds responsible for that trend were not identified.

The analysis of volatiles emitted from the dressings taken from cancer-associated fungative wounds demonstrated that dimethyl trisulfide is a major contributor in malodourous wounds and a potential indicator of bacterial infection[189][190]. Other compounds detected in these wounds include dimethyl disulfide, indole, 3-methylbutanal, and phenol[189] (Table 1.3). In a study investigating what chemicals attract screwworm flies to animal wounds, the volatiles of multiple wound-associated bacteria were analysed *in vitro* in blood cultures[191]. Dimethyl

31

disulfide and dimethyl trisulfide were subsequently labelled as major contributors to the volatilome of these cultures (Table 1.3).

Ex vivo wound and biofilm models present an alternative approach for studying the volatilome of the wound environment. Ashrafi et al. reported measurable shifts in the volatilome of ex vivo wound models in response to infection and biofilm production[83]; and antimicrobial therapy by way of antibiotics and electrical stimulation[84]. Most recently Slade et al.[95] used SESI-MS to discriminate the volatilomes of multiple strains of bacteria cultured in collagen-based biofilms. These studies highlight new ways in which microbial volatilomes can be investigated and may serve as a key bridging point between *in vitro* and *in vivo* volatilomics. Future ex vivo wound volatilomic studies should expand on this work by using broader sampling and analysis techniques to capture a more extensive range of volatiles in these systems.

Table 1.3: Discriminative compounds identified in clinical volatilomic studies and their relationship with microbial volatilomics

Compound	Chemical class	Potential production Mechanism	Sampling	Analysis	Confirmed microbes In vitro	In vivo
Indole	Indole	Tryptophan metabolism (shikimate pathway)	SPME [189] MCC-IMS[165]	GC-MS[191][189], MCC-IMS[165]	*E.coli* [165][42][45] *P. vulgaris*[191], *P. rettgeri* [191], *P. mirabilis*[190], *K. oxytoca*[191], *P. stuartii*	VAP[165], Wounds[191]
Dimethyl trisulfide	Sulfide	Sulfur metabolism	SPME[191], [189]/ sorption tubes [165]	GC-MS [191] [189]	*P. vulgaris*[191], *P. rettgeri* [191], *P. mirabilis*[190], *K. oxytoca*[191], *P. stuartii* [191], *P. aeruginosa* [10], *S. pneumoniae* [128], *H. influenzae* [128]	Wounds [191] [189], Oral malodour
Dimethyl disulfide	Sulfide	Sulfur metabolism	SPME [191] [189]	GC-MS [191] [189]	*P. vulgaris*[191], *P. rettgeri* [191], *P. mirabilis*[190], *K. oxytoca*[191], *P. stuartii* [191], *S. aureus* [10], *S. pneumoniae* [128], *H. influenzae* [128]	Wounds [191] [189], Oral malodour
Dimethyl sulfide	Sulfide	Sulfur metabolism	Sorption tube [14] [149]	GC-MS[14][149]	*C. albicans*[14], *M.tuberculosis* [149], *S. pneumoniae* [128], *H. influenzae* [128]	VAP[14]
Ethyl acetate	Acetate	Primary metabolism	Sorption tube[10][193]	GC-MS[10][193]	Many	VAP[10],CF infection[193]
Acetone	Ketone	Lipid metabolism	Sorption tubes [74] [162] [173]	GC-MS [74][162] [173]	Many	COVID-19 [74] VAP[162][192], CF infection[173]
2-Butanone	Ketone	Primary metabolism	Sorption tubes [74][173]	GC-MS[7474][173],	Many	COVID-19 [74], CF infection[173][193]

Acetoin	Ketone	Primary metabolism	Sorption tubes [74] [173] [10] , SPME [42][45]	GC-MS[74][10] [173][42][45]	*S. aureus* [10][42][45] *E. coli*[42][45]	COVID-19 [74]CF infection[173]
Methanol	Alcohol	Primary metabolism	Sorption tubes [82]	GC-MS[82]		COVID-19 [82]
2,2-dimethyl 1-propanol	Alcohol		Sorption tubes [74]74	GC-MS[74]7474		COVID-19 [74]74
1-Propanol	Alcohol	Primary metabolism	Sorption tubes [74]	GC-MS[74]74		COVID-19 [74]74
2-Butyl-1-octanol	Alcohol	Fatty acid metabolism	SPME / sorption tubes [15] [76]	GC-MS[15]	*A. baumanii*[15]	VAP[15], Tuberculosis[76]
Ethanol	Alcohol	Primary metabolism	Sorption tube [14] [162][173]	GC-MS[162][14] [173]	Many	VAP[162] [14], CF infection[173]
Isopropyl Alcohol	Alcohol	Primary metabolism	Sorption tube [162]	GC-MS[162]		VAP[162]
Ethanal	Aldehyde	Primary metabolism	Sorption tubes [74] [82]	GC-MS[74][82]		COVID-19 [74][82]
Propanal	Aldehyde	Primary metabolism	Sorption tube [14]	GC-MS[14]	*S. aureus*[14], *C. albicans*[14]	VAP[14]
3-Methylbutanal	Aldehyde	Primary metabolism	Sorption tube [14] [76], SPME [42][45]	GC-MS[14] [42][45]	*S. aureus* [14][42][45] *M. tuberculosis* [76]	VAP[14], Tuberculosis [76]
Methylpent-2-enal	Aldehyde	Fatty acid metabolism		PTR-MS [90]		COVID-19 [90]
Heptanal	Aldehyde	Fatty acid metabolism	Sorption tubes [75] [82] [76]	GC-MS[75] [82] [76], SIFT-MS[89]		COVID-19 [75] [82], Tuberculosis [76], Gastric cancer[89]
Octanal	Aldehyde	Fatty acid metabolism	Sorption tubes [75] [82]	GC-MS[75] [82], SIFT-MS[89]		COVID-19 [74] [82] [75], Gastric cancer[89]
Nonanal	Aldehyde	Fatty acid metabolism	Sorption tubes [75] [76]	GC-MS [76] [7474][75],PTR-MS[90], SIFT-MS[89]	*A. baumanii*[15]	COVID-19 [74] [90] [75], Tuberculosis[76], Gastric cancer[89]
Decanal	Aldehyde	Fatty acid metabolism	SPME / sorption tubes [15]	SIFT-MS[89]	*A. baumanii*[15]	Gastric cancer[89], VAP [15]
Tetradecanal	Aldehyde	Fatty acid metabolism	Sorption tube [162]	GC-MS[162][42]	*E.coli* [42]	VAP[162]
Benzaldehyde	Aldehyde	Primary metabolism	Sorption tubes [74]	GC-MS[74]		COVID-19 [74]

33

Compound	Class	Metabolism	Sampling	Analysis	Microorganism	Disease
Acetaldehyde	Aldehyde	Primary metabolism	Sorption tube [14] [76]	GC-MS[14][76]	*S. aureus*[14], *C. albicans*[14], *S. pneumoniae* [128], *H. influenzae* [128]	VAP[14], Tuberculosis[76]
Acrolein	Aldehyde		Sorption tube [162]	GC-MS[162]		VAP[162]
Phenol	Phenol		SPME[189][191]	GC-MS[189][191] SIFT-MS[88][89]	*P. vulgaris*[191], *P. rettgeri* [191], *P. mirabilis*[190], *K. oxytoca*[191], *P. stuartii* [191]	Wounds [191] [189], Gastric cancer[87] [88][89]
Methyl phenol	Phenol			SIFT-MS [88][89]		Gastric cancer [88][89]
Ethyl phenol	Phenol			SIFT-MS [88][89]		Gastric cancer [88][89]
4- (1,1-dimethylpropyl)phenol	Phenol		Sorption tube [16]	GC-MS[16]	*M.tuberculosis* [16]	Tuberculosis[16]
Cyclohexane	Hydrocarbon		Sorption tubes [76]	GC-MS[76]		Tuberculosis[76]
Butane	Hydrocarbon	Fatty acid metabolism	Sorption tube [14]	GC-MS[14]	*S. aureus* [14]	VAP[14]
Butane, 2-methyl	Hydrocarbon		Sorption tube [162]	GC-MS[162]		VAP[162]
Heptane	Hydrocarbon	Fatty acid metabolism	Sorption tube [162]	GC-MS[162]		VAP[162]
Octane	Hydrocarbon	Fatty acid metabolism	Sorption tube [192]	GC-MS[192]		VAP[192]
Nonane	Hydrocarbon	Fatty acid metabolism	Sorption tube [192]	GC-MS[192]		VAP[192]
4-ethyl-2,2,6,6-tetramethylheptane	Hydrocarbon	Fatty acid metabolism	Sorption tube [16]	GC-MS[16]	*M.tuberculosis* [16]	Tuberculosis[16]
5-methyl-5-propyl-nonane	Hydrocarbon	Fatty acid metabolism	SPME / sorption tubes [15]	GC-MS[15]	*A. baumanii*[15]	VAP[15]
Decane	Hydrocarbon	Fatty acid metabolism	Sorption tubes [75]	GC-MS[75]	*P. aeruginosa*[45][42], *S. epidermidis* [45] *E. coli* [42]	COVID-19 [75]
Carane	Hydrocarbon		Sorption tube [162]	GC-MS[162]		VAP[162]
Undecane	Hydrocarbon	Fatty acid metabolism	Sorption tube[192] [173]	GC-MS[192] [173],		CF infection[173], VAP[192]
2,6,10-trimethyl-dodecane	Hydrocarbon	Fatty acid metabolism	SPME / sorption tubes [15]	GC-MS[15]	*A. baumanii*[15]	VAP[15]
Tridecane	Hydrocarbon		Sorption tubes [74] [76] [192]	GC-MS[74][76] [192]	*P. aeruginosa*[45]	COVID-19 [74], Tuberculosis[76], VAP[192]

34

Tetradecane	Hydrocarbon	Fatty acid metabolism	SPME[15] sorption tubes [15] [14] [192]	GC-MS[15][192] [14]	*A. baumanii*[15], E. coli [45]	VAP[15] [14] [192]
Pentadecane	Hydrocarbon	Fatty acid metabolism	Sorption tubes [74] [192]	GC-MS[74] [192]	*P. aeruginosa*[45]	COVID-19 [74], VAP[192]
2-Methylpropene	Hydrocarbon	Fatty acid metabolism	Sorption tube [14]	GC-MS[14]	*S. aureus* [14]	VAP[14]
Cyclohexene	Hydrocarbon		Sorption tubes [74]	GC-MS[74]		COVID-19 [74]74
3-Heptene	Hydrocarbon	Fatty acid metabolism	Sorption tubes [74]	GC-MS[74]		COVID-19 [74]74
1-Octene	Hydrocarbon	Fatty acid metabolism	Sorption tubes [76]	GC-MS[76]		Tuberculosis[76]
4-Methyl-1-decene	Hydrocarbon	Fatty acid metabolism	Sorption tube [16]	GC-MS[[16]	*M.tuberculosis* [16]	Tuberculosis[16]
1-Undecene	Hydrocarbon	Fatty acid metabolism	SPME / sorption tubes [15]	GC-MS[15]	*A. baumanii* [15] *P. aeruginosa*[109][110] [42][45]	VAP[15]
Longifolene	Hydrocarbon	Fatty acid metabolism	SPME / sorption tubes [15]	GC-MS[15]	*A. baumanii*[15]	VAP[15]
1,3-Butadiene	Hydrocarbon	Fatty acid metabolism	Sorption tube [14]	GC-MS[14]	*S. aureus* [14], *S. pneumonae* [128]	VAP[14]
1,4-Pentadiene	Hydrocarbon		Sorption tube [173]	GC-MS[173]	*S. aureus* [173]	CF infection[173]
2,4-Octadiene	Hydrocarbon		PTR-MS [90]	PTR-MS [90]		COVID-19 [90]
1-Chloroheptane	Hydrocarbon		PTR-MS [90]	PTR-MS [90]		COVID-19 [90]
2,3,6-Trimethylnapthalene	Hydrocarbon	Fatty acid metabolism	Sorption tube [16]	GC-MS[16]	*M.tuberculosis* [16]	Tuberculosis[16]
Naphthalene, 1-methyl-	Hydrocarbon	Fatty acid metabolism	Sorption tubes [76]	GC-MS[76]		Tuberculosis[76]
2-Methyl naphthalene	Hydrocarbon		Sorption tube[173]	GC-MS[173]		CF infection[173]
2-Penthyl furan	Furan		Sorption tubes [75]	GC-MS[75]		COVID-19 [75]
Furan, tetrahydro-	Furan		Sorption tube [162]	GC-MS[162]		VAP[162]

Acetic acid	Acid	Primary metabolism	Sorption tube [14] [87]	GC-MS[14] [42]	*S. aureus* [14] [42][45], *E.coli* [42][45], *S. pneumoniae* [128], *H. influenzae* [128]	VAP[14], Gastric cancer[87]
Butyric acid	Acid	Primary metabolism	Sorption tube [14] [87]	GC-MS[14][87], PTR-MS[87]	*S. aureus*[14]	VAP [14] Gastric cancer[87]
Pentanoic acid	Acid	Primary metabolism	Sorption tube [87]	GC-MS[87], PTR-MS[87], SIFT-MS[88][89]		Gastric cancer [87][88][89]
Hexanoic acid	Acid	Primary metabolism	Sorption tube [87]	GC-MS[87],PTR-MS[87], SIFT-MS[88][89]		Gastric cancer[87][88][89]

1.3. Other challenges in clinical volatilomics

One of the major challenges in clinical volatilomes is the lack of standardised experimental procedures for sample collection, sampling and analysis of volatiles from clinical subjects and specimens. This challenge has been discussed in section 2.2. of this review. In this section, the challenges of background interferences and the detection of alkanes in clinical samples will be discussed. These issues are highly prevalent in clinical volatilomics and collectively can be responsible for acquisition of inaccurate results and conclusions.

1.3.1. The background issue

Eliminating background interferences is a complex issue in clinical volatilomics. Volatile organic compounds are generated from a huge variety of both endogenous and exogenous sources. Firstly, understanding the regular factors that influence the volatilome of healthy individuals day-to-day is critical for identifying disease-associated volatilomic shifts in the future. This involves identifying the factors that influence the regular fluctuations that characterise the baseline volatilome of healthy humans. Each analytical matrix whether it be breath, sputum, skin, urine, blood, faeces, sperm, sweat or a wound has a background volatilome[194] that must be characterised prior to the identification of infection-specific or inflammation-specific volatilomic shifts. This diversity of the "healthy" human volatilome is illustrated in a compendium review by Costello et al.[127] which reports a total of 1840 volatile emanating from various components of the human body. This compound number was recently updated to 2746 compounds in a updated version of the original compendium by Drabinska et al[195] – which emphasizes the critical importance of updating such reviews. As discussed in section 2.3.2. of this review, the detection of certain inflammatory diseases such as viral infections may require particular attention towards detecting abnormal shifts in the abundance of these regular compounds volatilome rather than the detection of unique biomarkers. However, there may be more targeted solutions available for volatilomic profiling of bacterial- and fungal-associated infectious diseases due to the unique metabolic pathways that are at play. The metabolic responses of these pathogens *in vivo* may give rise to unique compound emissions that can be quantified in the future.

Secondly, setting sufficient controls for the interferences introduced from the experimental set-up minimises contamination in the analysis. For example, prior to SPME-GC-MS analyses of samples, analysing a blank SPME fiber to identify background compounds that are present on the fiber and also potential contaminant compounds retained on the GC column. Background volatilomic interferences will also significantly depend on the sample collection procedure and must be included in the blank control analyses. Factors that influence

background interferences during sample collection include direct contact with the sample site; volatilome of the sample collection tool (e.g. swab, PDMS patch, TEDLAR bag etc.) ; relative pre-treatments of sample prior to analysis; and the volatile background of the indoor environment where the sample is taken[196][197][198]. The importance of this latter point is especially relevant when discussing the TB skin detection study [151], as the authors themselves highlight the substantial background indoor abundances of their analytes of interest. A recent study[199] of indoor air in clinical environments highlighted the complexity of the indoor volatilome with respect to the many exogenous and endogenous sources of compounds. Despite this complexity, however, the authors proposed threshold levels for potential compounds of interest. Compound abundances below this threshold level, are classified as background interference, and if they are above the threshold level they can be included in the study.

1.3.2.   Accurate detection of branched molecules

Although GC-MS analysis provides high sensitivity of detection as well as a vast reference library of compounds, it is limited in its ability to accurately differentiate between branched alkanes. The first challenge associated with accurately identifying branched alkanes is that many branched alkanes share highly similar mass spectrums making accurate manual interpretation and identification difficult. The degree of difficulty in accurately interpreting and identifying these compounds also increases as the molecular weight of the compounds increase due to the higher number of possible structural combinations. This challenge also holds through for a variety of chemical classes such as ketones, alcohols, and esters.

The second challenge is associated with the Kovats retention index, which is frequently employed in GC-MS investigations to validate the identification of chromatographic peaks. The retention index relates the retention time of an analyte peak to the retention time of linear alkanes where the index (RI) of a linear alkane equals 100 times the number of carbon atoms (e.g. octane, RI = 800). Due to highly similar molecular weights and chemical structures, the RI of branched alkanes will also be highly similar which then further reduces confidence in accurate detection.

Both of these peak identification techniques are frequently used in microbial and clinical volatilomic studies. This presents a challenge, particularly in clinical volatilomics as many studies (Table 3) report various branched alkanes as discriminating components between disease-associated and non-disease associated volatilomes. Aliphatic alkanes are produced via lipid peroxidation[200][201] in the human body and have been correlated with the production of reactive oxygen species (ROS) during inflammation[202]. Although these compounds hold potentially valuable information on the state-of-disease or -inflammation in a given sample, careful consideration should be applied when assigning specific identities to the observed chromatographic peaks. Higher degrees of discrimination of complex alkanes has been previously achieved using two dimensional gas chromatography (GC x GC)[203]. This technique has been employed in clinical volatilomics for higher resolution of analyte peaks[16][75] in breath samples. The data analysis workflows and operation behind this analytical approach have been recently described[204]. Alternatively, if using standard GC-MS analysis, it may be more reasonable to simply report the collective abundance of alkanes across the samples alongside suspected identifications.

1.4.   Conclusion

In the last 15 years, the clinical potential of volatilomics has been demonstrated in numerous studies through the detection of discriminative volatilomic patterns for a variety of infectious diseases. The successes of these analyses have come through the use of untargeted screening of whole volatilomes rather than targeted screening for specific metabolites. Frequently in these studies, volatilomic discriminations of patients with disease have been through the detection of shifts in the abundance of multiple normally-occurring volatile components of the human volatilome rather than the detection of disease-specific compounds. However, microbial volatile metabolites have also been frequently detected in these studies and also contribute to these discriminations. Microbial *in vitro* volatilomics therefore plays a critical role alongside clinical volatilomics by validating the potential microbial cellular origin of compounds detected across *in vivo* studies. Studies that have used consistent experimental and data analysis workflows across both microbial and clinical volatilomics investigations have been successful at detecting various corresponding compounds across these systems. The broad standardisation of analysis techniques across the field is critical for accumulating and validating untargeted volatilomic data. Analysis techniques with large reference libraries such as GCMS should be employed alongside any other analyses to validate the identification of compounds. In addition to validating compound discoveries, wide spread standardisation of analysis techniques will also mitigate the current challenges posed by low sample sizes and background interferences. The gradual accumulation of non-specific volatilomic data will allow the identification of potential target analytes for future identifications - this is where the value of direct-quantification MS instruments will be realised. Within microbial volatilomics, the study of volatile-mediated microbial interactions allows a broader understanding of community dynamics within microbiomes to be obtained. As the knowledge of these highly specific volatile-mediated interactions grows, novel opportunities in future probiotic and antibiotic research could potentially emerge.

1.5.   References

[1] Shang, Y., Li, H. and Zhang, R., 2021. Effects of Pandemic Outbreak on Economies: Evidence From Business History Context. *Frontiers in Public Health*, 9. https://doi.org/10.3389/fpubh.2021.632043

[2] Verikios, G., 2020. The dynamic effects of infectious disease outbreaks: The case of pandemic influenza and human coronavirus. *Socio-Economic Planning Sciences*, 71, p.100898. https://doi.org/10.1016/j.seps.2020.100898

[3] Brancalion, P., Broadbent, E., de-Miguel, S., Cardil, A., Rosa, M., Almeida, C., Almeida, D., Chakravarty, S., Zhou, M., Gamarra, J., Liang, J., Crouzeilles, R., Hérault, B., Aragão, L., Silva, C. and Almeyda-Zambrano, A., 2020. Emerging threats linking tropical deforestation and the COVID-19 pandemic. *Perspectives in Ecology and Conservation*, 18(4), pp.243-246. https://doi.org/10.1016/j.pecon.2020.09.006

[4] Wolfe, N., Daszak, P., Kilpatrick, A. and Burke, D., 2005. Bushmeat Hunting, Deforestation, and Prediction of Zoonotic Disease. *Emerging Infectious Diseases*, 11(12), pp.1822-1827. http://dx.doi.org/10.3201/eid1112.040789

[5] Frieri, M., Kumar, K. and Boutin, A., 2017. Antibiotic resistance. *Journal of Infection and Public Health*, 10(4), pp.369-378. https://doi.org/10.1016/j.jiph.2016.08.007

[6] Chen, J., Ying, G. and Deng, W., 2019. Antibiotic Residues in Food: Extraction, Analysis, and Human Health Concerns. *Journal of Agricultural and Food Chemistry*, 67(27), pp.7569-7586. https://doi.org/10.1021/acs.jafc.9b01334

[7] Hoshi, K., Yamano, Y., Mitsunaga, A., Shimizu, S., Kagawa, J. and Ogiuchi, H., (2002). Gastrointestinal diseases and halitosis: association of gastric Helicobacter pyloriinfection. *International Dental Journal*, 52(5), pp.207-211. https://doi.org/10.1002/j.1875-595X.2002.tb00926.x

[8] Budd, M., Tanaka, K., Holmes, L., Efron, M., Crawford, J. and Isselbacher, K., (1967). Isovaleric Acidemia. *New England Journal of Medicine*, 277(7), pp.321-327. https://doi.org/10.1056/NEJM196708172770701

[9] Insam, H. and Seewald, M., (2010). Volatile organic compounds (VOCs) in soils. *Biol. Fertil. Soils*, 46(3), pp.199-213. https://doi.org/10.1007/s00374-010-0442-3 .

Pre-examination copy
Submitted 27th May 2022

[10] Filipiak, W., Sponring, A., Baur, M., Filipiak, A., Ager, C., Wiesenhofer, H., Nagl, M., Troppmair, J. and mann, A. (2012). Molecular analysis of volatile metabolites released specifically by staphylococcus aureus and pseudomonas aeruginosa. *BMC Microbiol.* , 12(113). https://doi.org/10.1186/1471-2180-12-113.

[11] Chen, J., Tang, J., Shi, H., Tang, C. and Zhang, R. (2016).Characteristics of volatile organic compounds produced from five pathogenic bacteria by headspace-solid phase micro-extraction/gas chromatography-mass spectrometry. *J. Basic Microbiol*, 57(3), pp.228-237. https://doi.org/10.1002/jobm.201600505.

[12] Misztal, P., Lymperopoulou, D., Adams, R., Scott, R., Lindow, S., Bruns, T., Taylor, J., Uehling, J., Bonito, G., Vilgalys, R. and Goldstein, A., (2018). Emission Factors of Microbial Volatile Organic Compounds from Environmental Bacteria and Fungi. *Environ. Sci. Technol.* , 52(15), pp.8272-8282. https://doi.org/10.1021/acs.est.8b00806.

[13] Schulz-Bohm, K., Martín-Sánchez, L. and Garbeva, P., (2017).Microbial Volatiles: Small Molecules with an Important Role in Intra- and Inter-Kingdom Interactions. *Frontiers in Microbiology*, 8. https://doi.org/10.3389/fmicb.2017.02484.

[14] Filipiak, W., Beer, R., Sponring, A., Filipiak, A., Ager, C., Schiefecker, A., Lanthaler, S., Helbok, R., Nagl, M., Troppmair, J. and Amann, A., Breath analysis for in vivo detection of pathogens related to ventilator-associated pneumonia in intensive care patients: a prospective pilot study. *J. Breath Res.*, 9(1), p.016004. https://doi.org/10.1088/1752-7155/9/1/016004 . (2015).

[15] Gao, J., Zou, Y., Wang, Y., Wang, F., Lang, L., Wang, P., Zhou, Y. and Ying, K., Breath analysis for noninvasively differentiating Acinetobacter baumannii ventilator-associated pneumonia from its respiratory tract colonization of ventilated patients. *J. Breath Res.*, 10(2), p.027102. https://doi.org/10.1088/1752-7155/10/2/027102 . (2016).

[16] Mellors, T., Nasir, M., Franchina, F., Smolinska, A., Blanchet, L., Flynn, J., Tomko, J., O'Malley, M., Scanga, C., Lin, P., Wagner, J. and Hill, J., Identification of Mycobacterium tuberculosis using volatile biomarkers in culture and exhaled breath. *J. Breath Res.*, 13(1), p.016004. https://doi.org/10.1088/1752-7163/aacd18 . (2018).

[17] Weisskopf, L., Schulz, S. and Garbeva, P., (2021). Microbial volatile organic compounds in intra-kingdom and inter-kingdom interactions. *Nature Reviews Microbiology*,. https://doi.org/10.1038/s41579-020-00508-1

[18] Elmassry, M. & Piechulla, B. (2020). Volatilomes of bacterial infections in humans. *Front. Neurosci.* https://doi.org/10.3389/fnins.2020.00257

[19] Kai, M., (2020). Diversity and Distribution of Volatile Secondary Metabolites Throughout Bacillus subtilis Isolates. *Frontiers in Microbiology*, 11. https://doi.org/10.3389/fmicb.2020.00559 .

[20] Schulz, S. & Dickschat, J. (2007). Bacterial volatiles: The smell of small organisms. *ChemInform.* https://doi.org/10.1039/b507392h

[21] Ryu, C., Weisskopf, L. and Piechulla, B., (2020). *Bacterial volatile compounds as mediators of airborne interactions*. Singapore: Springer.

[22] Lemfack,M. C. *et al*. (2017).mVOC 2.0: a database of microbial volatiles. *Nucleic Acids Res*. https://doi.org/10.1093/nar/gkx1016..

[23] Pott, D., Osorio, S. and Vallarino, J., 2019. From Central to Specialized Metabolism: An Overview of Some Secondary Compounds Derived From the Primary Metabolism for Their Role in Conferring Nutritional and Organoleptic Characteristics to Fruit. *Frontiers in Plant Science*, 10. https://doi.org/10.3389/fpls.2019.00835

[24] Anraku, Y., 1988. BACTERIAL ELECTRON TRANSPORT CHAINS. *Annual Review of Biochemistry*, 57(1), pp.101-132. https://doi.org/10.1146/annurev.bi.57.070188.000533

[25] Kracke, F., Vassilev, I. and Kramer, J., 2015. Microbial electron transport and energy conservation - the foundation for optimizing bioelectrochemical systems. *Frontiers in Microbiology*, 6. https://doi.org/10.3389/fmicb.2015.00575

[26] Basan, M., Hui, S., Okano, H., Zhang, Z., Shen, Y., Williamson, J. and Hwa, T., 2015. Overflow metabolism in Escherichia coli results from efficient proteome allocation. *Nature*, 528(7580), pp.99-104. https://doi.org/10.1038/nature15765

[27] Wolfe, A., 2005. The Acetate Switch. *Microbiology and Molecular Biology Reviews*, 69(1), pp.12-50. https://doi.org/10.1128/MMBR.69.1.12-50.2005

[28] Huberts, D., Niebel, B. and Heinemann, M., 2011. A flux-sensing mechanism could regulate the switch between respiration and fermentation. *FEMS Yeast Research*, 12(2), pp.118-128. https://doi.org/10.1111/j.1567-1364.2011.00767.x

[29] Pfeiffer, T. and Morley, A., 2014. An evolutionary perspective on the Crabtree effect. *Frontiers in Molecular Biosciences*, https://doi.org/10.3389/fmolb.2014.00017

[30] Szenk, M., Dill, K. and de Graff, A., 2017. Why Do Fast-Growing Bacteria Enter Overflow Metabolism? Testing the Membrane Real Estate Hypothesis. *Cell Systems*, 5(2), pp.95-104. https://doi.org/10.1016/j.cels.2017.06.005

[31] Zhuang, K., Vemuri, G. and Mahadevan, R., 2011. Economics of membrane occupancy and respiro-fermentation. *Molecular Systems Biology*, 7(1), p.500. https://doi.org/10.1038/msb.2011.34

[32] Lemfack, M., Ravella, S., Lorenz, N., Kai, M., Jung, K., Schulz, S. and Piechulla, B., Novel volatiles of skin-borne bacteria inhibit the growth of Gram-positive bacteria and affect quorum-sensing controlled phenotypes of Gram-negative bacteria. *Syst. Appl. Microbiol.*, 39(8), pp.503-515. https://doi.org/10.1016/j.syapm.2016.08.008. (2016).

[33] Díaz-Pérez, A., Díaz-Pérez, C. and Campos-García, J., 2015. Bacterial l-leucine catabolism as a source of secondary metabolites. *Reviews in Environmental Science and Bio/Technology*, 15(1), pp.1-29. https://doi.org/10.1007/s11157-015-9385-3

[34] Dansette, P. and Azerad, R., 1974. The shikimate pathway. *Biochimie*, 56(5), pp.751-755. https://doi.org/10.1146/annurev.arplant.50.1.473

[35] Zhang, H., Cao, M., Jiang, X., Zou, H., Wang, C., Xu, X. and Xian, M., 2014. De-novo synthesis of 2-phenylethanol by Enterobactersp. CGMCC 5087. *BMC Biotechnology*, 14(1). https://doi.org/10.1186/1472-6750-14-30

[36] Roager, H. and Licht, T., 2018. Microbial tryptophan catabolites in health and disease. *Nature Communications*, 9(1). https://doi.org/10.1038/s41467-018-05470-4

[37] Mann, S., 1967. Chinazolinderivate bei Pseudomonaden. *Archiv für Mikrobiologie*, 56(4), pp.324-329.

[38] Flemming, H., Wingender, J., Szewzyk, U., Steinberg, P., Rice, S. and Kjelleberg, S., 2016. Biofilms: an emergent form of bacterial life. *Nature Reviews Microbiology*, 14(9), pp.563-575. https://doi.org/10.1038/nrmicro.2016.94

[39] Zhu, Z., Wang, H., Shang, Q., Jiang, Y., Cao, Y. and Chai, Y., 2012. Time Course Analysis of Candida albicans Metabolites during Biofilm Development. *Journal of Proteome Research*, 12(6), pp.2375-2385. https://doi.org/10.1021/pr300447k

[40] Fujita, Y., Matsuoka, H. and Hirooka, K., 2007. Regulation of fatty acid metabolism in bacteria. *Molecular Microbiology*, 66(4), pp.829-839.

[41] Clayden, J., Greeves, N. and Warren, S., 2012. *Organic chemistry*. 2nd ed. Oxford: Oxford University Press, pp.1160-1162.

[42] Fitzgerald, S., Holland, L. and Morrin, A., 2021. An Investigation of Stability and Species and Strain-Level Specificity in Bacterial Volatilomes. *Frontiers in Microbiology*, 12. https://doi.org/10.3389/fmicb.2021.693075

[43] Brock, N., Menke, M., Klapschinski, T. and Dickschat, J., 2014. Marine bacteria from the Roseobacter clade produce sulfur volatiles via amino acid and dimethylsulfoniopropionate catabolism. *Organic & Biomolecular Chemistry*, 12(25), p.4318.

Pre-examination copy
Submitted 27th May 2022

44 Schäfer, H. and Eyice, Ö., 2019. Microbial Cycling of Methanethiol. *Current Issues in Molecular Biology*, pp.173-182.

45 Fitzgerald, S., Duffy, E., Holland, L. and Morrin, A., (2020). Multi-strain volatile profiling of pathogenic and commensal cutaneous bacteria. *Scientific Reports*, 10(1). https://doi.org/10.1038/s41598-020-74909-w .

46 Landaud, S., Helinck, S. and Bonnarme, P., 2008. Formation of volatile sulfur compounds and metabolism of methionine and other sulfur compounds in fermented food. *Applied Microbiology and Biotechnology*, 77(6), pp.1191-1205.https://doi.org/10.1007/s00253-007-1288-y

47 Lee, H., Kho, H., Chung, J., Chung, S. and Kim, Y., 2006. Volatile Sulfur Compounds Produced by Helicobacter pylori. *Journal of Clinical Gastroenterology*, 40(5), pp.421-426. https://doi.org/10.1097/00004836-200605000-00011

48 Rees, C., Shen, A. and Hill, J., 2016. Characterization of the Clostridium difficile volatile metabolome using comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry. *Journal of Chromatography B*, 1039, pp.8-16. https://doi.org/10.1016/j.jchromb.2016.11.009

49 Ahmed, W., Lawal, O., Nijsen, T., Goodacre, R. and Fowler, S., 2017. Exhaled Volatile Organic Compounds of Infection: A Systematic Review. *ACS Infectious Diseases*, 3(10), pp.695-710. https://doi.org/10.1021/acsinfecdis.7b00088

50 Lawal, O., Ahmed, W., Nijsen, T., Goodacre, R. and Fowler, S., 2017. Exhaled breath analysis: a review of 'breath-taking' methods for off-line analysis. *Metabolomics*, 13(10).

51 Ara, K., Hama, M., Akiba, S., Koike, K., Okisaka, K., Hagura, T., Kamiya, T. and Tomita, F., Foot odor due to microbial metabolism and its control. *Can. J. Microbiol.*, 52(4), pp.357-364. https://doi.org/10.1139/w05-130 (2006).

52 M. Gallagher, C.J. Wysocki, J.J. Leyden, A.I. Spielman, X. Sun, G. Preti, Analyses of volatile organic compounds from human skin, Br. J. Dermatol. 159 . 780e791. https://doi.org/10.1111/j.1365-2133.2008.08748.x. (2008).

53 Duffy, E., Jacobs, M., Kirby, B. and Morrin, A., Probing skin physiology through the volatile footprint: Discriminating volatile emissions before and after acute barrier disruption. *Exp. Dermatol.*, 26(10), pp.919-925. https://doi.org/10.1111/exd.13344 (2017).

54 Wu, L., Yuan, Z., Li, Z., Huang, Z. and Hu, B., 2020. In vivo solid-phase microextraction swab sampling of environmental pollutants and drugs in human body for nano-electrospray ionization mass spectrometry analysis. *Analytica Chimica Acta*, 1124, pp.71-77.

55 Haalboom, M., Gerritsen, J. and van der Palen, J., Differentiation between infected and non-infected wounds using an electronic nose. *Clin. Microbiol. Infect.*, 25(10), pp.1288.e1-1288.e6. https://doi.org/10.1016/j.cmi.2019.03.018 (2019).

56 Shestivska, V., Dryahina, K., Nunvář, J., Sovová, K., Elhottová, D., Nemec, A., Smith, D. and Španěl, P., Quantitative analysis of volatile metabolites released in vitro by bacteria of the genus Stenotrophomonas for identification of breath biomarkers of respiratory infection in cystic fibrosis. *J. Breath Res.*, 9(2), p.027104. https://doi.org/10.1088/1752-7155/9/2/027104 . (2015).

57 Gallego, E., Roca, F., Perales, J. and Guardino, X. Comparative study of the adsorption performance of a multi-sorbent bed (Carbotrap, Carbopack X, Carboxen 569) and a Tenax TA adsorbent tube for the analysis of volatile organic compounds (VOCs). *Talanta*, 81(3), pp.916-924. https://doi.org/10.1016/j.talanta.2010.01.037 (2010).

58 Hübschmann, H. *Handbook of GC-MS*. 3rd ed. Weinheim, Germany: Wiley, pp.57 - 65. (2015).

59 Pawliszyn, J., 2012. *Handbook of Solid phase microextraction*. 1st ed. Elseiver, p.105.

60 Piñeiro-García, A., González-Alatorre, G., Tristan, F., Fierro-Gonzalez, J. and Vega-Díaz, S., 2018. Simple preparation of reduced graphene oxide coatings for solid phase micro-extraction (SPME) of furfural to be detected by gas chromatography/mass spectrometry. *Materials Chemistry and Physics*, 213, pp.556-561.

61 Jenkins, C. and Bean, H., (2020). Dependence of the Staphylococcal Volatilome Composition on Microbial Nutrition. *Metabolites*, 10(9), p.347. **https://doi.org/10.3390/metabo10090347**

62 Bean, H., Rees, C. and Hill, J., Comparative analysis of the volatile metabolomes of Pseudomonas aeruginosa clinical isolates *J. Breath Res.*, 10(4), p.047102. https://doi.org/10.1088/1752-7155/10/4/047102 . (2016).

63 Ratiu, I., Bocos-Bintintan, V., Monedeiro, F., Milanowski, M., Ligor, T. and Buszewski, B., An Optimistic Vision of Future: Diagnosis of Bacterial Infections by Sensing Their Associated Volatile Organic Compounds. *Crit. Rev. Anal. Chem.*, pp.1-12. https://doi.org/10.1080/10408347.2019.1663147 (2019).

64 Belluomo, I., Boshier, P., Myridakis, A., Vadhwana, B., Markar, S., Spanel, P. and Hanna, G., 2021. Selected ion flow tube mass spectrometry for targeted analysis of volatile organic compounds in human breath. *Nature Protocols*, 16(7), pp.3419-3438. https://doi.org/10.1038/s41596-021-00542-0

65 Smith, D. and Španěl, P., Selected ion flow tube mass spectrometry (SIFT-MS) for on-line trace gas analysis. *Mass Spectrom. Rev.*, 24(5), pp.661-700. https://doi.org/10.1002/mas.20033 (2005).

66 Španěl, P. and Smith, D., Progress in SIFT-MS: Breath analysis and other applications. *Mass Spectrom. Rev.*, 30(2), pp.236-267. https://doi.org/10.1002/mas.20303 (2010).

67 Blake, R., Monks, P. and Ellis, A., Proton-Transfer Reaction Mass Spectrometry. *Chemical Reviews*, 109(3), pp.861-896. https://doi.org/10.1021/cr800364q (2009).

68 Ashrafi, M., Bates, M., Baguneid, M., Alonso-Rasgado, T., Rautemaa-Richardson, R. and Bayat, A.,. Volatile organic compound detection as a potential means of diagnosing cutaneous wound infections. *Wound Repair and Regeneration*, 25(4), pp.574-590. https://doi.org/10.1111/wrr.12563 (2017).

69 Mohan, D., Keir, H., Richardson, H., Mayhew, D., Boyer, J., van der Schee, M., Allsworth, M., Miller, B., Tal-Singer, R. and Chalmers, J., 2021. Exhaled volatile organic compounds and lung microbiome in COPD: a pilot randomised controlled trial. *ERJ Open Research*, 7(4), pp.00253-2021.

70 Grote, C. and Pawliszyn, J., 1997. Solid-Phase Microextraction for the Analysis of Human Breath. *Analytical Chemistry*, 69(4), pp.587-596.

71 Yuan, Z., Li, W., Wu, L., Huang, D., Wu, M. and Hu, B., 2020. Solid-Phase Microextraction Fiber in Face Mask for in Vivo Sampling and Direct Mass Spectrometry Analysis of Exhaled Breath Aerosol. *Analytical Chemistry*, 92(17), pp.11543-11547.

72 Bond, A., Vernon, A., Reade, S., Mayor, A., Minetti, C., Wastling, J., Lamden, K. and Probert, C., 2015. Investigation of Volatile Organic Compounds Emitted from Faeces for the Diagnosis of Giardiasis. *Journal of Gastrointestinal and Liver Diseases*, 24(3), pp.281-286. https://doi.org/10.15403/JGLD.2014.1121.243.ABO

73 Jenkins, C. and Bean, H., (2020). Dependence of the Staphylococcal Volatilome Composition on Microbial Nutrition. *Metabolites*, 10(9), p.347. **https://doi.org/10.3390/metabo10090347**

74 Ibrahim, W., Cordell, R., Wilde, M., Richardson, M., Carr, L., Devi Dasi, A., Hargadon, B., Free, R., Monks, P., Brightling, C., Greening, N. and Siddiqui, S., 2021. Diagnosis of COVID-19 by exhaled breath analysis using gas chromatography-mass spectrometry. *ERJ Open Research*, pp.00139-2021.

75 Berna, A., Akaho, E., Harris, R., Congdon, M., Korn, E., Neher, S., M'Farrej, M., Burns, J. and Odom John, A., 2021. Reproducible Breath Metabolite Changes in Children with SARS-CoV-2 Infection. *ACS Infectious Diseases*, 7(9), pp.2596-2603.

Pre-examination copy
Submitted 27th May 2022

76 Phillips, M., Cataneo, R., Condos, R., Ring Erickson, G., Greenberg, J., La Bombardi, V., Munawar, M. and Tietje, O., 2007. Volatile biomarkers of pulmonary tuberculosis in the breath. *Tuberculosis*, 87(1), pp.44-52.

77 Patel, M., Fowler, D., Sizer, J. and Walton, C., 2019. Faecal volatile biomarkers of Clostridium difficile infection. *PLOS ONE*, 14(4), p.e0215256. https://doi.org/10.1371/journal.pone.0215256

78 Zhang, J., Tian, Y., Luo, Z., Qian, C., Li, W. and Duan, Y., 2021. Breath volatile organic compound analysis: an emerging method for gastric cancer detection. *Journal of Breath Research*, 15(4), p.044002. https://doi.org/10.1088/1752-7163/ac2cde .

79 Boots, A., Smolinska, A., van Berkel, J., Fijten, R., Stobberingh, E., Boumans, M., Moonen, E., Woulters, E., Dallinga, J. and Van Schooten, F., Identification of microorganisms based on headspace analysis of volatile organic compounds by gas chromatography–mass spectrometry. *J. Breath Res.*, 8(2). https://doi.org/10.1088/1752-7155/8/2/027106 (2014).

80 Elgaali, H., Hamilton-Kemp, T., Newman, M., Collins, R., Yu, K. and Archbold, D., Comparison of long-chain alcohols and other volatile compounds emitted from food-borne and related Gram positive and Gram negative bacteria. *J. Basic Microbiol.*, 42(6), pp.373-380. https://doi.org/10.1002/1521-4028(200212)42:6<373::AID-JOBM373>3.0.CO;2-4 (2002).

81 Traxler, S., Bischoff, A., Saß, R., Trefz, P., Gierschner, P., Brock, B., Schwaiger, T., Karte, C., Blohm, U., Schröder, C., Miekisch, W. and Schubert, J., 2018. VOC breath profile in spontaneously breathing awake swine during Influenza A infection. *Scientific Reports*, 8(1).

82 Ruszkiewicz, D., Sanders, D., O'Brien, R., Hempel, F., Reed, M., Riepe, A., Bailie, K., Brodrick, E., Darnley, K., Ellerkmann, R., Mueller, O., Skarysz, A., Truss, M., Wortelmann, T., Yordanov, S., Thomas, P., Schaaf, B. and Eddleston, M., 2020. Diagnosis of COVID-19 by Analysis of Breath with Gas Chromatography-Ion Mobility Spectrometry - A Feasibility Study. *SSRN Electronic Journal*,.

83 Ashrafi, M., Novak-Frazer, L., Bates, M., Baguneid, M., Alonso-Rasgado, T., Xia, G., Rautemaa-Richardson, R. and Bayat, A. Validation of biofilm formation on human skin wound models and demonstration of clinically translatable bacteria-specific volatile signatures. *Sci. Rep.*, 8(1). https://doi.org/10.1038/s41598-018-27504-z. (2018).

84 Ashrafi, M., Novak-Frazer, L., Morris, J., Baguneid, M., Rautemaa-Richardson, R. and Bayat, A. Electrical stimulation disrupts biofilms in a human wound model and reveals the potential for monitoring treatment response with volatile biomarkers. *Wound Repair and Regeneration*, 27(1), pp.5-18. https://doi.org/10.1111/wrr.12679. (2018).

85 Mochalski, P., King, J., Unterkofler, K., Hinterhuber, H. and Amann, A., Emission rates of selected volatile organic compounds from skin of healthy volunteers. *Journal of Chromatography B*, 959, pp.62-70. https://doi.org/10.1016/j.jchromb.2014.04.006 (2014).

86 King, J., Mochalski, P., Kupferthaler, A., Unterkofler, K., Koc, H., Filipiak, W., Teschl, S., Hinterhuber, H. and Amann, A., Dynamic profiles of volatile organic compounds in exhaled breath as determined by a coupled PTR-MS/GC-MS study. *Physiological Measurement*, 31(9), pp.1169-1184. https://doi.org/10.1088/0967-3334/31/9/008 (2010).

87 Adam, M., Fehervari, M., Boshier, P., Chin, S., Lin, G., Romano, A., Kumar, S. and Hanna, G., 2019. Mass-Spectrometry Analysis of Mixed-Breath, Isolated-Bronchial-Breath, and Gastric-Endoluminal-Air Volatile Fatty Acids in Esophagogastric Cancer. *Analytical Chemistry*, 91(5), pp.3740-3746. https://doi.org/10.1021/acs.analchem.9b00148

88 Kumar, S., Huang, J., Abbassi-Ghadi, N., Mackenzie, H., Veselkov, K., Hoare, J., Lovat, L., Španěl, P., Smith, D. and Hanna, G., 2015. Mass Spectrometric Analysis of Exhaled Breath for the Identification of Volatile Organic Compound Biomarkers in Esophageal and Gastric Adenocarcinoma. *Annals of Surgery*, 262(6), pp.981-990.

89 Kumar, S., Huang, J., Abbassi-Ghadi, N., Španěl, P., Smith, D. and Hanna, G., 2013. Selected Ion Flow Tube Mass Spectrometry Analysis of Exhaled Breath for Volatile Organic Compound Profiling of Esophago-Gastric Cancer. *Analytical Chemistry*, 85(12), pp.6121-6128.

90 Grassin-Delyle, S., Roquencourt, C., Moine, P., Saffroy, G., Carn, S., Heming, N., Fleuriet, J., Salvator, H., Naline, E., Couderc, L., Devillier, P., Thévenot, E. and Annane, D., 2021. Metabolomics of exhaled breath in critically ill COVID-19 patients: A pilot study. *EBioMedicine*, 63, p.103154.

91 Bunge, M., Araghipour, N., Mikoviny, T., Dunkl, J., Schnitzhofer, R., Hansel, A., Schinner, F., Wisthaler, A., Margesin, R. and Mark, T., On-Line Monitoring of Microbial Volatile Metabolites by Proton Transfer Reaction-Mass Spectrometry. *Appl. Environ. Microbiol.*, 74(7), pp.2179-2186. https://doi.org/10.1128/AEM.02069-07, (2008)

92 Thorn, R., Reynolds, D. and Greenman, J., Multivariate analysis of bacterial volatile compound profiles for discrimination between selected species and strains in vitro. *J. Microbiol. Methods*, 84(2), pp.258-264. https://doi.org/10.1016/j.mimet.2010.12.001 . (2011).

93 Shestivska, V., Španěl, P., Dryahina, K., Sovová, K., Smith, D., Musílek, M. and Nemec, A., Variability in the concentrations of volatile metabolites emitted by genotypically different strains of Pseudomonas aeruginosa. *J. Appl. Microbiol.*, 113(3), pp.701-713. https://doi.org/10.1111/j.1365-2672.2012.05370.x (2012).

94 Shestivska, V., Nemec, A., Dřevínek, P., Sovová, K., Dryahina, K. and Španěl, P., 2011. Quantification of methyl thiocyanate in the headspace of Pseudomonas aeruginosa cultures and in the breath of cystic fibrosis patients by selected ion flow tube mass spectrometry. *Rapid Communications in Mass Spectrometry*, 25(17), pp.2459-2467.

95 Slade, E., Thorn, R., Young, A. and Reynolds, D., 2021. Real-time detection of volatile metabolites enabling species-level discrimination of bacterial biofilms associated with wound infection. *Journal of Applied Microbiology*,.

96 Zhu, J., Bean, H., Kuo, Y. and Hill, J., Fast Detection of Volatile Organic Compounds from Bacterial Cultures by Secondary Electrospray Ionization-Mass Spectrometry. *J. Clin. Microbiol.*, 49(2), pp.769-769. https://doi.org/ 10.1128/JCM.00392-10. (2011).

97 Cookson, B., Five decades of MRSA: controversy and uncertainty continues. *The Lancet*, 378(9799), pp.1291-1292.(2011). https://doi.org/10.1016/S0140-6736(11)61566-3

98 Baptista, I., Santos, M., Rudnitskaya, A., Saraiva, J., Almeida, A. and Rocha, S., A comprehensive look into the volatile exometabolome of enteroxic and non-enterotoxic Staphylococcus aureus strains. *The International Journal of Biochemistry & Cell Biology*, 108, pp.40-50. https://doi.org/10.1016/j.biocel.2019.01.007 (2019).

99 Filipiak, W., Sponring, A., Filipiak, A., Baur, M., Ager, C., Wiesenhofer, H., Margesin, R., Nagl, M., Troppmair, J., Amann, A., Volatile organic compounds (VOCs) released by pathogenic microorganisms in vitro: potential breath biomarkers for early- stage diagnosis of disease. In: Amann, A., Smith, D. (Eds.), Volatile Biomarkers: Non- Invasive Diagnosis in Physiology and Medicine. Elsevier B.V., Oxford, UK, pp. 463–512. https://doi.org/10.1016/B978-0-44-462613-4.00023-4 . (2013).

100 Xiao, Z. and Xu, P. Acetoin Metabolism in Bacteria. *Crit. Rev. Microbiol.*, 33(2), pp.127-140. https://doi.org/10.1080/10408410701364604. (2007).

101 Ji, X., Huang, H. and Ouyang, P., Microbial 2,3-butanediol production: A state-of-the-art review. *Biotechnol. Adv.*, 29(3), pp.351-364. https://doi.org/10.1016/j.biotechadv.2011.01.007 . (2011)

102 Gardner, S., Hillis, S., Heilmann, K., Segre, J. & Grice, E. (2012). The neuropathic diabetic foot ulcer microbiome is associated with clinical factors. *Diabetes* 62(3), 923–930. https://doi.org/10.2337/db12-0771 .

103 Ndosi, M. *et al.* (2017). Prognosis of the infected diabetic foot ulcer: A 12-month prospective observational study. *Diabetic Med.* 35(1), 78–88. https://doi.org/10.1111/dme.13537 .

41

[104] Lyczak, J., Cannon, C. and Pier, G., Establishment of Pseudomonas aeruginosa infection: lessons from a versatile opportunist. *Microbes Infect.*, 2(9), pp.1051-1060. https://doi.org/10.1016/S1286-4579(00)01259-4 . (2000).

[105] Abdulrazak, A., Ibrahim Bitar, Z., Ayesh Al-Shamali, A. and Ahmed Mobasher, L., Bacteriological study of diabetic foot infections. *Journal of Diabetes and its Complications*, 19(3), pp.138-141. https://doi.org/10.1016/j.jdiacomp.2004.06.001 . (2005).

[106] Bradbury, J., Epidemic Pseudomonas aeruginosa strain sequenced. *The Lancet Infectious Diseases*, 9(1), p.14. (2009). https://doi.org/10.1016/S1473-3099(08)70298-6

[107] Høiby, N., Ciofu, O. and Bjarnsholt, T., Pseudomonas aeruginosa biofilms in cystic fibrosis. *Future Microbiol.*, 5(11). https://doi.org/10.2217/fmb.10.125 (2010).

[108] Pang, Z., Raudonis, R., Glick, B., Lin, T. and Cheng, Z., 2019. Antibiotic resistance in Pseudomonas aeruginosa: mechanisms and alternative therapeutic strategies. *Biotechnology Advances*, 37(1), pp.177-192.

[109] Senecal, A., Magnone, J., Yeomans, W. and Powers, E., Rapid detection of pathogenic bacteria by volatile organic compound (VOC) analysis. *Chemical and Biological Early Warning Monitoring for Water, Food, and Ground*, Proceedings volume 4575. https://doi.org/10.1117/12.456915 . (2002).

[110] Bos, L., Sterk, P. and Schultz, M., Volatile Metabolites of Pathogens: A Systematic Review. *PLoS Pathog.*, 9(5), p.e1003311. https://doi.org/10.1371/journal.ppat.1003311 . (2013).

[111] Zscheppank, C., Wiegand, H., Lenzen, C., Wingender, J. and Telgheder, U., Investigation of volatile metabolites during growth of Escherichia coli and Pseudomonas aeruginosa by needle trap-GC-MS. *Anal. Bioanal. Chem.*, 406(26), pp.6617-6628. https://doi.org/10.1007/s00216-014-8111-2 . (2014).

[112] Phan, J., Ranjbar, S., Kagawa, M., Gargus, M., Hochbaum, A. and Whiteson, K., 2020. Thriving Under Stress: Pseudomonas aeruginosa Outcompetes the Background Polymicrobial Community Under Treatment Conditions in a Novel Chronic Wound Model. *Frontiers in Cellular and Infection Microbiology*, 10.

[113] Lawal, O., Muhamadali, H., Ahmed, W., White, I., Nijsen, T., Goodacre, R. and Fowler, S., 2018. Headspace volatile organic compounds from bacteria implicated in ventilator-associated pneumonia analysed by TD-GC/MS. *Journal of Breath Research*, 12(2), p.026002.

[114] Jarvis, W. and Martone, W., Predominant pathogens in hospital infections. *J. Antimicrob. Chemother.*, 29(suppl A), pp.19-24. https://doi.org/10.1093/jac/29.suppl_A.19 .(1992).

[115] Wang, J., Hsueh, P., Wang, J., Lee, L., Yang, P. and Luh, K., Recurrent Infections and Chronic Colonization by an Escherichia coli Clone in the Respiratory Tract of a Patient with Severe Cystic Bronchiectasis. *J. Clin. Microbiol.*, 38(7), pp.2766-2767. (2000).

[116] Bonacorsi, S. and Bingen, E., Molecular epidemiology of Escherichia coli causing neonatal meningitis. *Int. J. Med. Microbiol.*, 295(6-7), pp.373-381. https://doi.org/10.1016/j.ijmm.2005.07.011 . (2005).

[117] Berstad, A., Raa, J. and Valeur, J. Indole the scent of a healthy 'inner soil'. *Microb. Ecol. Health Dis.*, 26. https://doi.org/10.3402/mehd.v26.27997. (2015)

[118] Lee, J. and Lee, J. Indole as an intercellular signal in microbial communities. *FEMS Microbiol. Rev.*, 34(4), pp.426-444. https://doi.org/10.1111/j.1574-6976.2009.00204.x. (2010).

[119] Rodriguez, G., Tashiro, Y. and Atsumi, S., 2014. Expanding ester biosynthesis in Escherichia coli. *Nature Chemical Biology*, 10(4), pp.259-265.

[120] McKenna, R. and Nielsen, D., Styrene biosynthesis from glucose by engineered E. coli. *Metab. Eng.*, 13(5), pp.544-554. https://doi.org/10.1016/j.ymben.2011.06.005 (2011).

[121] Liu, T., Vora, H. and Khosla, C., Quantitative analysis and engineering of fatty acid biosynthesis in E. coli. *Metab. Eng.*, 12(4), pp.378-386. https://doi.org/10.1016/j.ymben.2010.02.003. (2010).

[122] Rodriguez, G., Tashiro, Y. and Atsumi, S., . Expanding ester biosynthesis in Escherichia coli. *Nat. Chem. Biol.*, 10(4), pp.259-265. https://doi.org/10.1038/nchembio.1476 (2014).

[123] Guo, D., Pan, H. and Li, X., Metab. Eng. of Escherichia coli for production of biodiesel from fatty alcohols and acetyl-CoA. *Appl. Microbiol. Biotechnol.*, 99(18), pp.7805-7812. https://doi.org/10.1007/s00253-015-6809-5 (2015).

[124] Ko, W., Community-Acquired Klebsiella pneumoniae Bacteremia: Global Differences in Clinical Patterns. *Emerging Infect. Dis.*, 8(2), pp.160-166. https://doi.org/10.3201/eid0802.010025 (2002).

[125] Madhi, S., Klugman, K. and The Vaccine Trialist Group, A role for Streptococcus pneumoniae in virus-associated pneumonia. *Nature Medicine*, 10(8), pp.811-813. https://doi.org/10.1038/nm1077. (2004).

[126] van de Beek, D., de Gans, J., Tunkel, A. and Wijdicks, E., Community-Acquired Bacterial Meningitis in Adults. *N. Engl. J. Med.*, 354(1), pp.44-53. https://doi.org/10.1056/NEJMra052116 . (2006).

[127] Ku, Y., Chuang, Y., Chen, C., Lee, M., Yang, Y., Tang, H. and Yu, W., Klebsiella pneumoniae Isolates from Meningitis: Epidemiology, Virulence and Antibiotic Resistance. *Sci. Rep.*, 7(1). https://doi.org/10.1038/s41598-017-06878-6 (2017).

[128] Filipiak, W., Sponring, A., Baur, M., Ager, C., Filipiak, A., Wiesenhofer, H., Nagl, M., Troppmair, J. and Amann, A., Characterization of volatile metabolites taken up by or released from Streptococcus pneumoniae and Haemophilus influenzae by using GC-MS. *Microbiology*, 158(12), pp.3044-3053. https://doi.org/10.1099/mic.0.062687-0 (2012).

[129] Rees, C., Franchina, F., Nordick, K., Kim, P. and Hill, J., Expanding the Klebsiella pneumoniae volatile metabolome using advanced analytical instrumentation for the detection of novel metabolites. *J. Appl. Microbiol.*, 122(3), pp.785-795. https://doi.org/10.1111/jam.13372 (2017).

[130] Jünger, M., Vautz, W., Kuhns, M., Hofmann, L., Ulbricht, S., Baumbach, J., Quintel, M. and Perl, T., Ion mobility spectrometry for microbial volatile organic compounds: a new identification tool for human pathogenic bacteria. *Appl. Microbiol. Biotechnol.*, 93(6), pp.2603-2614. https://doi.org/10.1007/s00253-012-3924-4 . (2012).

[131] Carolus, H., Van Dyck, K. and Van Dijck, P., 2019. Candida albicans and Staphylococcus Species: A Threatening Twosome. *Frontiers in Microbiology*, 10. https://doi.org/10.3389/fmicb.2019.02162

[132] Costa, C., Bezerra, A., Almeida, A. and Rocha, S., 2020. Candida Species (Volatile) Metabotyping through Advanced Comprehensive Two-Dimensional Gas Chromatography. *Microorganisms*, 8(12), p.1911.

[133] López-Ramos, J., Bautista, E., Gutiérrez-Escobedo, G., Mancilla-Montelongo, G., Castaño, I., González-Chávez, M. and De Las Peñas, A., 2021. Analysis of Volatile Molecules Present in the Secretome of the Fungal Pathogen Candida glabrata. *Molecules*, 26(13), p.3881.

[134] Perl, T., Jünger, M., Vautz, W., Nolte, J., Kuhns, M., Borg-von Zepelin, M. and Quintel, M., 2011. Detection of characteristic metabolites of Aspergillus fumigatus and Candida species using ion mobility spectrometry - metabolic profiling by volatile organic compounds. *Mycoses*, 54(6), pp.e828-e837.

Pre-examination copy
Submitted 27th May 2022

[135] Rodrigues, C. and Černáková, L., 2020. Farnesol and Tyrosol: Secondary Metabolites with a Crucial quorum-sensing Role in Candida Biofilm Development. *Genes*, 11(4), p.444.

[136] Lindsay, A., Deveau, A., Piispanen, A. and Hogan, D., 2012. Farnesol and Cyclic AMP Signaling Effects on the Hypha-to-Yeast Transition in Candida albicans. *Eukaryotic Cell*, 11(10), pp.1219-1225.

[137] Hornby, J., Jensen, E., Lisec, A., Tasto, J., Jahnke, B., Shoemaker, R., Dussault, P. and Nickerson, K., 2001. Quorum Sensing in the Dimorphic Fungus Candida albicans Is Mediated by Farnesol. *Applied and Environmental Microbiology*, 67(7), pp.2982-2992.

[138] Jabra-Rizk, M., Meiller, T., James, C. and Shirtliff, M., 2006. Effect of Farnesol on Staphylococcus aureus Biofilm Formation and Antimicrobial Susceptibility. *Antimicrobial Agents and Chemotherapy*, 50(4), pp.1463-1469.

[139] Pammi, M., Liang, R., Hicks, J., Barrish, J. and Versalovic, J., 2011. Farnesol Decreases Biofilms of Staphylococcus epidermidis and Exhibits Synergy With Nafcillin and Vancomycin. *Pediatric Research*, 70(6), pp.578-583.

[140] Chauhan, N. and Mohan Karuppayil, S., 2021. Dual identities for various alcohols in two different yeasts. *Mycology*, 12(1), pp.25-38.

[141] Lei, X., Deng, B., Ruan, C., Deng, L. and Zeng, K., 2022. Phenylethanol as a quorum sensing molecule to promote biofilm formation of the antagonistic yeast Debaryomyces nepalensis for the control of black spot rot on jujube. *Postharvest Biology and Technology*, 185, p.111788.

[142] Pu, L., Jingfan, F., Kai, C., Chao-an, L. and Yunjiang, C., 2014. Phenylethanol promotes adhesion and biofilm formation of the antagonistic yeast Kloeckera apiculate for the control of blue mold on citrus. *FEMS Yeast Research*, 14(4),

[143] Han, T., Tumanov, S., Cannon, R. and Villas-Boas, S., 2013. Metabolic Response of Candida albicans to Phenylethyl Alcohol under Hyphae-Inducing Conditions. *PLoS ONE*, 8(8), p.e71364.

[144] Nobile, C. and Johnson, A., 2015. Candida albicans Biofilms and Human Disease. *Annual Review of Microbiology*, 69(1), pp.71-92. https://doi.org/10.1146/annurev-micro-091014-104330

[145] *Global Tuberculosis Report 2013*. Geneva: World Health Organization. (2013).

[146] World Health Organization, Global Tuberculosis Report 2019, World Health Organization, Licence: CC BY-NC-SA 3.0 IGO, Geneva, Switzerland 2019.

[147] Lu, C., Liu, Q., Sarma, A., Fitzpatrick, C., Falzon, D. and Mitnick, C., 2013. A Systematic Review of Reported Cost for Smear and Culture Tests during Multidrug-Resistant Tuberculosis Treatment. *PLoS ONE*, 8(2), p.e56074. https://doi.org/10.1371/journal.pone.0056074

[148] Syhre, M. and Chambers, S., The scent of Mycobacterium tuberculosis. *Tuberculosis*, 88(4), pp.317-323. https://doi.org/10.1016/j.tube.2008.01.002 (2008).

[149] McNerney, R., Mallard, K., Okolo, P. and Turner, C., Production of volatile organic compounds by mycobacteria. *FEMS Microbiol. Lett.* 328(2), pp.150-156. https://doi.org/10.1111/j.1574-6968.2011.02493.x (2012).

[150] Zetola, N., Modongo, C., Matsiri, O., Tamuhla, T., Mbongwe, B., Matlhagela, K., Sepako, E., Catini, A., Sirugo, G., Martinelli, E., Paolesse, R. and Di Natale, C., Diagnosis of pulmonary tuberculosis and assessment of treatment response through analyses of volatile compound patterns in exhaled breath samples. *Journal of Infection*, 74(4), pp.367-376. https://doi.org/10.1016/j.jinf.2016.12.006. (2017).

[151] Vishinkin, R., Busool, R., Mansour, E., Fish, F., Esmail, A., Kumar, P., Gharaa, A., Cancilla, J., Torrecilla, J., Skenders, G., Leja, M., Dheda, K., Singh, S. and Haick, H., 2021. Profiles of Volatile Biomarkers Detect Tuberculosis from Skin. *Advanced Science*, 8(15), p.2100235.

[152] Wieslander, G., Kumlin, A. and Norbäck, D., 2010. Dampness and 2-Ethyl-1-hexanol in Floor Construction of Rehabilitation Center: Health Effects in Staff. *Archives of Environmental & Occupational Health*, 65(1), pp.3-11. https://doi.org/10.180/19338240903390248

[153] Purcaro, G., Rees, C., Wieland-Alter, W., Schneider, M., Wang, X., Stefanuto, P., Wright, P., Enelow, R. and Hill, J., 2018. Volatile fingerprinting of human respiratory viruses from cell culture. *Journal of Breath Research*, 12(2), p.026015.

[154] Sanchez, E. and Lagunoff, M., 2015. Viral activation of cellular metabolism. *Virology*, 479-480, pp.609-618.

[155] Ratcliffe, N., Wieczorek, T., Drabińska, N., Gould, O., Osborne, A. and De Lacy Costello, B., 2020. 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. *Journal of Breath Research*, 14(3), p.034001.

[156] Fuchs, P., Loeseken, C., Schubert, J. and Miekisch, W., 2010. Breath gas aldehydes as biomarkers of lung cancer. *International Journal of Cancer,* 126.

[157] Callol-Sanchez, L., Munoz-Lucas, M., Gomez-Martin, O., Maldonado-Sanz, J., Civera-Tejuca, C., Gutierrez-Ortega, C., Rodriguez-Trigo, G. and Jareno-Esteban, J., 2017. Observation of nonanoic acid and aldehydes in exhaled breath of patients with lung cancer. *Journal of Breath Research*, 11(2), p.026004.

[158] Chastre, J. and Fagon, J., 2002. Ventilator-associated Pneumonia. *American Journal of Respiratory and Crit. Care Med.*, 165(7), pp.867-903. https://doi.org/10.1164/ajrccm.165.7.2105078 (2002).

[159] Hunter, J., 2012. Ventilator associated pneumonia. *BMJ*, 344, pp.e3325-e3325. https://doi.org/10.1136/bmj.e3325 . (2012).

[160] Kollef, M., Chastre, J., Fagon, J., François, B., Niederman, M., Rello, J., Torres, A., Vincent, J., Wunderink, R., Go, K. and Rehm, C., 2014. Global Prospective Epidemiologic and Surveillance Study of Ventilator-Associated Pneumonia due to Pseudomonas aeruginosa*. *Crit. Care Med.*, 42(10), pp.2178-2187. https://doi.org/10.1097/CCM.0000000000000510 . (2014).

[161] Agbaht, K., Diaz, E., Muñoz, E., Lisboa, T., Gomez, F., Depuydt, P., Blot, S. and Rello, J., 2007. Bacteremia in patients with ventilator-associated pneumonia is associated with increased mortality: A study comparing bacteremic vs. nonbacteremic ventilator-associated pneumonia*. *Critical Care Medicine*, 35(9), pp.2064-2070

[162] Schnabel, R., Fijten, R., Smolinska, A., Dallinga, J., Boumans, M., Stobberingh, E., Boots, A., Roekaerts, P., Bergmans, D. and van Schooten, F., Analysis of volatile organic compounds in exhaled breath to diagnose ventilator-associated pneumonia. *Sci. Rep.*, 5(1). https://doi.org/10.1038/srep17179. (2015).

[163] Bradbury, J., Epidemic Pseudomonas aeruginosa strain sequenced. *The Lancet Infectious Diseases*, 9(1), p.14. (2009). https://doi.org/10.1016/S1473-3099(08)70298-6

[164] Kunze-Szikszay, N., Walliser, K., Luther, J., Cambiaghi, B., Reupke, V., Dullin, C., Vautz, W., Bremmer, F., Telgheder, U., Zscheppank, C., Quintel, M. and Perl, T., 2019. Detecting Early Markers of Ventilator-Associated Pneumonia by Analysis of Exhaled Gas. *Critical Care Medicine*, 47(3), pp.e234-e240.

[165] Kunze, N., Göpel, J., Kuhns, M., Jünger, M., Quintel, M. and Perl, T., 2013. Detection and validation of volatile metabolic patterns over different strains of two human pathogenic bacteria during their growth in a complex medium using multi-capillary column-ion mobility spectrometry (MCC-IMS). *Applied Microbiology and Biotechnology*, 97(8), pp.3665-3676.

[166] Lawal, O., Knobel, H., Weda, H., Nijsen, T., Goodacre, R. and Fowler, S., 2018. TD/GC–MS analysis of volatile markers emitted from mono- and co-cultures of Enterobacter cloacae and Pseudomonas aeruginosa in artificial sputum. *Metabolomics*, 14(5).

Pre-examination copy
Submitted 27th May 2022

[167] Maes, M., Higginson, E., Pereira-Dias, J., Curran, M., Parmar, S., Khokhar, F., Cuchet-Lourenco, D., Lux, J., Sharma-Hajela, S., Ravenhill, B., Hamed, I., Heales, L., Mahroof, R., Soderholm, A., Forrest, S., Sridhar, S., M. Brown, N., Baker, S., Navapurkar, V., Dougan, G., Bartholdson Scott, J. and Conway Morris, A., 2021. Ventilator-associated pneumonia in critically ill patients with COVID-19. *Critical Care*, 25.

[168] Heijerman, H., 2005. Infection and inflammation in cystic fibrosis: A short review. *Journal of Cystic Fibrosis*, 4, pp.3-5.

[169] O'Sullivan, B. and Freedman, S., Cystic Fibrosis. *The Lancet*, 373(9678). https://doi.org/10.1016/S0140-6736(09)60327-5 . (2009).

[170] Gilchrist, F., Španěl, P., Smith, D. and Lenney, W., The in vitro identification and quantification of volatile biomarkers released by cystic fibrosis pathogens. *Anal. Methods*, 7(3), pp.818-824. https://doi.org/10.1039/C4AY02981J (2015).

[171] Dryahina, K., Sovová, K., Nemec, A. and Španěl, P., 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(3), p.037102. https://doi.org/10.1088/1752-7155/10/3/037102 (2016).

[172] Kramer, R., Sauer-Heilborn, A., Welte, T., Guzman, C., Höfle, M. and Abraham, W., . A rapid method for breath analysis in cystic fibrosis patients. *Eur. J. Clin. Microbiol. Infect. Dis.*, 34(4), pp.745-751. https://doi.org/10.1007/s10096-014-2286-5 . (2014)

[173] Neerincx, A., Geurts, B., van Loon, J., Tiemes, V., Jansen, J., Harren, F., Kluijtmans, L., Merkus, P., Cristescu, S., Buydens, L. and Wevers, R., Detection of Staphylococcus aureus in cystic fibrosis patients using breath VOC profiles. *J. Breath Res.*, 10(4), p.046014. https://doi.org/10.1088/1752-7155/10/4/046014 (2016).

[174] Robroeks, C., van Berkel, J., Dallinga, J., Jöbsis, Q., Zimmermann, L., Hendriks, H., Wouters, M., van der Grinten, C., van de Kant, K., van Schooten, F. and Dompeling, E., Metabolomics of Volatile Organic Compounds in Cystic Fibrosis Patients and Controls. *Pediatr. Res*, 68(1), pp.75-80. https://doi.org/10.1203/PDR.0b013e3181df4ea0 (2010).

[175] Brawner, K., Morrow, C. and Smith, P., Gastric Microbiome and Gastric Cancer. *The Cancer Journal*, 20(3), pp.211-216. https//doi.org/10.1097/PPO.0000000000000043 . (2014).

[176] Zhan, X., Duan, J. and Duan, Y., 2012. Recent developments of proton-transfer reaction mass spectrometry (PTR-MS) and its applications in medical research. *Mass Spectrometry Reviews*, 32(2), pp.143-165. https://doi.org/10.1002/mas.21357

[177] Armstrong, D., Boulton, A. and Bus, S. Diabetic Foot Ulcers and Their Recurrence. *N. Engl. J. Med.*, 376(24), pp.2367-2375. https://doi.org/10.1056/NEJMra1615439 . (2017).

[178] Jeffcoate, W. and Harding, K. Diabetic foot ulcers. *The Lancet*, 361(9368), pp.1545-1551. https://doi.org/10.1016/S0140-6736(03)13169-8. (2003).

[179] Ndosi, M., Wright-Hughes, A., Brown, S., Backhouse, M., Lipsky, B., Bhogal, M., Reynolds, C., Vowden, P., Jude, E., Nixon, J. and Nelson, E. Prognosis of the infected diabetic foot ulcer: a 12-month prospective observational study. *Diabetic Med.*, 35(1), pp.78-88. https://doi.org/10.1111/dme.13537 (2017).

[180] Gardner, S., Hillis, S., Heilmann, K., Segre, J. and Grice, E., The Neuropathic Diabetic Foot Ulcer Microbiome Is Associated With Clinical Factors. *Diabetes*, 62(3), pp.923-930. https://doi.org/10.2337/db12-0771 (2012).

[181] Kalan, L., Meisel, J., Loesche, M., Horwinski, J., Soaita, I., Chen, X., Uberoi, A., Gardner, S. and Grice, E., Strain- and Species-Level Variation in the Microbiome of Diabetic Wounds Is Associated with Clinical Outcomes and Therapeutic Efficacy. *Cell Host Microbe*, 25(5), pp.641-655.e5. https://doi.org/10.1016/j.chom.2019.03.006 (2019).

[182] Lipsky, B., Aragón-Sánchez, J., Diggle, M., Embil, J., Kono, S., Lavery, L., Senneville, É., Urbančič-Rovan, V., Van Asten, S., Peters, E., Abbas, Z. and Malone, M. IWGDF guidance on the diagnosis and management of foot infections in persons with diabetes. *IWGDF*. (2019).

[183] Park, J., Suh, D., Kim, H., Lee, Y., Kwak, I. and Choi, G. *Diabetes Res. Clin. Pract.*, 128, pp.51-57. https://doi.org/10.1016/j.diabres.2017.04.008 (2017).

[184] Thomas, A., Riazanskaia, S., Cheung, W., Xu, Y., Goodacre, R., Thomas, C., Baguneid, M. and Bayat, A. Novel noninvasive identification of biomarkers by analytical profiling of chronic wounds using volatile organic compounds. *Wound Repair and Regeneration*, 18(4), pp.391-400. https://doi.org/10.1111/j.1524-475X.2010.00592.x. (2010)

[185] Gontcharova, V., . A Comparison of Bacterial Composition in Diabetic Ulcers and Contralateral Intact Skin. *The Open Microbiology Journal*, 4(1), pp.8-19. https://doi.org/10.2174/1874285801004010008 . (2010).

[186] Schneider, L., Korber, A., Grabbe, S. and Dissemond, J., Influence of pH on wound-healing: a new perspective for wound-therapy?. *Arch. Dermatol. Res.*, 298(9), pp.413-420. https://doi.org/10.1007/s00403-006-0713-x (2006).

[187] Wlaschek, M. and Scharffetter-Kochanek, K.,. Oxidative stress in chronic venous leg ulcers. *Wound Repair and Regeneration*, 13(5), pp.452-461. https://doi.org/10.1111/j.1067-1927.2005.00065.x (2005).

[188] Broza, Y., Mochalski, P., Ruzsanyi, V., Amann, A. and Haick, H., Hybrid Volatolomics and Disease Detection. *Angew. Chem. International Edition*, 54(38), pp.11036-11048 https://doi.org/10.1002/anie.201500153 . (2015).

[189] SHIRASU, M., NAGAI, S., HAYASHI, R., OCHIAI, A. and TOUHARA, K., 2009. Dimethyl Trisulfide as a Characteristic Odor Associated with Fungating Cancer Wounds. *Bioscience, Biotechnology, and Biochemistry*, 73(9), pp.2117-2120.

[190] Thuleau, A., Dugay, J., Dancremont, C., Jemmali, Z., Elard, J., Ricke, Y., Cassoux, N., Watson, S., Escande, M. and Fromantin, I., 2018. Volatile Organic Compounds of Malignant Breast Cancer Wounds: Identification and Odors. *INDEX WOUNDS*, 30(11), pp.337-344.

[191] Zhu, J., Chaudhury, M., Durso, L., Sagel, A., Skoda, S., Jelvez-Serra, N. and Santanab, E., 2017. Semiochemicals released from five bacteria identified from animal wounds infested by primary screwworms and their effects on fly behavioral activity. *PLOS ONE*, 12(6), p.e0179090.

[192] van Oort, P., White, I., Ahmed, W., Johnson, C., Bannard-Smith, J., Felton, T., Bos, L., Goodacre, R., Dark, P. and Fowler, S., 2021. Detection and quantification of exhaled volatile organic compounds in mechanically ventilated patients – comparison of two sampling methods. *The Analyst*, 146(1), pp.222-231. https://doi.org/10.1039/C9AN01134J

[193] Kos, R., Brinkman, P., Neerincx, A., Paff, T., Gerritsen, M., Lammers, A., Kraneveld, A., Heijerman, H., Davies, J., Janssens, H., Majoor, C., Weersink, E., Sterk, P., Haarman, E., Bos, L. and Maitland-van der Zee, A., 2020. Targeted analysis of volatile organic compounds for detection of Pseudomonas aeruginosa in cystic fibrosis patients by exhaled breath analysis. *Journal of Cystic Fibrosis*, 19, p.S52. https://doi.org/10.1016/j.jcf.2021.04.015

[194] de Lacy Costello, B., Amann, A., Al-Kateb, H., Flynn, C., Filipiak, W., Khalid, T., Osborne, D. and Ratcliffe, N., 2014. A review of the volatiles from the healthy human body. *Journal of Breath Research*, 8(1), p.014001.

[195] Drabińska, N., Flynn, C., Ratcliffe, N., Belluomo, I., Myridakis, A., Gould, O., Fois, M., Smart, A., Devine, T. and Costello, B., 2021. A literature survey of all volatiles from healthy human breath and bodily fluids: the human volatilome. *Journal of Breath Research*, 15(3), p.034001.

[196] Duffy, E. and Morrin, A., Endogenous and microbial volatile organic compounds in cutaneous health and disease. *TrAC, Trends Anal. Chem*, 111, pp.163-172. https://doi.org/10.1016/j.trac.2018.12.012 (2019).

44

[197] Bernier, U., Kline, D., Barnard, D., Schreck, C. and Yost, R., 2000. Analysis of Human Skin Emanations by Gas Chromatography/Mass Spectrometry. 2. Identification of Volatile Compounds That Are Candidate Attractants for the Yellow Fever Mosquito (Aedes aegypti). *Analytical Chemistry*, 72(4), pp.747-756.

[198] Jiang, R., Cudjoe, E., Bojko, B., Abaffy, T. and Pawliszyn, J., 2013. A non-invasive method for in vivo skin volatile compounds sampling. *Analytica Chimica Acta*, 804, pp.111-119.

[199] Salman, D., Ibrahim, W., Kanabar, A., Joyce, A., Zhao, B., Singapuri, A., Wilde, M., Cordell, R., McNally, T., Ruszkiewicz, D., Hadjithekli, A., Free, R., Greening, N., Gaillard, E., Beardsmore, C., Monks, P., Brightling, C., Siddiqui, S. and Thomas, C., 2022. The variability of volatile organic compounds in the indoor air of clinical environments. *Journal of Breath Research*, 16(1). https:doi.org/10.1088/1752-7163/ac3565

[200] Aghdassi, E. and Allard, J., 2000. Breath alkanes as a marker of oxidative stress in different clinical conditions. *Free Radical Biology and Medicine*, 28(6), pp.880-886.

201 Phillips, M., Greenberg, J. and Cataneo, R., 2000. Effect of age on the profile of alkanes in normal human breath. *Free Radical Research*, 33(1), pp.57-63.

202 Mittal, M., Siddiqui, M., Tran, K., Reddy, S. and Malik, A., Reactive Oxygen Species in Inflammation and Tissue Injury. *Antioxidants & Redox Signaling*, 20(7), pp.1126-1167. https://doi.org/10.1089/ars.2012.5149 . (2014).

[203] Dallüge, J., Beens, J. and Brinkman, U., 2003. Comprehensive two-dimensional gas chromatography: a powerful and versatile analytical tool. *Journal of Chromatography A*, 1000(1-2), pp.69-108. https://doi.org/10.1016/S0021-9673(03)00242-5

[204] Nolvachai, Y., McGregor, L., Spadafora, N., Bukowski, N. and Marriott, P., 2020. Comprehensive Two-Dimensional Gas Chromatography with Mass Spectrometry: Toward a Super-Resolved Separation Technique. *Analytical Chemistry*, 92(18), pp.12572-12578. https://doi.org/10.1021/acs.analchem.0c02522

[197] Bernier, U., Kline, D., Barnard, D., Schreck, C. and Yost, R., 2000. Analysis of Human Skin Emanations by Gas Chromatography/Mass Spectrometry. 2. Identification of Volatile Compounds That Are Candidate Attractants for the Yellow Fever Mosquito (Aedes aegypti). *Analytical Chemistry*, 72(4), pp.747-756.

[198] Jiang, R., Cudjoe, E., Bojko, B., Abaffy, T. and Pawliszyn, J., 2013. A non-invasive method for in vivo skin volatile compounds sampling. *Analytica Chimica Acta*, 804, pp.111-119.

[199] Salman, D., Ibrahim, W., Kanabar, A., Joyce, A., Zhao, B., Singapuri, A., Wilde, M., Cordell, R., McNally, T., Ruszkiewicz, D., Hadjithekli, A., Free, R., Greening, N., Gaillard, E., Beardsmore, C., Monks, P., Brightling, C., Siddiqui, S. and Thomas, C., 2022. The variability of volatile organic compounds in the indoor air of clinical environments. *Journal of Breath Research*, 16(1). https:doi.org/10.1088/1752-7163/ac3565

[200] Aghdassi, E. and Allard, J., 2000. Breath alkanes as a marker of oxidative stress in different clinical conditions. *Free Radical Biology and Medicine*, 28(6), pp.880-886.

201 Phillips, M., Greenberg, J. and Cataneo, R., 2000. Effect of age on the profile of alkanes in normal human breath. *Free Radical Research*, 33(1), pp.57-63.

202 Mittal, M., Siddiqui, M., Tran, K., Reddy, S. and Malik, A., Reactive Oxygen Species in Inflammation and Tissue Injury. *Antioxidants & Redox Signaling*, 20(7), pp.1126-1167. https://doi.org/10.1089/ars.2012.5149 . (2014).

[203] Dallüge, J., Beens, J. and Brinkman, U., 2003. Comprehensive two-dimensional gas chromatography: a powerful and versatile analytical tool. *Journal of Chromatography A*, 1000(1-2), pp.69-108. https://doi.org/10.1016/S0021-9673(03)00242-5

[204] Nolvachai, Y., McGregor, L., Spadafora, N., Bukowski, N. and Marriott, P., 2020. Comprehensive Two-Dimensional Gas Chromatography with Mass Spectrometry: Toward a Super-Resolved Separation Technique. *Analytical Chemistry*, 92(18), pp.12572-12578. https://doi.org/10.1021/acs.analchem.0c02522

Pre-examination copy
Submitted 27th May 2022

Chapter 2: Multi -strain volatile profiling of pathogenic and commensal cutaneous bacteria

Declaration of Authorship

Candidates are required to submit a separate **Declaration of Authorship** form for each co-authored paper submitted for examination as part of a PhD by Publication thesis. Further information is available from the [accompanying guideline document](#).[1]

Section 1: Candidate's details	
Candidate's Name	SHANE FITZGERALD
DCU Student Number	14482908
School	School of Chemical Sciences
Principal Supervisor	Aoife Morrin
Title of PhD by Publication Thesis	Microbial Volatile Organic Compounds – A Path Towards Rapid Non-Invasive Detection of Wound Infections
Section 2: Paper details	
Title of co-authored paper included in the thesis under examination	Multi-strain volatile profiling of pathogenic and commensal cutaneous bacteria.
Publication Status	Published
ISSN and link to URL (where available)	https://doi.org/10.1038/s41598-020-74909-w

This paper is one of ☐3☐ co-authored papers to be submitted as part of the PhD by publication thesis submitted for examination

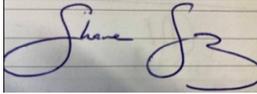| Section 3: Candidate's contribution to the paper |
|---|
| Provide details below of the nature and extent of your contribution to the paper (include both your intellectual and practical contributions) and your overall contribution in percentage terms: |
| I conceptualised and planned the experiments and I carried out the experimental investigation, data analysis, data visualisation, and writing of manuscripts. |

Pre-examination copy
Submitted 27th May 2022

Where a paper has joint or multiple authors, list the names of all other authors who contributed to the work (this can be appended in a separate document, where necessary)**:**
Dr. Aoife Morrin
Dr. Linda Holland

Section 4: Signature and Validation

I confirm that the following statements are true:
the information I have provided in this form is correct
this paper is based on research undertaken during my candidature at DCU

Signature of PhD Candidate: Date: 24th May 2022

I confirm that the information provided by the candidate is correct:

Signature of Principal Supervisor: Date: 25th May 2022

In some cases, it may be appropriate for verification to be given by both the principal supervisor <u>and</u> the lead/corresponding author of the work (where the lead/corresponding author of the work is not the candidate or the principal supervisor):

Signature of Lead/Corresponding Author: Date: 25th May 2022

Home institution: Dublin City University

Pre-examination copy
Submitted 27th May 2022

Abstract

Bacteria emit volatile organic compounds (VOCs) as by-products of metabolism. The detection of VOCs emitted by pathogenic bacteria has been proposed as a potential non-invasive approach for characterising various infectious diseases as well as wound infections. Species- and strain-level diversity in specific infections are associated with poor outcomes. In this study, multiple strains of pathogenic and commensal cutaneous bacteria were analysed using headspace (HS) solid phase micro-extraction (SPME) coupled with gas chromatography - mass spectrometry (GC-MS). We investigated species- and strain-level variation of recovered VOC profiles using multivariate analysis techniques. A kinetic study was also carried out to assess the relationship between bacterial VOC profiles and the respective growth phase of cells. Comprehensive VOC profiles for each strain were obtained. Species-level discrimination was clearly observed across all species tested, while there was limited evidence of strain-level discrimination between respective strains. However, certain VOCs appeared to be specific to individual strains of *E. coli* and *P. aeruginosa* samples. The VOC profiles of particular compound groups were proportional to the respective growth phase for individual *S. aureus* and *P. aeruginosa* strains. This work serves to identify species- and strain-level VOC profiles as well as metabolomic trends that can potentially support and aid interpretation of observed trends in future untargeted studies.

49

2.1. Introduction

The production of VOCs by microorganisms in different media and biological fluids has been investigated for many years with the aim of characterising various disease-specific odours [1]. VOC profiling studies of pure bacterial cultures are needed to identify the cellular origin of metabolites associated with specific phenotypes of pathogens under specific conditions[2]. Untargeted profiling analyses investigating respiratory infections in patients with cystic fibrosis[3][4], pneumonia[5], and tuberculosis[6] have demonstrated the discriminatory power and diagnostic potential of VOCs. These analyses rely on the identification of trends in VOC profiles between various disease-associated subjects and healthy subjects. The 'top down' workflow generally consists of analysing whole VOC profiles via multi-variate analysis techniques such as principal component analysis (PCA) to discriminate between the two groups[5][6]. The same approaches are typically employed in bacterial VOC profiling studies to investigate species-level diversity. Studying VOCs emitted from bacteria under controlled conditions identifies potential mechanisms behind projected VOC profiles in infectious disease-associated individuals.

Bacteria produce VOCs as side-products of primary metabolism and secondary metabolism[7]. The aim of primary metabolism is to simply metabolise all available glucose and derive as much adenosine triphosphate (ATP) as possible, which occurs during the exponential growth phase[8]. Secondary metabolism occurs in the stationary growth phase – under resource-limited conditions - and involves the further metabolism of primary metabolites[9] and fermentation processes that generate alcohols and acetate[8]. The biosynthesis and subsequent metabolism of fatty acids (FAs) are both multi step processes that can generate VOCs at each individual step[9]. FAs are produced from acetyl CoA (or propionyl-, isobutyryl-, isovaleryl-, or 2-methylbutyryl-CoA), which are extended with malonate units to assemble various fatty acids. They are then metabolised by the β-oxidation pathway[10]. The processes consist of multiple decarboxylation, hydrolysis, and reductions which generate a variety of alkanes[11], 1-alkenes[12][11], methyl ketones[13][14], and 1-alkanols[15][9]. Microbes can also metabolise amino acids to produce volatile short-chain FAs[9][16][17] such as 3-methylbutyric acid[18]. Proposed microbial metabolic pathways of VOC production are discussed further elsewhere[9].

Previous studies have shown that bacteria have species-specific VOC profiles that are directly influenced by growth parameters such as growth media[19][20], incubation time[21][22], oxygen content (headspace volume)[23], temperature[24] and pH[25]. The results from studies are also influenced by the sampling techniques employed. Frequently used sampling techniques coupled with GC-MS include SPME[26][27], thermal desorption tubes[10], direct syringes[28][29]. Direct detection methods such as SIFT-MS[26], SESI-MS[30] and PTR-MS[31] have been previously employed for real-time analysis of VOCs, however, the resulting VOC profiles obtained from these methods typically contain low numbers of compounds. The variation in growth parameters and instrumental techniques across studies make inter-study comparisons difficult and highlights the need for more supporting literature and comprehensive data. The mVOC database[32] (updated to mVOC 2.0[33]) contains thousands of logged VOCs from a wide range of microbes, as well as proposed metabolomic pathways. Databases such as this could potentially emerge as a valuable tool in the field, however mVOC 2.0 is still in the early stages of development and requires continuous updating and additional data to make it as comprehensive and reliable as possible.

It is estimated that around one in four people with diabetes will develop a diabetic foot ulcer (DFU) in their lifetime[34]. Infections of DFUs are directly associated with poor outcomes[35] and it has recently been demonstrated that strain-level diversity in wound infections is associated with infection severity[36]. Using a rapid HS-SPME-GCMS workflow, we obtained the VOC profiles of multiple strains of prevalent wound pathogens[37][38] *S. aureus, P. aeruginosa,* and *E. coli*, as well as multiple strains of the skin commensal *S. epidermidis* and media controls. We investigated species- and strain-level variation using multi-variate analysis techniques. We also monitored VOC emissions of individual strains of *S. aureus* and *P. aeruginosa* over 48 h to assess the relationship between the bacterial VOC profile and growth phase. The results of our study demonstrate that comprehensive datasets are critical for the interpretation of strain-level variation in VOC profiles and highlights that further work is required in this area. Profiling the emission of certain compound classes over time visualised multiple relationships between the bacterial VOC profiles and the respective growth phase of the cells and showed that bacterial growth can potentially be metabolically tracked. Untargeted bacterial VOC profiling has the potential to be employed as a non-invasive diagnostic tool for a range of infections including cutaneous disorders and chronic wounds. This work serves to identify species- and strain- specific cellular metabolites and metabolomic trends that could potentially support and aid interpretation of observed trends in future untargeted studies.

2.2. Methods

2.2.1. Growth of Bacteria

The following bacterial strains were studied: *S. aureus (*DSM2569 and DSM799)*; P. aeruginosa* (DSM105372 and DSM25642)*; E. coli* (DSM30083 and DSM103372; and *S. epidermidis* (CSF41498 & RP62A)*.* All *S. aureus, P. aeruginosa, and E. coli* isolates were obtained from Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures GmbH; *S. epidermidis* strains were provided by Prof. O'Gara at NUI Galway. Each strain was streaked individually on tryptone soy broth (TSB) agar media plates and incubated at 37°C overnight. Overnight liquid cultures were prepared in 4 mL of TSB broth and incubated at 37°C overnight with shaking (180 rpm). The samples are referred throughout the text using the following acronyms*: **EC.A:** E. coli DSM103372,* **EC.B:** *E. coli DSM30083,* **PA.A**: *P. aeruginosa DSM105372,* **PA.B**: *P. aeruginosa DSM25642,* **SA.A:** *S. aureus DSM2569,* **SA.B:** *S. aureus DSM799,* **SEP.A:** *S. epidermidis, CSF41498,* **SEP.B:** *S. epidermidis RP62A and* **TSB***:* growth media control. Samples for VOC analysis were prepared in 20 mL sterile headspace (HS) vials (Merck, Cork, Ireland). Overnight cultures were diluted in 5 mL of TSB media to a final cell count of approximately 10^7-10^{10} colony forming units (CFU)/ mL in the HS vials which were then sealed with magnetic Polytetrafluoroethylene /silicone septum screw caps (Merck, Cork, Ireland). Samples were set up in triplicate and incubated at 37°C with shaking for a set period of time. Nine black media controls were also sampled and analysed at these conditions. Growth curve analysis was performed using *S. aureus* and *P. aeruginosa.* Samples were set up in triplicate and incubated at 37°C with shaking. A spectrometer is used to measure the optical density of a given culture at a 600 nm (OD_{600}). OD_{600} was taken at 2, 4, 6, 8, 24, 32, 48 h.

2.2.2. Sampling Procedure

SPME fibers were used for sampling VOCs and comprised of 85µm Carboxen/PDMS Stableflex (2 cm) assemblies (Supelco Corp., Bellefonte, PA, USA). Prior to sampling, each bacterial or control sample was removed from the shaking incubator and placed in a standard incubator at 37°C. The SPME needle was pierced through the septum of the HS vial, and the fibre was exposed to the HS of the sample for 20 min, after which, the fibre was retracted and the SPME assembly removed from the vial. The SPME fibre was then inserted into the GC inlet and thermally desorbed at 250°C for 2 min for subsequent separation and detection by mass spectrometry.

Background subtraction was carried out by sampling blank headspace vials, and blank media samples. Compounds recovered from these blank analyses were individually assessed. Compounds recovered samples with signal-to-noise ratios greater than 3:1 were considered for inclusion in the study. Compounds recovered from blank analyses are summarised in Table 2.1.

2.2.3. Gas chromatography–mass spectrometry

An Agilent 6890 GC connected to an Agilent 5973 mass selective detector (Agilent Technologies, Inc., Santa Clara, CA, USA) was used for all analyses. Separations were performed on a DB-WAX column (Agilent Technologies Ireland, Cork) (30 m × 0.25 mm × 0.32 µm). The carrier gas used was helium, with a constant flow rate of 1.3 mL /min For manual injections of SPME fibers, the system was equipped with a SPME Merlin Microseal (Merlin Instrument Company, Newark, DE, USA), and the inlet was maintained at a temperature of

52

250°C. Split-less injection was used for all samples, with a gas purge being activated after 2 min. Each SPME fibre was desorbed for 2 min within a SPME inlet liner (Supelco). The initial GC oven temperature was 40°C for 5 min and was programmed to increase at a rate of 10°C min^{-1} to 240°C, with a final hold for 5 min at this temperature, giving an overall running time of 29 min. The MS was operated at a scan range of 35-400 *m/z*, scan rate of 3.94 s^{-1}, ion source temperature 230°C and ionising energy of 70 eV. Identification of compounds relied on a three phase protocol whereby National Institute of Standards and Technology (NIST) library (2017) - match factors of >70% were initially used to assess potential ID matches; fragmentation patterns of potential matches were then manually interpreted before being validated using retention index matching. Retention index (RI) matching was used to support the identification of these compounds. Any compound found to have an RI value ≤12 RI units of the RI values found in the NIST database were deemed acceptable matches. A standard mixture of saturated alkanes (C$_7$-C$_{30}$; Merck, Cork, Ireland) was used for RI matching.

2.2.4. Data Analysis

The open source software OpenChrom[39] was used to analyse raw chromatographic data. Chromatographic peaks were compared using the NIST Chemistry WebBook. Peaks found to be from exogenous sources such as the SPME fiber, glass vial, and column were removed from the dataset. RStudio was used for data exploration and visualisation. Raw bacterial VOC data was standardised using centering and scaling [40]. Centering converts all the values in the dataset to fluctuations around zero rather than fluctuations around the mean VOC abundance. It adjusts for differences in the offset between low and high abundances. Scaling converts the values in the dataset into ratios relative to the difference in abundances between the VOCs, which allows each VOC to be equally represented in the subsequent data analysis. Hierarchical clustering and principal component analysis (PCA) were carried out on the dataset using the R packages : 'FactoMineR', 'factoextra', 'pheatmap', 'egg' and 'cluster' . Other R packages used included: 'tidyverse', 'ggplot2', 'ggfortify'.

Pre-examination copy
Submitted 27th May 2022

2.3. Results

2.3.1. Comparative analysis of volatiles emitted from planktonic bacteria cultures

GC-MS analysis of the VOCs recovered from the SPME fibers showed that there was a wide variety of compound classes present in the HS of the bacterial cultures. Numerous blank samples were collected and analysed to identify and exclude exogenous compounds from the SPME fiber, glass vial, and column. A total of 65 identified compounds were identified from the bacterial and control samples (see Figure 2.1). Of these, 19 compounds were found in the HS of media control samples. Following 24 h incubation, *S. aureus, P. aeruginosa, E. coli*, and *S. epidermidis* all generated characteristic VOCs. Two individual strains of each species of bacteria were cultured and analysed in triplicate (see Methods for specific information on species and strains). Compound identification was performed and RI matched VOC profiles were established for each bacterial strain and integrated into a complete dataset, which incorporated all species, strains and controls tested and the compounds identified. An initial visual inter-strain comparison was performed by overlaying the chromatograms (Figure S2.1-S2.4) which demonstrate a high degree of similarity between the VOC profiles of respective strains.

Hierarchical clustering was performed to visualise the similarities across the VOC data. It is a statistical method used to classify multiple objects into groups (clusters) based on similarities between them. The results are visualised as a dendrogram (Figure 2.1). Dendrograms are bottom-up representations of the clustering procedure; each object is initially assigned to its own cluster, and these individual clusters are grouped together based on their similarity. The clustering algorithm then progressively joins similar clusters together until all objects are grouped by a single cluster. The length of an edge between a cluster and its split is proportional to the dissimilarity between the split clusters[41]. Figure 2.1 visualises the clustering results coupled with a heatmap to show the different patterns in individual VOC abundances across all bacterial strains. The heatmap contains 65 rows, which are labelled by each VOC, and listed in order of increasing retention time. In Figure 2.1, the relative VOC abundances are visualised using a gradient red / yellow colour scale, where dark orange/red represents a high abundance; and orange/yellow represents a low abundance. All bacterial strains – except for *SEP.A* - were successfully clustered with their respective species. The Euclidean distance between the *E. coli* strains and the other bacteria tested was the greatest, verifying that the two *E. coli* strains had the most discriminative VOC signatures of all samples tested. S. *epidermidis* strains emitted a lower number of VOCs in comparison to the other species and were clustered close to the media control as a result. *S. aureus* and *P. aeruginosa* both have clearly differentiated VOC signatures, which can be seen from the heatmap by the relatively high number of unique 'red' values.

Principal component analysis (PCA) was performed to summarise the dissimilarities in the data. PCA reduces the dimensionality of the data by identifying characteristic VOCs, and using them to construct new linear variables called principal components (PCs), along which the variation is maximal. The PCs are variables consisting of linear combinations of the original variables; which can then be visualised using scores plots. Scores plots show inter-unit distances and visualise species- and sample- like patterns revealed by the PCA to identify groups that characterise the overall dataset[42]. Initially a cluster number of 5 was used to summarise the 4 bacterial species and the control group (Figure S2.5). However, the two *E.coli*

54

strains form two individual clusters rather than clustering together and the *S. epidermidis* samples were clustered with the control media samples (Figure 2.5). A cluster number of 4 was subsequently chosen as it produced the best summary of the data (Figure 2.2.). *S. aureus, P. aeruginosa*, and *E. coli* strains were all successfully clustered to their respective species; *S. epidermidis* strains were clustered to the control samples.
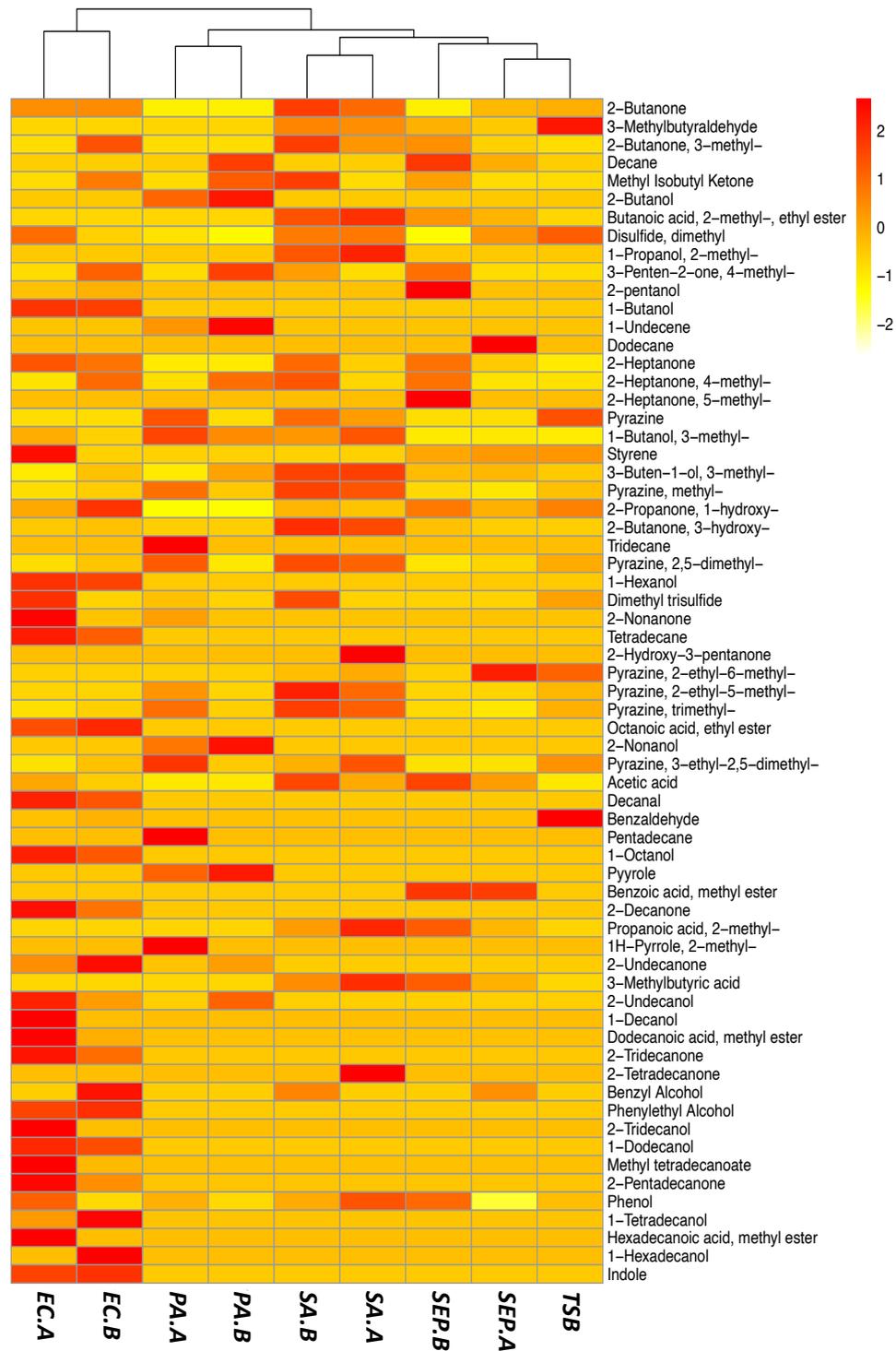
Pre-examination copy
Submitted 27th May 2022

Figure 2.1: Heatmap showing the relative abundance of VOCs recovered (rows) from each bacterial strain (columns). Values were scaled and centred by their respective rows, with highly abundant VOCs being coloured red, and less abundant VOCs being marked orange - yellow .
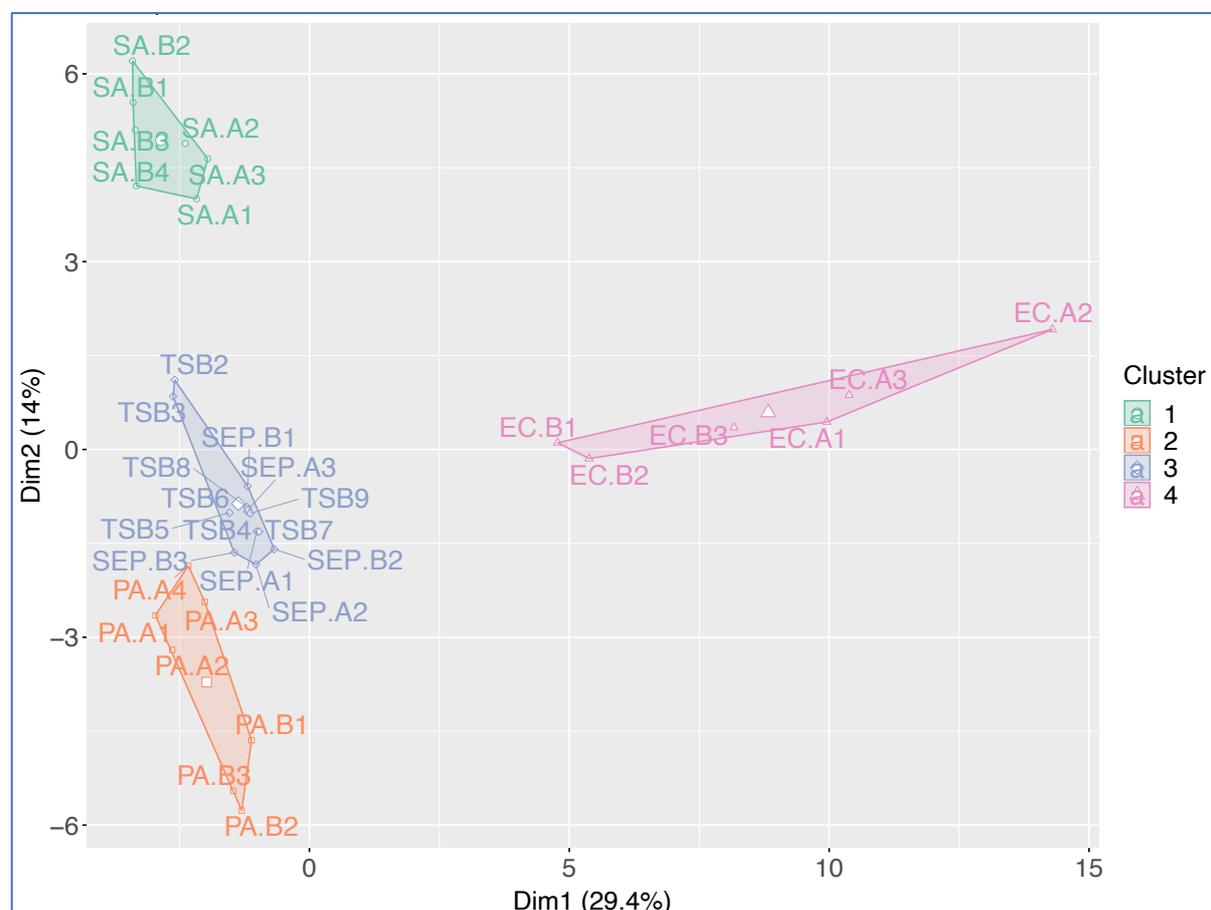

Figure 2.2: Scores plot representation of scores of bacterial samples. No. of clusters (n=4). PC1 and PC2 summarised 43.4% of the variance of the overall dataset, with 29.4% being summarised by PC1 and 14% being summarised by PC2.

The distribution of the major compound classes recovered from all tested species of bacteria and controls is summarised in Figure 2.3. Ketones were recovered from all bacterial samples (Figure 2.3) and the controls. The largest number of ketones were recovered from *E. coli* samples. The relative abundance of individual ketones can be seen in Figure 2.1. Long chain methyl ketones such as 2-undecanone, 2-tridecanone, and 2-pentadecanone were recovered from *E. coli*. High abundances of 3-hydroxy-2-butanone (acetoin) were observed in all *S. aureus* chromatograms. Acetoin was also detected in *S. epidermidis* and *E. coli* samples, but in lower amounts compared to that of *S. aureus*. *P. aeruginosa* was found to emit a lower number of ketones than other species tested. Both *P. aeruginosa* strains were found to emit low abundances of 2-undecanone, while 2-nonanone was only detected in *PA.B* (Figure 2.1). 2-butanone, 3-methyl-2-butanone, and 1-hydroxy-2-propanone were all detected in the media controls.

Alcohols were recovered in high abundances from all bacterial samples and low abundances from the controls. In Figure 2.3, the median abundance value for alcohols was similar for *E.*

coli, *S. aureus,* and *P. aeruginosa* samples; the abundance of alcohols in *S. epidermidis* and control samples was relatively lower. Out of the bacteria tested, some alcohols were shared between species, and others were unique to individual species. 3-methyl-1-butanol (isoamyl alcohol) was detected in every strain of bacteria and was particularly abundant in all *S. aureus* and *PA.A* samples (Figure 2.1). 3–methyl-3-buten-1-ol was detected in each strain except for *PA.B* and *EC.A*. 2-butanol and 2-nonanol were extracted from both *P. aeruginosa* strains, though 2-undecanol was only extracted from the *PA.B* strain. We identified various 1- and 2 - alcohols from both *E. coli* strains, these included 1-hexanol, 1-octanol, 1-decanol (only *EC.A*), 2-undecanol, 2-tridecanol (*EC.A)*, 1-tetradecanol, and 1-hexadecanol (only *EC.B*). Low abundances of 1-dodecanol and benzyl alcohol were the only alcohols detected in the blank growth media.

High abundances of acids were detected in all *S. aureus* and *S. epidermidis* samples, and to a lesser extent in the *E. coli* samples; while none were detected in the control or a in *P. aeruginosa* samples (Figure 2.3)*.* High abundances of 3-methylbutyric acid and acetic acid; and relatively lower abundances of propanoic acid, 2-methyl- and were observed in all *S. aureus* and *S. epidermidis* chromatograms. High abundances of acetic acid were also detected in *E. coli* samples (Figure 2.1). No acids were detected in the control samples.

Aldehydes were detected in low abundances in *S. aureus, E. coli,* and *S. epidermidis* samples, and in high abundances in the controls (Figure 2.3). 3-methylbutyraldehyde was detected in moderate abundances in *S. aureus* samples and lower abundances in *S. epidermidis* samples (Figure 2.1). Low abundances of decanal and benzaldehyde were detected in *E. coli* samples (Figure 2.1). No aldehydes were recovered from any *P. aeruginosa* samples. A high abundance of aldehydes such as 3-methylbutyraldehyde and benzaldehyde were detected in the control samples (Figure 2.3).

Fatty acid ethyl esters were detected in *E. coli, S. aureus* and *S. epidermidis* samples. The highest number of individual fatty acid ethyl esters were detected in *E. coli* samples (Figure 2.1) , whereas *S. aureus* samples had the highest median abundance of fatty acid ethyl esters (Figure 2.3). Butanoic acid, 2-methyl, ethyl ester was detected in *S. aureus* (relatively high abundance) and *S. epidermidis* (relatively mid abundance); long-chain compounds such as dodecanoic acid, methyl ester, methyl tetradecanoate, and hexadecenoic acid, methyl ester were detected in *E. coli* samples. No fatty acid ethyl esters were detected in the control samples.

Pyrazine compounds were detected in all bacteria samples. All pyrazines were also detected in all media controls. Variation seen in Figure 2.3 could be a result of batch variation, as it can be seen that the error bars of the growth media control box covers the interquartile range of the other species tested.

There were multiple characteristic compounds detected that didn't fall into the compound classes discussed above. 1-undecene and pyrrole in *P. aeruginosa* samples (2-methyl-1H-pyrrole was only present in the HS of *PA.B)*. Indole (the most abundant compound detected out of all the bacterial samples) was detected in all *E. coli* samples. Styrene was detected in *E. coli* samples, *S. epidermidis* samples and in very low abundances in control samples.

Pre-examination copy
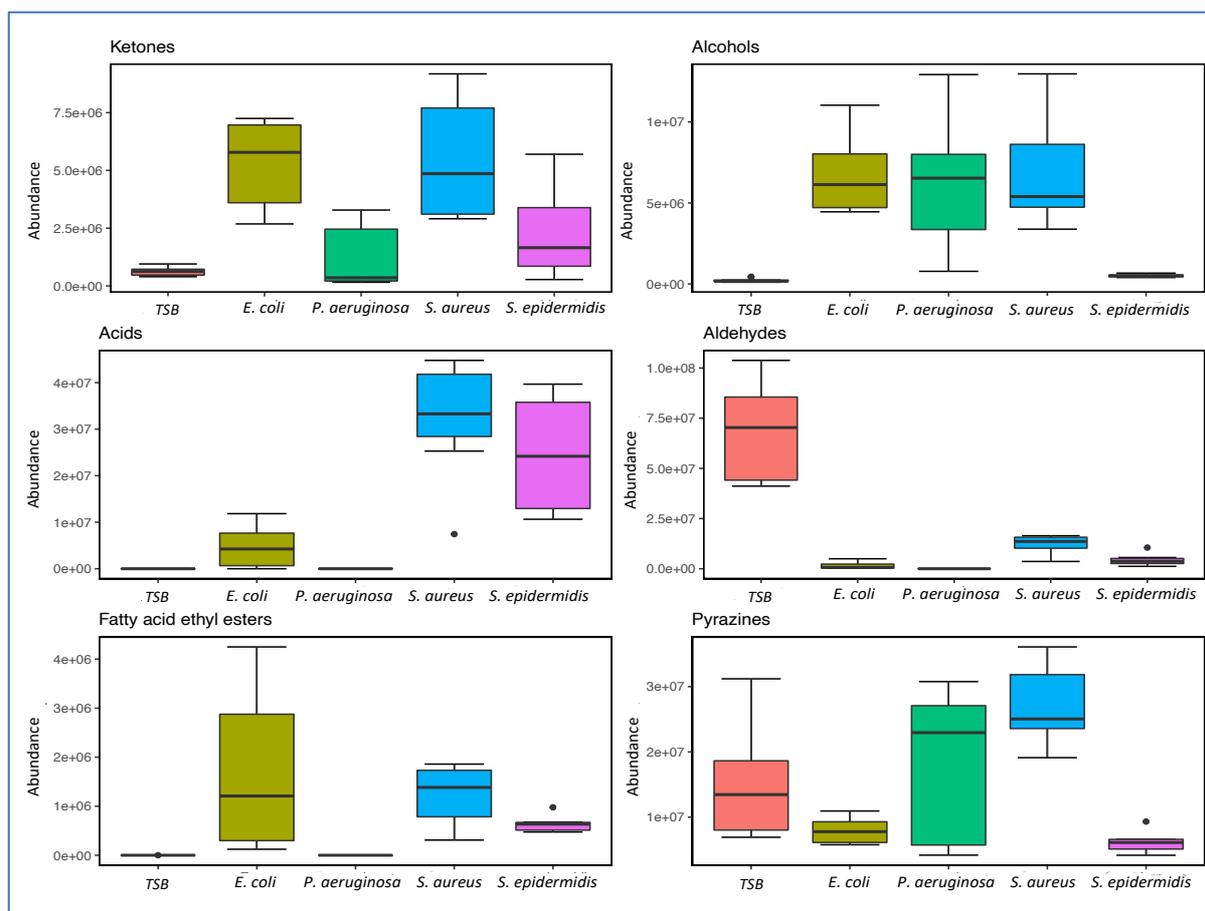Submitted 27th May 2022

Figure 2.3: Grouped boxplots showing the abundance of ketones (top left), alcohols (top right), acids (middle left), aldehydes (middle right), fatty acid ethyl esters (bottom left), and pyrazines (bottom right) recovered from the control and each species of bacteria (t = 24 h). Respective strains were grouped together to clearly summarise the species-level discrimination across the data. The boxes represent the interquartile range: the line running across each box represents the 50th percentile (median), the top of the box represents the 75th percentile, and the bottom of the box represents the 25th percentile. The whiskers (error bars) represent either the smallest or largest value within 1.5 times the interquartile range above the 75th percentile or below the 25th percentile. The black dots above or below the boxes represent outliers that are greater than 1.5 times the interquartile range beyond either end of the box.

2.3.2. Kinetic study of *S. aureus* and *P. aeruginosa* VOC production

To investigate the relationship between VOC emission and bacterial growth stage, growth curves for individual *S. aureus (SA.A)* and *P. aeruginosa (PA.B)* strains were constructed (Figure 2.4a & b) and compared with VOC emission abundances of specific compound classes over time (Figure 2.4 (c) & (d)). Growth curves were constructed from OD_{600} readings taken at defined time points over 48 h. Emission kinetics plots for individual compounds can be found for *S. aureus* (Figure S2.6) and *P. aeruginosa* (Figure S2.7) in the supplementary information. VOCs were sampled using SPME at equivalent time points over this same period. As SPME is an equilibration process, it was necessary to allow samples equilibrate at 37°C for 0.5-1 h before VOC sampling. Due to this, the 0 h data points used for VOC abundances (Figure 2.4 (c) & (d)) were taken from a set of blank media controls equilibrated at 37°C.

58

The constructed growth curves allow the visualisation of the growth stage of an organism. In this case, OD_{600} as a measure of the turbidity of a bacterial liquid culture – as bacterial cells multiply, the liquid culture becomes more turbid – is used to assess bacterial growth stage[43]. Typically, organisms proceed along the path set by the bacterial growth curve, passing through four characteristic stages: lag, log (exponential growth), stationary, and decline. The lag phase is a distinct growth phase whereby the organisms are adapting to their new environment and preparing for rapid growth[44]. The first OD_{600} measurement for these experiments was taken after 2 h and indicated that the bacteria were already in the log phase of growth. This can be clearly seen in both *S. aureus* and *P. aeruginosa* samples as a rapid increase in the OD between 0–8 h (Figure 2.4a & b). During this log phase all available resources such as glucose and fatty acids are consumed. When energy sources become limited, the bacteria then enter the stationary phase of growth, and activate reserve pathways that enable the metabolism of secondary substrates to survive[45]. *S. aureus* cell growth decreased between 8–24 h, therefore the stationary phase of growth is not visible in Figure 2.4(a); while *P. aeruginosa* remained in the stationary growth phase from 8–48 h (Figure 2.4(b)).

Overall, Figure 4 shows that the VOC abundances for certain compound classes change with respect to growth phase for both species. For example, aldehyde abundances decreased in *S. aureus* samples in the first 8 h (exponential growth phase) of incubation, and then continued to decrease over the following 40 h but at a slower rate (Figure 2.4(c)). A similar trend was observed in *P. aeruginosa* with aldehyde abundances decreasing to minimal levels by approximately 24 h (Figure 2.4(d)). The initial rapid decrease in aldehyde abundances likely indicates a rapid consumption of the aldehydic compounds present in the growth media during the exponential growth phase. In *S. aureus* samples, the exponential growth phase was also marked by proportional exponential increases in abundances of acids, ketones, and alcohols (Figure 2.4(c)). In *P. aeruginosa* samples, an increase in alcohol abundance was observed in the exponential growth phase, as well as an emission of pyrrole after 2 h incubation (Figure 2.4(d)). Abundances of pyrazines did not change significantly over 48 h for either species which indicated that no pyrazine compounds were consumed or produced by either bacteria over the course of the experiment.

The final phase of the bacterial growth curve is the death phase which is characterised by the net loss of bacterial cells, where the rate of cell death is greater than the rate of cell production due to unsuitable conditions such as exhausted nutrients and lack of oxygen[46]. The death phase of *S. aureus* cells occurred between 8–24 h and was marked by a depletion in aldehyde abundances. In *S. aureus* samples, between 8–24 h, alcohol abundance increased as the aldehyde abundance decreased, suggesting that aldehydes were potentially metabolised into alcohols. From 24–32 h, a reduction in alcohol abundance simultaneously occurred with a comparable increase in acid abundance, which indicated that the alcohols were further metabolised into acids. Abundances of acids, alcohols, aldehydes, and ketones all subsequently declined from 32–48 h. In *P. aeruginosa* samples, the growth curve shows that *P. aeruginosa* cells remained in the stationary growth phase from approx. 8–48 h. During this period there was no significant change in emission of alcohols and ketones. The cumulative abundance of pyrrole compounds appeared to increase and decrease in the early

growth phase between 2–4 h, which was followed by another increase from 4–8 h, before gradually declining in the later stages of the experiment (24–48 h).
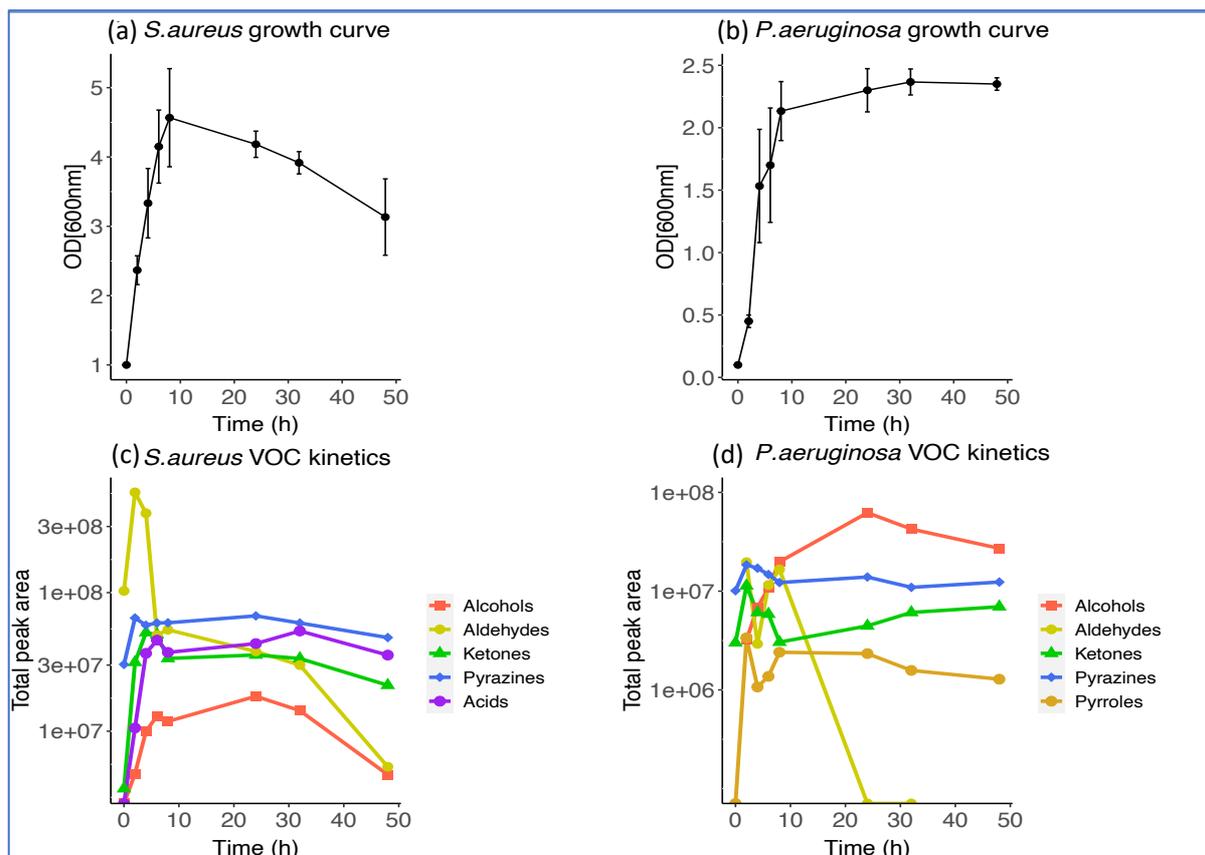

Figure 2.4: Bacterial growth curves of (a) S. aureus (n=3) and (b) P. aeruginosa (n=3) over an incubation period of 48 h OD$_{600}$ measurements were taken at 2, 4, 6, 8 , 24, 32, 48 h. Kinetic plots of (c) S. aureus and (d) P. aeruginosa showing the changes in VOC abundances of different compound classes over an incubation period of 48 h.

2.4. Discussion

In this study we used HS SPME coupled with GC-MS to collect and identify VOCs emitted from multiple strains of pathogenic and commensal species of bacteria. Using multi-variate analysis techniques such as hierarchical clustering (Figure 2.1) and PCA (Figure 2.2) to reveal patterns in the data, each species was successfully discriminated based on their respective VOC profile. The skin commensal, *S. epidermidis*, emitted a relatively low number of VOCs. Acetic acid, 3-methylbutyric acid and 3-methylbutyraldehyde were among the VOCs recovered from its HS. These compounds were also highly abundant in all *S. aureus* samples. Filipiak et al. proposed that *Staphylococcal* species catabolise amino acids found in the growth media to 3-methylbutyraldehyde, which is then oxidised by an aldehyde dehydrogenase to form 3-methylbutyric acid[18] (isovaleric acid). 3-methyllbutyric acid is a characteristic VOC emitted by various *Staphylococcal* species[47][48] and is strongly associated with the generation of body odour[16][49]. *S. aureus, S. epidermidis,* and *E. coli* also produced the high amounts of 3-hydroxy-2-butanone (acetoin). Acetoin is an uncharged molecule - produced by bacteria through the conversion of pyruvate and is known to prevent the over-acidification of the intracellular environment [50]. It can be seen in Figure 2.1 that the *S. epidermidis* strains did generate characteristic VOC profiles, but they were clustered close to the media control due to the limited number of species-specific VOCs recovered from the HS of the samples. Coagulase-negative *Staphylococci* such as *S. epidermidis* have been previously reported to exhibit a relatively slow metabolism of carbohydrates when compared to pathogenic bacteria[51][52]. Despite this, in a recent comparative study, the VOC profiles of *S. epidermidis* were reported to be highly dependent on the growth media used[20]. *S. epidermidis* cultures emit a relatively lower number of VOCs when cultured in TSB media[20], compared to other growth media; this would suggest that our choice of growth media was a potential factor that influenced the limited VOC profiles of *S. epidermidis* observed here.

P. aeruginosa emitted a stable set of compounds that allowed it to be clearly discriminated from the other species tested, whereby the notable VOCs found here have all been previously reported. The VOC profiles obtained from the *P. aeruginosa* strains are similar to those reported by Bean et al.[53] and Filipiak et al.[10]. The most abundant compounds recovered from the *P. aeruginosa* HS were 3-methyl-1-butanol and 1-undecene. 3-methyl-1-butanol was common to all bacteria tested while 1-undecene was unique to *P. aeruginosa* and has been previously reported to be produced through the fatty acid metabolism[9]. Acetoin has been previously reported to be emitted by *Pseudomonads*[19], and was detected in relatively low abundances during the early growth phase of *P. aeruginosa* (Figure S2.7). Pyrrole is a unique nitrogen-containing compound that was detected and has also been previously reported[53][42]. 2-nonanone (only detected in PA.A) and 2-undecanone were detected and have been previously reported as being potentially specific to *P. aeruginosa* biofilms[48]. However the results obtained from this study and other studies[54][53] show that these specific compounds are also emitted by planktonic cultures. We did not detect 2-aminoacetophenone - an odorous VOC frequently reported in *P. aeruginosa* VOC profiles[53][55][56] - in any of the samples in either strain.

The *E.coli* strains emitted the largest number of VOCs and had the most distinctive VOC profiles of the four species tested (Figures 2.1 & 2.2). Acetic acid was present in all *E.coli* samples; which has been previously reported to be a product of anaerobic respiration of

carbohydrates[57]. Compounds such as styrene and a variety of fatty acid ethyl esters were extracted from the HS of the *E.coli* samples. This is in agreement with the literature reports on the biosynthesis of these compounds by *E. coli*, which has recently gained interest in the biofuel industry due to the petrochemical properties of these compounds[58][59]. Indole was the most abundant compound recovered from the HS of *E. coli*. It is commonly found in human faeces as a product of *E. coli* activity in the human gut[60], and its high abundance is likely responsible for the characteristic foul odour of the culture. The detection of 1-alcohols such as 1-butanol, 1-hexanol, 1-octanol, and 1-decanol was in agreement with existing literature [61][62]. Indole is the major VOC produced by *E. coli*[9] and has been documented as an intercellular signal molecule amongst diverse bacteria[63]. It is produced by over eighty species of bacteria, though very few produce comparable abundances to *E. coli* [63]. From this, it may be possible to identify the presence of *E. coli* in a real sample from the abundance of indole recovered.

There was limited strain-level diversity observed in VOC profiles. Differences between the two *E. coli* strain profiles included the presence of 1-decanol and 2-tridecanol in *EC.A* samples only; and 1-hexadecanol being present only in *EC.B* samples. Another example of strain-level diversity was between the two *P. aeruginosa* strains, where 2-nonanone was only present in PA.A samples, and 2-undecanol was only present in PA.B samples. Strain-level diversity observed in *S. aureus* and *S. epidermidis* was primarily due to varying abundances of compounds emitted between strains (Figure 2.1). Quantitative strain-level discrimination of bacterial VOC profiles via SIFT-MS has been previously reported for *E. coli* and *Proteus Mirabilis*[64]. However, the number of compounds detected via SIFT-MS appears to be limited across studies[64][65], and analyses of more comprehensive VOC profiles are required to confirm the prospect of strain-specificity. Bean et al. identified a total of 391 compounds across 24 clinical *P. aeruginosa* isolates taken from 8 different sites of the body[53]. They assessed strain-level diversity via hierarchical clustering of the VOC profiles and found that although 4 of isolates taken from the eye clustered together, there was not enough evidence across the rest of the data to suggest that *P. aeruginosa* strains can be differentiated. In our study, overall, there were no significant differences in the whole VOC profiles between strains (*E. coli*, p = 0.484; *S. aureus*, p = 0.472; *P. aeruginosa*, p = 0.434; *S. epidermidis*, p = 0.113). However, the above examples of *E. coli* and *P. aeruginosa* highlight potential measurable differences in specific VOCs between strains of specific species.

The plots shown in Figure 2.4 demonstrate the relationship between the emission of particular compound classes and the growth of *S. aureus* and *P. aeruginosa* cells. The plots shown in Figure S2.6 and S2.7 visualise the emission of individual compounds against time. It can be seen in both kinetic plots (Figure 2.4 (c) and (d)) that aldehyde abundances sharply decreased following the incubation of both *S. aureus* and *P. aeruginosa* samples. As aldehydes were predominantly found in the TSB growth media (Figure 2.4 (c) and (d) (t=0)), it is highly likely that compounds such as benzaldehyde and 3-methylbutyraldehyde were rapidly metabolised by the bacteria and reduced to alcohols[9][66]. Aldehydes (e.g. 3-methylbutyralde) can be reduced to alcohols (e.g. 3-methyl-1-butanol) via alcohol dehydrogenases, or oxidized to acids (e.g. 3-methylbutyric acid) via aldehyde dehydrogenases[18]. *P. aeruginosa* has been reported to metabolise aldehydes very efficiently[18]. 3-methylbutyraldehyde has been described as a marker of *S. aureus* growth - and is a known precursor of 3-methylbutyric acid[18], however it was also found in the HS of blank media controls. We observed an 8-fold increase in the abundance of 3-methylbutyraldehyde between 0–2 h which indicated that *S.*

62

aureus was emitting this compound in the early phase of growth (Figure S2.6), which was followed by a steady decline of 3-methylbutyraldehyde from 2–48 h.

Secondary metabolism of alcohols, aldehydes, fatty acids, and ketones generate many volatile intermediary compounds via reversible reactions, and generate various lipids, alkanes and alkenes as irreversible end products of these pathways[9][66].As the abundance of viable *S. aureus* cells decreases (Figure 2.4 (a)), there are indications of secondary metabolism in Figure 2.4 (c) at 8–24 h, where aldehyde abundances decrease further as the alcohol abundances increase, and then at 24–32 h where alcohol abundances decrease as the acid abundances increase. In this study, the decline of acids, aldehydes, alcohols and ketones from 32-48 h could suggest that these metabolic pathways have been exhausted and that these compounds have been gradually degraded into lipid or hydrocarbon end products[66].

In Figure 2.4 (d), it can be seen that as *P. aeruginosa* cell growth stagnates in the stationary phase between 24–48 h, the emission rate of alcohols, pyrrole, and ketones is arrested and there is no further net increase in any of these chemical classes. Volatile nitrogen-containing compounds such as pyrrole have been previously reported to reach a maximum abundance after a short period of incubation, and then gradually degrade over time[10]**.**In our study, the cumulative abundance of pyrrole compounds (2-methyl-1H-pyrrole and pyrrole) reached a maximum after 2 h incubation. 2-methyl-1H-pyrrole was only detected at 2 and 8 h, while pyrrole was detected at every timepoint and reached maximum abundance at 24 h (Figure S2.7). Degradation mechanisms for pyrrole are not described in the literature. However, the cumulative abundance of pyrrole was essentially halved between 24–48 h, which is in agreement with the finding reported by Filipiak et al.[21] The overall patterns observed in bacterial VOC profiles in response to the growth and death of cells suggests that the detection and monitoring of VOCs could potentially provide a non-invasive means of metabolically tracking bacterial growth.

2.5. Conclusion

The aims of this study were to obtain the VOC profiles of multiple strains of four prevalent bacterial species present in infected wounds; to investigate species- and strain-level diversity in the VOC profiles obtained; and to assess how VOC profiles of *S. aureus* and *P. aeruginosa* were affected with respect to cell growth. Comprehensive VOC profiles for each strain were obtained using HS-SPME GC-MS. Each strain tested emitted a variety of compound classes that allowed clear species-level discrimination. *E. coli* strains emitted the greatest diversity of VOCs, with long chain alcohols, ketones, and indole being the most characteristic VOCs recovered. *S. epidermidis* emitted a relatively low number of VOCs and had the least discriminative VOC profile.   Strain-level variation in VOC profiles was limited, however, differences in the VOC profiles of *E. coli* and *P. aeruginosa* strains highlight specificity in certain bacterial species. Profiling the emission of certain compound classes by *S. aureus* and *P. aeruginosa* over time demonstrated a proportional relationship between the emission of particular compound classes and the respective growth phase of the cells. The results obtained using this robust HS-SPME GCMS workflow are comprehensive with high numbers of identified compounds being recovered, giving high levels of discriminatory power to the method, highlighting its strong potential application for future untargeted microbial studies.

2.6. References

[1] Shirasu, M. and Touhara, K., The scent of disease: volatile organic compounds of the human body related to disease and disorder. *J. Biochem.*, 150(3), pp.257-266. https://doi.org/10.1093/jb/mvr090. (2011).

[2] Filipiak, W. and Bojko, B., SPME in clinical, pharmaceutical, and biotechnological research – How far are we from daily practice?. *TrAC, Trends Anal. Chem.*, 115, pp.203-213. https://doi.org/10.1016/j.trac.2019.02.029. (2019).

[3] Neerincx, A., Geurts, B., van Loon, J., Tiemes, V., Jansen, J., Harren, F., Kluijtmans, L., Merkus, P., Cristescu, S., Buydens, L. and Wevers, R., Detection of Staphylococcus aureus in cystic fibrosis patients using breath VOC profiles. *J. Breath Res.*, 10(4), https://doi.org/10.1088/1752-7155/10/4/046014. (2016).

[4] Robroeks, C., van Berkel, J., Dallinga, J., Jöbsis, Q., Zimmermann, L., Hendriks, H., Wouters, M., van der Grinten, C., van de Kant, K., van Schooten, F. and Dompeling, E., Metabolomics of Volatile Organic Compounds in Cystic Fibrosis Patients and Controls. *Pediatr. Res.*, 68(1), pp.75-80. https://doi.org/10.1203/00006450-201011001-00143. (2010).

[5] Schnabel, R., Fijten, R., Smolinska, A., Dallinga, J., Boumans, M., Stobberingh, E., Boots, A., Roekaerts, P., Bergmans, D. and van Schooten, F., Analysis of volatile organic compounds in exhaled breath to diagnose ventilator-associated pneumonia. *Sci. Rep.*, 5(1). https://doi.org/10.1038/srep17179. (2015).

[6] Zetola, N., Modongo, C., Matsiri, O., Tamuhla, T., Mbongwe, B., Matlhagela, K., Sepako, E., Catini, A., Sirugo, G., Martinelli, E., Paolesse, R. and Di Natale, C., Diagnosis of pulmonary tuberculosis and assessment of treatment response through analyses of volatile compound patterns in exhaled breath samples. *Journal of Infection*, 74(4), pp.367-376. https://doi.org/10.1016/j.jinf.2016.12.006. (2017).

[7] Korpi, A., Järnberg, J. and Pasanen, A. Microbial Volatile Organic Compounds. *Critical Reviews in Toxicology*, 39(2), pp.139-193. https://doi.org/10.1080/10408440802291497. (2009).

[8] Elmassry, M. and Piechulla, B., Volatilomes of Bacterial Infections in Humans. *Front. Neurosci.*, 14. https://doi.org/10.3389/fnins.2020.00257. (2020).

[9] Schulz, S. and Dickschat, J. Bacterial Volatiles: The Smell of Small Organisms. *ChemInform*, 38(45). https://doi.org/10.1039/b507392h. (2007).

[10] Fujita, Y., Matsuoka, H. and Hirooka, K., Regulation of fatty acid metabolism in bacteria. *Mol. Microbiol.*, 66(4), pp.829-839. https://doi.org/10.1111/j.1365-2958.2007.05947.x. (2007).

[11] Schirmer, A., Rude, M., Li, X., Popova, E. and del Cardayre, S., Microbial Biosynthesis of Alkanes. *Science*, 329(5991), pp.559-562. https://doi.org/10.1126/science.1187936. (2010).

[12] Choi, Y. and Lee, S., Microbial production of short-chain alkanes. *Nature*, 502(7472), pp.571-574. https://doi.org/10.1038/nature12536. (2013).

[13] Goh, E., Baidoo, E., Keasling, J. and Beller, H., Engineering of Bacterial Methyl Ketone Synthesis for Biofuels. *Appl. Environ. Microbiol.*, 78(1), pp.70-80. https://doi.org/10.1128/AEM.06785-11 . (2011).

[14] Forney, F. and Markovetz, A.,. The biology of methyl ketones. *J. Lipid Res.*, 12(4), pp.383-395. (1971).

[15] Akhtar, M., Turner, N. and Jones, P., Carboxylic acid reductase is a versatile enzyme for the conversion of fatty acids into fuels and chemical commodities. *Proceedings of the National Academy of Sciences*, 110(1), pp.87-92. https://doi.org/10.1073/pnas.1216516110. (2012).

[16] James, A., Casey, J., Hyliands, D. and Mycock, G., Fatty acid metabolism by cutaneous bacteria and its role in axillary malodour. *World J. Microbiol. Biotechnol.*, 20(8), pp.787-793. https://doi.org/10.1007/s11274-004-5843-8. (2004).

[17] James, A., Cox, D. and Worrall, K., Microbiological and biochemical origins of human foot malodour. *Flavour Fragrance J.*, 28(4), pp.231-237. https://doi.org/10.1002/ffj.3136. (2012)

[18] Filipiak, W., Sponring, A., Filipiak, A., Baur, M., Ager, C., Wiesenhofer, H., Margesin, R., Nagl, M., Troppmair, J., Amann, A., Volatile organic compounds (VOCs) released by pathogenic microorganisms in vitro: potential breath biomarkers for early- stage diagnosis of disease. In: Amann, A., Smith, D. (Eds.), Volatile Biomarkers: Non- Invasive Diagnosis in Physiology and Medicine. Elsevier B.V., Oxford, UK, pp. 463–512. https://doi.org/10.1016/B978-0-44-462613-4.00023-4. (2013).

[19] Zareian, M., Silcock, P. and Bremer, P. Effect of medium compositions on microbially mediated volatile organic compounds release profile. *J. Appl. Bacteriol.*, 125(3), pp.813-827. https://doi.org/10.1111/jam.13908. (2018).

[20] Jenkins, C. and Bean, H., Influence of media on the differentiation of Staphylococcus spp. by volatile compounds. *J. Breath Res*, 14(1), p.016007. https://doi.org/10.1088/1752-7163/ab3e9d. (2019).

[21] Filipiak, W., Sponring, A., Baur, M., Filipiak, A., Ager, C., Wiesenhofer, H., Nagl, M., Troppmair, J. and Amann, A. Molecular analysis of volatile metabolites released specifically by staphylococcus aureus and pseudomonas aeruginosa. *BMC Microbiol.*, 12(113). https://doi.org/10.1186/1471-2180-12-113. (2012).

[22] Chen, J., Tang, J., Shi, H., Tang, C. and Zhang, R. Characteristics of volatile organic compounds produced from five pathogenic bacteria by headspace-solid phase micro-extraction/gas chromatography-mass spectrometry. *J. Basic Microbiol*, 57(3), pp.228-237. https://doi.org/10.1002/jobm.201600505. (2016).

[23] Insam, H. and Seewald, M., Volatile organic compounds (VOCs) in soils. *Biol. Fertil. Soils*, 46(3), pp.199-213. https://doi.org/10.1007/s00374-010-0442-3 . (2010).

[24] Misztal, P., Lymperopoulou, D., Adams, R., Scott, R., Lindow, S., Bruns, T., Taylor, J., Uehling, J., Bonito, G., Vilgalys, R. and Goldstein, A., Emission Factors of Microbial Volatile Organic Compounds from Environmental Bacteria and Fungi. *Environ. Sci. Technol.*, 52(15), pp.8272-8282. https://doi.org/10.1021/acs.est.8b00806. (2018).

[25] Schulz-Bohm, K., Martín-Sánchez, L. and Garbeva, P., Microbial Volatiles: Small Molecules with an Important Role in Intra- and Inter-Kingdom Interactions. *Frontiers in Microbiology*, 8. https://doi.org/10.3389/fmicb.2017.02484. (2017).

[26] Shestivska, V., Dryahina, K., Nunvář, J., Sovová, K., Elhottová, D., Nemec, A., Smith, D. and Španěl, P., 2015. Quantitative analysis of volatile metabolites released in vitro by bacteria of the genus Stenotrophomonas for identification of breath biomarkers of respiratory infection in cystic fibrosis. *J. Breath Res*, 9(2), p.027104. https://doi.org/10.1088/1752-7155/9/2/027104. (2015).

[27] Tait, E., Perry, J., Stanforth, S. and Dean, J. Identification of Volatile Organic Compounds Produced by Bacteria Using HS-SPME-GC–MS. *J. Chromatogr. Sci.*, 52(4), pp.363-373. https://doi.org/10.1093/chromsci/bmt042. (2013).

[28] Ashrafi, M., Novak-Frazer, L., Bates, M., Baguneid, M., Alonso-Rasgado, T., Xia, G., Rautemaa-Richardson, R. and Bayat, A. Validation of biofilm formation on human skin wound models and demonstration of clinically translatable bacteria-specific volatile signatures. *Sci. Rep.*, 8(1). https://doi.org/10.1038/s41598-018-27504-z. (2018).

Pre-examination copy
Submitted 27th May 2022

[29] Ashrafi, M., Novak-Frazer, L., Morris, J., Baguneid, M., Rautemaa-Richardson, R. and Bayat, A. Electrical stimulation disrupts biofilms in a human wound model and reveals the potential for monitoring treatment response with volatile biomarkers. *Wound Repair and Regeneration*, 27(1), pp.5-18. https://doi.org/10.1111/wrr.12679. (2018).

[30] Zhu, J., Bean, H., Kuo, Y. and Hill, J., Fast Detection of Volatile Organic Compounds from Bacterial Cultures by Secondary Electrospray Ionization-Mass Spectrometry. *J. Clin. Microbiol.*, 49(2), pp.769-769. https://doi.org/ 10.1128/JCM.00392-10. (2011).

[31] Bunge, M., Araghipour, N., Mikoviny, T., Dunkl, J., Schnitzhofer, R., Hansel, A., Schinner, F., Wisthaler, A., Margesin, R. and Märk, T., On-Line Monitoring of Microbial Volatile Metabolites by Proton Transfer Reaction-Mass Spectrometry. *Appl. Environ. Microbiol.*, 74(7), pp.2179-2186. https://doi.org/10.1128/AEM.02069-07 . (2008).

[32] Lemfack, M. C., Nickel, J., Dunkel, M., Preissner, R. & Piechulla, B. mVOC: a database of microbial volatiles. *Nucleic Acids Res.* **42**, D744–8. https://doi.org/10.1093/nar/gkt1250. (2014).

[33] Lemfack, M., Gohlke, B., Toguem, S., Preissner, S., Piechulla, B. and Preissner, R., mVOC 2.0: a database of microbial volatiles. *Nucleic Acids Res.* , 46(D1), pp.D1261-D1265.https://doi.org/10.1093/nar/gkx1016. (2017).

[34] Armstrong, D., Boulton, A. and Bus, S. Diabetic Foot Ulcers and Their Recurrence. *N. Engl. J. Med.* , 376(24), pp.2367-2375. https://doi.org/10.1056/NEJMra1615439. (2017).

[35] Ndosi, M., Wright-Hughes, A., Brown, S., Backhouse, M., Lipsky, B., Bhogal, M., Reynolds, C., Vowden, P., Jude, E., Nixon, J. and Nelson, E. . Prognosis of the infected diabetic foot ulcer: a 12-month prospective observational study. *Diabetic Med.*, 35(1), pp.78-88. https://doi.org/10.1111/dme.13537. (2017).

[36] Kalan, L., Meisel, J., Loesche, M., Horwinski, J., Soaita, I., Chen, X., Uberoi, A., Gardner, S. and Grice, E., 2019. Strain- and Species-Level Variation in the Microbiome of Diabetic Wounds Is Associated with Clinical Outcomes and Therapeutic Efficacy. *Cell Host Microbe*, 25(5), pp.641-655.e5. https://doi.org/10.1016/j.chom.2019.03.006. (2019).

[37] Gardner, S., Hillis, S., Heilmann, K., Segre, J. and Grice, E., The Neuropathic Diabetic Foot Ulcer Microbiome Is Associated With Clinical Factors. *Diabetes*, 62(3), pp.923-930. https://doi.org/10.2337/db12-0771. (2012).

[38] Abdulrazak, A., Ibrahim Bitar, Z., Ayesh Al-Shamali, A. and Ahmed Mobasher, L., Bacteriological study of diabetic foot infections. *Journal of Diabetes and its Complications*, 19(3), pp.138-141. https://doi.org/10.1016/j.jdiacomp.2004.06.001 (2005).

[39] Wenig, P. and Odermatt, J. OpenChrom: a cross-platform open source software for the mass spectrometric analysis of chromatographic data. *BMC Bioinf.* , 11(1). https://doi.org/10.1186/1471-2105-11-405. (2010).

[40] A van den Berg, R., CJ Hoefsloot, H., A Westerhuis, J., K Smilde, A. and J van der Werf, M. Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics*, 7(142). https://doi.org/10.1186/1471-2164-7-142. (2006).

[41] Wittek, P. *Unsupervised Learning In Quantum Machine Learning,*. 1st ed. Elsevier Science, pp.60-61.(2014).

[42] Ringnér, M. What is principal component analysis?. *Nat. Biotechnol.*, 26(3), pp.303-304. https://doi.org/10.1038/nbt0308-303. (2008).

[43] Tortara, G., Funke, B. and Case, C. Microbiology. 13th ed. Pearson Education UK, pp. 172 - 173 (2018).

[44] Rolfe, M., Rice, C., Lucchini, S., Pin, C., Thompson, A., Cameron, A., Alston, M., Stringer, M., Betts, R., Baranyi, J., Peck, M. and Hinton, J. (2011). Lag Phase Is a Distinct Growth Phase That Prepares Bacteria for Exponential Growth and Involves Transient Metal Accumulation. *J. Bacteriol,* . 194(3), pp.686-701. https://doi.org/10.1128/JB.06112-11. (2011).

[45] Thorn, R. and Greenman, J. Microbial volatile compounds in health and disease conditions. *J. Breath Res.*, 6(2), p.024001. https://doi.org/10.1088/1752-7155/6/2/024001. (2012).

[46] Prado Barragán, L., Figueroa, J., Rodríguez Durán, L., Aguilar González, C. and Hennigs, C. Fermentative Production Methods in *Biotransformation of Agricultural Waste and By-Products*. Elsevier Science, pp.189-217. https://doi.org/10.1016/B978-0-12-803622-8.00007-0. (2016).

[47] Lemfack, M., Ravella, S., Lorenz, N., Kai, M., Jung, K., Schulz, S. and Piechulla, B., Novel volatiles of skin-borne bacteria inhibit the growth of Gram-positive bacteria and affect quorum-sensing controlled phenotypes of Gram-negative bacteria. *Syst. Appl. Microbiol.*, 39(8), pp.503-515. https://doi.org/10.1016/j.syapm.2016.08.008. (2016).

[48] Ara, K., Hama, M., Akiba, S., Koike, K., Okisaka, K., Hagura, T., Kamiya, T. and Tomita, F., Foot odor due to microbial metabolism and its control. *Can. J. Microbiol.*, 52(4), pp.357-364. https://doi.org/10.1139/w05-130 (2006).

[49] Zeng, X., Leyden, J., Lawley, H., Sawano, K., Nohara, I. and Preti, G., Analysis of characteristic odors from human male axillae. *J. Chem. Ecol.*, 17(7), pp.1469-1492. https://doi.org/10.1007/BF00983777. (1991).

[50] Xiao, Z. and Xu, P. Acetoin Metabolism in Bacteria. *Crit. Rev. Microbiol.*, 33(2), pp.127-140. https://doi.org/10.1080/10408410701364604. (2007).

[51] Sánchez Mainar, M., Stavropoulou, D. and Leroy, F. Exploring the metabolic heterogeneity of coagulase-negative staphylococci to improve the quality and safety of fermented meats: a review. *Int. J. Food Microbiol.* , 247, pp.24-37. https://doi.org/10.1016/j.ijfoodmicro.2016.05.021 . (2017).

[52] Sivakanesan, R. and Dawes, E. Anaerobic Glucose and Serine Metabolism in Staphylococcus epidermidis. *Microbiology*, 118(1), pp.143-157. https://doi.org/10.1099/00221287-118-1-143. (1980).

[53] Bean, H., Rees, C. and Hill, J., 2016. Comparative analysis of the volatile metabolomes of Pseudomonas aeruginosa clinical isolates. *J. Breath Res.*, 10(4), p.047102. https://doi.org/10.1088/1752-7155/10/4/047102. (2016).

[54] Zechman, J. and Labows Jr., J. Volatiles of Pseudomonas aeruginosa and related species by automated headspace concentration – gas chromatography. *Can. J. Microbiol.,* 31(3), pp.232-237. (1985).

[55] Shestivska, V., Španěl, P., Dryahina, K., Sovová, K., Smith, D., Musílek, M. and Nemec, A., Variability in the concentrations of volatile metabolites emitted by genotypically different strains of Pseudomonas aeruginosa. *J. Appl. Microbiol.*, 113(3), pp.701-713. https://doi.org/10.1111/j.1365-2672.2012.05370.x. (2012).

[56] Preti, G., Thaler, E., Hanson, C., Troy, M., Eades, J. and Gelperin, A., Volatile compounds characteristic of sinus-related bacteria and infected sinus mucus: Analysis by solid-phase microextraction and gas chromatography–mass spectrometry. *Journal of Chromatography B*, 877(22), pp.2011-2018. https://doi.org/10.1016/j.jchromb.2009.05.028. (2009).

[57] Han, K., Lim, H. and Hong, J. Acetic Acid Formation in Escherichia coli Fermentation. *Biotechnol. Bioeng.*, 39, pp.663-671. (1991).

[58] McKenna, R. and Nielsen, D. Styrene biosynthesis from glucose by engineered E. coli. *Metab. Eng.*, 13(5), pp.544-554. https://doi.org/10.1016/j.ymben.2011.06.005. (2011)

[59] Kalscheuer, R. Microdiesel: Escherichia coli engineered for fuel production. *Microbiology*, 152(9), pp.2529-2536. https://doi.org/10.1099/mic.0.29028-0. (2006).

[60] Berstad, A., Raa, J. and Valeur, J. Indole the scent of a healthy 'inner soil'. *Microb. Ecol. Health Dis.*, 26. https://doi.org/10.3402/mehd.v26.27997. (2015)

Pre-examination copy
Submitted 27th May 2022

[61] Hang, Y., Chingin, K., Liang, J., Wang, X. and Hu, L., Fast detection of volatile organic compounds from Staphylococcal blood cultures by CDI-MS. *RSC Adv.* , 7(40), pp.24789-24794. https://doi.org/10.1039/C7RA01815K. (2017).

[62] Senecal, A., Magnone, J., Yeomans, W. and Powers, E., Rapid detection of pathogenic bacteria by volatile organic compound (VOC) analysis. *Chemical and Biological Early Warning Monitoring for Water, Food, and Ground*, Proceedings volume 4575. https://doi.org/10.1117/12.456915. (2002).

[63] Lee, J. and Lee, J. Indole as an intercellular signal in microbial communities. *FEMS Microbiol. Rev.*, 34(4), pp.426-444. https://doi.org/10.1111/j.1574-6976.2009.00204.x. (2010).

[64] Thorn, R., Reynolds, D. and Greenman, J., .Multivariate analysis of bacterial volatile compound profiles for discrimination between selected species and strains in vitro. *J. Microbiol. Methods*, 84(2), pp.258-264. https://doi.org/10.1016/j.mimet.2010.12.001 . (2011).

[65] Chippendale, T., Gilchrist, F., Španěl, P., Alcock, A., Lenney, W. and Smith, D., Quantification by SIFT-MS of volatile compounds emitted by in vitro cultures of S. aureus, S. pneumoniae and H. influenzae isolated from patients with respiratory diseases. *Anal. Methods*, 6(8), p.2460. https://doi.org/10.1039/C4AY00209A . (2014)

[66] Schmidt, R., Cordovez, V., de Boer, W., Raaijmakers, J. and Garbeva, P. Volatile affairs in microbial interactions. *The ISME Journal*, 9(11), pp.2329-2335. https://doi.org/10.1038/ismej.2015.42. (2015).

Pre-examination copy
Submitted 27th May 2022

Chapter 3: An investigation of stability and species and strain-level specificity in bacterial volatilomes

Declaration of Authorship

Candidates are required to submit a separate **Declaration of Authorship** form for each co-authored paper submitted for examination as part of a PhD by Publication thesis. Further information is available from the accompanying guideline document.[2]

Section 1: Candidate's details	
Candidate's Name	SHANE FITZGERALD
DCU Student Number	14482908
School	School of Chemical Sciences
Principal Supervisor	Aoife Morrin
Title of PhD by Publication Thesis	Microbial Volatile Organic Compounds – A Path Towards Rapid Non-Invasive Detection of Wound Infections
Section 2: Paper details	
Title of co-authored paper included in the thesis under examination	An Investigation of Stability and Species and Strain-Level Specificity in Bacterial Volatilomes.
Publication Status	Published
ISSN and link to URL (where available)	https://doi.org/10.3389/fmicb.2021.693075

This paper is one of [3] co-authored papers to be submitted as part of the PhD by publication thesis submitted for examination

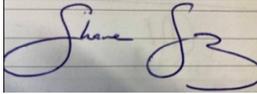| Section 3: Candidate's contribution to the paper |
|---|
| Provide details below of the nature and extent of your contribution to the paper (include both your intellectual and practical contributions) and your overall contribution in percentage terms: |
| I conceptualised and planned the experiments and I carried out the experimental investigation, data analysis, data visualisation, and writing of manuscripts. |

Where a paper has joint or multiple authors, list the names of all other authors who contributed to the work (this can be appended in a separate document, where necessary):

Dr. Aoife Morrin

Dr. Linda Holland

Section 4: Signature and Validation

I confirm that the following statements are true:

the information I have provided in this form is correct

this paper is based on research undertaken during my candidature at DCU

Signature of PhD Candidate: Date: 24th May 2022

I confirm that the information provided by the candidate is correct:

Signature of Principal Supervisor: Date: 25th May 2022

In some cases, it may be appropriate for verification to be given by both the principal supervisor and the lead/corresponding author of the work (where the lead/corresponding author of the work is not the candidate or the principal supervisor):

Signature of Lead/Corresponding Author: Date: 25th May 2022

Home institution: Dublin City University

Pre-examination copy
Submitted 27th May 2022

Abstract

Microbial volatilomics is a rapidly growing field of study and has shown great potential for applications in food, farming, and clinical sectors in the future. Due to the varying experimental methods and growth conditions employed in microbial volatilomic studies as well as strain-dependent volatilomic differences, there is limited knowledge regarding the stability of microbial volatilomes. Consequently, cross-study comparisons and validation of results and data can be challenging. In this study, we investigated the stability of the volatilomes of multiple strains of *Staphylococcus aureus, Pseudomonas aeruginosa* and *Escherichia coli* across three frequently used nutrient-rich growth media. Volatilomic stability was assessed based on media-, time- and strain-dependent variation across the examined bacterial volatilomes. Strain-level specificity of the observed volatilomes of *E. coli* and *P. aeruginosa* strains was further investigated by comparing the emission of selected compounds at varying stages of cell growth. Headspace solid phase microextraction (HS-SPME) sampling coupled with gas chromatography mass spectrometry (GC-MS) was used to analyse the volatilome of each strain. The whole volatilomes of the examined strains demonstrate a high degree of stability across the three examined growth media. At the compound-level, media dependent differences were observed particularly when comparing the volatilomes obtained in glucose-containing brain heart infusion (BHI) and tryptone soy broth (TSB) growth media with the volatilomes obtained in glucose-free Lysogeny broth (LB) media. These glucose-dependent volatilomic differences were primarily seen in the emission of primary metabolites such as alcohols, ketones, and acids. Strain-level differences in the emission of specific compounds in *E. coli* and *P. aeruginosa* samples were also observed across the media. These strain-level volatilomic differences were also observed across varying phases of growth of each strain, therefore confirming that these strains had varying core and accessory volatilomes. Our results demonstrate that, at the species-level, the examined bacteria have a core volatilome that exhibits a high-degree of stability across frequently-used growth media. Media-dependent differences in microbial volatilomes offer valuable insights into identifying the cellular origin of individual metabolites. The observed differences in the core and accessory volatilomes of the examined strains illustrate the complexity of microbial volatilomics as a study while also highlighting the need for more strain-level investigations to ultimately elucidate the whole volatilomic capabilities of microbial species in the future.

Pre-examination copy
Submitted 27th May 2022

3.1. Introduction

For many years the occurrence of disease-specific volatiles have been used as a supporting factor in the clinical diagnoses of various disorders, e.g. the sulfide emission in the breath of *Helicobacter pylori*-positive patients with gastrointestinal issues[1]; and the sweaty feet odour of patients with isovaleric acidemia[2]. As a result, the study of volatile organic compounds (VOCs) produced by commensal and pathogenic microorganisms has emerged as a path to characterising these disease-specific volatiles. In particular, the last 15 years has seen a significant rise in the study of microbial VOCs due to the universal implementation of improved analytical methodology and data analysis techniques. Comprehensive sampling and analytical methods have broadened the spectrum of compounds that can be investigated while the incorporation of dimension reduction and clustering methods has enabled the identification of discriminatory trends across the microbial VOC data. Most studies have been primarily focused on the investigation of *in vitro* microbial cultures and have been critical in identifying metabolic and cellular pathways of particular compounds. These studies have demonstrated that species-level differences in VOC production do exist between pathogenic and commensal microbial species and highlight the need for further study.

The diversity and mechanisms behind microbial volatilomes have been recently illustrated in several comprehensive review papers[3][4][5][6] and books[7]. Growth parameters such as growth media[8][9], growth phase[10][11], oxygen content [12], temperature[13] and pH[14], all influence microbial volatilomes. Another less studied factor in the overall variation seen across the volatilomes of microbial species is the occurrence of strain-level specificity in volatilomic emission within a given species[15][3]. In addition to this, the variation in sampling (SPME[16][17], thermal desorption tubes[10][18], direct syringe[19][20]) and analytical techniques (GC-MS[17][10][17][16], selected ion-flow-tube (SIFT) MS[21], proton transfer reaction (PTR) MS[22])) employed across the field also have a direct influence on the VOC profiles reported in the literature. Consequently, cross-study validation of reported microbial VOC profiles remains a major challenge in the field. However, the recent establishment of the mVOC 2.0 database[23] has allowed for some qualitative comparisons of microbial volatilomes and will evolve to be a valuable tool in the field. The database contains thousands of logged compounds along with growth conditions and analytical methods used to acquire the volatilomes of a wide range of microbes. With such a platform available, a community-wide effort is required to build on it and to ensure that the microbial volatilomic profiles available on the database are as comprehensive as possible with respect to the literature. This will ultimately allow the full examination of individual microbial volatilomic profiles relative to all the conditions in which they have been previously examined. Therefore, in order to elucidate the full spectrum of microbial volatilomes, there is a strong need to investigate them under varying conditions in controlled settings.

S. aureus, P. aeruginosa, and *E. coli* are highly prevalent wound pathogens and are responsible for particularly severe infections in diabetic foot ulcers (DFUs)[24][25] . It is estimated that around one in four people with diabetes will develop a diabetic foot ulcer (DFU) in their lifetime[26]. Infections of DFUs highly increase the risk of poor outcomes such as amputation[27]. The duration of the DFU is proportional to its severity and is closely associated with species- and strain-level microbial diversity within the infection[27][28]. Currently in clinics, time consuming techniques such as blood tests and traditional plate-based techniques are

71

employed to detect potential infections[29]. As early detection is critical in preventing severe infections, rapid non-invasive detection of pathogenic bacterial volatiles in wounds and wound samples could potentially speed up the turnover of clinical information and greatly contribute to the clinical workflow. Our group is currently working on detecting pathogen-specific volatile compounds in DFU swab samples. Volatilomic profiling of pure microbial cultures has played a critical role in our preparation for clinical volatilomic work and also for the interpretation of the data obtained.

The stability of bacterial volatilomes both in different nutritional conditions and strain-to-strain remains relatively understudied. In this study, we examined multiple strains of *E. coli, P. aeruginosa,* and *S. aureus* volatilomes across different growth media. By examining microbial volatilomic variability across different strains and different media, the core and accessory volatilomes of these bacteria can be elucidated. These terms were introduced by Bean et al.[30] but for the context of this study, core compounds refer to compounds that are emitted by both strains across all media; accessory compounds are compounds emitted by at least one strain in at least one medium. Our key objectives of this work were 1) to obtain comprehensive volatilomic data for multiple strains of *S. aureus, P. aeruginosa,* and *E. coli* in three different growth media (BHI, LB, TSB); 2) to assess the stability and variation of the observed bacterial volatilomes; and 3) to temporally investigate strain-level specificity within the selected volatilomes by comparing the emission of specific compounds at progressive stages of growth and development of the cells.

3.2.   Methods

3.2.1. Growth of Bacteria

The following bacterial strains were examined: *S. aureus* (DSM2569 and DSM799)*; P. aeruginosa* (DSM19880 and DSM25642)*; E. coli* (DSM30083 and DSM105372). All *S. aureus, P. aeruginosa,* and *E. coli* isolates were obtained from Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures GmbH. Each strain was streaked individually on tryptone soy broth (TSB) agar media plates. For each replicate, a single colony was inoculated in 4 mL of TSB, BHI, or LB broth and incubated at 37°C overnight. Each replicate overnight culture was individually incubated in a 50 mL conical centrifuge tube. Each overnight culture was diluted to a total volume of 5 mL in growth media (BHI,TSB,LB,NB) to a cell count of approximately 10^8-10^9 colony forming units (CFU)/ mL in the 20 mL headspace vials which were then sealed with magnetic Polytetrafluoroethylene /silicone septum screw caps (Merck, Cork, Ireland). For each of the examined growth media, five samples of each strain were incubated at 37°C and shaking for 24h – after which point the headspace (HS) of each sample was directly sampled and analysed (described below)

The samples are referred throughout the text using the following acronyms*: **EC.A:** E. coli DSM103372, **EC.B:** E. coli DSM30083, **PA.A**: P. aeruginosa DSM105372, **PA.B**: P. aeruginosa DSM25642, **SA.A:** S. aureus DSM2569, **SA.B:** S. aureus DSM799, **TSB**: Tryptone soy broth (OXOID : CM0129) , **BHI**: Brain heart infusion (OXOID : CM1135), **LB**: Lysogeny broth (SIGMA : L3022; NaCl 5g/L), and NB: Nutrient broth (OXOID : CM0001) .*

3.2.2. Growth curve analysis

Growth curve analysis was performed on *P. aeruginosa* and *E. coli* samples (n = 3)*.* Bacterial samples were diluted to an initial OD$_{600}$ of 0.1 which corresponded to a cell count of 10^8-10^9

72

cfu/ml. Prior to each round of Solid phase microextraction (SPME) sampling, the OD_{600} of each sample was measured by extracting 20 μL from the culture using a stainless steel needle and syringe. This was done by piercing the needle through the septum of the HS vial and tilting the vial to extract the small volume of culture. OD_{600} was measured at 1, 2, 3, 4, 5, 6, 7, 8, 24 h.

3.2.3. VOC sampling Procedure

SPME fibers were used for sampling VOCs and consisted of 85 μm Carboxen/Polydimethylsiloxane Stableflex (2 cm) assemblies (Supelco Corp., Bellefonte, PA, USA). Prior to sampling, each bacterial or control sample was removed from the shaking incubator and placed in a standard incubator at 37°C. The SPME needle was pierced through the septum of the HS vial, and the fibre was exposed to the HS of the sample for 20 min while agitated. Following this, the fibre was retracted and the SPME assembly removed from the vial. The SPME fibre was then inserted into the GC inlet and thermally desorbed at 250°C for 2 min for subsequent separation and detection by mass spectrometry. During the temporal analysis, the magnetic screw caps of each sample were also tightly covered with parafilm following each round of sampling to minimise any loss of VOCs.

Background subtraction was carried out by sampling blank headspace vials, and blank media samples. Compounds recovered from these blank analyses were individually assessed. Compounds recovered samples with signal-to-noise ratios greater than 3:1 were considered for inclusion in the study. Compounds recovered from blank analyses are summarised in Table 3.1.

3.2.4. Gas chromatography–mass spectrometry (GC-MS)

An Agilent 6890 GC connected to an Agilent 5973 mass selective detector (Agilent Technologies, Inc., Santa Clara, CA, USA) was used for all analyses. Separations were performed on a DB-WAX column (Agilent Technologies Ireland, Cork) (30 m × 0.25 mm × 0.32 μm). The carrier gas used was helium, with a constant flow rate of 1.3 mL /min For manual injections of SPME fibers, the system was equipped with a SPME Merlin Microseal (Merlin Instrument Company, Newark, DE, USA), and the inlet was maintained at a temperature of 250°C. Split-less injection was used for all samples, with a gas purge being activated after 2 min. Each SPME fibre was desorbed for 2 min within a SPME inlet liner (Supelco). The initial GC oven temperature was 40°C for 5 min and was programmed to increase at a rate of 10°C min^{-1} to 240°C, with a final hold for 5 min at this temperature, giving an overall running time of 29 min. The transfer line temperature was set at 230°C. The MS was operated at a scan range of 35-400 m/z, scan rate of 3.94 s^{-1}, ion source temperature 230°C and ionising energy of 70 eV. Identification of compounds relied on a three phase protocol whereby National Institute of Standards and Technology (NIST) library (2017) - match factors of >70% were initially used to assess potential ID matches; fragmentation patterns of potential matches were then manually interpreted before being validated using retention index matching. Retention index (RI) values for polar columns provided by the NIST Chemistry WebBook, SRD 69, was used to support the identification of these compounds. Any compound found to have an RI value ≤12 RI units of the RI values found in the NIST database were deemed acceptable matches. A external standard mixture of saturated alkanes (C7-C30; Merck, Cork, Ireland) was injected into the GC-MS under the same temperature conditions as the samples and used for

RI matching. This was done by rapidly dipping an exhausted SPME needle into the mixture once and injecting it into the GC-MS. A fully functional SPME fiber was not used for this because exposure to hexane degrades the fiber integrity.

3.2.5. Data Analysis

Agilent MassHunter Qualitative Analysis 10.0 software was used to analyse raw chromatographic data. Peak acquisition and the respective peak area data were calculated by employing the chromatogram deconvolution compound mining algorithm. Chromatographic peaks were compared using the NIST Chemistry WebBook. Peaks found to be from exogenous sources such as the SPME fiber, glass vial, and column were removed from the dataset. Only peaks that could be accurately identified and that were detected in over one replicate sample were included in the final peak list. R (version 1.2.5033) was used for data exploration and visualisation. Raw bacterial VOC data was standardised using centering and scaling [31]. Centering converts all the values in the dataset to fluctuations around zero rather than fluctuations around the mean VOC abundance. It adjusts for differences in the offset between low and high abundances. Scaling converts the values in the dataset into ratios relative to the difference in abundances between the VOCs, which allows each VOC to be equally represented in the subsequent data analysis. For compounds that were present in some replicate samples (of a given strain in a given media) and absent from others, these missing values were imputed as zero. For compounds that were absent from all replicates (of a given strain in a given media), these missing values remained missing values. Hierarchical clustering and principal component analysis (PCA) were carried out on the dataset using the R packages: 'FactoMineR' (version: 2.4), 'factoextra' (version: 1.0.7), 'pheatmap'(version: 1.0.12), 'egg' (version: 0.4.5) and 'cluster' (version:2.1.0). For the hierarchical clustering analysis, Euclidean distance was used as the measure of (dis)similarity. Other R packages used for the graphics in this study were: 'tidyverse' (version:1.3.1), 'ggplot2' (version: 3.3.5), 'ggfortify' (version:0.4.12) .

3.3. Results

3.3.1. Stability of core VOC profile across nutrient-rich media

The PCA scores plots shown in Figure 3.1 and the heatmaps shown in Figure S3.1-3.3 visualise the similarities and dissimilarities between the bacterial samples across three different nutrient-rich growth media. Across the three examined growth media, a total of 64 compounds were used to investigate the overall discrimination of observed VOC profiles at the species-level. For the unsupervised analyses, whole bacterial volatilomes were analysed based on 55 compounds in BHI (Figure 3.1 top left); 57 compounds in LB (Figure 1 top right); and 49 compounds in TSB (Figure 3.1 bottom left). In Figure 3.1, each sample is colour coded based on its respective species. Similar to our previous results[17], the variation between *S. aureus* and *P. aeruginosa* samples was summarised by PC2 (y-axis), while the variation between *E. coli* and the other bacteria was summarised by PC1 (x-axis). *S. aureus* samples appear to have the most stable volatilome across the three media as they are tightly clustered together in the bottom left corner of the plot. *P. aeruginosa* has a slightly higher degree of media-dependent distribution of samples as it can be seen that the VOC profiles of samples cultured in TSB appear to be less variable than that of samples cultured in BHI and LB. In contrast to this, in *E. coli* samples, a relatively high degree of variability was observed in

samples grown in LB compared to that of samples grown in BHI and TSB. This sample-level stability and variability was quantitated using Euclidean distances and plotted as matrices (Figure S3.4-3.6) to clearly illustrate the sample-, strain- and species- level volatilomic differences. Hierarchical clustering coupled with heatmaps were also employed to analyse the similarities across the whole volatilomes of these bacterial samples. These plots are available in the Supplementary Information (SI; Figures S3.1-3.3). In these heatmaps, samples were clustered based off Euclidean distance (dissimilarity). In these figures, to illustrate what compounds were responsible for the clustering of the bacterial samples, across the different media, hierarchical clustering was also performed on the compound abundances. Across the three media, the bacterial samples were generally successfully clustered to their respective species. There were some exceptions: EC.B_TSB_E in TSB and PA.A_LB_E in LB were incorrectly clustered; and some *E. coli* samples in LB formed a secondary E. coli cluster, this can also be seen in the PCA plot (Figure 3.1). These volatilomic differences between the E. coli samples in LB can be clearly seen in the heatmap shown in Figure S3.3 and appear to be due to differences in the emission of accessory compounds. The results shown in the PCA plots (Figure 3.1), hierarchical clustering heatmaps (Figure S3.1-3.3) clearly demonstrate that same compounds were responsible for the discrimination of the examined bacterial volatilomes across the growth media.
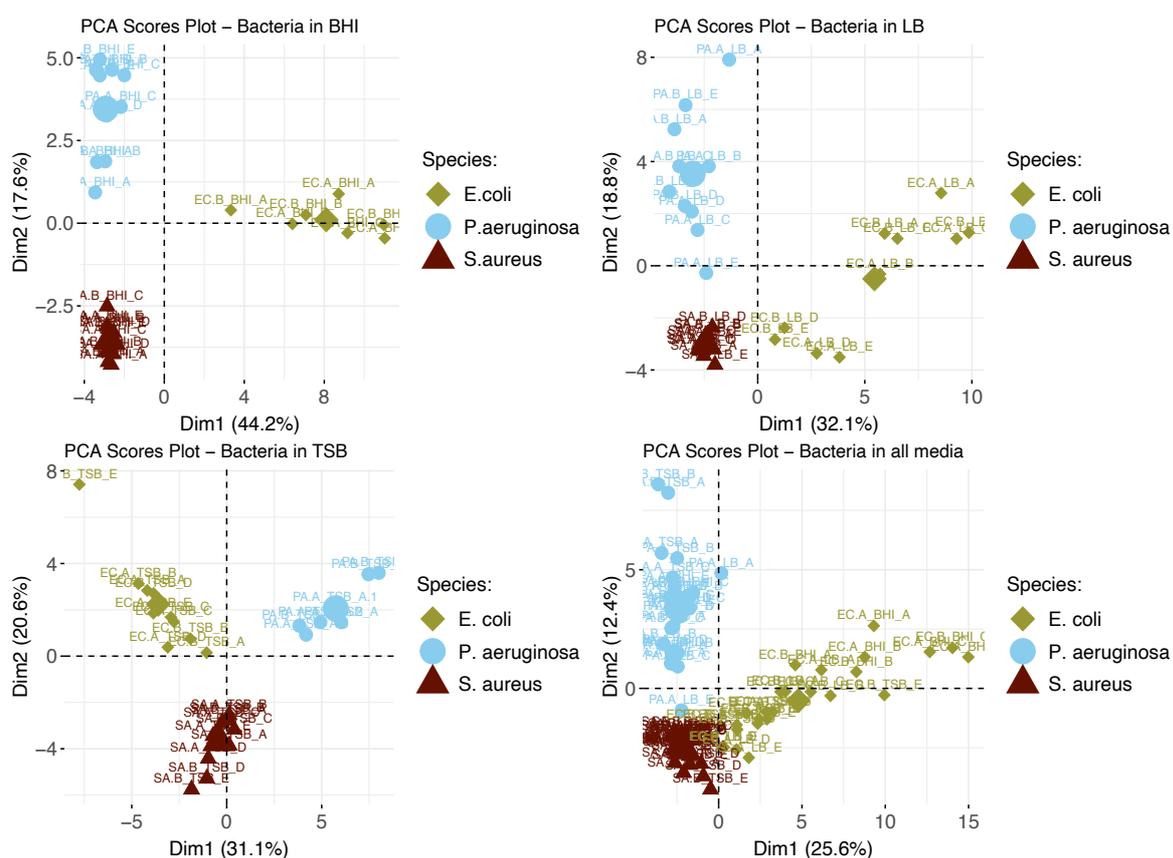

*Figure 3.1: Labelled scores plot representations of scores from the PCA analyses of bacterial samples in BHI (top left), LB (top right), TSB (bottom left), and all examined media (bottom right). The large symbols in each plot are the geometric means for each species. The corresponding strain names to the abbreviated titles of the bacterial samples shown in this plot are as follows: **EC.A:** E. coli DSM103372, **EC.B:** E. coli DSM30083, **PA.A**: P. aeruginosa*

*DSM105372, **PA.B**: P. aeruginosa DSM25642, **SA.A:** S. aureus DSM2569, **SA.B:** S. aureus DSM799*

3.3.2. Chemical composition of bacterial volatilomes

The bar plots shown in Figure 3.2 illustrate the difference in abundance of each chemical class in BHI, TSB and LB media for the species. Across each of the examined species it can be seen that for the majority of chemical classes, the lowest abundances of compounds were detected in LB media. The stacked percentage bar plots shown in the bottom row of Figure 3.2 illustrate the media-dependent chemical composition differences at both the species- and strain-level. Across the three nutrient-rich growth media, overall, there were not major variations in the chemical composition of the bacterial VOC profiles, and in this regard, were considered relatively stable. The results do suggest that for the three media examined, the bacterial volatilomes were species-dependent rather than media-dependent. Additional information about each individual compound identified can be found in the boxplots provided in the SI (Figures S3.7-3.12).
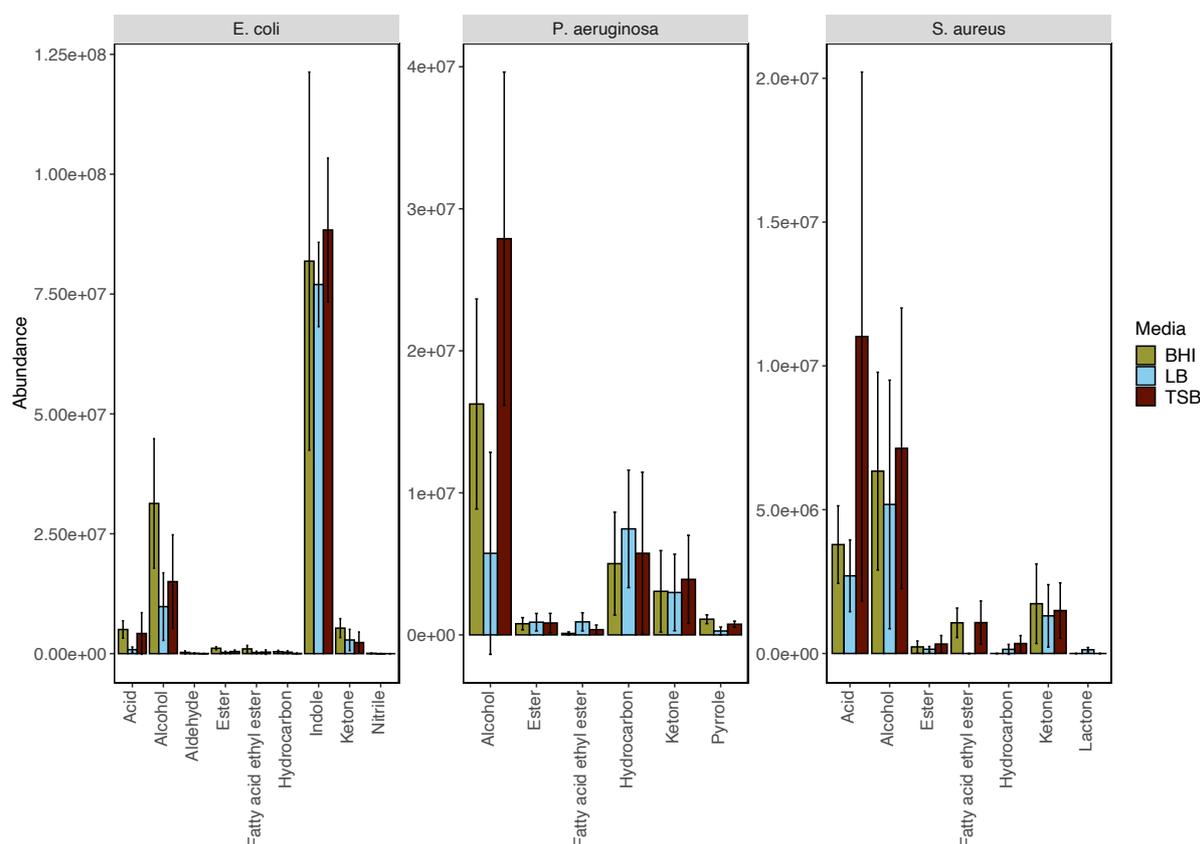

Figure 3.2: Grouped bar plot illustrating the differences in emission of individual chemical classes in BHI, LB and TSB growth media by E. coli, P. aeruginosa and S. aureus. This bar plot was obtained by summing the mean abundance of each chemical class detected in each of the examined bacteria. Significant media-dependent differences are illustrated in the corresponding grouped boxplot (Figure S3.14).
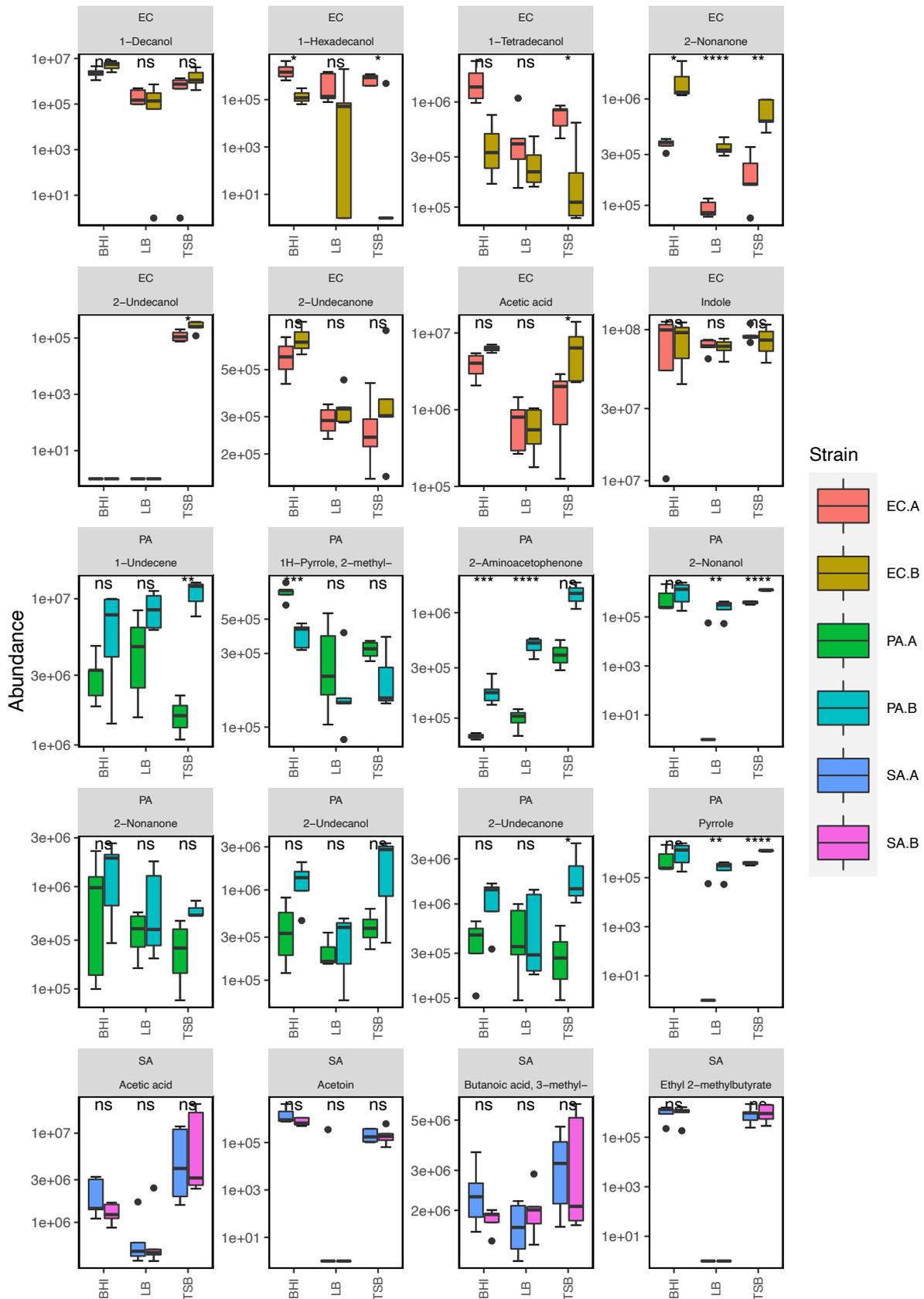
76

*Figure 3.3: Box plot representation of selected compounds emitted by S. aureus, P. aeruginosa, and E. coli strains in BHI (n=5), TSB (n=5) , and LB media (n = 5). *Five replicates were analysed for each strain in each media except for E. coli (EC.A and EC.B) in BHI (n = 4)*

77

E. coli produced the highest number of VOCs in all media and was highly active metabolically in nutrient-rich environments as it emitted a diverse volatilome in all media. Across all of the media was mainly characterised by the heavy emission of indole, which was up to 500-fold more abundant than all other compounds in the volatilome. The grouped bar plots (Figure 3.2) and box plots (Figure 3.3) show that this hyper-emission was relatively uniform across the different media and confirm that this compound is an essential byproduct of *E. coli* metabolism. Other major chemical classes emitted by the *E. coli* strains were alcohols, acids, and ketones. These chemical classes did vary across the media as the abundance of 1-alcohols was relatively lower in LB compared to the BHI and TSB. Interestingly, in the majority of compounds, the emission of various ketones was slightly higher in LB samples compared to the BHI and TSB samples. This could be due to the higher dependence on the fatty acid metabolic pathway for energy rather than the primary metabolism of glucose that gives rise to a wide variety of acids, alcohols, and fatty acid ethyl esters. In contrast to *S. aureus*, the lack of glucose in the LB medium had a significant influence (Figure S3.14) on the acid profile of the *E. coli* samples as the abundance of acetic acid saw up to a 12-fold reduction compared to samples cultured in TSB and BHI (Figure 3.3).

The primary chemical classes recovered in the *P. aeruginosa* strains were alcohols, ketones, hydrocarbons and pyrroles. Across the three media, *P. aeruginosa* emitted a variety of alcohols including 2-nonanol, 2-undecanol and 3-methyl-1-butanol, which were among the most abundant compounds produced (Figure 3.3). High abundances of alkene were emitted by both strains across all of the examined media highlighting it as one of the integral components of the *P. aeruginosa* volatilome. A significant reduction (Figure S3.14) in alcohol abundance was observed in the samples cultured in LB medium. Higher abundances of 1-undecene were detected in *P. aeruginosa* LB samples. The chemical composition of the *P. aeruginosa* VOC profile in BHI and TSB is dominated by alcohols (Figure 3.2), this was the most radical media-induced shift that was observed. A possible explanation for this is that due to the lack of glucose in LB medium, fatty acid metabolic pathways were alternatively utilised to give rise to a relatively higher abundance of compounds such as 1-undecene. Pyrrole-like compounds were emitted by both strains in all media, with the highest abundances being emitted by the samples cultured in BHI medium (Figure 3.2/Figure 3.3). The characteristic amine-containing ketone, 2-aminoacetophenone, was also emitted to varying degrees by both strains of *P. aeruginosa* (Figure 3.3). These results demonstrate the dual nature of the *P. aeruginosa* volatilome, as across different media, it exhibits stability on a qualitative level while exhibiting high variation on a quantitative level.

Across the examined media, 80% of the chemical composition of the observed *S. aureus* VOC profiles were acids and alcohols (Figure 3.2). Key compounds within these chemical groups were 3-methylbutyric acid, acetic acid, acetoin and 3-methyl-1-butanol (Figure 3.3). The

relationship between acids and alcohols did however vary across the three media, this can be seen in the grouped bar plots (Figure 3.2). In LB medium, overall acid abundance was reduced, for example, we observed a 3-fold reduction in acetic acid abundance between LB and BHI, and a 10-fold reduction between LB and TSB. Similarly, the emission of the majority of chemical classes by *S. aureus* samples was lowest in LB medium, primarily due to the lack of available glucose in the media. The influence of glucose on *S. aureus* volatilomes was also illustrated through the low abundances of key acids emitted in glucose-free nutrient broth (NB) (Nutrient Broth - Figure S3.15). Less abundant chemical classes such as fatty acid ethyl esters, lactones, hydrocarbons and aldehydes demonstrated high variation across the media. In LB medium, low abundances of the closely associated compounds 1,4-butanediol and butyrolactone were recovered from both *S. aureus* strains - these compounds were not detected in BHI or TSB *S. aureus* samples. Conversely, butanoic acid, 2-methyl-, ethyl ester (ethyl 2-methylbutyrate) was emitted in relatively high abundances in BHI and TSB but wasn't emitted in LB medium (Figure 3.3). These results further illustrate the significant influence that glucose has on the *S. aureus* volatilome. Additional information about all of the compounds identified across all the strains can be accessed in the SI (Figure S3.8 – S3.12).

### 3.3.3. Strain-dependent differences in VOC emission

Across the different growth media, we observed measurable differences in the emission of particular compounds between the strains of *P. aeruginosa* and *E. coli*. In Figure 3.3, compounds emitted by *E. coli* strains that had strain-dependent variation included 1-hexadecanol ($p = 0.005$ in TSB, $p = 0.05$ in BHI, $p = 0.34$ in LB), 1-tetradecanol ($p = 0.003$ in TSB, $p = 0.01$ in BHI, $p = 0.14$ in LB, acetic acid ($p = 0.02$ in TSB, $p = 0.04$ in BHI, $p = 0.31$ in LB, and 2-nonanone ($p = 0.002$ in TSB, $p = 0.05$ in BHI, $p = 0.0001$ in LB. In *P. aeruginosa* samples, the abundances of 1-undecene ($p = 0.01$ in TSB, $p = 0.05$ in BHI, $p = 0.02$ in LB), 2-aminoacetophenone ($p = 0.008$ in TSB, $p = 0.0003$ in BHI, $p = 0.000005$ in LB, 2-nonanol ($p = 0.0003$ in TSB, $p = 0.193$ in BHI, $p = 0.008$ in LB, 2-nonanone ($p = 0.01$ in TSB, $p = 0.18$ in BHI, $p = 0.10$ in LB, 2-undecanone ($p = 0.09$ in TSB, $p = 0.01$ in BHI, $p = 0.33$ in LB, and 2-undecanol ($p = 0.10$ in TSB, $p = 0.06$ in BHI, $p = 0.33$ in LB showed a variety of differences between the two strains. The differences were consistent across the samples from each examined media indicating that these emission differences were due to strain-level specificity in VOC emission. To further investigate this, we analysed the two *E. coli* and *P. aeruginosa* strains (n=3) individually at progressive points in their growth in TSB medium to map the emission kinetic profile of these compounds and to ultimately determine if these strain-specific differences were consistent at varying time points. Growth curves for each strain were also constructed based on OD_{600} measured from the same samples (Figure 3.4). Of the aforementioned compounds, clear kinetic differences were observed in the compounds shown in Figure 3.4. The volatilomes of the two *S. aureus* strains were highly stable with respect to each other and therefore will not be discussed in this section.

In both *E. coli* strains, hyperproliferation of cells was observed between 0 – 4 h which corresponded to a proportionate emission of acetic acid, alcohols, and indole (Figure 3.4). The growth rate of EC.B cells was faster than that of EC.A. Of the alcohols emitted, 1-hexadecanol and 1-tetradecanol were only emitted by EC.A (Figure 3.4). Although these compounds were previously recovered from some of the EC.B samples in the previous investigation (Figure 3.3), they were emitted in relatively low abundances and there was

variation in their occurrence sample-to-sample. Conversely, 2-nonanone was emitted in significantly higher abundances by EC.B than EC.A across the three growth media (p = 0.002 in TSB; p = 0.049 in BHI; p = 0.0001 in LB) (Figure S3.4). This strain-dependent difference was further confirmed by the kinetic profiles of both strains (Figure 3.4), as 2-nonanone was not recovered from any EC.A samples while being emitted proportionately with the growth of EC.B cells (1 – 4 h). Cell growth of both strains stagnated for approximately 4 h (between 4 – 8 h), this stagnation was reflected by an overall reduction in the emission of VOCs from both strains between 5 – 6 h. From 8 – 24 h, cell numbers of both *E. coli* strains steadily grew again, however, abundances of compounds such as 1-alcohols collectively declined to varying degrees in both strains. Significantly high abundances of indole were emitted by both strains and correlated with the incubation time and growth of the cells.

In contrast to the growth of *E. coli* cells, both *P. aeruginosa* strains demonstrated a slower increase over the first 8 h of incubation. There was a marked difference in the volatilome activity of both strains (Figure 3.4). In agreement with the multi-media results shown in Figure 3.3, the kinetic plots shown in Figure 3.4 demonstrate that PA.B was metabolically more active than PA.A. High sample-to-sample variance in the occurrence of 2-undecanol, 2-undecanone, 2-nonanol, and 2-aminoacetophenone in PA.A samples across the media (Figure 3.3) and kinetically (Figure 3.4) indicated that these compounds were irregular accessory compounds to the PA.A volatilome. In contrast to this, 2-aminoacetophenone was a correlative marker of progressive phases of cell growth in PA.B samples, whereas 2-nonanol, 2-undecanol, and 2-undecanone marked the latter phase of PA.B growth as they were emitted at some point between 8 – 24 h. Although 3-methy-1-butanol was an abundant correlative growth marker of both *P. aeruginosa* strains, it was emitted at a 3-fold higher abundance in PA.B samples between 8 – 24 h (p = 0.05 at 24 h). Clear differences in the emission of 1-undecene (p = 0.01 in TSB; p = 0.05 in BHI; p = 0.02 in LB) across the three media (Figure 3.3) were observed between the two *P. aeruginosa* strains. This was also observed in the kinetic experiments as in the first 8 h of incubation, the abundances of 1-undecene recovered from PA.B were consistently 10-fold higher than the abundances recovered from PA.A (p = 0.009 at 4 h, p = 0.009 at 5 h, p = 0.004 at 6 h, p =0.02 at 7 h). After 8 h, in PA.B samples, 1-undecene abundances sharply decreased.
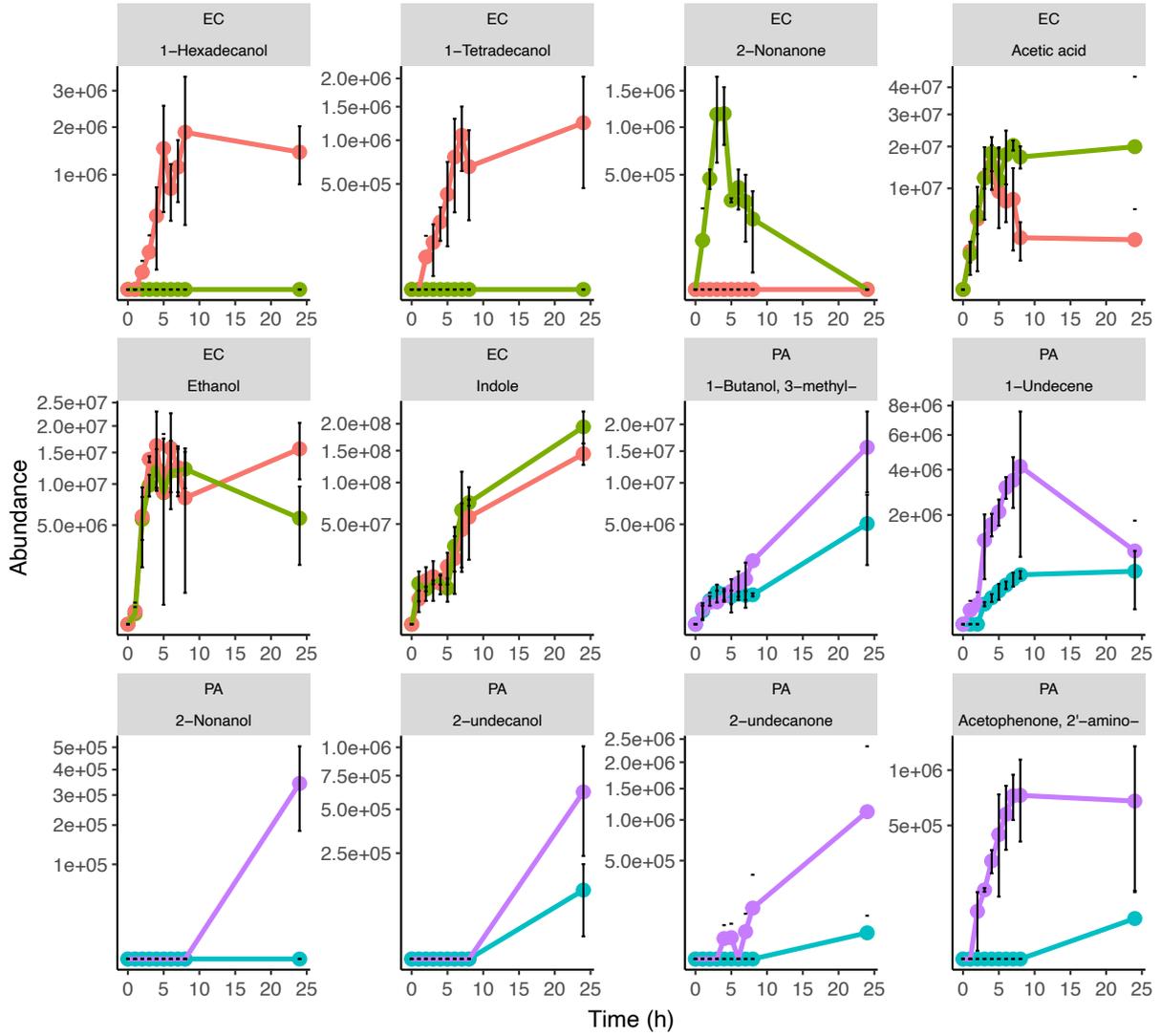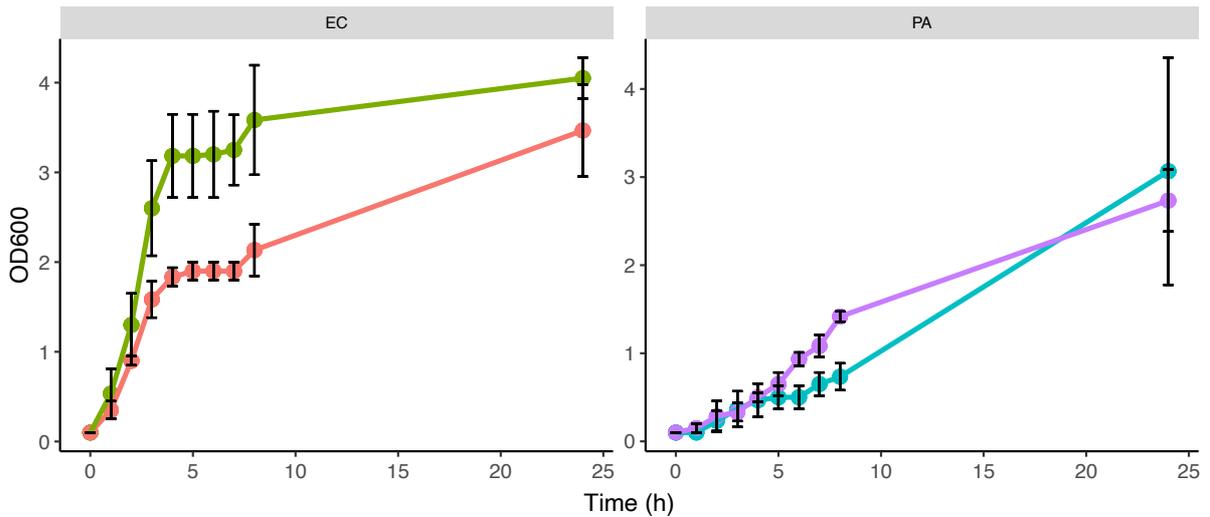
Figure 3.4: Kinetic profiles of strain-dependent emission from EC.A, EC.B, PA.A, and PA.B (for all examined strains, n=3) for selected compounds sampled at specific phases of growth of bacterial samples in TSB growth media. Error bars represent the standard deviation around

81

the mean abundance values. Y-axis labels on all VOC kinetic plots are scaled by square root. Corresponding cell growth curves based on OD$_{600}$ measurements measured from each replicate at the respective sampling time. The corresponding strain names to the abbreviated titles of the bacterial samples shown in this plot are as follows: **EC.A:** *E. coli DSM103372,* **EC.B:** *E. coli DSM30083,* **PA.A**: *P. aeruginosa DSM105372,* **PA.B**: *P. aeruginosa DSM25642.*

3.4.   Discussion

Microorganisms produce a diverse range of volatile metabolites that have different physico-chemical properties and biological activities. These volatile metabolites serve an important role in inter-species and inter-kingdom communication and are involved in both beneficial and deleterious interactions between microorganisms[1].   Microbial VOCs have been proposed as potential biomarkers of disease[4] that can be non-invasively analysed to support clinical workflows in the future. Clinical studies involving untargeted breath profiling analyses have demonstrated the potential discriminatory power and diagnostic potential of VOCs for diseases such as pneumonia[32], cystic fibrosis[33][34], tuberculosis[35], and COVID-19[36]. It has also been well-established that microbial diversity at the species-level is associated with poor outcomes and longer durations of diabetic ulcer wounds[27]. A recent study also reported that diversity at the strain-level of specific pathogenic species is a large contributor to infection severity[28]. Consequently, the development of a non-invasive rapid sampling platform potentially capable of discriminating microbes at the species- and strain-level is highly desirable in clinic settings. In the conclusion of our previous study[17], we outlined that our future work will involve the volatilomic analysis of wound samples to identify infection-specific markers. Early unpublished results of this work suggests that volatile markers detected in this study such as 3-methylbutyric acid, acetic acid, 3-methyl-1-butanol, propanoic acid and ethanol do heavily persist in samples obtained from severe wound infections. However, this work is currently ongoing and a higher number of samples are required from a varying spectrum of infected wounds to further support these early findings.

Volatilomic profiling of pure microbial cultures has played a critical role in identifying the cellular origins of metabolites associated with specific species and strains of pathogens. In this study we employed HS-SPME-GC-MS to investigate the stability of the volatilome of multiple strains of prominent wound-associated pathogens across different nutrient-rich growth media. Oxygen availability is expected to have a strong influence on volatile compound formation by bacteria as it dictates whether respiration or fermentation pathways are utilised for metabolism[3]. Each bacterial sample was made up to a total volume of 5 mL, leaving a 15 mL headspace above the culture. It is predicted that the bacteria utilised the available oxygen in this headspace to carry out aerobic respiration in the early stages of growth. The examined pathogens are capable of anaerobic and/or fermentation under specific conditions, and may be capable of shifting their metabolism from aerobic respiration in oxygen-exhausted conditions[41]. It was expected that this occurred during the 24 h incubation and consequently the volatilomes reported in this study are the result of both aerobic and anaerobic metabolism. Strain-level differences that were observed in the examined bacterial volatilomes were further investigated by analysing and comparing the volatilome of each strain at progressive phases of cell growth and development.

82

Dimension reduction methods such as principal component analysis (PCA)[17][11][37] and clustering methods such as hierarchical clustering (HC)[17][30] are frequently employed to identify species-specific trends across the bacterial volatilomic samples. PCA identifies species- and strain-specific volatilomic differences across the bacterial volatilomes and amplifies these differences by constructing new linear variables called principal components (PCs), along which the variation is maximal. The PCs can then be visualised using scores plots. Scores plots show inter-sample distances that clearly illustrate patterns across the data which can be used to identify groups that characterise the overall dataset[38]. In this study we used these techniques to perform an unsupervised analysis of the volatiomes obtained from the examined strains across each growth media. The PCA scores plots shown in Figure 3.1 illustrate that across the media the examined bacterial volatilomes demonstrated similar differentiation from each other which allowed clear and consistent species-level discrimination. These scores plots also indicate a certain degree of stability of the whole volatilomes of the examined species across the media. Similar volatilomic stability across different media was previously demonstrated by Dryahina et al.[39] – in this study, PCA analysis of SIFT-MS volatilomic data showed that *S. aureus* and *P. aeruginosa* samples cultured in BHI, NB and Mueller Hinton Broth (MHB) exhibited species-dependent clustering rather than media-dependent clustering. Hierarchical clustering is another common statistical method used to classify multiple samples into groups (clusters). The results are visualised as dendrograms. The length of an edge in a dendrogram between a cluster and its split is proportional to the dissimilarity (Euclidean distance) between the split clusters[40]. Dendrograms were used for two purposes in this study, to investigate the volatilomic similarities between the examined bacterial strains, and to determine what compounds were responsible for the discrimination of the bacterial volatilomes across the different media. Heatmaps coupled with dendrograms complement the PCA results (Figure 3.1) and demonstrate clear discrimination of the bacterial species in each of the examined media (Figure S3.1-3.3). The heatmaps also offer minor insights at the strain-level differences that are present in the volatilomes as the samples from *P. aeruginosa* (Figure S3.1 & S3.2) are roughly clustered together. In each of the examined media, it appeared that the same compounds were responsible for the discrimination of the bacterial volatilomes (Figure S3.1-3.3). Following the analyses, the chemical classes and compounds responsible for the observed media-, species- and strain-dependent differences were further investigated.

Although the chemical compositions of each bacterial species and strain demonstrated a high degree of stability (Figure 3.2 – bottom row), the abundances of chemical classes varied across BHI, TSB and LB growth medium (Figure 3.2). The volatilomes of both *E. coli* strains were dominated primarily by the emission of indole and fatty alcohols (collectively ~ 85%), with the remaining 15% of the volatilome being mostly acids and ketones (Figure 3.2). Indole is produced in a one-step reaction by the enzymatic catalysation of the amino acid tryptophan[41]. Growth media free of glucose such as LB medium contain an abundance of amino acids[42] that can be alternatively metabolised to form indole which explains why indole abundances were relatively uniform across the media (Figure 3.2). *E. coli* primarily produces fatty alcohols via the fatty acid metabolic pathways[43], however, fatty alcohols can also be derived from glucose metabolism[44]. A variety of fatty alcohols were identified at varying abundances across the examined media (Figure 3.3 and Figure S3.5) between the two *E. coli* strains. Despite EC.B cells having a higher growth rate, EC.A emitted significantly higher

83

abundances of 1-tetradecanol and 1-hexadecanol than EC.B across the media (Figure 3.3) and temporally (Figure 3.4). This supports the fact that the emission of these compounds is not just simply dependent on the progressive growth of cells, but rather the different metabolic pathways each strain utilises as the cells proliferate. Fatty alcohols ($C_3 - C_{16}$) have been frequently reported[31][11][17][45] to be emitted by *E. coli* with varying combinations being emitted strain-to-strain. Acetic acid abundances were high in both strains (Figure 3.3) and demonstrated a proportionate increase with cell proliferation in the first 5 h of incubation (Figure 3.4).

P. aeruginosa is a common gram-negative bacteria that has been associated with severe infections of burns[46] and diabetic foot ulcers[27], and it has also been previously labelled as the most common cause of persistent, fatal respiratory infections in cystic fibrosis patients[47][48]. The chemical composition of the observed *P. aeruginosa* volatilomes was primarily made up of alcohols and hydrocarbons (~75%) across the media (Figure 3.2). This appears to be in agreement with recent results published by Davis et al.[48], although there are differences in the *P. aeruginosa* volatilomes at the compound-level. The remaining 25% of the volatilomes were composed of ketones, pyrroles and fatty acid esters (Figure 3.2). Characteristic compounds (Figure 3.3) emitted by both strains across the three media that have also been frequently detected in *P. aeruginosa* volatilomic analyses included 2-aminoacetophenone[37.][50][49], 2-undecanol[17][10], 2-undecanone[16][17][37], 1-undecene[37][17][50][10], 2-nonanone[48][10][17], and pyrrole[37][17][10]. Production of 1-undecene by both strains of *P. aeruginosa* was enhanced when the samples were cultured in LB medium most likely due to the absence of glucose (Figure 3.2/Figure 3.3). Although significant differences in 1-undecene emission were observed between both strains across the media (Figure 3.3) and kinetically (Figure 3.4), 1-undecene was shown to be a temporal growth marker of both strains. This long chain alkene is derived from the metabolism of fatty acids and is a key component of the volatilomes of various *Pseudomonas spp*[51][52][53]. 2-aminoacetophenone is produced in the amino acid degeneration (shikimate) pathway via the loss of a hydroxyl group on anthranilic acid[3] (derived from chorismate via chorismate lyase[54]). Despite being widely reported as a characteristic marker for *P. aeruginosa*[49], we observed a high degree of strain-dependent variation in both its emission across the different media (Figure 3.3) and kinetically (Figure 3.4). Our results suggest that the amino acid degradation pathway may potentially be more active in some *P. aeruginosa* strains than others. Differences in 2-aminoacetophenone production could also be due to strain-level differences in the regulation of quorum sensing [55].

Since the emergence of antibiotic-resistant strains of *S. aureus* that caused epidemics in the 1950s and 1960s there have been large global efforts to develop early detection systems[56]. In the last 15 years, the understanding of the core and accessory components of the *S. aureus* volatilome has been steadily growing due to a growing number of *in vitro* volatilomic profiling studies[10][17][9][22][57]. The clinical value of *in vitro* volatilomic profiling has been demonstrated by Filipiak et al. who reported detection of VOCs, known to be emitted by *S. aureus*, in the breath of patients with *S. aureus*-positive respiratory infections[58]. In this study here, the *S. aureus* volatilome demonstrated the highest degree of stability across the media out of the three examined species. This was illustrated by the tight clusters of 'SA' samples in each score plot in Figure 3.1 and in the heatmaps (Figure S3.1-3.3). Across TSB, BHI and LB media, the *S.*

aureus volatilome was composed primarily of acids and alcohols (~80% total volatilome), at different ratios in each. Other chemical classes recovered in all media included ketones and fatty acid esters. The characteristic compounds that were emitted by both *S. aureus* strains were not surprising and included 3-methylbutyric acid, acetic acid, ethyl 2-methylbutyrate and acetoin – all of which have been frequently reported in the literature[9][10][11][57]. These key compounds all originate from primary metabolic pathways[3]. 3-methylbutyric acid is derived from the metabolism of amino acids (leucine) while acetic acid, ethyl 2-methylbutyrate and acetoin arise from different stages of the fermentation process in glucose metabolism[15][3]. The different primary metabolic pathways from which these compounds are derived are reflected in Figure 3.3, as in the glucose-free LB medium, acetic acid, ethyl 2-methylbutyrate and acetoin abundances are significantly lower than abundances observed in TSB and BHI while 3-methylbutyric acid emission was comparable across the media due to presence of amino acid substrates in all media. Similar reductive effects on the emission of acids and esters by *S. aureus* in glucose-free media was recently reported by Jenkins et al[9].

Comprehensive bacterial volatilomic data will only be obtained through the analysis of a high number of strains. Strain-to-strain volatilomic variability has been relatively under studied, however, studies that have been carried out have highlighted potential strain-specific differences in bacterial volatilomes. Purcaro et al[59] reported discrimination of multiple strains of *P. aeruginosa* in the breath of infected mice. In our previous study[17], when analysing triplicate samples of the same strains, we observed small differences between both *P. aeruginosa* and *E. coli* strains and concluded that further work is needed incorporating a higher number of samples to comprehensively resolve these differences. Analysing different microbial strains at specific stages of cell growth and development has been recently proposed as an effective approach to comprehensively elucidate the variation in microbial volatilomes at the strain-level[5]. In this study, we employed this approach in a more targeted manner to further investigate the strain-dependent emission variation of specific compounds – observed across multiple media - in *P. aeruginosa* and *E. coli* samples. The resulting kinetic plots shown in Figure 3.4 confirm that there was consistent strain-level specificity in the temporal emission of individual compounds. These results were in agreement with our earlier observed differences between the strains for the same compounds in Figure 3.3. The results obtained from this study also validate observed differences in VOC emission previously reported[17] between *P. aeruginosa* strains. Although the volatilomes obtained from all of the *E. coli* samples in this study are mostly in agreement with those previously reported[17], the same strain-level differences were not observed. It should also be noted, that when analysing the full spectrum of VOCs recovered from these strains in an unsupervised manner via PCA (Figure 3.1), neither *P. aeruginosa* nor *E. coli* strains were discriminated from each other. Our results show that when investigating the full volatilomes of these strains, clear discrimination of the strains will not be achieved, it is only when the volatilomes are investigated at the compound level with corresponding emission kinetic data that these strain-level differences can be fully elucidated. While our data clearly shows strain-level volatilomic differences exist within specific species of bacteria, it also highlights the complexity of strain-level volatilomic discrimination. Strain-level differences were not observed between the examined *S. aureus* strains in this study, however, volatilomic discrimination of different *S. aureus* strains (MSSA and MRSA) has been previously reported[60] indicating that strain-dependent antibiotic sensitivities could also be potential factor in strain-level volatilomic diversity. Baptista et al.

85

reported a high number of compounds and also demonstrated strain-level volatilomic differences across *S. aureus* samples based on enterotoxicity[57]. Microbial volatilomic variability across different media has been frequently reported in the literature [8][37][61][62]. This was particularly highlighted by Rees et al.[62] who detected a total of 365 compounds from 9 *Klebsiella pneumoniae* clinical isolates across TSB, LB, BHI and MHB media – only 10% of the compounds were common across all examined media. Many of the studies discussed in this text have highlighted the complexity and high degree of specificity of microbial volatilomes. Future work should focus on investigating the strain-level volatilomic diversity present within specific microbial species by analysing a higher number of strains obtained from ecologically varying environments across different media. It is through, firstly, comprehensively investigating the factors that govern the emission of pathogen-specific metabolites that will allow volatilomics to be employed in clinics in the future.

3.5. Conclusion

Bacterial volatilomes are influenced by different nutritional environments and strain-level differences. By investigating this stability in this work a comprehensive understanding of the volatilomes of the examined bacterial species was achieved. Our objectives in this study were: 1) to obtain comprehensive volatilomic data from multiple strains of the wound-associated pathogens *S. aureus, P. aeruginosa* and *E. coli* in three different growth media; 2) to assess the stability and variation of the observed bacterial volatilomes; and 3) to assess strain-level specificity within the examined strains by comparing the emission of specific compounds at progressive stages of growth and development of the cells. Using HS-SPME GC-MS analysis we successfully analysed the VOCs produced from each strain to obtain characteristic species-specific volatilomes. The observed volatilomes demonstrated a high degree of stability across the examined media, however, glucose-free media had a reductive effect on the emission of various primary metabolites. Strain-level variation was observed in *P. aeruginosa* and *E. coli* samples across the examined media in the emission of particular compounds. Comparative temporal volatilomic analysis of these strains confirmed that there were differences in the emission of individual compounds between the examined strains. Moving forward in microbial volatilomics, performing similar multi-strain kinetic experiments will provide a more comprehensive view of the capabilities of microbial volatilomes. Additionally, analysing microbial volatilomes in different growth media allows specific metabolic pathways responsible for VOC production to be intimately investigated. Building a comprehensive understanding of the limits and possibilities of microbial volatilomics will elucidate what can be ultimately achieved in future applications.

On a clinical level, these pathogens pose a significant challenge as they cause sever wound infections. We are currently working on the study of the diabetic foot ulcer volatilome and have seen that some characteristic compounds (e.g. 3-methyl-1-butanol, 3-methylbutyric acid, ethanol) also persist in wound samples. Our clinical work is still in the early stages, however, we hope to identify infection-specific volatilomic patterns that will potentially allow the detection of wound infections earlier.

Pre-examination copy
Submitted 27th May 2022

3.6. References

1 Hoshi, K., Yamano, Y., Mitsunaga, A., Shimizu, S., Kagawa, J. and Ogiuchi, H., (2002). Gastrointestinal diseases and halitosis: association of gastric Helicobacter pyloriinfection. *International Dental Journal*, 52(5), pp.207-211. https://doi.org/10.1002/j.1875-595X.2002.tb00926.x

2 Budd, M., Tanaka, K., Holmes, L., Efron, M., Crawford, J. and Isselbacher, K., (1967). Isovaleric Acidemia. *New England Journal of Medicine*, 277(7), pp.321-327. https://doi.org/10.1056/NEJM196708172770701

3 Weisskopf, L., Schulz, S. and Garbeva, P., (2021). Microbial volatile organic compounds in intra-kingdom and inter-kingdom interactions. *Nature Reviews Microbiology*,. https://doi.org/10.1038/s41579-020-00508-1

4Elmassry, M. & Piechulla, B. (2020). Volatilomes of bacterial infections in humans. *Front. Neurosci.* https://doi.org/10.3389/fnins.2020.00257

5 Kai, M., (2020). Diversity and Distribution of Volatile Secondary Metabolites Throughout Bacillus subtilis Isolates. *Frontiers in Microbiology*, 11. https://doi.org/10.3389/fmicb.2020.00559 .

6 Schulz, S. & Dickschat, J. (2007). Bacterial volatiles: The smell of small organisms. *ChemInform.* https://doi.org/10.1039/b507392h

7 Ryu, C., Weisskopf, L. and Piechulla, B., (2020). *Bacterial volatile compounds as mediators of airborne interactions*. Singapore: Springer.

8 Jenkins, C. and Bean, H., (2019). Influence of media on the differentiation of Staphylococcus spp. by volatile compounds. *J. Breath Res*, 14(1), p.016007. https://doi.org/10.1088/1752-7163/ab3e9d.

9 Jenkins, C. and Bean, H., (2020). Dependence of the Staphylococcal Volatilome Composition on Microbial Nutrition. *Metabolites*, 10(9), p.347. **https://doi.org/10.3390/metabo10090347**

10 Filipiak, W., Sponring, A., Baur, M., Filipiak, A., Ager, C., Wiesenhofer, H., Nagl, M., Troppmair, J. and Amann, A. (2012). Molecular analysis of volatile metabolites released specifically by staphylococcus aureus and pseudomonas aeruginosa. *BMC Microbiol.* , 12(113). https://doi.org/10.1186/1471-2180-12-113.

11 Chen, J., Tang, J., Shi, H., Tang, C. and Zhang, R. (2016).Characteristics of volatile organic compounds produced from five pathogenic bacteria by headspace-solid phase micro-extraction/gas chromatography-mass spectrometry. *J. Basic Microbiol*, 57(3), pp.228-237. https://doi.org/10.1002/jobm.201600505.

12 Insam, H. and Seewald, M., (2010). Volatile organic compounds (VOCs) in soils. *Biol. Fertil. Soils*, 46(3), pp.199-213. https://doi.org/10.1007/s00374-010-0442-3 .

13 Misztal, P., Lymperopoulou, D., Adams, R., Scott, R., Lindow, S., Bruns, T., Taylor, J., Uehling, J., Bonito, G., Vilgalys, R. and Goldstein, A., (2018). Emission Factors of Microbial Volatile Organic Compounds from Environmental Bacteria and Fungi. *Environ. Sci. Technol.* , 52(15), pp.8272-8282. https://doi.org/10.1021/acs.est.8b00806.

14 Schulz-Bohm, K., Martín-Sánchez, L. and Garbeva, P., (2017).Microbial Volatiles: Small Molecules with an Important Role in Intra- and Inter-Kingdom Interactions. *Frontiers in Microbiology*, 8. https://doi.org/10.3389/fmicb.2017.02484.

15 Schulz, S., Schlawis, C. , Koteska, D., Harig, T., and Biwer, P., (2020). *Chapter 3:* Structural Diversity of Bacterial Volatiles' *in Bacterial Volatile Compounds as Mediators of Airborne Interactions*. 1st ed. [S.l.]: SPRINGER VERLAG, SINGAPOR, pp. 93 - 116.

16 Timm, C., Lloyd, E., Egan, A., Mariner, R. and Karig, D., (2018). Direct Growth of Bacteria in Headspace Vials Allows for Screening of Volatiles by Gas Chromatography Mass Spectrometry. *Frontiers in Microbiology*, 9. https://doi.org/10.3389/fmicb.2018.00491

17 Fitzgerald, S., Duffy, E., Holland, L. and Morrin, A., (2020). Multi-strain volatile profiling of pathogenic and commensal cutaneous bacteria. *Scientific Reports*, 10(1). https://doi.org/10.1038/s41598-020-74909-w .

18 Filipiak, W., Sponring, A., Baur, M., Ager, C., Filipiak, A., Wiesenhofer, H., Nagl, M., Troppmair, J. and Amann, A., (2012). Characterization of volatile metabolites taken up by or released from Streptococcus pneumoniae and Haemophilus influenzae by using GC-MS. *Microbiology*, 158(12), pp.3044-3053. https://doi.org/10.1099/mic.0.062687-0

19 Ashrafi, M., Novak-Frazer, L., Bates, M., Baguneid, M., Alonso-Rasgado, T., Xia, G., Rautemaa-Richardson, R. and Bayat, A. (2018). Validation of biofilm formation on human skin wound models and demonstration of clinically translatable bacteria-specific volatile signatures. *Sci. Rep.*, 8(1). https://doi.org/10.1038/s41598-018-27504-z.

20 Ashrafi, M., Novak-Frazer, L., Morris, J., Baguneid, M., Rautemaa-Richardson, R. and Bayat, A. (2018). Electrical stimulation disrupts biofilms in a human wound model and reveals the potential for monitoring treatment response with volatile biomarkers. *Wound Repair and Regeneration*, 27(1), pp.5-18. https://doi.org/10.1111/wrr.12679.

21 Shestivska, V., Dryahina, K., Nunvář, J., Sovová, K., Elhottová, D., Nemec, A., Smith, D. and Španěl, P., (2015). Quantitative analysis of volatile metabolites released in vitro by bacteria of the genus Stenotrophomonas for identification of breath biomarkers of respiratory infection in cystic fibrosis. *J. Breath Res.*, 9(2), p.027104. https://doi.org/10.1088/1752-7155/9/2/027104 .

22 Bunge,M. *et al.* (2008). On-line monitoring of microbial volatile metabolites by proton transfer reaction-mass spectrometry. *Appl.Environ. Microbiol.* **74**(7), 2179–2186. https://doi.org/10.1128/AEM.02069-07

23 Lemfack,M. C. *et al*. (2017).mVOC 2.0: a database of microbial volatiles. *Nucleic Acids Res*. https://doi.org/10.1093/nar/gkx1016..

24 Gardner, S., Hillis, S., Heilmann, K., Segre, J. & Grice, E. (2012). The neuropathic diabetic foot ulcer microbiome is associated with clinical factors. *Diabetes* **62**(3), 923–930. https://doi.org/10.2337/db12-0771 .

25 Ndosi, M. *et al.* (2017). Prognosis of the infected diabetic foot ulcer: A 12-month prospective observational study. *Diabetic Med.* **35**(1), 78–88. https://doi.org/10.1111/dme.13537 .

26 Armstrong, D., Boulton, A. & Bus, S. (2017). Diabetic foot ulcers and their recurrence. *N. Engl. J. Med.* **376**(24), 2367–2375. https://doi.org/10.1056/NEJMra1615439 .

[27] Abdulrazak, A., Bitar, Z. I., Al-Shamali, A. A. & Mobasher, L. A. (2005). Bacteriological study of diabetic foot infections. *J. Diabetes Complications* **19**(3), 138–141. https://doi.org/10.1016/j.jdiacomp.2004.06.001

[28] Kalan, L. *et al.* (2019). Strain- and species-level variation in the microbiome of diabetic wounds is associated with clinical outcomes and therapeutic efficacy. *Cell Host Microbe* **25**(5), 641 655.e5. https://doi.org/10.1016/j.chom.2019.03.006

[29] Lipsky, B., Aragón-Sánchez, J., Diggle, M., Embil, J., Kono, S., Lavery, L., Senneville, É., Urbančič-Rovan, V., Van Asten, S., Peters, E., Abbas, Z. and Malone, M. (2019). IWGDF guidance on the diagnosis and management of foot infections in persons with diabetes. *IWGDF*.

[30] Bean, H., Rees, C. and Hill, J., (2016). Comparative analysis of the volatile metabolomes of Pseudomonas aeruginosa clinical isolates *J. Breath Res.*, 10(4), p.047102. https://doi.org/10.1088/1752-7155/10/4/047102 .

[31] A van den Berg, R., CJ Hoefsloot, H., A Westerhuis, J., K Smilde, A. and J van der Werf, M., (2006). Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics*, 7(142). https://doi.org/10.1186/1471-2164-7-142.

[32] Schnabel, R. *et al.* (2015). Analysis of volatile organic compounds in exhaled breath to diagnose ventilator-associated pneumonia. *Sci.Rep.* https://doi.org/10.1038/srep17179

[33] Neerincx, A. *et al.* (2016).Detection of *Staphylococcus aureus* in cystic fibrosis patients using breath VOC profiles. *J. Breath Res.* https://doi.org/10.1088/1752-7155/10/4/046014

[34] Robroeks, C. *et al.* (2010). Metabolomics of volatile organic compounds in cystic fibrosis patients and controls. *Pediatr. Res.* **68**(1), 75–80. https://doi.org/10.1203/00006450-201011001-00143 .

[35] Zetola, N. *et al.* (2017).Diagnosis of pulmonary tuberculosis and assessment of treatment response through analyses of volatile compound patterns in exhaled breath samples. *J. Infect.* **74**(4), 367–376. https://doi.org/10.1016/j.jinf.2016.12.006

[36] Ruszkiewicz, D., Sanders, D., O'Brien, R., Hempel, F., Reed, M., Riepe, A., Baillie, J., Brodrick, E., Darnley, K., Ellerkmann, R., Mueller, O., Skarysz, A., Truss, M., Wortelmann, T., Yordanov, S., Thomas, P., Schaaf, B. and Eddleston, M., (2020). Diagnosis of COVID-19 by Analysis of Breath with Gas Chromatography-Ion Mobility Spectrometry: A Feasibility Study. *SSRN Electronic Journal*,. https://doi.org/10.1016/j.eclinm.2020.100609

[37] Tait, E., Perry, J., Stanforth, S. & Dean, J. (2013).Identification of volatile organic compounds produced by bacteria using HS-SPME-GC–MS. *J. Chromatogr. Sci.* **52**(4), 363–373. https://doi.org/10.1093/chromsci/bmt042

[38] Ringnér, M. (2008). What is principal component analysis?. *Nat. Biotechnol.* **26**(3), 303–304. https://doi.org/10.1038/nbt0308-303 .

[39] Dryahina, K., Sovová, K., Nemec, A. and Španěl, 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. *Journal of Breath Research*, 10(3), p.037102. https://doi.org/10.1088/1752-7155/10/3/037102

[40] Wittek, P. *Unsupervised Learning In Quantum Machine Learning* 1st edn, 60–61 (Elsevier Science, New York, 2014).

[41] Lemfack, M., Bahl, H., Piechulla, B. and Magnus, N., (2020). *Chapter 1: 'The Domain of Bacteria and Their Volatile Metabolic Potential' in Bacterial Volatile Compounds as Mediators of Airborne Interactions*. 1st ed. SPRINGER VERLAG, SINGAPOR, pp. 24 - 28. https://doi.org/10.1371/journal.pone.0158200

[42] Sezonov, G., Joseleau-Petit, D. and D'Ari, R., (2007). Escherichia coli Physiology in Luria-Bertani Broth. *Journal of Bacteriology*, 189(23), pp.8746-8749. https://doi.org/10.1128/JB.01368-07

[43] Liu, Y., Chen, S., Chen, J., Zhou, J., Wang, Y., Yang, M., Qi, X., Xing, J., Wang, Q. and Ma, Y., (2016). High production of fatty alcohols in Escherichia coli with fatty acid starvation. *Microbial Cell Factories*, 15(1). https://doi.org/10.1186/s12934-016-0524-5

[44] Youngquist, J., Schumacher, M., Rose, J., Raines, T., Politz, M., Copeland, M. and Pfleger, B., (2013). Production of medium chain length fatty alcohols from glucose in Escherichia coli. *Metabolic Engineering*, 20, pp.177-186. https://doi.org/10.1016/j.ymben.2013.10.006

[45] Bos, L., Sterk, P. and Schultz, M., (2013). Volatile Metabolites of Pathogens: A Systematic Review. *PLoS Pathogens*, 9(5), p.e1003311. https://doi.org/10.1371/journal.ppat.1003311

[46] Lyczak, J., Cannon, C. and Pier, G., (2000). Establishment of Pseudomonas aeruginosa infection: lessons from a versatile opportunist. *Microbes Infect.*, 2(9), pp.1051-1060. https://doi.org/10.1016/S1286-4579(00)01259-4 .

[47] Bradbury, J., (2009). Epidemic Pseudomonas aeruginosa strain sequenced. *The Lancet Infectious Diseases*, 9(1), p.14. https://doi.org/10.1016/S1473-3099(08)70298-6

[48] Davis, T., Karanjia, A., Bhebhe, C., West, S., Richardson, M. and Bean, H., (2020). Pseudomonas aeruginosa Volatilome Characteristics and Adaptations in Chronic Cystic Fibrosis Lung Infections. *mSphere*, 5(5). https://doi.org/10.1128/mSphere.00843-20

[49] Scott-Thomas, A., Syhre, M., Pattemore, P., Epton, M., Laing, R., Pearson, J. and Chambers, S., (2010). 2-Aminoacetophenone as a potential breath biomarker for Pseudomonas aeruginosa in the cystic fibrosis lung. *BMC Pulmonary Medicine*, 10(1). https://doi.org/10.1186/1471-2466-10-56

[50] Shestivska, V., Španěl, P., Dryahina, K., Sovová, K., Smith, D., Musílek, M. and Nemec, A., (2012). Variability in the concentrations of volatile metabolites emitted by genotypically different strains of Pseudomonas aeruginosa. *J. Appl. Microbiol.*, 113(3), pp.701-713. https://doi.org/10.1111/j.1365-2672.2012.05370.x

[51] Dandurishvili, N., Toklikishvili, N., Ovadis, M., Eliashvili, P., Giorgobiani, N., Keshelava, R., Tediashvili, M., Vainstein, A., Khmel, I., Szegedi, E. and Chernin, L., (2010). Broad-range antagonistic rhizobacteria Pseudomonas fluorescens and Serratia plymuthica suppress Agrobacterium crown gall tumours on tomato plants. *Journal of Applied Microbiology*, 110(1), pp.341-352. https://doi.org/10.1111/j.1365-2672.2010.04891.x

[52] Kanchiswamy, C., Malnoy, M. and Maffei, M., (2015). Chemical diversity of microbial volatiles and their potential for plant growth and productivity. *Frontiers in Plant Science*, 6. https://doi.org/10.3389/fpls.2015.00151

[53] Lo Cantore, P., Giorgio, A. and Iacobellis, N., (2015). Bioactivity of volatile organic compounds produced by Pseudomonas tolaasii. *Frontiers in Microbiology*, 6. https://doi.org/10.3389/fmicb.2015.01082

[54] Parthasarathy, A., Cross, P., Dobson, R., Adams, L., Savka, M. and Hudson, A., (2018). A Three-Ring Circus: Metabolism of the Three Proteogenic Aromatic Amino Acids and Their Role in the Health of Plants and Animals. *Frontiers in Molecular Biosciences*, 5. https://doi.org/10.3389/fmolb.2018.00029

[55] Kesarwani, M., Hazan, R., He, J., Que, Y., Apidianakis, Y., Lesic, B., Xiao, G., Dekimpe, V., Milot, S., Deziel, E., Lépine, F. and Rahme, L., 2011. A Quorum Sensing Regulated Small Volatile Molecule Reduces Acute Virulence and Promotes Chronic Infection Phenotypes. *PLoS Pathogens*, 7(8), p.e1002192

[56] Cookson, B., (2011). Five decades of MRSA: controversy and uncertainty continues. *The Lancet*, 378(9799), pp.1291-1292. https://doi.org/10.1016/S0140-6736(11)61566-3

[57] Baptista, I., Santos, M., Rudnitskaya, A., Saraiva, J., Almeida, A. and Rocha, S., (2019). A comprehensive look into the volatile exometabolome of enteroxic and non-enterotoxic Staphylococcus aureus strains. *The International Journal of Biochemistry & Cell Biology*, 108, pp.40-50. https://doi.org/10.1016/j.biocel.2019.01.007

[58] Filipiak, W., Beer, R., Sponring, A., Filipiak, A., Ager, C., Schiefecker, A., Lanthaler, S., Helbok, R., Nagl, M., Troppmair, J. and Amann, A., (2015). Breath analysis for in vivo detection of pathogens related to ventilator-associated pneumonia in intensive care patients: a prospective pilot study. *J. Breath Res.*, 9(1), p.016004. https://doi.org/10.1088/1752-7155/9/1/016004 .

[59] Purcaro, G., Nasir, M., Franchina, F., Rees, C., Aliyeva, M., Daphtary, N., Wargo, M., Lundblad, L. and Hill, J., (2019). Breath metabolome of mice infected with Pseudomonas aeruginosa. *Metabolomics*, 15(1). https://doi.org/10.1111/2049-632X.12107

[60] Boots, A. *et al.* (2014). Identification of microorganisms based on headspace analysis of volatile organic compounds by gas chromatography–mass spectrometry. *J. Breath Res.* https://doi.org/10.1088/1752-7155/8/2/027106.

[61] Zareian, M., Silcock, P. and Bremer, P. (2018). Effect of medium compositions on microbially mediated volatile organic compounds release profile. *J. Appl. Bacteriol.*, 125(3), pp.813-827. https://doi.org/10.1111/jam.13908.

[62] Rees, C., Nordick, K., Franchina, F., Lewis, A., Hirsch, E. and Hill, J., (2017). Volatile metabolic diversity of Klebsiella pneumoniae in nutrient-replete conditions. *Metabolomics*, 13(2). https://doi.org/10.1007/s11306-016-1161-z

Pre-examination copy
Submitted 27th May 2022

Chapter 4: Multi-Strain and -Species Investigation of Volatile Metabolites emitted from Planktonic and Biofilm *Candida* cultures

Declaration of Authorship

Candidates are required to submit a separate **Declaration of Authorship** form for each co-authored paper submitted for examination as part of a PhD by Publication thesis. Further information is available from the [accompanying guideline document](#).[3]

Section 1: Candidate's details	
Candidate's Name	SHANE FITZGERALD
DCU Student Number	14482908
School	School of Chemical Sciences
Principal Supervisor	Aoife Morrin
Title of PhD by Publication Thesis	Microbial Volatile Organic Compounds – A Path Towards Rapid Non-Invasive Detection of Wound Infections
Section 2: Paper details	
Title of co-authored paper included in the thesis under examination	Multi-Strain and -Species Investigation of Volatile Metabolites Emitted from Planktonic and Biofilm Candida Cultures.
Publication Status	Published
ISSN and link to URL (where available)	https://doi.org/10.3390/metabo12050432

This paper is one of ⌊3⌋ co-authored papers to be submitted as part of the PhD by publication thesis submitted for examination

Section 3: Candidate's contribution to the paper
Provide details below of the nature and extent of your contribution to the paper (include both your intellectual and practical contributions) and your overall contribution in percentage terms:
I conceptualised and planned the experiments and I carried out the experimental investigation, data analysis, data visualisation, and writing of manuscripts.

Pre-examination copy
Submitted 27th May 2022

Where a paper has joint or multiple authors, list the names of all other authors who contributed to the work (this can be appended in a separate document, where necessary):
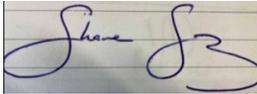
Dr. Aoife Morrin

Dr. Linda Holland

Section 4: Signature and Validation

I confirm that the following statements are true:

the information I have provided in this form is correct

this paper is based on research undertaken during my candidature at DCU

Signature of PhD Candidate: Date: 24th May 2022

I confirm that the information provided by the candidate is correct:

Signature of Principal Supervisor: Date: 25th May 2022

In some cases, it may be appropriate for verification to be given by both the principal supervisor <u>and</u> the lead/corresponding author of the work (where the lead/corresponding author of the work is not the candidate or the principal supervisor):

Signature of Lead/Corresponding Author: Date: 25th May 2022

Home institution: Dublin City University

Pre-examination copy
Submitted 27th May 2022

Abstract

Candida parapsiliosis is a prevalent neonatal pathogen that attains its virulence through its strain-specific ability to form biofilms. The use of volatilomics, the profiling of volatile metabolites from microbes is a non-invasive, simple way to identify and classify microbes and has shown great potential for pathogen identification. Although *C. parapsiliosis* is one of the most common clinical fungal pathogens, its volatilome has never been characterised. In this work, planktonic volatilomes of ten clinical strains of *C. parapsilosis* were analysed, along with a single strain of *Candida albicans.* Headspace-solid-phase microextraction coupled with gas chromatography-mass spectrometry was employed to analyse the samples. Species-, strain-, and media- influences on the fungal volatilomes were investigated. 24 unique metabolites from the examined *Candida spp.* (22 from *C. albicans*; 18 from *C. parapsilosis*) were included in this study. Chemical classes detected across the samples included alcohols, fatty acid esters, acetates, thiols, sesquiterpenes and nitrogen-containing compounds. *C. albicans* volatilomes were most clearly discriminated from *C. parapsilosis* based on the detection of unique sesquiterpene compounds. The effect of biofilm formation on the *C. parapsilosis* volatilomes was investigated for the first time by comparing volatilomes of a biofilm-positive strain and a biofilm-negative strain over time (0 – 48 hours) using a novel sampling approach. Volatilomic shifts in the profiles of alcohols, ketones, acids, and acetates were observed specifically in the biofilm-forming samples and attributed to biofilm maturation.

This work highlights species-specificity of *Candida* volatilomes and also marks the clinical potential for volatilomics for non-invasively detecting fungal pathogens. Additionally, the range of biofilm-specificity across microbial volatilomes is potentially far-reaching and therefore characterising these volatilomic changes in pathogenic fungal and bacterial biofilms could lead to novel opportunities for detecting severe infections early.

Keywords: Volatilomics; *Candida*; Biofilm; Fungi; Pathogen; Gas chromatography; Volatile metabolites; Solid phase microextraction; Mass spectrometry

Pre-examination copy
Submitted 27th May 2022

4.1. Introduction

Candida parapsilosis is an emerging pathogen that typically resides as a human commensal with limited pathogenicity. However, *C. parapsilosis* has been highlighted as a growing infectious burden due to growing prevalence of infections arising from blood and indwelling medical devices[1] , as well as its growing prevalence in neonatal sepsis[2]. *C. parapsilosis* also has a strain-specific ability to produce biofilms which determines differences in human pathogenicity across the species[3]. *Candida albicans* is a more clinically prevalent organism within the *Candida* genus. It is a polymorphic fungus that can grow either as ovoid-shaped budding yeast; as an elongated ellipsoid cell with pseudohyphae; or a parallel-walled true hyphal form – which is associated with increased virulence. This polymorphism allows *C. albicans* to infect a wide variety of host niches as it shifts its metabolism, expression of adhesions, ability to form biofilm and virulence with each morphological transition. Morphological transitions are mediated by environmental cues[4] such as temperature, pH, O_2 and CO_2 content, quorum sensing interactions with neighbouring microbes, and the emission of volatile signalling compounds[5]. As the fungus transitions from the yeast cell to the ellipsoid cell to the true hyphal form, its virulence increases – this is also reversible. In contrast to *C. albicans, C. parapsilosis* cannot form true hyphae, it instead physiologically exists as an ovoid-shaped budding yeast or pseudohyphal form[1] . As a result, *C. parapsilosis* infections are typically less diverse and less severe than *C. albicans* infections.

Microorganisms have evolved to utilise a wide variety of metabolic pathways to survive in constantly changing environments. These pathways include the metabolism of sugars, amino acids, fatty acids, sulfur- and nitrogen-containing compounds, and terpenes[6]. Volatile organic compounds (VOCs) are produced as biproducts at each stage of each respective pathway[7]. The species- and strain-specific ways in which microbes regulate their metabolism significantly contributes to the complexity of characterising their volatilomes. For this reason, in order to obtain comprehensive data for a specific microbe or microbial group, it is necessary to track species- and strain-level volatilomic diversity across the genus. Many metabolites are commonly emitted across microbial species[8], however, the whole array of compounds that are emitted, and the abundances by which they are emitted are species-specific and are collectively referred to as its volatilome. The composition of volatilomes depends on multiple factors such as nutritional substrates[9][10][8], strain-to-strain metabolic variation[11][12], growth phase of cells[13], and pH[14]. As a result, broad characterisation of microbial volatilomes is one of the major challenges of the field. Significant progress has been made in the last five years in building awareness of the field with the publishing of comprehensive reviews[6][15][16][17] and books[18]. Tackling challenges in the field will require comprehensive and standardised experimental workflows that support broad untargeted screening and identification of metabolites across a wide range of microbial species. Although there have been recent investigations of the volatilomes of several prevalent *Candida* genus[19][20], there is a need for more studies to support and validate these works. The examined *Candida* volatilomes have demonstrated low specificity in the early stages of growth but develop discriminative features in the latter stages of cellular development. The chemical composition of these volatilomes have been shown to be rich in acids, aldehydes, alcohols, hydrocarbons, esters, terpenic compounds, sulfur-containing compounds and phenols. *C. albicans, Candida tropicalis*, and *Candida glabrata* [19][20] are among the genus that have been investigated. However, despite being one of the most commonly isolated *Candida*

96

genus in clinics, volatilomic data on *C. parapsilosis* is very limited - the only study available[21]reported the detection of just 3 compounds (ethanol, 2-phenylethyl alcohol and 3-octanone).

Similar to the majority of microbes, very little is known about how biofilm formation affects the emission of volatile compounds. Volatilomic discrimination of biofilms of wound-associated pathogens was recently demonstrated using a highly effcient open flow system that was coupled with both direct mass spectrometry and headspace-solid-phase microextraction coupled with gas chromatography-mass spectrometry (HS-SPME GCMS) to allow real-time analysis of steady state biofilms[22]. Although this study demonstrated a comprehensive experimental workflow for investigating biofilms, the volatilomic emissions were compared to uninoculated media controls, therefore the specificity of VOCs emitted from the biofilm itself was not determined. We hypothesise that this specificity can be determined through the dual investigation of biofilm-positive and biofilm-negative strains of a respective microbe. We also hypothesise that biofilm development will cause kinetic shifts in the volatilome of the biofilm-forming microbe. Following the production of biofilm, bacteria and fungi slow their metabolism to regulate the use of available substrates. This allows them to survive in challenging environments and can also render them more tolerant to antimicrobial drugs and stimuli as the reduction in central metabolic flux reduces the intake of toxins[23]. Up to now, it has yet to be investigated whether this reduction in central metabolic activity during biofilm development has measurable effects on the volatilome of microbes.

In this work, we obtain comprehensive multi-strain volatilomes for planktonic *C. parapsilosis* (10 different strains) using a standardised headspace solid-phase microextraction (SPME) with GC-MS analysis workflow[11][13] and compare it to the volatilome of a *C. albicans* strain obtained under the same conditions to understand species and strain specific differences across the genus. Comprehensive volatilomic data is reported across varying growth conditions for *C. parapsilosis*, an understudied clinical pathogen. Furthermore, we also apply our workflow in a volatilomic study of *Candida* biofilms for the first time to identify metabolites emitted in biofilm-specific metabolic pathways by comparing biofilm-positive and biofilm-negative strains of *C. parapsilosis.* This final aspect of work makes a significant contribution to the study of biofilm volatilomics by demonstrating specificity and highlights new opportunities for further study in this research area.

4.2. Methods

4.2.1. Growth of *Candida* planktonic samples

The following *Candida* strains were examined: *CA*: *C. albicans* (*DSM 1386*) ; *C. parapsilosis* (***CP1**: CLIB214; **CP2**: CDC317; **CP3**: CDC173; **CP4**: 711701; **CP5**: CDC167; **CP6**: J961250; **CP7**: CDC179; **CP8**: J930733; **CP9**: 103; **CP10**: J930631/1.)* - See Table S1 for reference list. *Growth media: **YPD**: Yeast peptone dextrose ; **TSB**: Tryptone soy broth*. Each strain was streaked individually on Yeast peptone dextrose (YPD) agar plates and incubated at 30°C overnight. For each replicate, a single colony was inoculated in 4 mL of YPD or TSB broth and incubated at 37°C overnight shaking at 180 rpm. Each replicate overnight culture was individually incubated in a 50 mL conical centrifuge tube. Each overnight culture was diluted to a total volume of 5 mL in growth media (YPD or TSB) to a cell count of approximately 10^7 colony

forming units (CFU)/ mL in the 20 mL headspace vials which were then sealed with magnetic Polytetrafluoroethylene /silicone septum screw caps (Merck, Cork, Ireland). For samples cultured in YPD, samples of each CP strain (n=3) and CA samples (n=5) were incubated at 37°C and shaking for 24h – after which point the headspace (HS) of each sample was directly sampled and analysed (described below).

4.2.2. Growth of *Candida* biofilm samples

CP6 (J961250), CP1 (CLIB214) and *C. albicans* (DSM 1386) were investigated in this section of the study. For each replicate, a single colony was inoculated in 4 mL of YPD broth and incubated at 37°C overnight, shaking at 180 rpm. Each replicate culture was individually incubated in a 50 mL conical centrifuge tube overnight. The following day, each replicate was centrifuged at 4000 rpm for 4 min and washed with 1 mL PBS twice. The cells were then diluted to an OD of 1 (10^7 CFU/mL). In a 6 well plate, 2.5 mL of YPD was initially added to each well. 2.5 mL of cells was then deposited in each well to bring the total volume of each well to 5 mL (media:cells, 50:50). Note: To minimise the risk of contamination, it was important to ensure that the 6 well plate was covered in between each liquid transfer. The lid was placed on the 6 well plate and the plate was carefully placed in the sampling container ('Good For You' 850 cm^3 Borosilicate glass containers, dimensions : 19.3 x 13.6 x 6.7 cm). The lids of these containers have a snap lock with a silicone seal to enhance sterility. The lid of the 6 well plate was removed to momentarily expose the cultures (see Figure S4.3) before the lid of the sampling container was closed over the system. A layer of parafilm was wrapped around the lid of the sampling container. The container was then placed in a static incubator at 37°C for 24 hours.
Following SPME sampling of the HS of the sampling container (see next section) at 24 h, exhausted media and waste was manually removed from each well using an auto-pipette and they were washed twice with 1 mL PBS. Fresh media (5 mL) was then slowly deposited in each well and the plate was re-sealed in the sampling container and re-incubated at 37°C for 24 h. Again to avoid contamination, the system was covered in between liquid transferrals.

4.2.3. Crystal violet staining of biofilm samples

Following the HS sampling of the 6-well plate samples, the waste media was removed from each well before being washed with 1 mL PBS (each well) twice. The plates were then left overnight to dry at room temperature. 1 mL of 0.4% crystal violet was added to each well and left for 15 min. The crystal violet was then removed and the wells were washed with 1 mL PBS three times. The plates were then left to dry at room temperature.

4.2.4. HS-SPME sampling

SPME fibers were used for sampling VOCs and consisted of 85 μm Carboxen/Polydimethylsiloxane Stableflex (2 cm) assemblies (Supelco Corp., Bellefonte, PA, USA). For the planktonic cultures, the SPME needle was pierced through the septum of the HS vial, and the fibre was exposed to the HS of the sample for 20 min while agitated. Following this, the fibre was retracted and the SPME assembly removed from the vial. The SPME fibre was then inserted into the GC inlet and thermally desorbed at 250°C for 2 min for subsequent separation and detection by mass spectrometry. During the temporal analysis, the magnetic screw caps of each sample were also tightly covered with parafilm following each round of sampling to minimise any loss of VOCs. For the biofilm samples, a septum was fixed to the lid

of the sampling container with multiple layers of tape (see Figure S4.3). The SPME fiber was pierced through this septum and exposed to the headspace of sampling container containing the 6-well plate samples for 30 min while static in an incubator at 37°C .

Background subtraction was carried out by sampling blank headspace vials, and blank media samples. Compounds recovered from these blank analyses were individually assessed. Compounds recovered samples with signal-to-noise ratios greater than 3:1 were considered for inclusion in the study. Compounds recovered from blank analyses are summarised in Figure 4.3.

4.2.5. Gas chromatography–mass spectrometry

An Agilent 6890 GC connected to an Agilent 5973 mass selective detector (Agilent Technologies, Inc., Santa Clara, CA, USA) was used for all analyses. Separations were performed on a DB-WAX column (Agilent Technologies Ireland, Cork) (30 m × 0.25 mm × 0.32 µm). The carrier gas used was helium, with a constant flow rate of 1.3 mL /min For manual injections of SPME fibers, the system was equipped with a SPME Merlin Microseal (Merlin Instrument Company, Newark, DE, USA), and the inlet was maintained at a temperature of 250°C. Split-less injection was used for all samples, with a gas purge being activated after 2 min. Each SPME fibre was desorbed for 2 min within a SPME inlet liner (Supelco). The initial GC oven temperature was 40°C for 5 min and was programmed to increase at a rate of 10°C min^{-1} to 240°C, with a final hold for 5 min at this temperature, giving an overall running time of 29 min. The transfer line temperature was set at 230°C. The MS was operated at a scan range of 35-400 *m/z*, scan rate of 3.94 s^{-1}, ion source temperature 230°C and ionising energy of 70 eV. Identification of compounds relied on a three phase protocol whereby National Institute of Standards and Technology (NIST) library (2017) - match factors of >70% were initially used to assess potential ID matches; fragmentation patterns of potential matches were then manually interpreted before being validated using retention index matching. Retention index (RI) values for polar columns provided by the NIST Chemistry WebBook, SRD 69, was used to support the identification of these compounds. Any compound found to have an RI value ≤12 RI units of the RI values found in the NIST database were deemed acceptable matches. See Table S2 for the chromatographic retention and mass spectral validation of each compound. An external standard mixture of saturated alkanes (C$_7$-C$_{30}$; Merck, Cork, Ireland) was injected into the GC-MS under the same temperature conditions as the samples and used for RI matching. This was done by rapidly dipping an exhausted SPME needle into the mixture once and injecting it into the GC-MS. A fully functional SPME fiber was not used for this because exposure to hexane degrades the fiber integrity.

4.2.6. Data Analysis

Agilent MassHunter Qualitative Analysis 10.0 software was used to analyse raw chromatographic data. Peak acquisition and the respective peak area data were calculated by employing the chromatogram deconvolution compound mining algorithm. Chromatographic peaks were compared using the NIST Chemistry WebBook. Peaks found to be from exogenous sources such as the SPME fiber, glass vial, and column were removed from the dataset. Only peaks that could be accurately identified and that were detected in more than one replicate sample were included in the final peak list. R (version 1.2.5033) was used for data exploration and visualisation. Raw VOC data was standardised using scaling [24]. Scaling converts the

values in the dataset into ratios relative to the difference in abundances between the VOCs, which allows each VOC to be equally represented in the subsequent data analysis. For compounds that were present in some replicate samples (of a given strain in a given media) and absent from others, these missing values were imputed as zero. For compounds that were absent from all replicates (of a given strain in a given media), these missing values remained missing values. Hierarchical clustering and principal component analysis (PCA) were carried out on the datasets using the R packages: 'pheatmap'(version: 1.0.12), 'egg' (version: 0.4.5) and 'cluster' (version:2.1.0). For the hierarchical clustering analysis, Euclidean distance was used as the measure of (dis)similarity. Hierarchical clustering is a bottom-up unsupervised learning method that characterises samples within a dataset based on their similarity to each other. Each sample initially represents its own cluster and is then subsequently clustered to similar clusters until all of the samples fall under one large cluster, this large complex cluster is called a dendrogram. The individual arms of the dendrogram represent the clusters and the length of each arm represents the Euclidean distance or dissimilarity between the samples. Therefore, the longer the arm of one cluster, the more dissimilar it is from the rest of the samples. Autoscaling of the abundance values was employed as the normalisation technique for this analysis. Autoscaling normalises the abundance values for each emitted compound with respect to their occurrence across the fungal samples. Although it is an effective normalisation method, it is limited by amplifying variances in the data due to inflation of extremely low and high abundance values.

Other R packages used for the graphics in this study were: 'tidyverse' (version:1.3.1), 'ggplot2' (version: 3.3.5), 'ggfortify' (version:0.4.12) . Statistical analyses were facilitated by the R package 'ggpubr' , version: 0.4.0). Mean comparison p-values were calculated using the Wilcoxon test.

4.3. Results

4.3.1. Discriminative volatilomics of planktonic *Candida* spp. at the species- and strain-level

The heatmap shown in Figure 4.1 clearly illustrates the species-level discrimination of the volatilomes of one strain of *C. albicans* and 10 clinical strains of *C. parapsilosis* cultured in YPD media. The columns of the heatmap represent the mean abundance values of the single strain of *C. albicans* (n=5) and the 10 *C. parapsilosis* clinical strains (n=3). Following background subtraction of the whole sample volatilomes, a total of 25 unique compounds were recovered from the broad volatilomic screening of the *C. albicans* and *C. parapsilosis* isolates . Out of these 25 compounds, 23 were emitted by *C. albicans*, and 18 were emitted by *C. parapsilosis*. Therefore on a qualitative level, the volatilomes of these two fungal pathogens were highly similar. However, there were large differences in compound abundances recovered from *C. albicans* and *C. parapsilosis* as for most compounds, *C. albicans* emitted higher abundances than *C. parapsilosis*. This occurrence coupled with the fact that *C. albicans* emitted a greater number of compounds than *C. parapsilosis* would indicate that it is potentially more metabolically active than *C. parapsilosis*. In Figure 4.1, minor strain-level variation can be observed as CP6, CP4 and CP9 are not clustered with the other CP strains. These strains are discriminated from the other strains as they emit relatively higher abundances of several compounds. Overlayed and individual chromatograms for *C. parapsilosis* and *C. albicans* volatilomes are shown in Figure S4.1 and S4.2. The species-level and strain-level diversity of the *Candida* spp. volatilome will be further discussed with respect to the individual unique compounds in the next section.
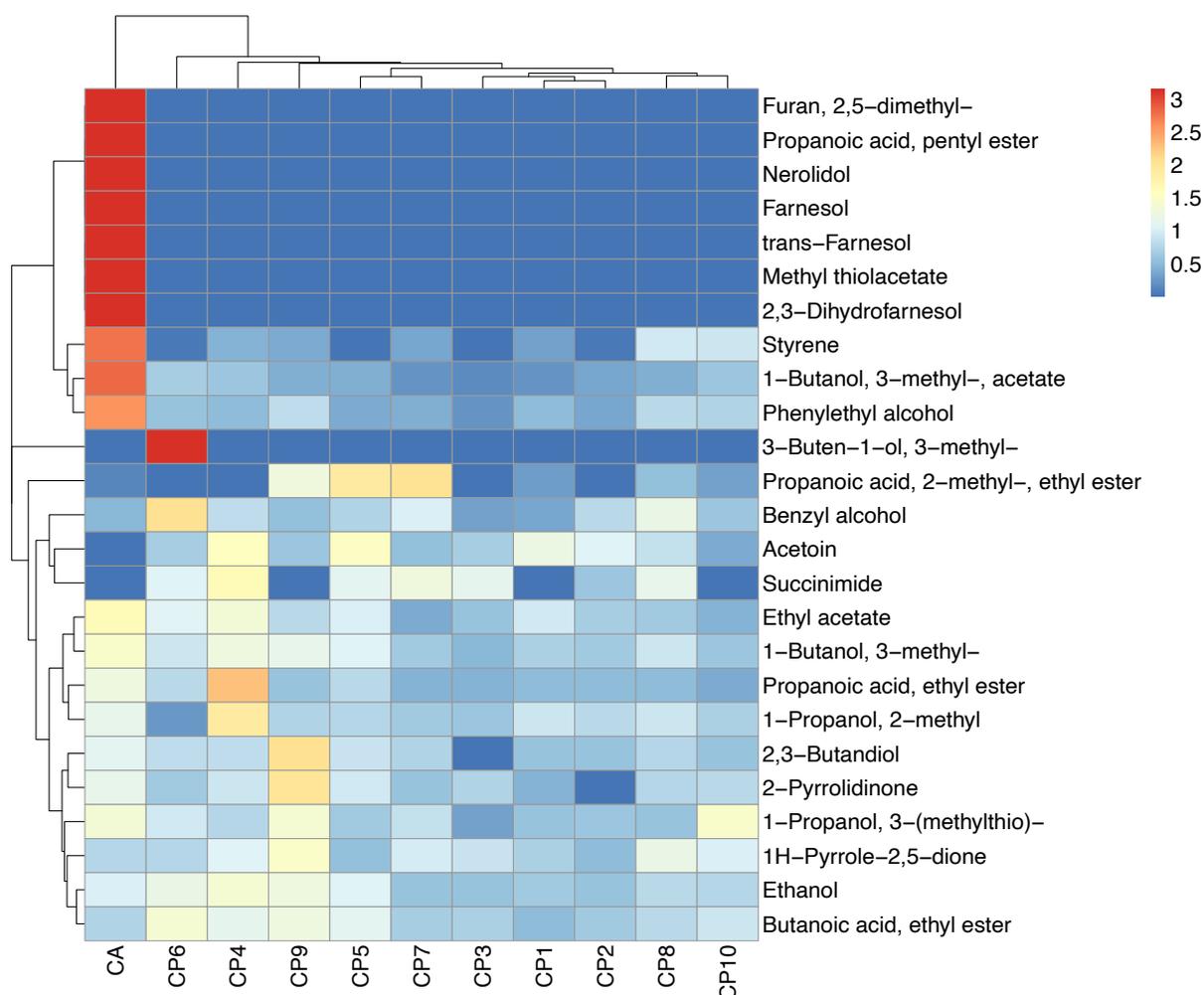
100

Figure 4.1: Heatmap plot illustrating hierarchical clustering characterisation of the planktonic *Candida* volatilomes in YPD media. Each column represents the mean abundance values from each examined *Candida* strain (*C. parapsilosis*, n = 3; *C. albicans*, n = 5). Abundance values were auto-scaled across each row, by each compound.

4.3.2. Chemical composition of planktonic *Candida* volatilomes

Extracted *Candida* volatilomes consisted of alcohols, fatty acid esters, pyrroles, ketones, and acetates. The boxplots shown in Figure 4.2 show the abundance of each compound recovered from the *C. albicans* samples and the *C. parapsilosis* clinical strains. The *C. parapsilosis* volatilome was primarily characterised by significant abundances of ethanol; 1-butanol, 3-methyl-; 1-butanol, 3-methyl-, acetate; and phenylethyl alcohol. Other highly abundant compounds included the fatty acid esters: butanoic acid, ethyl ester; propanoic acid, ethyl ester; and propanoic acid, 2-methyl, ethyl ester. These compounds were also highly abundant in the *C. albicans* volatilome, that of which also contained propanoic acid, pentyl ester. Several sesquiterpene compounds were emitted by the *C. albicans* samples, these were farnesol, nerolidol, trans-farnesol, and 2,3-dihydrofarnesol. These compounds were emitted in relatively moderate to high abundances in these samples and significantly contributed to the discrimination of the *C. albicans* volatilome from the *C. parapsilosis* volatilome. Another

notable difference between the two fungal volatilomes was the emission of the ketone acetoin, which was highly abundant in all *C. parapsilosis* samples and significantly less in the *C. albicans* samples. Methyl thiolacetate is a sulfur-containing compound that was only recovered in *C. albicans* samples. Within the 10 clinical strains of *C. parapsilosis*, there were varying emissions of certain compounds such as ethyl acetate, phenylethyl alcohol, propanoic acid, 2-methyl, ethyl ester, and 1-butanol, 3-methyl-, acetate. Succinimide and 3-buten-1-ol, 3-methyl- were only recovered from specific *C. parapsilosis* strains and were absent from *C. albicans* samples. These results highlight notable differences across species as well as a high degree of stability across the *C. parapsilosis* volatilome at the strain-level.
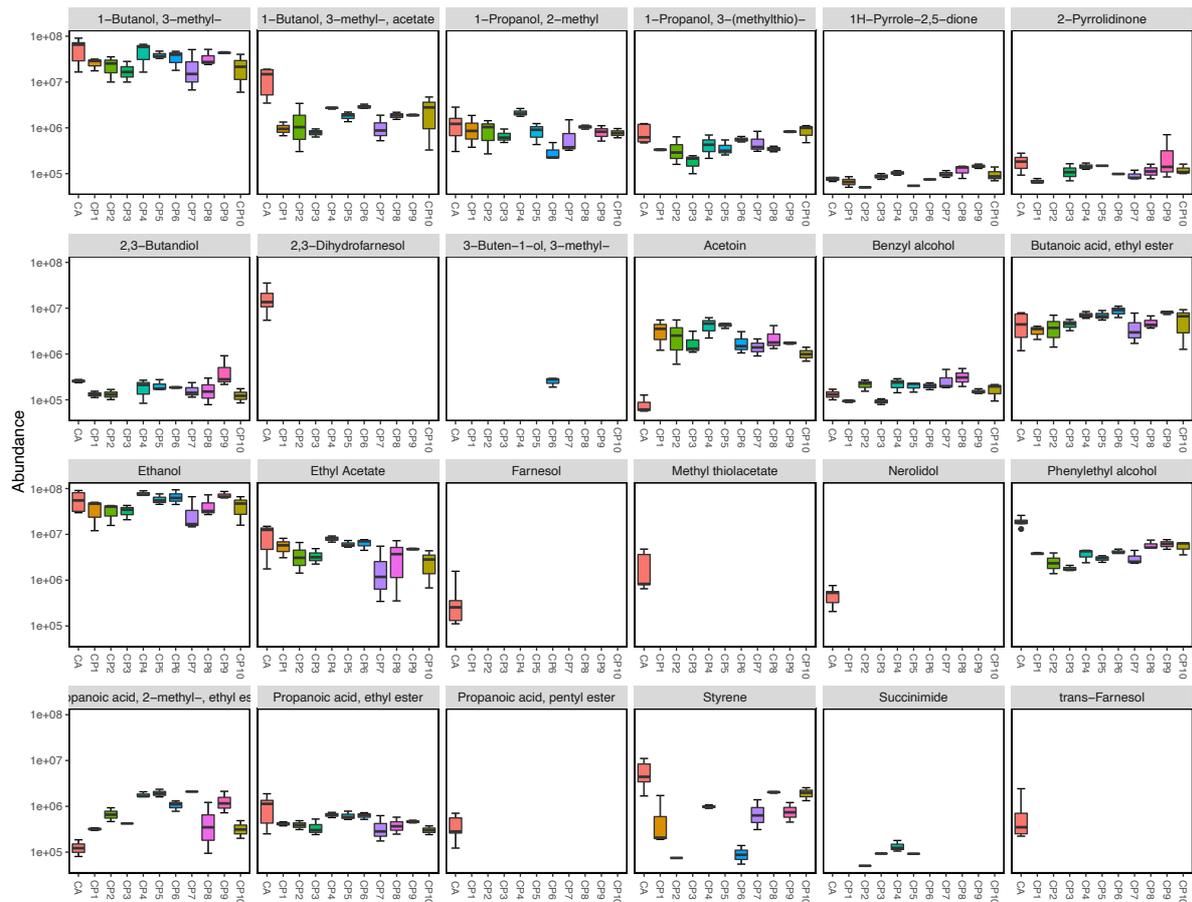

Figure 4.2: Individual comparative boxplots illustrating the emission of each compound by each of the examined planktonic *Candida* strains (CP 1-10 n=3; CA, n=5). The y-axis of each plot was scaled by log10 to improve visibility and interpretability of the plots. Names of compounds that are not fully visible: row 4, 1st : propanoic acid,2-methyl-,ethyl, ester.

4.3.3. Media-dependent influences on planktonic *Candida* volatilomes

Significant differences were observed in the compositions of the planktonic *Candida* volatilomes when the culture media was changed from YPD to TSB media. CP1 and *C. albicans* were investigated to assess how stable their volatilomes were across the different nutritional media. When cultured in TSB media at 37°C, following background subtraction, only 12 unique compounds were recovered from *C. albicans* samples (compared to 22 in YPD) and 7 unique compounds were recovered from the CP1 samples (compared to 18 in YPD). The

Pre-examination copy
Submitted 27th May 2022

grouped barcharts shown in Figure 4.3 visualise the compositional influence varying media had on the *Candida* volatilomes. In the case of the *C. albicans*, 11 out of 12 of the compounds recovered from the TSB were also present in its volatilome when cultured in YPD with the exception being 1-pentanol, 2-methyl-. In the case of the CP1 TSB samples, 6 of the 7 compounds that were recovered from samples were also recovered from the YPD samples with very low abundances of benzeneacetaldehyde being the exception. For both *C. albicans* and CP1, relatively high emissions of ethanol; 1-butanol, 3-methyl-; and phenylethyl alcohol persist across the two examined media. However, significantly less 1-propanol, 2-methyl and acetoin were recovered in CP1 TSB samples than in CP1 YPD samples, and similarly, significantly less of 1-butanol, 3-methyl-, acetate; 2,3-dihydrofarnesol; methyl thiolacetate; and 1-propanol, 2-methyl were recovered in *C. albicans* TSB samples than in *C. albicans* YPD samples. The observed higher number of compounds recovered from the YPD samples of both *C. albicans* and CP1, as well as the significantly higher abundances of volatile compounds emitted in YPD media confirm that both *Candida spp.* are more metabolically active when grown in YPD.
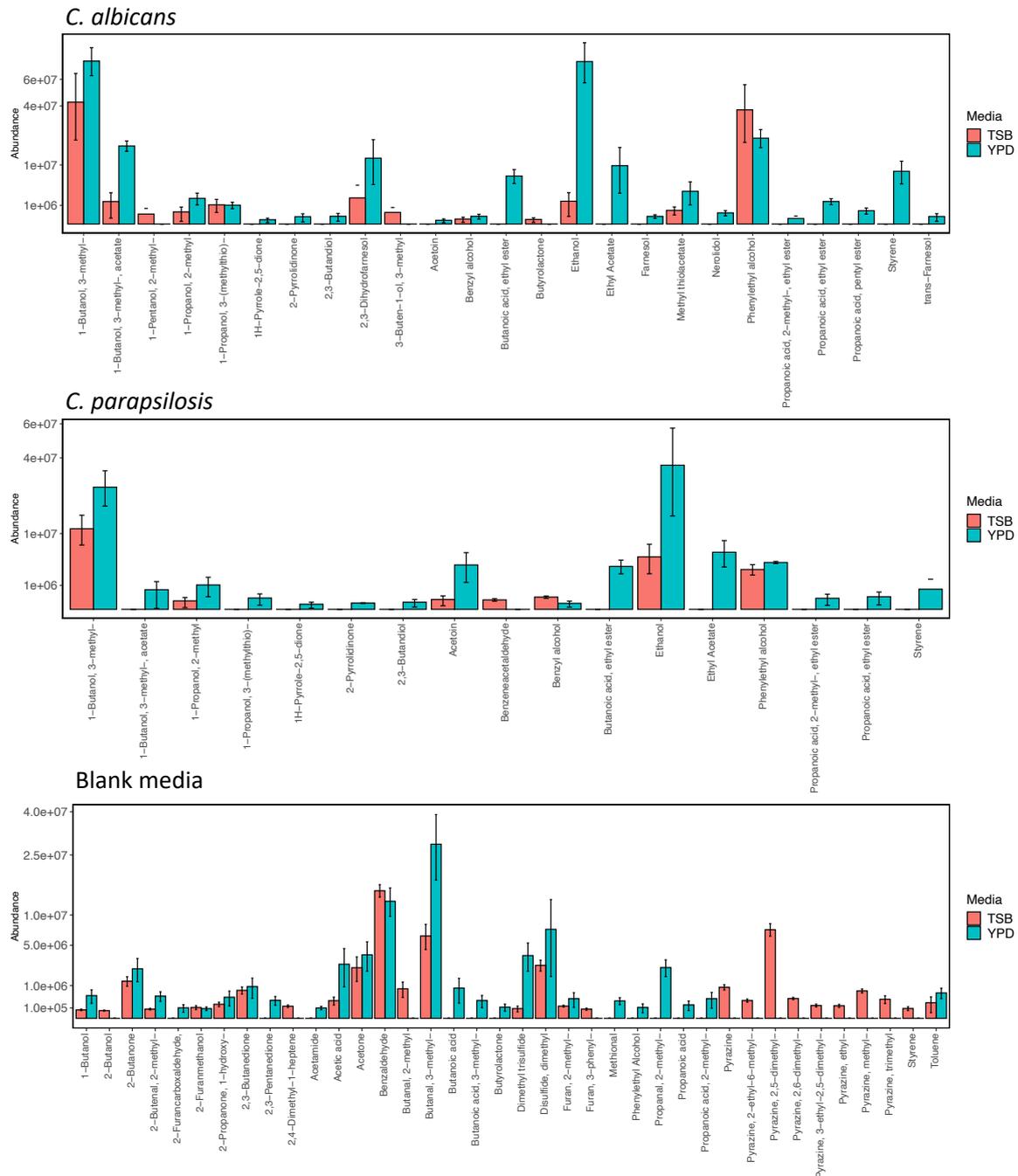
Pre-examination copy
Submitted 27th May 2022

Figure 4.3*: Cross-media comparative bar charts illustrating the volatilomes of *C. albicans* (top, n = 3) and *C. parapsilosis* (middle, n=3) cultures in TSB (n=5) and YPD (n=5) media, and TSB and YPD blank samples (bottom).

Figure 4.3 (bottom) shows the background volatilome of TSB and YPD blank samples. Key differences in the volatilome of the two media included significantly higher abundances of acid and sulfide compounds in the YPD blank samples; and a higher variety of pyrazine compounds in the TSB blank samples. There was an absence of fatty acid ester compounds in both *C. albicans* and CP1 TSB samples. YPD broth contains relatively high abundances of a

104

variety of available acids such as acetic acid, propanoic acid, propanoic acid, butanoic acid and butanoic acid, 3-methyl (isovaleric acid) (Figure 4.3, bottom), which can act as the precursors for the esterification and formation of butanoic acid, ethyl ester; propanoic acid, ethyl ester; and propanoic acid, 2-methyl, ethyl ester. The relatively high abundances of acetic acid in YPD are also responsible for the significant abundances of 1-butanol, 3-methyl-, acetate present in all *Candida* samples in YPD. This compound was likely formed through the esterification of acetic acid with 1-butanol, 3-methyl-. Low abundances of 1-butanol, 3-methyl-, acetate were recovered in the *C. albicans* TSB samples and it was completely absent in the CP1 TSB samples (Figure 4.3). Similarly, the absence of ethyl acetate in the TSB samples of both *C. albicans* and CP1 indicate that a high abundance of acetic acid may be required to form ethyl acetate via esterification with ethanol. Overall from the volatilomes obtained from the blank media (Figure 4.3, bottom) and the observed differences in *Candida* volatilomes across these media (Figure 4.3, top and middle), YPD provides an environment richer in substrates and volatilomic precursors that give rise to *Candida* volatilomes with higher degrees of complexity.

4.3.4. Biofilm-dependent influences on *Candida* volatilomes

The effects of biofilm formation and maturity on *Candida* volatile emissions were examined. Out of the *C. parapsilosis* strains, CP6 was chosen to be investigated due to strong ability to produce biofilms – this is clearly shown by the crystal violet assay of the CP strains in Figure S4.5. Biofilm-negative CP1 (Figure S4.5) was analysed as a control in this investigation as it shares a highly similar planktonic volatilome to CP6 but is unable to form a biofilm in YPD. Biofilm-negative *C. albicans* (Figure S4.4) was also analysed as a relative comparison to the CP samples. In order to carry out sampling of biofilm volatilomes, the experimental set-up for planktonic cultures needed to be adopted for monitoring and characterising the volatilome of maturing biofilms (see Figure S4.3). Volatile emissions at 24 and 48 h of biofilm growth were sampled and analysed. Unbound cells and exhausted media were washed away following sample collection at 24 h and fresh YPD was introduced to sustain biofilm growth between 24 – 48 h. Although biofilm formation was not observed in the CP1 and CA samples, there was a degree of cell adhesion to the wells following PBS washes at 24 h as the cell cultures regenerated between 24 - 48 h. This was clearly indicated through the turbid appearance and the volatilome of the samples at 48 h. Individual compound abundances over 48 h for biofilm-negative *C. albicans* and CP1 are shown in Figure S4.6 (*C. albicans*) and Figure S4.7 (CP1). The comparative boxplots in Figure 4.6 illustrate the emission of key compounds from the biofilm-positive CP6 strain and the biofilm-negative CP1 strain. There were clear significant differences (Figure S4.8) between CP1 and CP6 in the emission of primary metabolites such as ethanol ($p < 0.01$), 3-methyl-1-butanol ($p < 0.01$), 2-methyl-1-propanol ($p < 0.01$) and acetoin ($p < 0.001$). In biofilm-positive CP6 samples, between 24 – 48 h, the short chain fatty acids: propanoic acid, 2-methyl propanoic acid, butanoic acid, and 3-methyl butanoic acid were completely consumed across all samples (Figure 4.5). Significant decreases ($p < 0.01$) in the abundances of acetic acid were also observed at 48 h across the biofilm-positive CP6 samples. Small abundances of butanoic acid, 2-methyl- were only detected in the CP6 samples at 48 h biofilm growth. In biofilm-negative CP1 samples, significant decreases in acid abundances were observed in butanoic acid, propanoic acid and 2-methyl propanoic acid (Figure S4.7). No significant differences in acid abundance were observed in biofilm-negative *C. albicans* cultures (Figure S4.6). Significant increases ($p < 0.01$)

of phenylethyl alcohol were detected in the biofilm-positive CP6 between 24 and 48 h growth. This increase was not observed in either biofilm-negative *C. albicans* (Figure S4.6) and CP1 (Figure S4.7) samples. Methyl thiolacetate was detected in biofilm-positive CP6 samples at 48 h, this compound was not detected in biofilm-negative CP1 samples at any time points but was detected in biofilm-negative *C. albicans* (Figure S4.6).
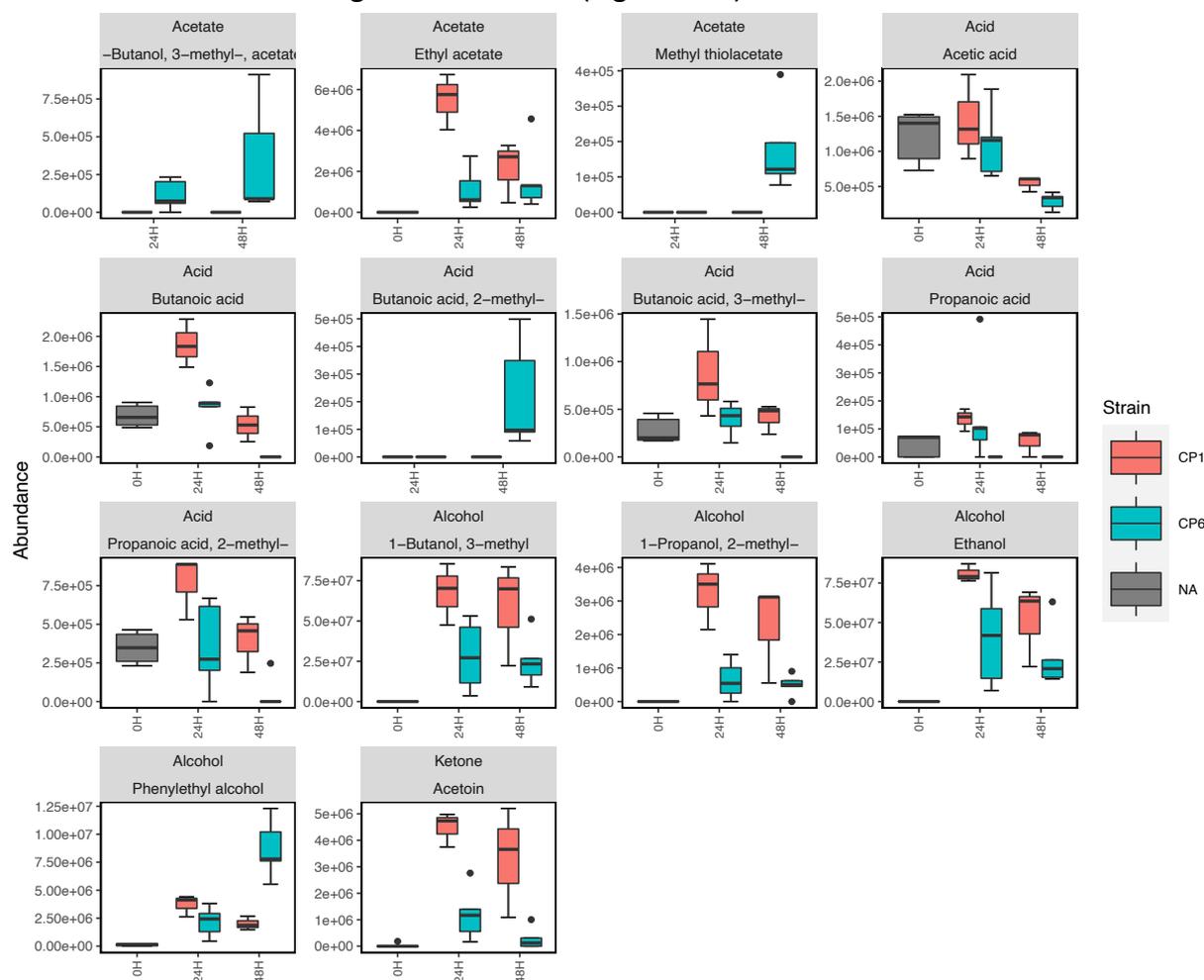

Figure 4.4: Grouped boxplots comparing the abundances of compounds emitted from biofilm-negative forming CP1 (n = 3) and biofilm-positive CP6 (n = 5). In both cases, the volatilomes were sampled and analysed using the experimental set-up shown in Figure S4.3. Statistically significant differences in the abundance of compounds at each time point are illustrated through the star system where * = p < 0.05 and ** = p < 0.01.
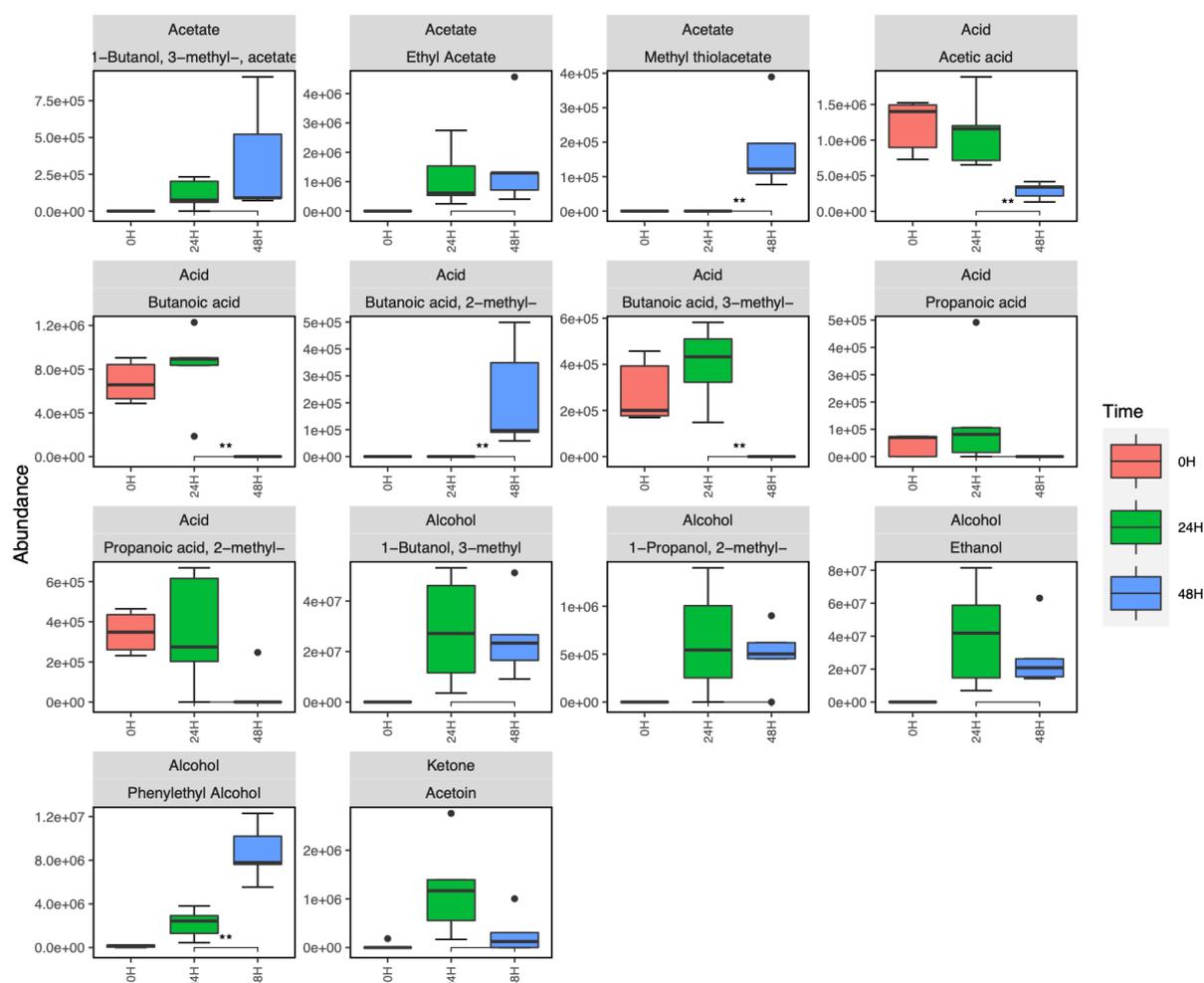
Figure 4.5: Compound boxplots illustrating the abundances of various compounds detected in the headspace of CP6 biofilm cultures at 0, 24 and 48 h growth. Statistically significant differences in the abundance of compounds at each time point are illustrated through the star system where * = p < 0.05 and ** = p < 0.01.

4.4. Discussion

In this work, we characterised the previously understudied volatilome of the emerging fungal pathogen, *C. parapsilosis,* across varying planktonic growth parameters and compared it to the volatilome of *C. albicans*. Our primary objective for this study was to investigate species-, media-, time-, and biofilm-dependent variation across *C. albicans* and *C. parapsilosis* . The volatilomes of 10 clinical strains of *C. parapsilosis* and one strain of *C. albicans* were analysed using HS-SPME-GCMS. Clear discrimination of the dissimilarities across the planktonic *C. parapsilosis* and *C. albicans* volatilomes was achieved using hierarchical clustering (Figure 4.1). Among the 25 unique compounds identified across both species in YPD media, 23 compounds were recovered from *C. albicans* and 18 compounds were recovered from both *Candida* species. The most abundant compounds were alcohols: ethanol; 1-butanol, 3-methyl-; and 2-phenylethyl alcohol – all of which are produced through primary metabolism of glucose and amino acids. Ethanol is typically derived from glucose metabolism and fermentation while 1-butanol, 3-methyl- is produced from the downstream breakdown of leucine[25]. 2-phenylethyl alcohol is produced from the degradation of aromatic amino acids

107

[6] via the shikimate pathway[26]. These core alcohols that both *Candida* species emitted are not characteristic and are common metabolic products that are widely emitted across both bacterial and fungal kingdoms[11][6][27]. The emission of various fatty acid esters also contributed to all *Candida* volatilomes, these included the acetate esters: 1-butanol, 3-methyl-, acetate- and ethyl acetate; and the short chain ethyl esters: butanoic acid, ethyl ester; propanoic acid, ethyl ester; and propanoic acid, 2-methyl, ethyl ester. Esterification is a common metabolic process utilised by yeasts[28] which arises from the reaction of specific alcohols and carboxylic acids, mediated by acetyl coenzyme A – a primary metabolism product formed from the decarboxylation of pyruvate[29].

The emission of all compounds discussed above was shown to be highly dependent on the growth media used for the culture (Figure 4.3). Volatilomic screening of microbes across different nutritional environments allows a more comprehensive view of the range of compounds that can be emitted. Microbial biosynthetic pathways for the volatile metabolites detected can be somewhat elucidated by investigating the variation in substrates and precursors available in different growth media. However, these pathways can only be fully elucidated using highly specialised techniques such as ^{13}C labelling-metabolic fluxomic analyses[30][31]. YPD was the primary media examined here as it is widely used for the growth of yeasts; TSB is a less-complex universal media that can be used for both fungi and bacteria. Following broad analysis of all *Candida* strains in YPD, our goal was to determine potential synthetic pathways for the emitted compounds by analysing and comparing the *Candida* volatilomes from the acid-free and sulfide-limited TSB media. For *C. albicans* and *C. parapsilosis* (CP1), the alcohols mentioned above persisted in volatilomes recovered from both TSB and YPD media while the fatty acid esters were essentially absent in the TSB recovered volatilomes. Volatilomic analysis of the blank TSB and YPD media revealed that YPD contained a variety of available fatty acids from which both *Candida* strains could utilise to produce fatty acid esters via esterification. Interestingly bacteria such as *Staphylococcus aureus* and *Escherichia coli* have demonstrated the ability to produce fatty acid esters in fatty acid-free media such as TSB[11]; this is potentially due to the secondary metabolism of their primary acidic metabolites such as acetic acid, propanoic acid, and butanoic acid, 3-methyl-.

The unique set of sesquiterpene compounds: farnesol, 2,3-dihydrofarnesol, trans-farnesol and nerolidol clearly discriminated the *C. albicans* from the *C. parapsilosis* volatilome after 24 h incubation. These volatile sesquiterpenoid compounds have also previously discriminated *C. albicans* from *C. glabrata* and *C. tropicalis* volatilomes[20] . Farnesol is produced as a biproduct along the ergosterol synthesis pathway[32] and is a quorum sensing molecule critically used in *Candida* biofilm development. It's presence in the *C. albicans* samples and relative absence in the *C. parapsilosis* samples may also indirectly highlight the previously reported[5] biofilm formation differences between these species. It is primarily involved in the control of morphogenic transitions in *Candida*, mediating the transitions from a hypha-to-yeast[33] and inhibiting yeast-to-hypha transitions[34] – which is critical to *C. albicans* biofilm formation and virulence. Farnesol has also demonstrated strong antimicrobial properties showing anti-biofilm activity against several *Staphylococcal spp*[35][36]. The volatile nature of these bioactive farnesol-like compounds emitted by *Candida* also indicate their potential role in long-distance inter-microbial interactions. This supports the idea of microbial volatilomics being used in novel volatile antibiotic screening[37]. Farnesol compounds are not exclusively

produced by *C. albicans* and have been previously reported to be produced in small abundances by various *Candida spp.* including *C. parapsilosis*[38]. Other previously reported bioactive components that were detected in the *Candida* volatilomes included ethanol[37][39][40]; phenylethyl alcohol[41][42]; 1-butanol, 3-methyl- (isoamyl alcohol) [37][39][40][41][42]; 1-butanol, 3-methyl-, acetate- (isoamyl acetate) [37][39]40]; ethyl acetate-[38][40][41]; and propanoic acid, 2-methyl, ethyl ester [37][41]. However, although these compounds are highly common across the microbial kingdoms, it is suspected that only species-specific combinations and abundances of them elicit antimicrobial effects[43]. These effects are determined using volatilomic bioactivity assays, a field of study that is still in the early stages of development.

Volatilomic investigations involving microbial biofilms are currently lacking in the field. Notable studies describing systems for targeted volatilomic profiling of biofilms have been recently described by Slade et al.[44][23] using selected ion flow tube-MS. In this work, we demonstrate biofilm-specificity in *C. parapsilosis* volatilomes using a novel sampling approach via HS-SPME (see Figure S4.3) coupled with GC-MS analysis. However, there are limitations with this system that must be noted: 1) manual handling requirements for the transfer of liquids; 2) sampling containers are not specifically made for the task; and 3) increased risk of contamination. Although a standardised sampling container is not currently being constructed, manual handling requirements and contamination risks can be mitigated by exercising good aseptic technique procedures.

CP6 was confirmed as biofilm-positve using a crystal violet assay of CP1 – CP7 in YPD media (Figure S4.4, S4.5). CP1 and *C. albicans* did not form biofilms in YPD and were analysed across 48 h as comparative controls. It must be noted that the success of *C. albicans* biofilm formation is known to be significantly enhanced by coating the polystyrene wells with serum prior to inoculation[45]and using glucose-supplemented media[46] – neither of which were employed in this study. Following 48 h of biofilm development in CP6 samples, significant differences in the abundances of various compounds were observed, particularly between 24 and 48 h (Figure 4.5). Whereas in biofilm-negative CP1 and *C. albicans*, following the washing step at 24 h, the cells that remained in the wells grew in the same planktonic manner as they did from 0 – 24 h (Figure S4.6 and S4.7). Biofilm maturity was characterised by the metabolism of short-chain fatty acids as propanoic acid, 2-methyl propanoic acid, butanoic acid, and 3-methyl butanoic acid which all appeared to be completely consumed in CP6 samples after 24 h. Significant decreases in acetic acid also coincided with increases in the abundance of acetate molecules (Figure 4.5). Volatile acetic acid metabolism and the subsequent formation of acetate has previously been shown to enhance biofilm formation in bacterial cells[47]. In contrast, in non-biofilm forming CP1 and *C. albicans*, acids produced in the initial stages remained relatively stable. Fermentation metabolites such as ethanol, and short-chain amino acid (leucine) metabolites such as 2-methyl-1-propanol and 3-methyl-1-butanol were significantly more abundant in biofilm-negative CP1 samples compared to the biofilm-positive CP6 samples (Figure S4.8). Acetoin was also detected at significantly higher abundances in the non-biofilm forming CP1 samples. This primary metabolite is directly derived from the breakdown of pyruvate[6] and is produced to neutralise the extracellular environment to prevent over-acidification of the cells[48]. The relatively high abundances of acetoin and these primary alcohols in CP1 samples indicate a higher rate of primary metabolism in these cells. In contrast to this, the observed reductions in abundances of these compounds in the CP6

109

biofilm samples is in agreement with the reports of the down regulation of primary metabolic activity as the biofilm develops[49]. Interestingly, no significant differences in the abundances of these compounds were observed between the CP1 and CP6 when they were grown planktonically in the HS vials (Figure 4.2). Biofilm development in *Candida spp.* has also been associated with the upregulation of various amino acids[51] . In maturing CP6 biofilms, higher rates of aromatic amino acid degradation (shikimate pathway) were indicated through the detection of significant increases of phenylethyl alcohol detected at 48 h (Figure 4.5). Production of this molecule has also been linked to the stimulation of filamentous growth[50] and promotes biofilm formation in yeasts[51][52]. Interestingly, despite the filamentous growth this molecule stimulates in yeasts, it has demonstrated inhibition of hyphal formation in *C. albicans* cells[53][26] - which is a core step in *C. albicans* biofilm development[54]. Among the other compounds produced in the CP6 biofilms, one of the most notable was the sulfur-containing acetate, methyl thiolacetate, which was detected only after 48 h. Interestingly this molecule was not detected from any planktonic *C. parapsilosis* strains (Figure 4.2) – highlighting a potential biofilm-specific role in its emission in *C. parapsilosis*. In contrast to this, methyl thiolacetate was detected in *C. albicans* cultures in both HS vials (Figure 4.2) and the 6-well system (Figure S4.5). Although YPD media was found to be rich in dimethyl disulfide and dimethyl trisulfide (Figure 4.3), sulfides were not detected at either 24 or 48 h biofilm growth. Sulfur degradation is an integral metabolic pathway utilised by yeast cells for division[55] and sustaining growth[56].

Our novel sampling chamber HS-based experimental approach allowed biofilm-specific dynamics of microbial volatilomes to be explored using untargeted volatilomics. This technique could also be used to investigate potential biofilm-specificity in the volatilomes of other clinical fungal and bacterial pathogens. Another experimental application for this system in the future could be for volatilomic bioactivity assays to monitor volatile-mediated interactions between microbes - grown in separate wells, under a shared headspace. A major advantage of this method is that it can simply be adapted for a variety of sampling and analysis platforms (i.e. SIFT-MS and proton transfer tube-MS) and therefore can be used for qualitative and quantitative investigations.

4.5. Conclusion

In this work, the volatilome of multiple strains of *C. parapsilosis* was comprehensively investigated for the first time under varying growth conditions. These results were compared to volatilomic analyses of *C. albicans* under the same conditions to assess inter-species variation within the *Candida* genus. A total of 24 unique compounds were identified and allowed clear discrimination of the *C. parapsilosis* and *C. albicans* volatilomes from each other. Among these differences was the unique emission of sesquiterpene-type compounds (farnesol, 2,3-dihydrofarnesol, and nerolidol) by *C. albicans*. High degrees of stability were observed in the abundances of individual compounds detected across the 10 examined *C. parapsilosis* strains. However, volatilomic stability was not observed across different media as the effects of culturing in TSB significantly reduced the diversity of both *C. parapsilosis* and *C. albicans* cells compared to YPD cultures. While primary metabolites were detected in similar abundances across the two media due to relatively similar abundances of available glucose, high acid and sulfide contents in YPD enabled the generation of esters, acetates, and sulfur-containing compounds that were not recovered from TSB.

A novel biofilm-specificity study in the *C. parapsilosis* volatilome was also carried out. Following a comparative analyses, significantly less primary metabolites were detected in the biofilm-positive *C. parapsilosis* samples than the biofilm-negative cultures. Furthermore, in biofilm-positive samples, significant consumption of all short-chain fatty acids was observed while the unique increase in the abundance of phenylethyl alcohol correlated with biofilm maturity. This novel HS sampling set-up has a wide range of potential applications from biofilm volatilomic monitoring, characterisation of microbial co-cultures, and bioactivity assays. By examining the effects of strain-to-strain variation, media- and time-dependent emission, and biofilm formation; a more comprehensive view of the metabolic capabilities of microbes can be achieved. Comprehensive profiling of microbes in this manner will ultimately allow a simpler translation of microbial volatilomics workflows into clinical volatilomic applications.

4.6. References

[1] Trofa, D., Gácser, A. and Nosanchuk, J., 2008. *Candida parapsilosis* , an Emerging Fungal Pathogen. *Clinical Microbiology Reviews*, 21(4), pp.606-625. https://doi.org/10.1128/CMR.00013-08

[2] Pammi, M., Holland, L., Butler, G., Gacser, A. and Bliss, J., 2013. *Candida parapsilosis* Is a Significant Neonatal Pathogen. *Pediatric Infectious Disease Journal*, 32(5), pp.e206-e216. https://doi.org/10.1097/INF.0b013e3182863a1c

[3] Lafferty, S. and Butler, G., 2005. Phenotype switching affects biofilm formation by *Candida parapsilosis*. *Microbiology*, 151, pp.1073–1081. https://doi.org/10.1099/mic.0.27739-0

[4] Sudbery, P., 2011. Growth of *Candida albicans* hyphae. *Nature Reviews Microbiology*, 9(10), pp.737-748. https://doi.org/10.1038/nrmicro2636

[5] Martins, M., Henriques, M., Azeredo, J., Rocha, S., Coimbra, M. and Oliveira, R., 2007. Morphogenesis Control in *Candida albicans* and *Candida dubliniensis* through Signaling Molecules Produced by Planktonic and Biofilm Cells. *Eukaryotic Cell*, 6(12), pp.2429-2436. https://doi.org/10.1128/EC.00252-07

[6] Weisskopf, L., Schulz, S. and Garbeva, P., 2021. Microbial volatile organic compounds in intra-kingdom and inter-kingdom interactions. *Nature Reviews Microbiology*,. https://doi.org/10.1038/s41579-020-00508-1

[7] Schulz, S. and Dickschat, J. Bacterial Volatiles: The Smell of Small Organisms. *ChemInform*, 38(45). https://doi.org/10.1039/b507392h. (2007).

[8] Schulz, S., Schlawis, C., Koteska, D., Harig, T., and Biwer, P. (2020). "*Chapter 3:* Structural Diversity of Bacterial Volatiles", in *Bacterial Volatile Compounds as Mediators of Airborne Interactions*. 1st ed, eds C. M. Ryu, L. Weisskopf, and B. Piechulla (SINGAPOR: SPRINGER VERLAG), 93–116.

[9] Rees, C., Nordick, K., Franchina, F., Lewis, A., Hirsch, E., and Hill, J. (2017). Volatile metabolic diversity of *Klebsiella pneumoniae* in nutrient-replete conditions. *Metabolomics* 13:18. https://doi.org/10.1007/s11306-016-1161-z

[10] Jenkins, C., and Bean, H. (2019). Influence of media on the differentiation of Staphylococcus spp. by volatile compounds. *J. Breath Res.* 14:016007. https://doi.org/10.1088/1752-7163/ab3e9d

[11] Fitzgerald, S., Holland, L. and Morrin, A., 2021. An Investigation of Stability and Species and Strain-Level Specificity in Bacterial Volatilomes. *Frontiers in Microbiology*, 12. https://doi.org/10.3389/fmicb.2021.693075

[12] Boots, A. W., Smolinska, A., van Berkel, J. J., Fijten, R. R., Stobberingh, E. E., Boumans, M. L., et al. (2014). Identification of microorganisms based on headspace analysis of volatile organic compounds by gas chromatography–mass spectrometry. *J. Breath Res.* 8:027106 https://doi.org/10.1088/1752-7155/8/2/027106

[13] Fitzgerald, S., Duffy, E., Holland, L. and Morrin, A., (2020). Multi-strain volatile profiling of pathogenic and commensal cutaneous bacteria. *Scientific Reports*, 10(1). https://doi.org/10.1038/s41598-020-74909-w

[14] Insam, H. and Seewald, M., Volatile organic compounds (VOCs) in soils. *Biol. Fertil. Soils*, 46(3), pp.199-213. https://doi.org/10.1007/s00374-010-0442-3 (2010).

[15] Inamdar, A., Morath, S. and Bennett, J., 2020. Fungal Volatile Organic Compounds: More Than Just a Funky Smell?. *ANNUAL REVIEW OF MICROBIOLOGY*, 74, pp.101-116. https://doi.org/10.1146/annurev-micro-012420-080428

[16] Schmidt, R., Cordovez, V., de Boer, W., Raaijmakers, J. and Garbeva, P., Volatile affairs in microbial interactions. *The ISME Journal*, 9(11), pp.2329-2335. https://doi.org/10.1038/ismej.2015.42 (2015).

[17] Duffy, E. and Morrin, A., Endogenous and microbial volatile organic compounds in cutaneous health and disease. *TrAC, Trends Anal. Chem,* 111, pp.163-172. https://doi.org/10.1016/j.trac.2018.12.012 (2019).

[18] Ryu, C., Weisskopf, L., and Piechulla, B. (2020). *Bacterial volatile compounds as mediators of airborne interactions.* Singapore: Springer.

[19] Costa, C., Bezerra, A., Almeida, A. and Rocha, S., 2020. *Candida* Species (Volatile) Metabotyping through Advanced Comprehensive Two-Dimensional Gas Chromatography. *Microorganisms*, 8(12), p.1911. https://doi.org/10.3390/microorganisms8121911

[20] López-Ramos, J., Bautista, E., Gutiérrez-Escobedo, G., Mancilla-Montelongo, G., Castaño, I., González-Chávez, M. and De Las Peñas, A., 2021. Analysis of Volatile Molecules Present in the Secretome of the Fungal Pathogen *Candida glabrata*. *Molecules*, 26(13), p.3881. https://doi.org/10.3390/molecules26133881

[21] Perl, T., Jünger, M., Vautz, W., Nolte, J., Kuhns, M., Borg-von Zepelin, M. and Quintel, M., 2011. Detection of characteristic metabolites of Aspergillus fumigatus and *Candida* species using ion mobility spectrometry - metabolic profiling by volatile organic compounds. *Mycoses*, 54(6), pp.e828-e837. https://doi.org/10.1111/j.1439-0507.2011.02037.x

[22] Slade, E., Thorn, R., Young, A. and Reynolds, D., 2021. Real-time detection of volatile metabolites enabling species-level discrimination of bacterial biofilms associated with wound infection. *Journal of Applied Microbiology*,. https://doi.org/10.1111/jam.15313

[23] Bhargava, P. and Collins, J., 2015. Boosting Bacterial Metabolism to Combat Antibiotic Resistance. *Cell Metabolism*, 21(2), pp.154-155. http://dx.doi.org/10.1016/j.cmet.2015.01.012

[24] A van den Berg, R., CJ Hoefsloot, H., A Westerhuis, J., K Smilde, A. and J van der Werf, M., (2006). Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics*, 7(142). https://doi.org/10.1186/1471-2164-7-142.

[25] Serrazanetti, D., Ndagijimana, M., Sado-Kamdem, S., Corsetti, A., Vogel, R., Ehrmann, M. and Guerzoni, M., 2011. Acid Stress-Mediated Metabolic Shift in Lactobacillus sanfranciscensis LSCE1. *Applied and Environmental Microbiology*, 77(8), pp.2656-2666. 10.1128/AEM.01826-10

[26] Mitri, S., Koubaa, M., Maroun, R., Rossignol, T., Nicaud, J. and Louka, N., 2022. Bioproduction of 2-Phenylethanol through Yeast Fermentation on Synthetic Media and on Agro-Industrial Waste and By-Products: A Review. *Foods*, 11(1), p.109. https://doi.org/10.3390/foods11010109

[27] McNerney, R., Mallard, K., Okolo, P. and Turner, C., Production of volatile organic compounds by mycobacteria. *FEMS Microbiol. Lett.* 328(2), pp.150-156. https://doi.org/10.1111/j.1574-6968.2011.02493.x (2012).

[28] Saerens,, S., Delvaux, F., Verstrepen, K. and Thevelein, J., 2010. Production and biological function of volatile esters in Saccharomyces cerevisiae. *Microbial Biotechnology*, 3(2), pp.165-177. https://doi.org/10.1111/j.1751-7915.2009.00106.x

[29] Wolfe, A., 2005. The Acetate Switch. *Microbiology and Molecular Biology Reviews*, 69(1), pp.12-50. https://doi.org/10.1128/MMBR.69.1.12-50.2005

[30] Guo, W., Sheng, J. and Feng, X., 2015. 13C-Metabolic Flux Analysis: An Accurate Approach to Demystify Microbial Metabolism for Biochemical Production. *Bioengineering*, 3(1), p.3. https://doi.org/10.3390/bioengineering3010003

[31] Long, C. and Antoniewicz, M., 2019. High-resolution 13C metabolic flux analysis. *Nature Protocols*, 14(10), pp.2856-2877. https://doi.org/10.1038/s41596-019-0204-0

[32] Rodrigues, C. and Černáková, L., 2020. Farnesol and Tyrosol: Secondary Metabolites with a Crucial quorum-sensing Role in *Candida* Biofilm Development. *Genes*, 11(4), p.444. https://doi.org/10.3390/genes11040444

[33] Lindsay, A., Deveau, A., Piispanen, A. and Hogan, D., 2012. Farnesol and Cyclic AMP Signaling Effects on the Hypha-to-Yeast Transition in *Candida albicans*. *Eukaryotic Cell*, 11(10), pp.1219-1225. https://doi.org/10.1128/EC.00144-12

[34] Hornby, J., Jensen, E., Lisec, A., Tasto, J., Jahnke, B., Shoemaker, R., Dussault, P. and Nickerson, K., 2001. Quorum Sensing in the Dimorphic Fungus *Candida albicans* Is Mediated by Farnesol. *Applied and Environmental Microbiology*, 67(7), pp.2982-2992. https://doi.org/10.1128/AEM.67.7.2982-2992.2001

[35] Jabra-Rizk, M., Meiller, T., James, C. and Shirtliff, M., 2006. Effect of Farnesol on *Staphylococcus aureus* Biofilm Formation and Antimicrobial Susceptibility. *Antimicrobial Agents and Chemotherapy*, 50(4), pp.1463-1469. https://doi.org/10.1128/AAC.50.4.1463–1469.2006

[36] Pammi, M., Liang, R., Hicks, J., Barrish, J. and Versalovic, J., 2011. Farnesol Decreases Biofilms of Staphylococcus epidermidis and Exhibits Synergy With Nafcillin and Vancomycin. *Pediatric Research*, 70(6), pp.578-583. https://doi.org/10.1203/PDR.0b013e318232a984

[37] Strobel, G., Sears, J., Dirkse, E. and Markworth, C., 2001. Volatile antimicrobials from *Muscodor albus*, a novel endophytic fungus. *Microbiology*, 147(11), pp.2943-2950. https://doi.org/10.1099/00221287-147-11-2943

[38] Weber, K., Sohr, R., Schulz, B., Fleischhacker, M. and Ruhnke, M., 2009. Secretion of E,E -Farnesol and Biofilm Formation in Eight Different *Candida* Species. *Antimicrobial Agents and Chemotherapy*, 53(2), pp.848-848. https://doi.org/10.1128/AAC.01646-07

[39] Ezra, D., Hess, W. and Strobel, G., 2004. New endophytic isolates of Muscodor albus, a volatile-antibiotic-producing fungus. *Microbiology*, 150(12), pp.4023-4031. https://doi.org/10.1099/mic.0.27334-0

[40] Mitchell, A., Strobel, G., Moore, E., Robison, R. and Sears, J., 2010. Volatile antimicrobials from Muscodor crispans, a novel endophytic fungus. *Microbiology*, 156(1), pp.270-277. https://doi.org/10.1099/mic.0.032540-0

[41] Stinson, M., Ezra, D., Hess, W., Sears, J. and Strobel, G., 2003. An endophytic Gliocladium sp. of Eucryphia cordifolia producing selective volatile antimicrobial compounds. *Plant Science*, 165(4), pp.913-922. https://doi.org/10.1016/S0168-9452(03)00299-1

[42] Singh, S., Strobel, G., Knighton, B., Geary, B., Sears, J. and Ezra, D., 2011. An Endophytic Phomopsis sp. Possessing Bioactivity and Fuel Potential with its Volatile Organic Compounds. *Microbial Ecology*, 61(4), pp.729-739

[43] Hou, Q., Keren-Paz, A., Korenblum, E., Oved, R., Malitsky, S. and Kolodkin-Gal, I., 2021. Weaponizing volatiles to inhibit competitor biofilms from a distance. *npj Biofilms and Microbiomes*, 7(1). https://doi.org/10.1038/s41522-020-00174-4

[44] Slade, E., Thorn, R., Young, A. and Reynolds, D., 2019. An in vitro collagen perfusion wound biofilm model; with applications for antimicrobial studies and microbial metabolomics. *BMC Microbiology*, 19(1). https://doi.org/10.1186/s12866-019-1682-5

[45] Frade, J. and Arthington-Skaggs, B., 2010. Effect of serum and surface characteristics on *Candida albicans* biofilm formation. *Mycoses*, 54(4), pp.e154-e162. https://doi.org/10.1111/j.1439-0507.2010.01862.x

[46] Serrano-Fujarte, I., López-Romero, E., Reyna-López, G., Martínez-Gámez, M., Vega-González, A. and Cuéllar-Cruz, M., 2015. Influence of Culture Media on Biofilm Formation by *Candida* Species and Response of Sessile Cells to Antifungals and Oxidative Stress. *BioMed Research International*, 2015, pp.1-15. https://doi.org/10.1155/2015/783639

[47] Chen, Y., Gozzi, K., Yan, F. and Chai, Y., 2015. Acetic Acid Acts as a Volatile Signal To Stimulate Bacterial Biofilm Formation. *mBio*, 6(3). https://doi.org/10.1128/mBio.00392-15

[48] Xiao, Z. and Xu, P. Acetoin Metabolism in Bacteria. *Crit. Rev. Microbiol.*, 33(2), pp.127-140. https://doi.org/10.1080/10408410701364604. (2007).

[49] Zhu, Z., Wang, H., Shang, Q., Jiang, Y., Cao, Y. and Chai, Y., 2012. Time Course Analysis of *Candida albicans* Metabolites during Biofilm Development. *Journal of Proteome Research*, 12(6), pp.2375-2385 https://doi.org/10.1021/pr300447k

[50] Chauhan, N. and Mohan Karuppayil, S., 2021. Dual identities for various alcohols in two different yeasts. *Mycology*, 12(1), pp.25-38. https://doi.org/10.1080/21501203.2020.1837976

[51] Lei, X., Deng, B., Ruan, C., Deng, L. and Zeng, K., 2022. Phenylethanol as a quorum sensing molecule to promote biofilm formation of the antagonistic yeast Debaryomyces nepalensis for the control of black spot rot on jujube. *Postharvest Biology and Technology*, 185, p.111788. https://doi.org/10.1016/j.postharvbio.2021.111788
112

Pre-examination copy
Submitted 27th May 2022

52 Pu, L., Jingfan, F., Kai, C., Chao-an, L. and Yunjiang, C., 2014. Phenylethanol promotes adhesion and biofilm formation of the antagonistic yeast Kloeckera apiculate for the control of blue mold on citrus. *FEMS Yeast Research*, 14(4), https://doi.org/10.1111/1567-1364.12139

53 Han, T., Tumanov, S., Cannon, R. and Villas-Boas, S., 2013. Metabolic Response of *Candida albicans* to Phenylethyl Alcohol under Hyphae-Inducing Conditions. *PLoS ONE*, 8(8), p.e71364. https://doi.org/10.1371/journal.pone.0071364

54 Nobile, C. and Johnson, A., 2015. *Candida albicans* Biofilms and Human Disease. *Annual Review of Microbiology*, 69(1), pp.71-92. https://doi.org/10.1146/annurev-micro-091014-104330

55 Blank, H., Gajjar, S., Belyanin, A. and Polymenis, M., 2009. Sulfur Metabolism Actively Promotes Initiation of Cell Division in Yeast. *PLoS ONE*, 4(11), p.e8018. https://doi.org/10.1371/journal.pone.0008018

56 Castrillo, J., Zeef, L., Hoyle, D., Zhang, N., Hayes, A., Gardner, D., Cornell, M., Petty, J., Hakes, L., Wardleworth, L., Rash, B., Brown, M., Dunn, W., Broadhurst, D., O'Donoghue, K., Hester, S., Dunkley, T., Hart, S., Swainston, N., Li, P., Gaskell, S., Paton, N., Lilley, K., Kell, D. and Oliver, S., 2007. Growth control of the eukaryote cell: a systems biology study in yeast. *Journal of Biology*, 6(2), p.4. https://doi.org/10.1186/jbiol54

Pre-examination copy
Submitted 27th May 2022

Chapter 5: Non-Invasive Volatilomic Analysis of Infected and Non-infected Diabetic Foot Ulcers

Abstract

The growing occurrence of chronic wounds internationally has fast become a heavy economic burden on national healthcare systems. Diabetic foot ulcers, in particular, are a devastating complication that can arise in patients with diabetes mellitus. Infection of diabetic ulcers can result in significantly bad outcomes including amputation and increased risk of death. For this reason, detecting infections early in such wounds is a critical step in reducing their severity and duration. Currently, infections are detected using a combination of traditional microbiology plating techniques; testing of blood biomarkers; and the use of x-rays. Visual cues such as erythema, pus-formation, and colour changes are also employed by clinicians to aid diagnoses of infections. Among the challenges associated with these laboratory techniques are time constraints, requirements for highly specialised personnel and low specificity due to the influences of co-morbidities. The detection of volatile organic compounds emitted from the wound bed provides a major opportunity for rapid non-invasive localised analysis of the wound environment. In this work, we describe a simple experimental workflow that allows a rapid turnover of data that could potentially provide same-day diagnostic information in clinics. 23 participants (26 wounds total, 15 infected; 11 non-infected) were included in this work. Solid phase microextraction coupled with gas chromatography mass spectrometry were used to extract and analyse a broad range of volatile compounds from the swab samples. Among the chemical classes detected were aldehydes, alcohols, sulfides, acids, hydrocarbons, and ketones were detected. Severity of the wound infections was associated with the detection of higher numbers of compounds and higher degrees of chemical diversity in the swab samples. Between infected and non-infected samples, significant differences in the detection of 5 acids were consistently observed, these were butanoic acid, 3-methyl-, butanoic acid, propanoic acid, propanoic acid, 2-methyl, and acetic acid. We have previously demonstrated that these acids are generated through microbial metabolism. We therefore hypothesize that the increased abundance of these acids in infected wound samples indicate increased microbial load in the respective wounds. The results of this ongoing work provide clear insight into the potential of volatilomics for future clinical applications.

5.1. Introduction

Ulceration is caused by abnormal pressure applied to the foot – which diabetic ischaemia renders the skin less able to withstand – resulting in a break in the skin[1]. It is estimated that around one in four people with diabetes will develop a diabetic foot ulcer (DFU) in their lifetime[2]. Infections of DFUs are associated with poor outcomes, a 12-month observational study reported that out of 299 participants, ulcers only healed in 136 participants (45.5%) but recurred in 13; lower extremity amputations (LEA) were recorded in 52 (17.4%) participants and revascularization surgery in 18 (6.0%) participants; 45 (15.1%) of the participants died[3]. Early detection of infections results in early and more effective therapeutic interventions.

Diagnosing a foot infection can be complex as assessing the presence or severity of an infection, and/or differentiating soft tissue from bone infections are challenging[4]. Clinicians use visual and olfactory cues such as erythema, pus formation, colour and smell to guide diagnoses. Inflammatory blood biomarkers such as white blood cell (WBC), erythrocyte sedimentation rate (ESR), C-reactive protein (CRP) and procalcitonin (PCT) are also routinely measured to aid the diagnosis of infections. Out of these, CRP and PCT blood levels have showed the most reliable correlations with infections. However, these markers do not specifically indicate infection, they instead are upregulated in response to inflammation. Therefore, the challenges with using these markers include differentiating between infection and other inflammatory processes; and accounting for the influence that anti-inflammatory and antibiotic drugs. The analysis of volatile compounds being emitted from the wound bed may provide a more targeted and non-invasive avenue for the diagnosis of infection.

Volatilomic profiling is an emerging technique that has been used for characterising microbial metabolomes and different diseases. The recent publishing of comprehensive review papers[5][6][7][8] and books[9] have brought microbial and clinical volatilomics into focus and have highlighted major opportunities. The major benefit of using volatilomics for potential clinical diagnoses is that volatile compounds - by their nature – can be non-invasively collected and rapidly analysed. Disease-associated volatilomic shifts have been reported for a variety of diseases and maladies, however, there is currently wide variation in the results across studies due to population differences and no standardised methods. The collective move towards standardisation of clinical volatilomic sampling and analysis protocols – as was done with microbiome studies - will rapidly elevate the field and promote clinical applicability. Volatilomic profiling of chronic wounds and healthy skin controls have highlighted discriminatory patterns from the wound samples[10]. Although this would be expected as there are significant differences in the microbial composition[11] and chemical environment[12] between wounds and healthy skin. The instability of the skin volatilome has been demonstrated by measurable shifts following minor barrier disruption via several rounds of tape stripping. Open wound beds are also regions of high oxidative stress[13], which can generate volatiles through the molecular breakdown of the cellular components of the surrounding tissue[14]. From a clinical stand point, non-infected wounds may serve as more suitable controls than healthy skin for volatile profiling. Non-infected wounds are wounds that show no visual or microbiological indications of infection. Discrimination of infected wound volatilomes and non-infected wound volatilomes has been demonstrated using e-nose[55]. Although this study highlights the volatilomic discrimination between infected and non-infected wound samples, the compounds responsible for that trend were not identified.

116

Other studies have investigated the volatiles emitted from dressings taken from cancer-associated fungative wounds. These studies reported dimethyl trisulfide as a major contributor in malodourous wounds and a potential indicator of bacterial infection[15][16]. Other compounds detected in these wounds included dimethyl disulfide, indole, 3-methylbutanal, and phenol[190].

In this study, volatilomic characterisation of swab samples taken from infected and non-infected wounds was carried out using headspace solid phase microextraction coupled with gas chromatography (HS-SPME-GCMS). Infected wounds were classified by professional clinicians using the University of Texas diabetic wound classification system[17]. The first aim of this work was to obtain volatile profiles for infected wounds of varying severity, and for non-infected wounds. Our second aim was to use untargeted data analysis methods to identify discriminative trends between the whole profiles of each group. Finally, we aimed to demonstrate clear differences in the abundances of individual compounds between infected and non-infected wounds.

5.2. Methods

5.2.1. Participant profile

Diabetic patients attending the Diabetic clinic at Connolly Hospital (Blanchardstown, Dublin) were recruited for the study by Dr. Tommy Kyaw (RCSI) and the podiatrist Brid Cooney. All participant personal data was randomised in accordance with GDPR guidelines. No specific dietary, hygiene, or cosmetic regimes were applied. Participants were informed on the aim and purpose of the study through a patient information leaflet. Following this, participants were asked to provide written informed consent on the day of sampling. Dublin City University Research Ethics Committee and Connolly Hospital Research Ethics Committee approved the volatilomic study of swab samples taken from diabetic wounds. Wound infection status was classified based on the University of Texas classification system (Table 1). Microbiology plating results and Texas scores for each participant are shown in Table S2.

Table 1: University of Texas Wound Classification System[17]

		Grade			
		0	1	2	3
Stage	A	Pre or postulcerative lesion completely epithelialised	Superficial wound, not involving tendon, capsule, or bone	Wound penetrating to tendon or capsule	Wound penetrating to bone or joint
	B	with infection	with infection	with infection	with infection
	C	with ischemia	with ischemia	with ischemia	with ischemia
	D	with infection and ischemia	with infection and ischemia	with infection and ischemia	with infection and ischemia

117

5.2.2. HS-SPME sampling

SPME fibers were used for sampling VOCs and consisted of 85 μm Carboxen/Polydimethylsiloxane Stableflex (2 cm) assemblies (Supelco Corp., Bellefonte, PA, USA). Swab samples were placed in HS vials and incubated at 37°C for 2 hours. Following incubation, a SPME needle was pierced through the septum of the HS vial, and the fibre was exposed to the HS of the sample for 20 min while agitated. Following this, the fibre was retracted and the SPME assembly removed from the vial. The SPME fibre was then inserted into the GC inlet and thermally desorbed at 250°C for 2 min for subsequent separation and detection by mass spectrometry.

Background subtraction was carried out by sampling blank headspace vials, and blank swab samples (n=10). Compounds recovered from these blank analyses were individually assessed. Compounds recovered samples with signal-to-noise ratios greater than 3:1 were considered for inclusion in the study.

5.2.3. Gas chromatography–mass spectrometry (GC-MS)

An Agilent 6890 GC connected to an Agilent 5973 mass selective detector (Agilent Technologies, Inc., Santa Clara, CA, USA) was used for all analyses. Separations were performed on a DB-WAX column (Agilent Technologies Ireland, Cork) (30 m × 0.25 mm × 0.32 μm). The carrier gas used was helium, with a constant flow rate of 1.3 mL /min For manual injections of SPME fibers, the system was equipped with a SPME Merlin Microseal (Merlin Instrument Company, Newark, DE, USA), and the inlet was maintained at a temperature of 250°C. Split-less injection was used for all samples, with a gas purge being activated after 2 min. Each SPME fibre was desorbed for 2 min within a SPME inlet liner (Supelco). The initial GC oven temperature was 40°C for 5 min and was programmed to increase at a rate of 10°C min^{-1} to 240°C, with a final hold for 5 min at this temperature, giving an overall running time of 29 min. The transfer line temperature was set at 230°C. The MS was operated at a scan range of 35-400 m/z, scan rate of 3.94 s^{-1}, ion source temperature 230°C and ionising energy of 70 eV. Identification of compounds relied on a three phase protocol whereby National Institute of Standards and Technology (NIST) library (2017) - match factors of >70% were initially used to assess potential ID matches; fragmentation patterns of potential matches were then manually interpreted before being validated using retention index matching. Retention index (RI) values for polar columns provided by the NIST Chemistry WebBook, SRD 69, was used to support the identification of these compounds. Any compound found to have an RI value ≤12 RI units of the RI values found in the NIST database were deemed acceptable matches. A external standard mixture of saturated alkanes (C$_7$-C$_{30}$; Merck, Cork, Ireland) was injected into the GC-MS under the same temperature conditions as the samples and used for RI matching. This was done by rapidly dipping an exhausted SPME needle into the mixture once and injecting it into the GC-MS. A fully functional SPME fiber was not used for this because exposure to hexane degrades the fiber integrity.

5.2.4. Data Analysis

Agilent MassHunter Qualitative Analysis 10.0 software was used to analyse raw chromatographic data. Peak acquisition and the respective peak area data were calculated by employing the chromatogram deconvolution compound mining algorithm. Chromatographic peaks were compared using the NIST Chemistry WebBook. Peaks found to be from exogenous

118

sources such as the SPME fiber, glass vial, and column were removed from the dataset. Only peaks that could be accurately identified and that were detected in over one replicate sample were included in the final peak list. R (version 1.2.5033) was used for data exploration and visualisation. Raw bacterial VOC data was standardised using centering and scaling [18]. Centering converts all the values in the dataset to fluctuations around zero rather than fluctuations around the mean VOC abundance. It adjusts for differences in the offset between low and high abundances. Scaling converts the values in the dataset into ratios relative to the difference in abundances between the VOCs, which allows each VOC to be equally represented in the subsequent data analysis. For compounds that were present in some replicate samples (of a given strain in a given media) and absent from others, these missing values were imputed as zero. For compounds that were absent from all replicates (of a given strain in a given media), these missing values remained missing values. Hierarchical clustering and principal component analysis (PCA) were carried out on the dataset using the R packages: 'FactoMineR' (version: 2.4), 'factoextra' (version: 1.0.7), 'pheatmap'(version: 1.0.12), 'egg' (version: 0.4.5) and 'cluster' (version:2.1.0). For the hierarchical clustering analysis, Euclidean distance was used as the measure of (dis)similarity. Other R packages used for the graphics in this study were: 'tidyverse' (version:1.3.1), 'ggplot2' (version: 3.3.5), 'ggfortify' (version:0.4.12) .

5.3.   Results

5.3.1. The wound volatilome

Following background subtraction, a total of 42 compounds were deemed suitable for inclusion in the study. These included alcohols, aldehydes, acids, sulfides, ketones, esters and hydrocarbons. Detection frequencies for all other compounds are shown in Figure 5.1. The most frequently detected compounds were acetic acid, acetone, and ethanol. Low abundances of these compounds were also detected in various blank swab samples. Other compounds with a high detection frequency across infected samples in particular were butanoic acid, propanoic acid, propanoic acid, 2-methyl; butanoic acid, 3-methyl-; decane, undecane and phenylethyl alcohol. Various other alcohols, acetates, acids, sulfides, esters and ketones were detected less frequently across the samples, these are summarised in the heatmap in Figure 5.1. The heatmap shown in Figure 5.1 illustrates the chemical diversity of the volatile compounds recovered from 15 infected and 11 non-infected wound samples. Similar to the heatmaps that have been employed in Chapter 2, 3 and 4; this heatmap utilises hierarchical clustering to identify and cluster similar samples based on their volatilomes. Hierarchical clustering was also applied to the rows (compounds) of the heatmap to allow a clear visualisation of the compounds responsible for the discrimination of the swab samples. The results are visualised as dendrograms. The length of an leg in a dendrogram between a cluster and its split is proportional to the dissimilarity (Euclidean distance) between the split clusters[19]. Samples with the highest degree of dissimilarity were P19, P3, P4, P15, P22, P23, P10, P21 and P9. All of these samples except for P9 were collected from infected wounds. Out of these samples, P4, P15 and P19 were classified as 3B on the Texas score (Table 1) indicating severe infection penetrating the bone or joint; P21 and P22 were classified as 1D indicating superficial infected wound with ischemia; P9 was classified as 1A indicating a superficial non-infected wound; and P3 and P10 were classified as non-diabetic infected wounds. Using the Texas score system as a guide, it can be seen that a variety of acids, alcohols, ketones and

119

sulfides separate the severely infected wounds from the less infected (1B) and non-infected wounds. The compound-level comparative boxplots shown in the Supplementary information (Figure S5.2 – S5.7) further illustrate this significant discrimination. Non-infected samples were characterised by relatively less diverse and abundant volatilomes, which is visualised by the relatively short edges of the dendrogram (Figure 5.1). Compounds that were consistently detected in these samples included nonanal, octanal, ethanol, acetone, benzyl alcohol, and phenol. Less infected samples with a Texas score of 1B or 2B are less represented in Figure 5.1 compared to the 3B and 1D samples due to significant differences in the abundance and diversity of compounds detected. However, discriminative compounds were still detected in these samples and for this reason it was necessary to investigate these samples at the compound-level.
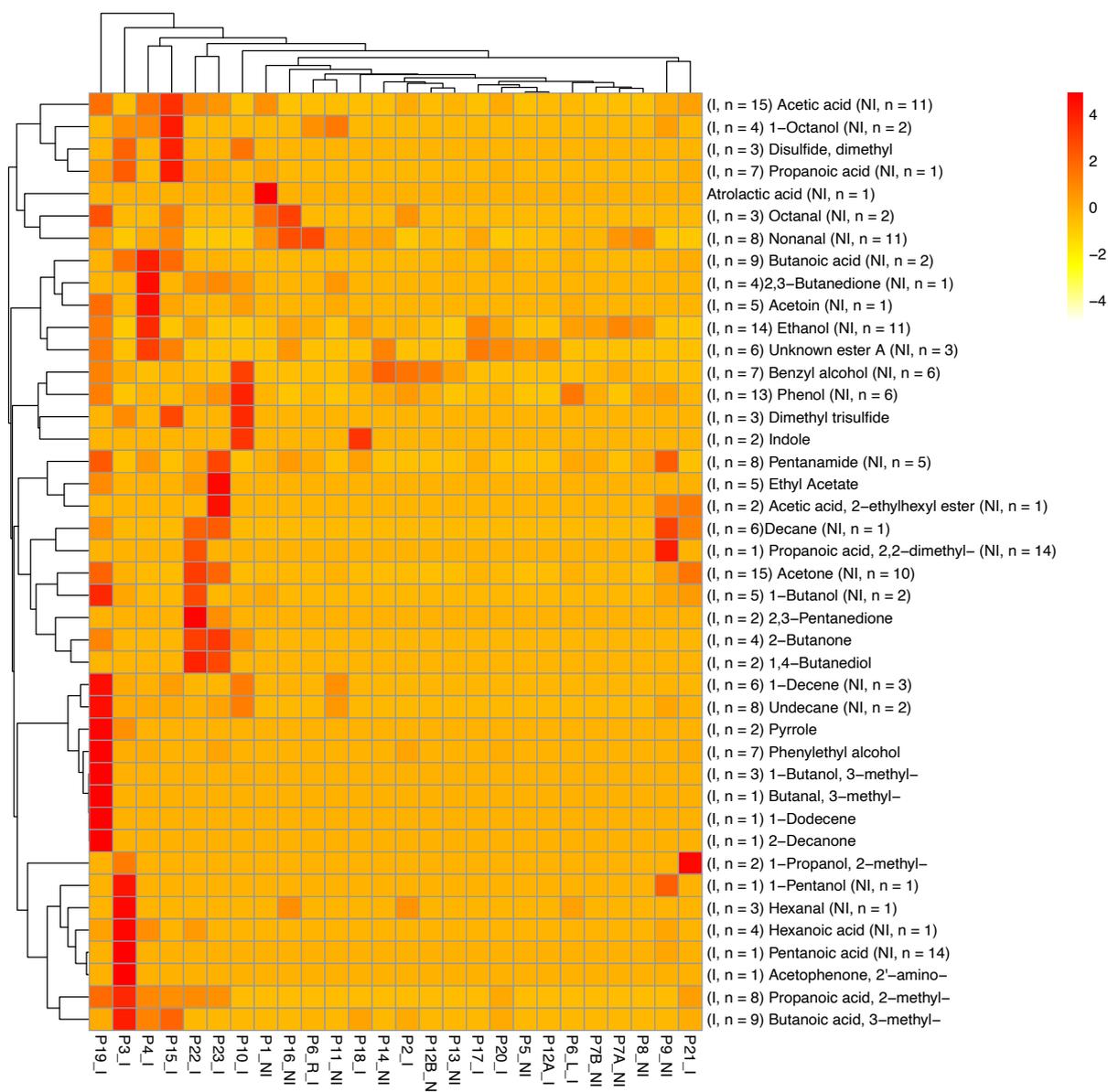

Figure 5.1: *Heatmap plot illustrating hierarchical clustering characterisation of swab samples taken from infected (n=15) and non-infected (n=11) wounds. Each column represents the mean abundance values from each sample. Abundance values were auto-scaled across each*

row. Each row represents the mean abundance values from each detected compound. The detection frequency of each compound in infected (I) and non-infected (NI) samples is given alongside each compound name.

5.3.2. Compound-level abundance differences between infected and non-infected wounds

Multiple significant differences in the recovery of specific compounds were observed across infected and non-infected wound samples. The grouped boxplots shown in Figure 5.3 illustrate the most notable infection-specific differences in the detected compounds across the analyses of 26 wound samples. Detection frequency of these compounds was also provided in Figure 5.3 at the top of each plot. The clearest and most frequent difference between infected and non-infected samples was the number and abundance of short-chain fatty acids detected in them. Infected samples had significantly higher abundances acetic acid, propanoic acid, propanoic acid, 2-methyl, butanoic acid, and butanoic acid, 3-methyl- . From Figure 5.3, it can be seen that for the majority of these short-chain fatty acids, the differences in abundance and diversity correlated with infection severity (characterised by the Texas score – Figure 5.3) of the wounds. Relatively low abundances of acetic acid were detected across all non-infected samples and in some blank swab samples. Butanoic acid, propanoic acid and hexanoic acid were also detected in relatively low abundance across a low number of non-infected samples. The significant differences in the abundance of short-chain fatty acids across the infected and non-infected wound samples are clearly illustrated in the acid-specific box plots shown in Figure 5.4. Among the other frequently detected compound-level differences across these samples were the detection of decane, undecane, and 1-decene – all of which were detected at a higher frequency and higher abundance in infected samples. The detection of compounds such as acetoin, 2-butanone, dimethyl disulfide, dimethyl trisulfide, phenylethyl alcohol, 1-butanol, 3-methyl- , 1-propanol, 2-methyl-, indole, and ethyl acetate were clear discriminators of infected wound samples. Although these compounds were individually detected at a relatively low frequency, they were detected across a wide range of infected samples while being comparatively absent in non-infected samples. As a result, in the case of many of these compounds the differences observed across the infected and non-infected samples were highly significant (indicated by **** : $p < 0.0001$ - Figure 5.3). It must be noted that these preliminary results are currently limited by a low number of samples and the differences presented here may not accurately represent the universal volatilomic differences between infected and non-infected wounds. However, as this work is currently ongoing and the number of samples being analysed increases, the observed statistical differences will be subject to change and – as a result - illustrate a higher degree of accuracy.

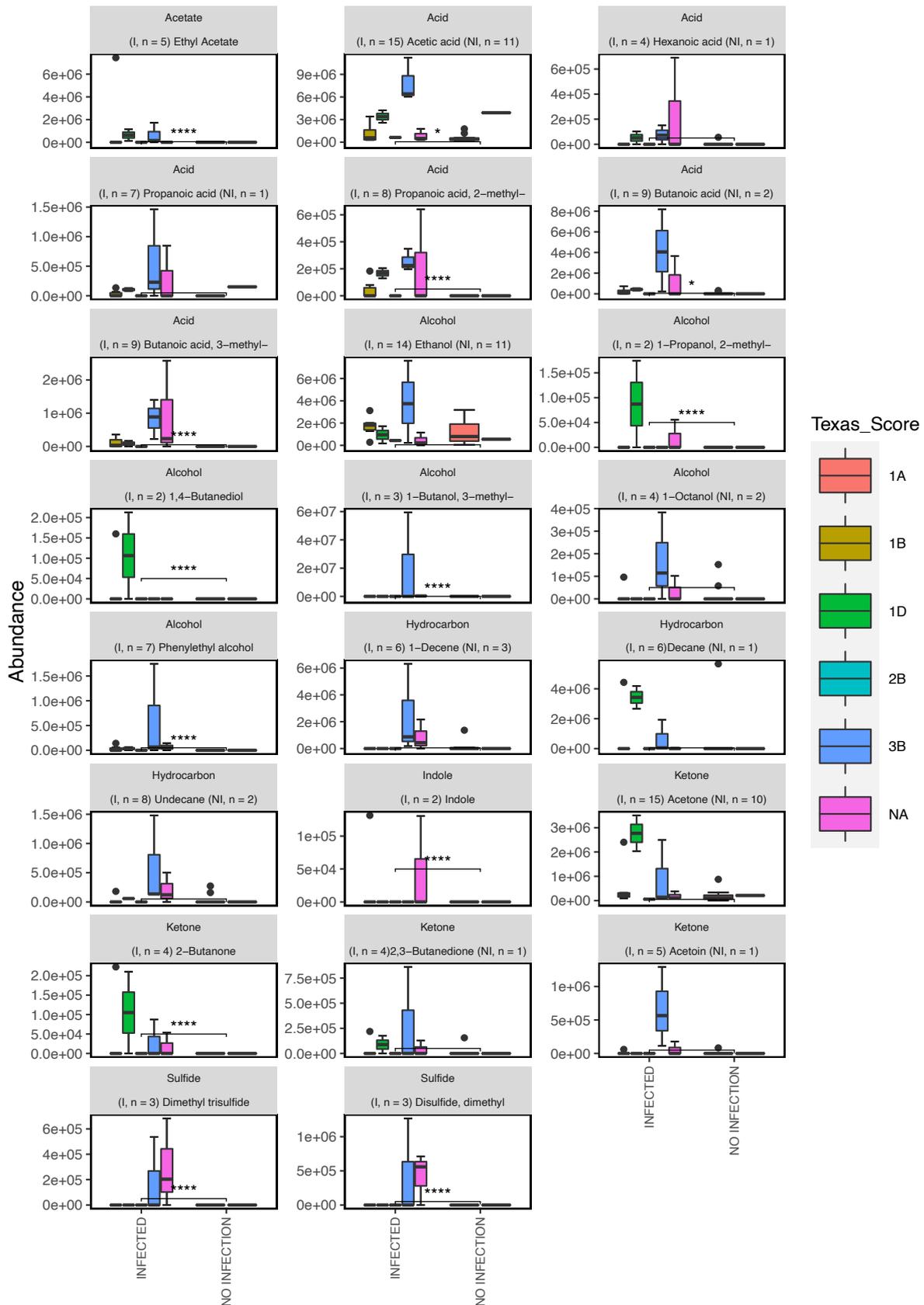Pre-examination copy
Submitted 27th May 2022

Figure 5.2 : Comparative boxplots illustrating differences in the abundance of various key compounds across infected and non-infected wound samples. Each plot illustrates the abundance of an individual ketone compound recovered from varying wounds classifications

*(Texas score). The Texas score is a common classification system used for diabetic ulcers and is visualised through the various colours of the boxes: 1A (non-infected superficial wound; RED); 1B (infected superficial wound; GREEN); 2B (infected wound penetrating tendon; BLUE); 3B (infected wound penetrating to bone or joint; PINK); NA (infected and non-infected, non-diabetic wounds; GREY). The detection frequency for each compound across the samples is shown above each plot alongside the name of each compound where I = infected and NI = non infected. Statistically significant differences are illustrated through the star system where * = p < 0.05; ** = p < 0.01; *** = p < 0.001; **** = p < 0.0001.*
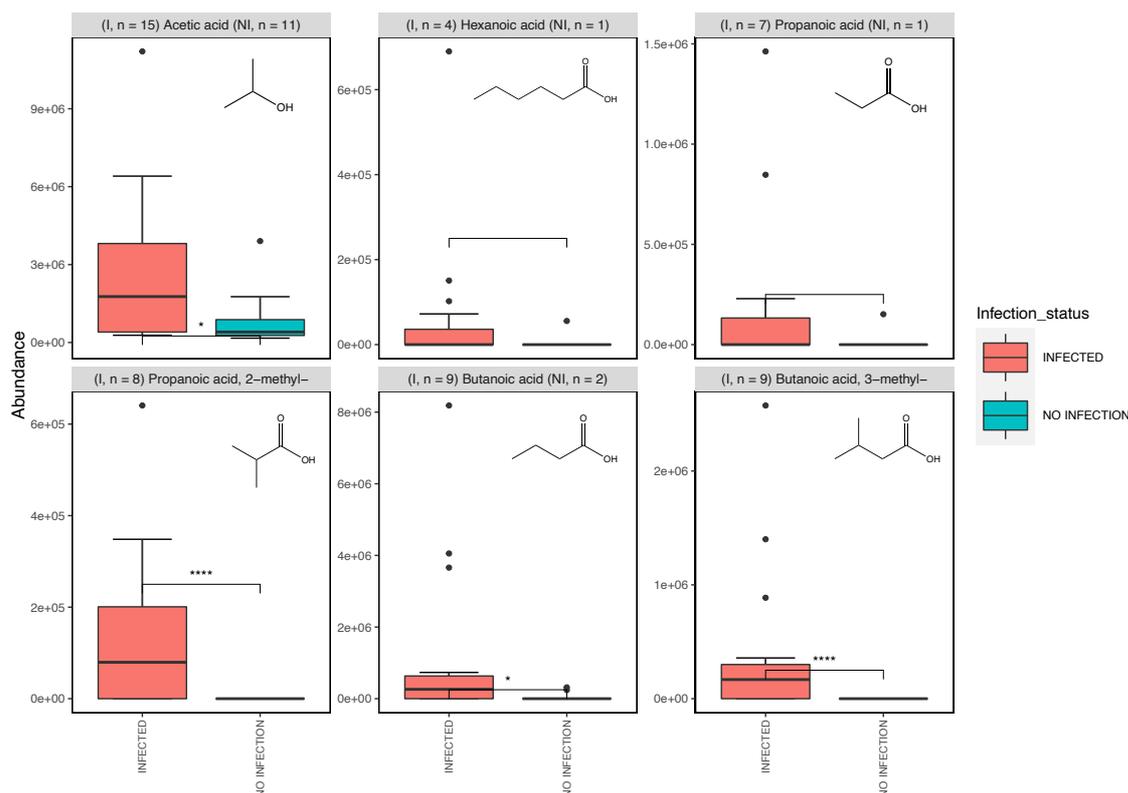

Figure 5.3 : Comparative box-plots illustrating significant differences in the recovery of six short-chain fatty acids across infected (red, n=15) and non-infected (blue, n=11) samples.

5.4.   Discussion

In this work, the volatilome of wound swab samples was investigated using an untargeted analytical workflow to identify discriminative infection-specific shifts. Wound infections are traditionally diagnosed in clinics using a combination of plating methods, blood tests, x-rays and visual cues. Blood biomarkers such as C-reactive protein, white blood cell counts, erythrocyte sedimentation rate and procalcitonin are frequently measured as indirect indicators of infection. The accuracy of these biomarkers for detecting infections is highly variable as they are also influenced by co-morbidities and the use of medications (antibiotics, anti-inflammatory drugs, etc.). Conventional plating methods are used to identify causative pathogens while x-rays are typically employed to confirm whether the infection has reached the bone (osteomyelitis). These methods are labour- and time-intensive and are require highly specialised personnel to carry them out. Rapid non-invasive volatilomic discrimination

of swab samples taken from infected and non-infected wounds was recently demonstrated using electronic-nose sampling[20]. Although this study demonstrated that infection-specific volatilomic differences exist, the compounds responsible for this discrimination were not identified. The aim of this work was to characterise the volatilome of swab samples taken from wounds of varying severity using HS-SPME-GCMS.

Wound severity was classified using the University of Texas diabetic wound classification system[17]. This system consists of a 4 x 4 matrix, according to depth (Grade 0,1,2,3) and presence of infection (Stage B), ischaemia (Stage C), or both (Stage D)[21]. Non infected wounds with a classification of 1A typically demonstrated low volatilomic diversity (Figure 5.2). In contrast, infected wounds with a classification of 3B showed high volatilomic diversity and abundances of compounds (Figure 5.2 and 5.3). Following background subtraction, in 26 wounds samples, a total of 42 compounds were included and classified as compounds-of-interest. These compounds had widely varying prevalence across these samples which is illustrated in Figure 5.1 and Figure 5.2. Severe wound infections have previously been associated with high species- and strain-level diversity[22]. In these cases, the diverse microbiome could potentially be the primary origin of the observed volatilomic diversity. Open wound beds are also regions of high oxidative stress[23], which can generate volatiles through the molecular breakdown of the cellular components of the surrounding tissue[24]. Among the most frequently detected compounds across both infected and non-infected samples were acetone, acetic acid, and ethanol. Although these compounds are produced by microbes through the metabolism of sugar and lipids[17], they were also detected in relatively low abundances in blank samples. Short-chain fatty acids were frequently detected in infected wound samples (Figure 5.1 and 5.2) and allowed the discrimination of severely infected wounds (Figure 5.3). The production of this class of compounds has been previously linked to antiphagocytic activity that impairs wound healing[25]. Interestingly, despite the significant differences in volatile fatty acid abundance between the infected and non-infected wounds, chronic wounds have been associated with an increase in pH of the wound environment[26][27]. Although knowledge around this phenomenon is limited[27], it is suspected that the alkalinity is caused by the liberation of ammonia from the breakdown of urea[28]. Urease enzymes hydrolyse urea to form ammonia and carbon dioxide – both of which have been recently targeted for the development of non-invasive sensors of infection based on pH[29] and CO_2 abundance[30]. Bacteria liberate amino acids and lipids from the biodegradation of necrotic tissue[31] which then act as the metabolic precursors for volatile metabolites. The presence of dimethyl trisulfide was reported to be the cause of malodour in cancer-associated fungating wounds[189]. In this work, the presence of dimethyl disulfide and dimethyl trisulfide clearly correlated with wound infections, however, these compounds were uncommon and were only detected in three participants – all of which had infected wounds (Figure 5.2). These compounds can arise from the oxidation of methanethiol[32] - a highly common compound that is associated with decaying biomass. Sulfur-containing volatile compounds can also be generated through the metabolism of the sulfur-containing amino acids, cysteine and methionine[33]. However, currently due to their detection infrequency, we do not propose these compounds as optimal target indicators of wound infections. Other infrequently detected but clear discriminators of infected and non-infected samples were the primary metabolites : ethyl acetate, 2-butanone, and acetoin (Figure 5.3). These compounds are

124

frequently emitted wound-associated pathogens[34][35] and are derived from the downstream metabolism of glucose[5].

This work is currently limited by a low number of participants which presents several challenges including: establishing baseline VOC abundances for the control group; accurately representing varying degrees of infection severity; refining the inclusion criteria for compounds; and validating potential target compounds. However, this work will be continuing into the future and we plan to recruit a further 50-75 patients. In addition to this, a longitudinal study that profiles the wounds of patients over the course of 6-12 months could potentially be highly beneficial to determine volatilomic shifts in response to wound healing or degradation. Although our current work clearly demonstrates the chemical diversity of the wound volatilome, swabbing at a single time point only provides a snap shot of the status of each wound. Despite this, the preliminary aims of this work were achieved in that specific compounds were identified that discriminate infected from non-infected wounds.

5.5.   Conclusion

In this work swab samples taken from infected and non-infected wounds were comprehensively analysed for the first time using HS-SPME-GCMS. The aims of this work were to identify discriminate trends between infected and non-infected wound samples; and then to identify the compounds responsible for these observed trends. Currently, there are a total of 42 compounds included in this work, this number will increase as the number of participants increases in the future. Non-infected wound samples typically had low volatilomic activity in contrast to infected samples. The degree of chemical diversity and compound abundance was closely associated with the severity of the infection, this was clearly demonstrated in the patients with osteomyelitis (Texas score: 3B). Samples taken from patients with less severe infections had varying degrees of chemical diversity and abundance. However, the abundance of short chain fatty acids was significantly more abundant in samples taken from less and severely infected patients compared to non-infected patients. This class of compounds will be a potential target group as this study progresses into the future. There were also a variety of ketones, alcohols, acetates and hydrocarbons that were detected less frequently in infected samples that have been previously determined to be microbial metabolites. This chemical diversity has again demonstrated the high metabolic activity of infected wounds. Another important point is that the microbes are breaking down necrotic tissue are utilising fundamental pathways to liberate and breakdown amino acids, lipids, and carbohydrates to produce the resulting metabolites. Our future work on this study will see the recruitment of 50-75 additional participants. This will allow more comprehensive conclusions to be made about this volatilomic data and potentially lead to further research avenues. Currently this preliminary data demonstrates the clear potential for the use of volatilomics for non-invasive wound infection diagnoses in the future.

5.6.   References

[1] Jeffcoate, W. and Harding, K. (2003). Diabetic foot ulcers. *The Lancet*, 361(9368), pp.1545-1551.

[2] Armstrong, D., Boulton, A. and Bus, S. (2017). Diabetic Foot Ulcers and Their Recurrence. *New England Journal of Medicine*, 376(24), pp.2367-2375.

[3] Ndosi, M., Wright-Hughes, A., Brown, S., Backhouse, M., Lipsky, B., Bhogal, M., Reynolds, C., Vowden, P., Jude, E., Nixon, J. and Nelson, E. (2017). Prognosis of the infected diabetic foot ulcer: a 12-month prospective observational study. *Diabetic Medicine*, 35(1), pp.78-88.

Pre-examination copy
Submitted 27th May 2022

[4] Lipsky, B., Aragón-Sánchez, J., Diggle, M., Embil, J., Kono, S., Lavery, L., Senneville, É., Urbančič-Rovan, V., Van Asten, S., Peters, E., Abbas, Z. and Malone, M. (2019). IWGDF guidance on the diagnosis and management of foot infections in persons with diabetes. *IWGDF*.

[5] Weisskopf, L., Schulz, S. and Garbeva, P., (2021). Microbial volatile organic compounds in intra-kingdom and inter-kingdom interactions. *Nature Reviews Microbiology*,. https://doi.org/10.1038/s41579-020-00508-1

[6] Elmassry, M. & Piechulla, B. (2020). Volatilomes of bacterial infections in humans. *Front. Neurosci.* https://doi.org/10.3389/fnins.2020.00257

[7] Kai, M., (2020). Diversity and Distribution of Volatile Secondary Metabolites Throughout Bacillus subtilis Isolates. *Frontiers in Microbiology*, 11. https://doi.org/10.3389/fmicb.2020.00559 .

[8] Schulz, S. & Dickschat, J. (2007). Bacterial volatiles: The smell of small organisms. *ChemInform.* https://doi.org/10.1039/b507392h

[9] Ryu, C., Weisskopf, L. and Piechulla, B., (2020). *Bacterial volatile compounds as mediators of airborne interactions*. Singapore: Springer.

[10] Thomas, A., Riazanskaia, S., Cheung, W., Xu, Y., Goodacre, R., Thomas, C., Baguneid, M. and Bayat, A. Novel noninvasive identification of biomarkers by analytical profiling of chronic wounds using volatile organic compounds. *Wound Repair and Regeneration*, 18(4), pp.391-400. https://doi.org/10.1111/j.1524-475X.2010.00592.x. (2010)

[11] Gontcharova, V., . A Comparison of Bacterial Composition in Diabetic Ulcers and Contralateral Intact Skin. *The Open Microbiology Journal*, 4(1), pp.8-19. https://doi.org/10.2174/1874285801004010008 . (2010).

[12] Schneider, L., Korber, A., Grabbe, S. and Dissemond, J., Influence of pH on wound-healing: a new perspective for wound-therapy?. *Arch. Dermatol. Res.*, 298(9), pp.413-420. https://doi.org/10.1007/s00403-006-0713-x (2006).

[13] Wlaschek, M. and Scharffetter-Kochanek, K.,. Oxidative stress in chronic venous leg ulcers. *Wound Repair and Regeneration*, 13(5), pp.452-461. https://doi.org/10.1111/j.1067-1927.2005.00065.x (2005).

[14] Broza, Y., Mochalski, P., Ruzsanyi, V., Amann, A. and Haick, H., Hybrid Volatolomics and Disease Detection. *Angew. Chem. International Edition*, 54(38), pp.11036-11048 https://doi.org/10.1002/anie.201500153 . (2015).

[15] SHIRASU, M., NAGAI, S., HAYASHI, R., OCHIAI, A. and TOUHARA, K., 2009. Dimethyl Trisulfide as a Characteristic Odor Associated with Fungating Cancer Wounds. *Bioscience, Biotechnology, and Biochemistry*, 73(9), pp.2117-2120.

[16] Thuleau, A., Dugay, J., Dancremont, C., Jemmali, Z., Elard, J., Ricke, Y., Cassoux, N., Watson, S., Escande, M. and Fromantin, I., 2018. Volatile Organic Compounds of Malignant Breast Cancer Wounds: Identification and Odors. *INDEX WOUNDS*, 30(11), pp.337-344.

[17] Lavery, L., Armstrong, D. and Harkless, L., 1996. Classification of diabetic foot wounds. *The Journal of Foot and Ankle Surgery*, 35(6), pp.528-531.

[18] A van den Berg, R., CJ Hoefsloot, H., A Westerhuis, J., K Smilde, A. and J van der Werf, M., (2006). Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics*, 7(142). https://doi.org/10.1186/1471-2164-7-142.

[19] Wittek, P. *Unsupervised Learning In Quantum Machine Learning* 1st edn, 60–61 (Elsevier Science, New York, 2014).

[20] Haalboom, M., Gerritsen, J. and van der Palen, J., Differentiation between infected and non-infected wounds using an electronic nose. *Clin. Microbiol. Infect.*, 25(10), pp.1288.e1-1288.e6. https://doi.org/10.1016/j.cmi.2019.03.018 (2019).

[21] Monteiro-Soares, M., Russell, D., Boyko, E., Jeffcoate, W., Mills, J., Morbach, S. and Game, F., 2020. Guidelines on the classification of diabetic foot ulcers (IWGDF 2019). *Diabetes/Metabolism Research and Reviews*, 36(S1).

[22] Kalan, L., Meisel, J., Loesche, M., Horwinski, J., Soaita, I., Chen, X., Uberoi, A., Gardner, S. and Grice, E., Strain- and Species-Level Variation in the Microbiome of Diabetic Wounds Is Associated with Clinical Outcomes and Therapeutic Efficacy. *Cell Host Microbe*, 25(5), pp.641-655.e5. https://doi.org/10.1016/j.chom.2019.03.006 (2019).

[23] Wlaschek, M. and Scharffetter-Kochanek, K.,. Oxidative stress in chronic venous leg ulcers. *Wound Repair and Regeneration*, 13(5), pp.452-461. https://doi.org/10.1111/j.1067-1927.2005.00065.x (2005).

[24] Broza, Y., Mochalski, P., Ruzsanyi, V., Amann, A. and Haick, H., Hybrid Volatolomics and Disease Detection. *Angew. Chem. International Edition*, 54(38), pp.11036-11048 https://doi.org/10.1002/anie.201500153 . (2015).

[25] Bowler, P., Duerden, B. and Armstrong, D., 2001. Wound Microbiology and Associated Approaches to Wound Management. *Clinical Microbiology Reviews*, 14(2), pp.244-269.

[26] Gethin, G., 2007. The significance of surface pH in chronic wounds. *Wound healing Science*, 3(3).

[27] Jones, E., Cochrane, C. and Percival, S., 2015. The Effect of pH on the Extracellular Matrix and Biofilms. *Advances in Wound Care*, 4(7), pp.431-439.

[28] Wallace, L., Gwynne, L. and Jenkins, T., 2019. Challenges and opportunities of pH in chronic wounds. *Therapeutic Delivery*, 10(11), pp.719-735.

[29] Qin, M., Guo, H., Dai, Z., Yan, X. and Ning, X., 2019. Advances in flexible and wearable pH sensors for wound healing monitoring. *Journal of Semiconductors*, 40(11), p.111607.

[30] Yusufu, D., Magee, E., Gilmore, B. and Mills, A., 2022. Non-invasive, 3D printed, colourimetric, early wound-infection indicator. *Chemical Communications*, 58(3), pp.439-442.

[31] King, J., KOERBER, A., Croft, J., Ward, J., Williams, P. and Sockett, R., 2003. Modelling host tissue degradation by extracellular bacterial pathogens. *Mathematical Medicine and Biology*, 20(3), pp.227-260.

[32] Schäfer, H. and Eyice, Ö., 2019. Microbial Cycling of Methanethiol. *Current Issues in Molecular Biology*, pp.173-182.

Pre-examination copy
Submitted 27th May 2022

[33] Landaud, S., Helinck, S. and Bonnarme, P., 2008. Formation of volatile sulfur compounds and metabolism of methionine and other sulfur compounds in fermented food. *Applied Microbiology and Biotechnology*, 77(6), pp.1191-1205.https://doi.org/10.1007/s00253-007-1288-y

[34] Fitzgerald, S., Holland, L. and Morrin, A., 2021. An Investigation of Stability and Species and Strain-Level Specificity in Bacterial Volatilomes. *Frontiers in Microbiology*, 12. https://doi.org/10.3389/fmicb.2021.693075

[35] Fitzgerald, S., Duffy, E., Holland, L. and Morrin, A., (2020). Multi-strain volatile profiling of pathogenic and commensal cutaneous bacteria. *Scientific Reports*, 10(1). https://doi.org/10.1038/s41598-020-74909-w .

Pre-examination copy
Submitted 27th May 2022

Conclusion and Future work
Shane Fitzgerald

In the last 4 years, with the support of my supervisors, Dr. Aoife Morrin and Dr. Linda Holland, I have explored microbial and clinical volatilomic profiling experimental and data analysis workflows. The published sections of this work have primarily focused on examining the factors surrounding microbial volatile emission. In these studies, microbial volatilomes were comprehensively characterised across varied conditions and discriminated from each other. These studies provided us with an understanding of the fundamental dynamics behind the microbial emission of volatile metabolites. Microbial volatilomes are highly dependent on species- and strain-level specificities, the respective growth phase of the cells, nutritional environment of the cells, and the whether the cells are growing planktonically or in a biofilm. The volatilomic profiling of planktonic cultures allows the potential source of compounds detected in clinical systems to be determined. It therefore serves a critical purpose in the accurate interpretation of clinical volatilomic data involving infectious diseases. This was clearly observed in the preliminary clinical work described in Chapter 5, as a variety of microbial metabolites that were detected in Chapter 2, 3, and 4 were also detected in the clinical samples.

The preliminary clinical work described in Chapter 5 will continue into the future. Currently there are 23 patients included in this work, 15 of which were infected and 11 of which were not infected at the time of sampling. Our current aim is to recruit a further 50 - 75 participants with a range of infection severity. This will allow a higher degree of accuracy to be achieved for the baseline abundances of compounds detected from non-infected controls. In addition to this, as the data becomes more robust, potential target compounds may emerge and the inclusion criteria for compounds can be refined. Currently, short chain fatty acid abundances significantly discriminate infected samples from non-infected samples. Further work is required to validate this observed trend. Although there have been wound volatilomic studies carried out previously (See Section 1.2.3.2. Chapter 1), we believe that we are the first to demonstrate the chemical diversity of the wound volatilome. This preliminary work could also potentially give rise to numerous follow up studies. Firstly, a longitudinal study that follows patients with chronic wounds over the course of 6 – 18 months would be of significant interest. Depending many factors, over the course of these periods, patients with chronic wounds acquire and lose infections. This provides a unique opportunity to profile the volatile output of these wounds at different stages of healing or degradation. The state of a given wound will then be compared against itself at a different time rather than with a wound from another individual. A comprehensive picture of this process over time will allow the identification of biomarkers of infection with a higher degree of certainty. Currently, we are limited in that we have identified compounds associated with wound infections. If these individuals were profiled over time it would provide more context on the potential mechanisms at work and the potential role these compounds play in the infected wound. If volatile wound infection biomarkers are identified and validated in the future. The next step would be towards quantifying these biomarkers via direct mass spectrometric methods such as SIFT-MS or PTR-MS, or a modified GC-MS method. Determining threshold concentration ranges for compounds at varying degrees of infection severity would be a long term goal.

However, this would require a significant number of samples to validate and would require wide collaboration across the world to account for population differences.

Another area of interest that has been generated from this work is the volatilomic profiling of microbial biofilms. From our work with Chapter 4, we determined that the volatilomic output of *C. parapsilosis* shifted when it formed a biofilm. The emission of primary metabolites was reduced and there was indicators that the metabolism of substrates such as aromatic amino acids was upregulated from the high abundances of phenylethyl alcohol. I would like to investigate the factors surrounding the biofilm-specific volatile metabolites of highly prevalent clinical pathogens such as Methicillin resistant *Staphylococcus aureus* (MRSA) versus that of Methicillin sensitive *Staphylococcus aureus* (MSSA). There are a wide variety of biofilm volatilomic experiments that could be carried out using our novel sampling set up described in Chapter 4. These include: examining the effect of varying concentrations of antibiotics on the volatile output of specific biofilm-positive pathogens; the volatile output of microbial cocultures; and identifying biofilm-specific metabolites across different microbial species. The volatilomic analysis of microbes in human extract media such as blood-based, keratinocyte-based, or other biological fluids derived from humans will also provide highly valuable and unique insights into the metabolic capabilities of pathogens when they are exposed to these environments. This could potentially improve the chances of identifying target pathogen-specific metabolites that aid diagnoses and enable the development of targeted volatilomic workflows. This future work could contribute to a new layer to microbial metabolomics and the study of infectious diseases.

130

Appendix

Chapter 2: Multi -strain volatile profiling of pathogenic and commensal cutaneous bacteria – Supplementary Information

Shane Fitzgerald[1], Emer Duffy[1], Linda Holland[2], Aoife Morrin*[1]
School of Chemical Sciences, National Centre for Sensor Research, Insight SFI Research Centre for Data Analytics, Dublin City University, Ireland
School of Biotechnology, Dublin City University, Ireland
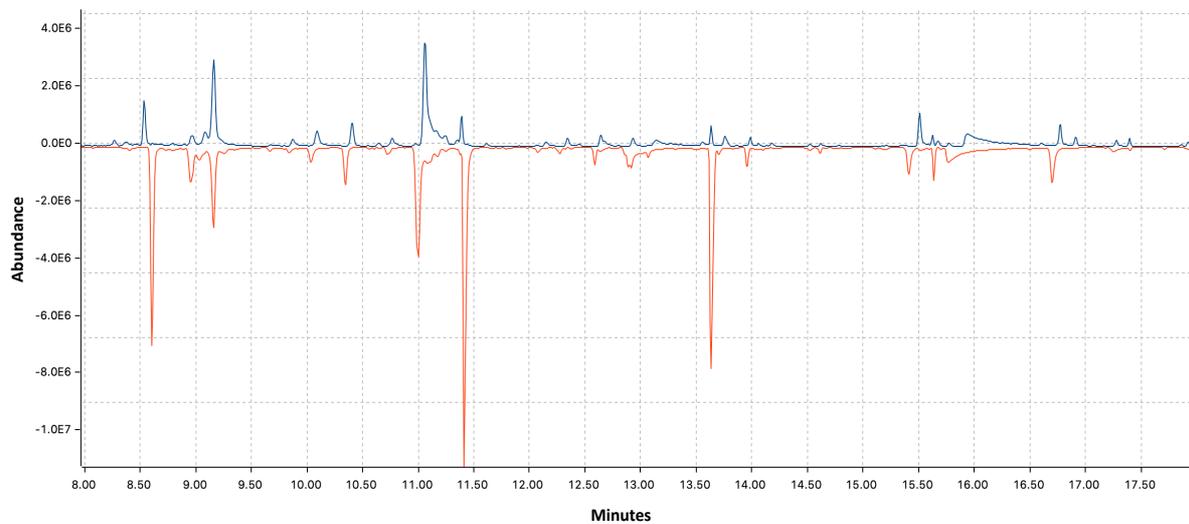
*aoife.morrin@dcu.ie

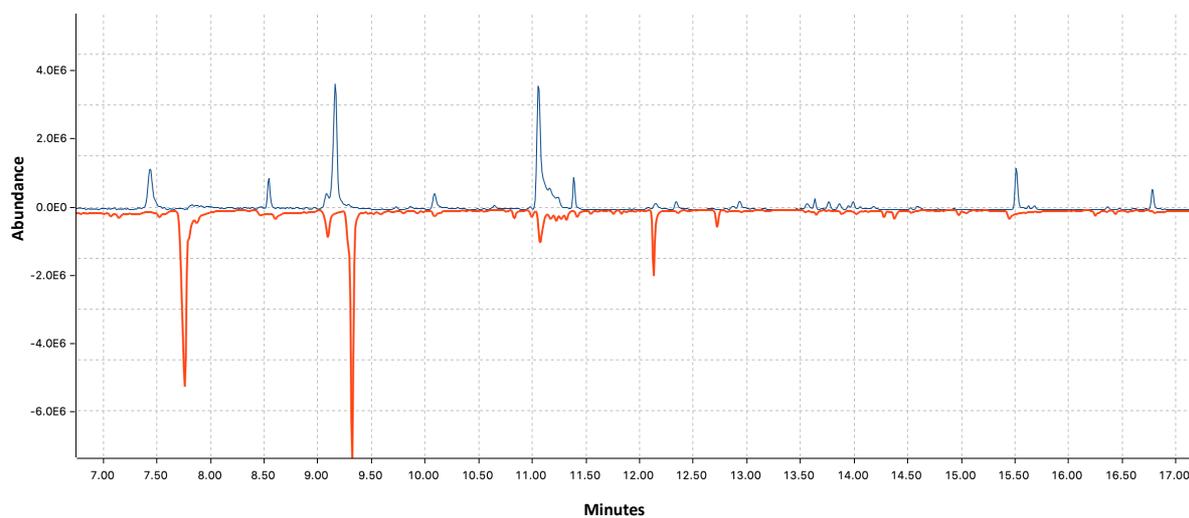*Figure S2.1: Overlaid chromatograms of SA.A (blue, top) and SA.B (red, bottom)*

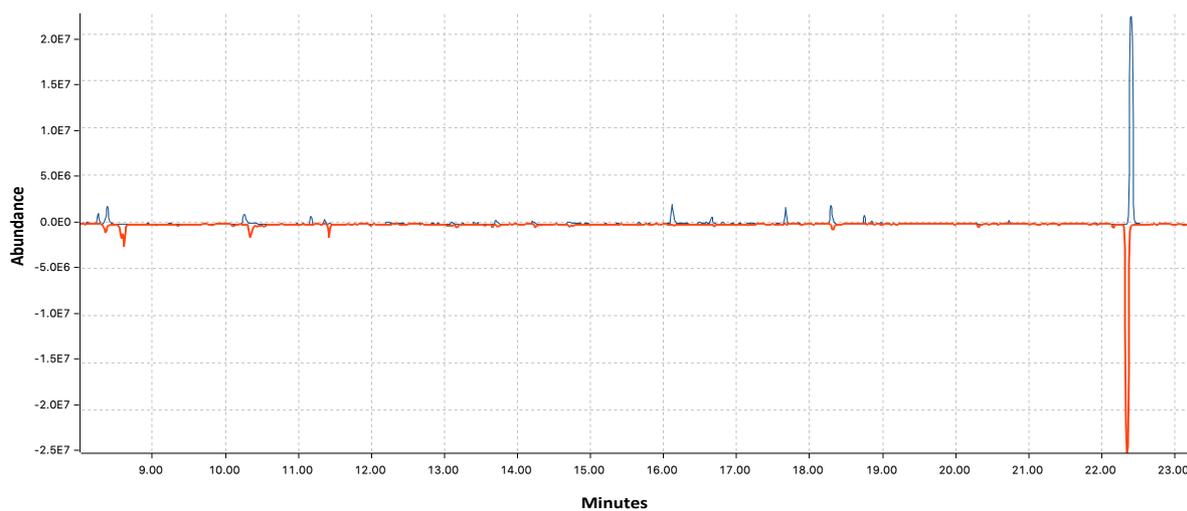*Figure S2.2: Overlaid chromatograms of PA.A (blue, top) and PA.B (red, bottom)*

Figure S2.3: Overlaid chromatograms of EC.A (blue, top) and EC.B (red, bottom)
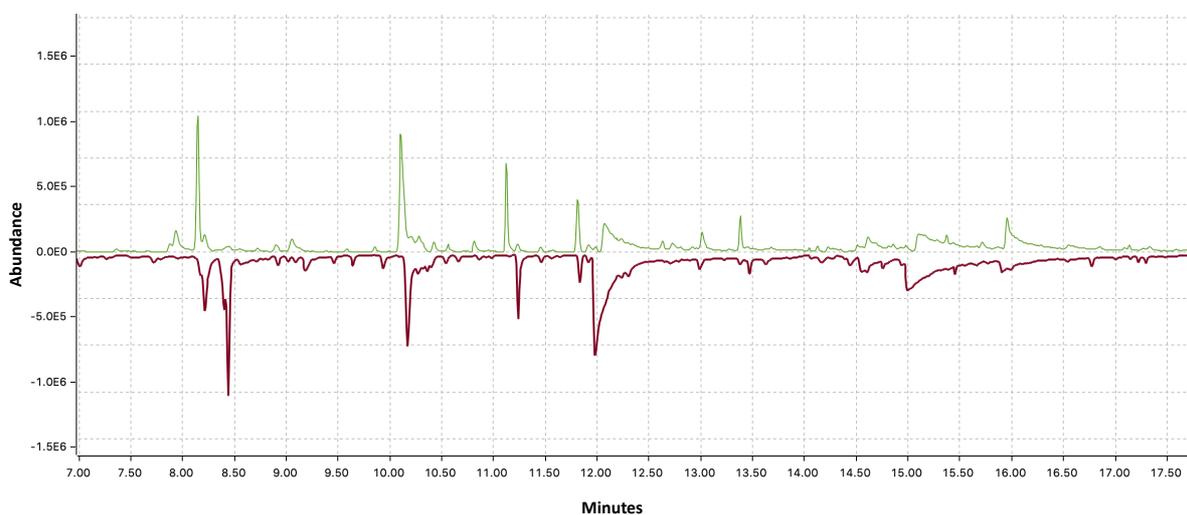

Figure S2.4: Overlaid chromatograms of SEP.A (blue, top) and SEP.B (red, bottom)
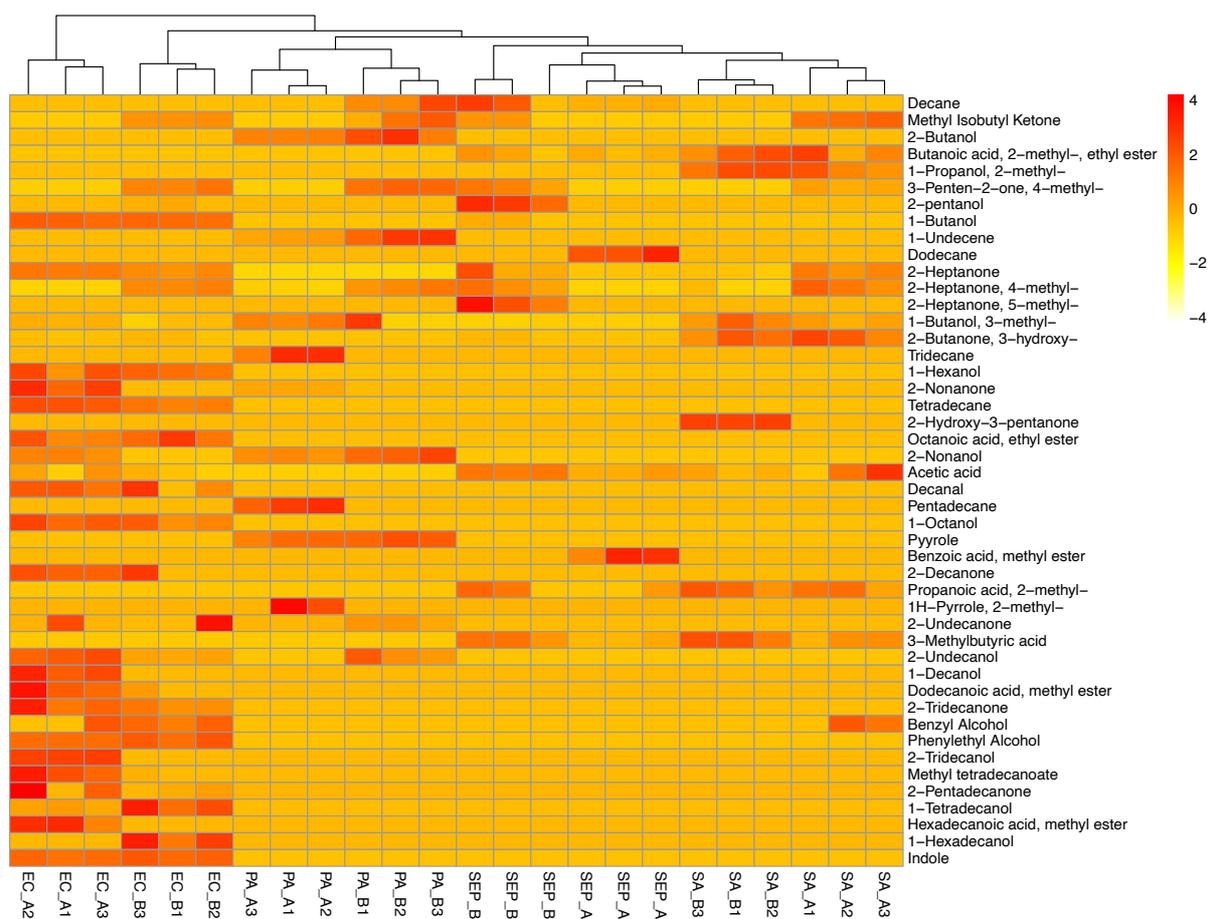
133

Figure S2.5: Heatmap showing the relative abundance of VOCs recovered (rows) from each bacterial strain (columns). Values were scaled and centred by their respective rows, with highly abundant VOCs being coloured red, and less abundant VOCs being marked orange - yellow .
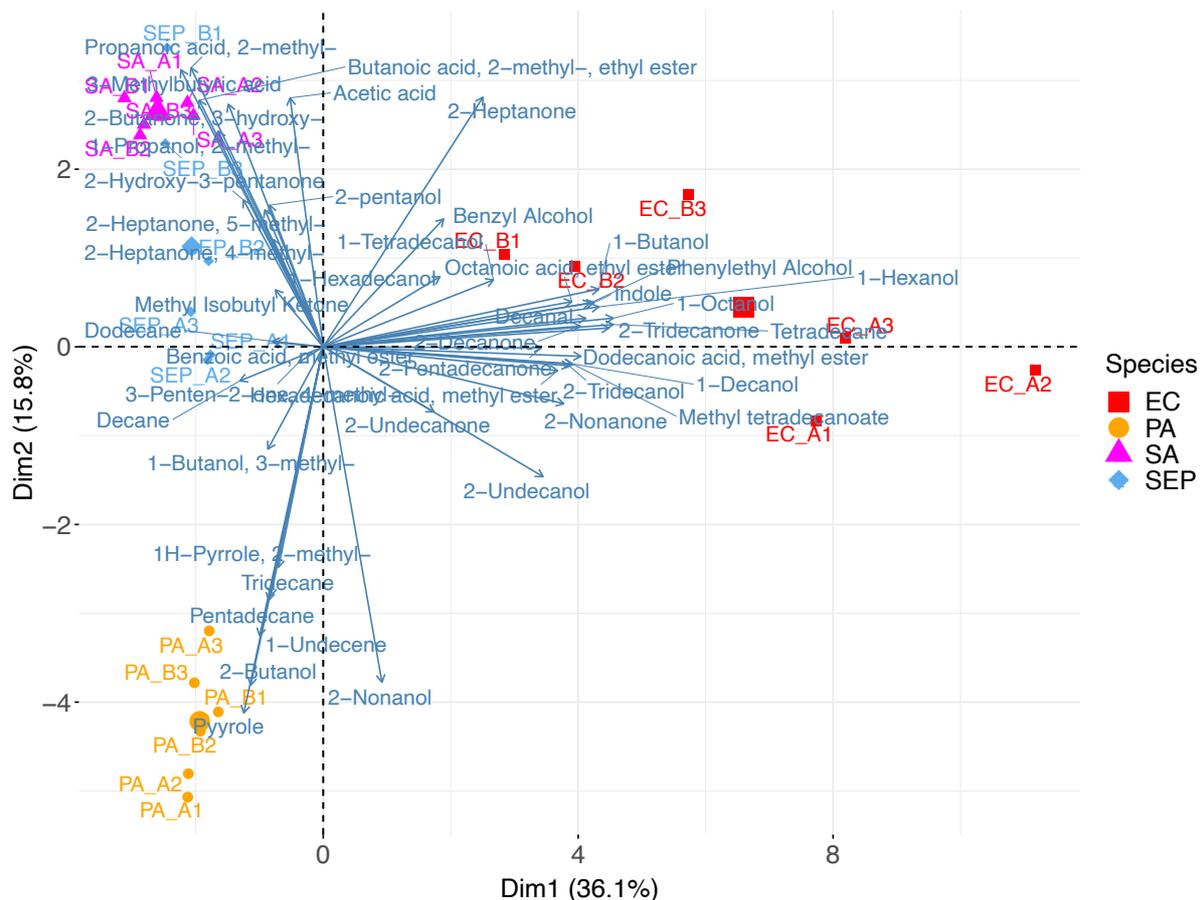
Figure S2.6: Loadings plot of bacteria only samples. Bacteria-specific VOCs are indicated with blue lines and are distributed across the plot with respect to their presence in specific species and the relative abundance emitted. Compounds detected in media control samples have been subtracted.

Pre-examination copy
Submitted 27th May 2022

Figure S2.7: Kinetic plots for individual VOC emissions from S. aureus samples (n=3) over 48 h.

Figure S2.8: Kinetic plots for individual VOC emissions from P. aeruginosa samples (n=3) over 48 h.

Table S2.1: Percentage normalised peak area values of each compound identified in the HS of liquid bacterial cultures (following 24 h incubation at 37°C) after 20 min sample collecti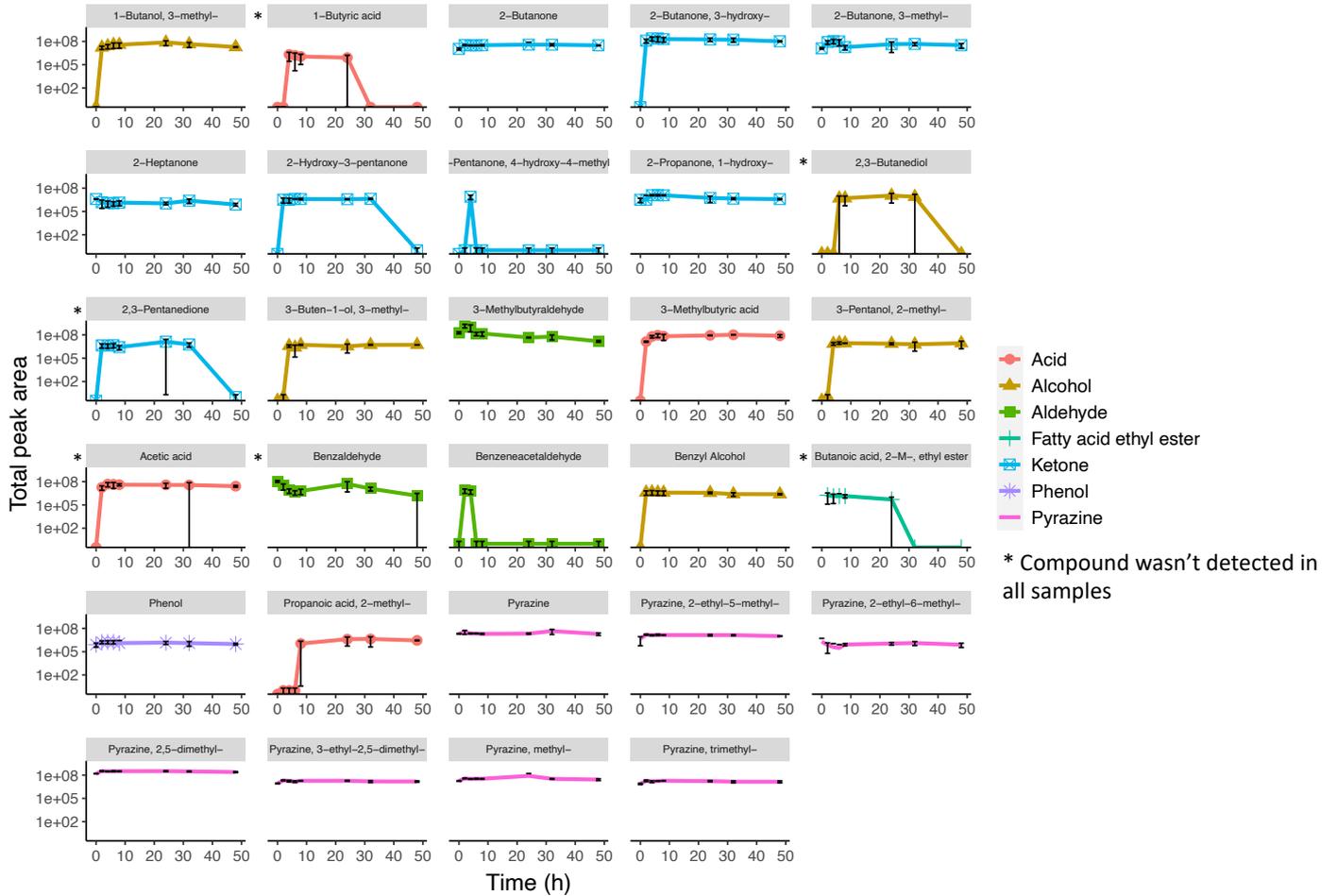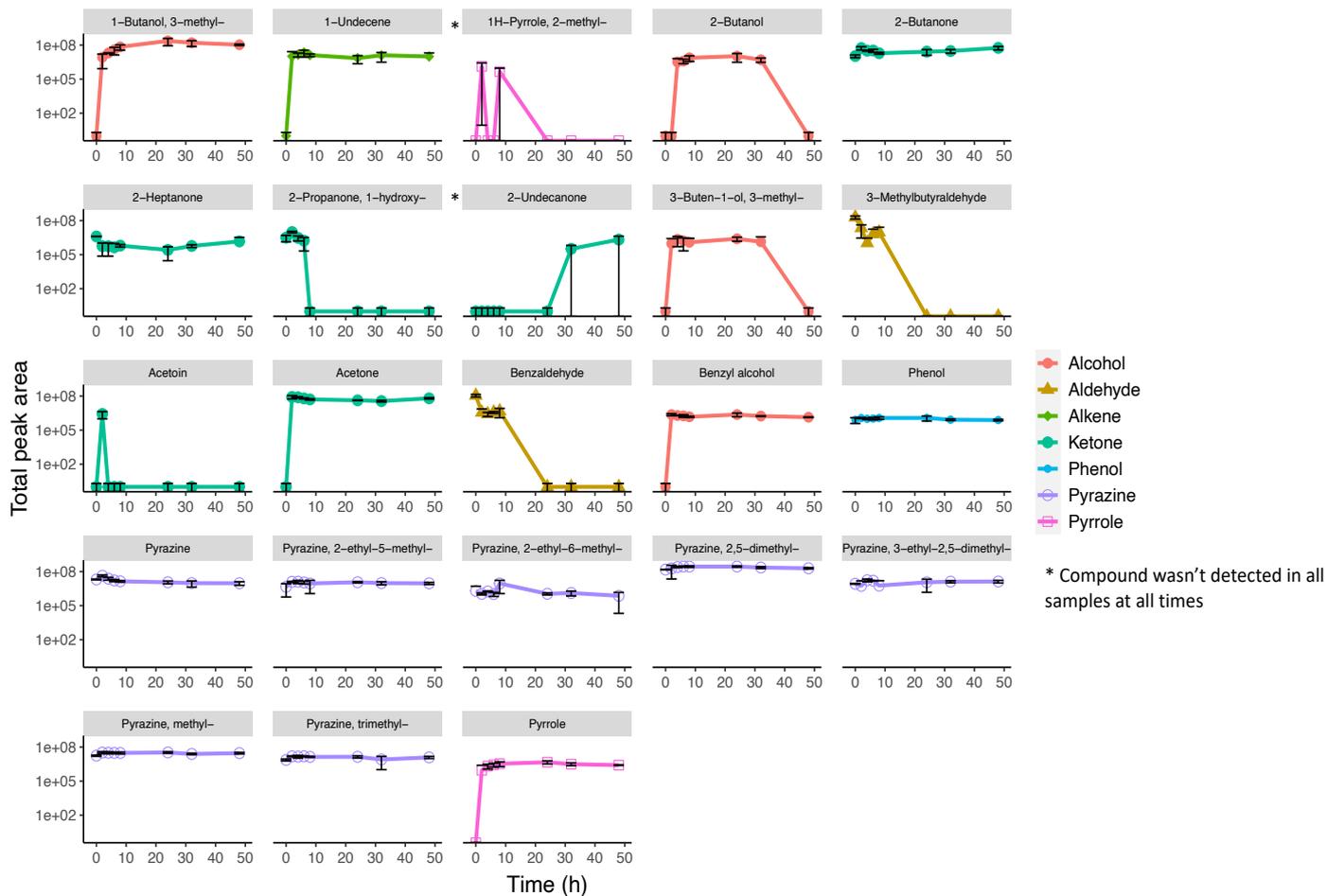on using the HS-SPME technique followed by thermal desorption to GC-MS. Compounds are listed in order of increasing retention time. Kovats retention index (RI) value range (polar column) for each compound is also provided.

Compound	CAS	RI	SA.A	SA.B	PA.A	PA.B	EC.A	EC.B	SEP.A	SEP.B	TSB
2-Butanone	78-93-3	900 ± 20	1.642	1.262	0.000	0.000	1.629	0.204	1.496	0.000	1.314
3-Methylbutyraldehyde	590-86-3	920 ± 20	3.229	2.635	0.000	0.000	0.000	0.000	1.789	2.580	18.521
2-Butanone, 3-methyl-	563-80-4	950 ± 20	2.910	1.419	0.000	0.000	0.000	1.875	0.000	2.013	1.221
Decane	124-18-5	-	0.000	0.000	0.000	1.725	0.000	0.000	1.056	2.533	0.000
Methyl Isobutyl Ketone	108-10-1	1010 ± 15	0.680	0.000	0.000	0.683	0.000	0.309	0.000	0.455	0.000
2-Butanol	78-92-2	1026 ± 15	0.000	0.000	0.167	0.263	0.000	0.000	0.000	0.000	0.000
Butanoic acid, 2-methyl-, ethyl ester	7452-79-1	1050 ± 25	0.444	0.701	0.000	0.000	0.000	0.000	0.383	0.291	0.000
Disulfide, dimethyl	624-92-0	1060 ± 30	1.655	2.170	1.261	0.766	3.210	0.705	4.897	1.036	4.505
1-Propanol, 2-methyl-	78-83-1	1090 ± 30	0.534	0.952	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3-Penten-2-one, 4-methyl-	141-79-7	1125 ± 25	0.230	0.000	0.000	0.720	0.000	0.343	0.000	0.703	0.000
2-pentanol	6032-29-7	1130 ± 25	0.000	0.000	0.000	0.000	0.000	0.015	0.000	0.459	0.000
1-Butanol	71-36-3	1125 ± 25	0.000	0.000	0.000	0.000	1.432	0.546	0.000	0.225	0.000
1-Butanol, 3-methyl-, acetate	821-95-4	1120 ± 15	0.000	0.179	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1-Undecene	112-40-3	1142 ± 12	0.000	0.000	3.975	14.926	0.000	0.000	0.000	0.000	0.000
Dodecane	110-43-0	-	0.000	0.000	0.000	0.000	0.000	0.000	0.410	0.000	0.000
2-Heptanone	6137-06-0	1184 ± 30	0.258	0.049	0.000	0.000	0.542	0.177	0.208	0.423	0.000
2-Heptanone, 4-methyl-	18217-12-4	1213 ± 10	2.354	0.281	0.131	2.436	0.000	1.497	0.000	3.292	0.074
2-Heptanone, 5-methyl-	290-37-9	1252	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.308	0.000
Pyrazine	123-51-3	1214 ± 22	0.664	0.320	1.095	0.000	0.000	0.000	0.000	0.000	1.887
1-Butanol, 3-methyl-	100-42-5	1199 ± 20	4.720	11.476	12.873	6.688	6.265	1.119	1.066	0.291	0.000

138

Styrene	763-32-6	1265 ± 35	0.000	0.000	0.000	0.000	0.341	0.000	0.000	0.000	0.105
3-Buten-1-ol, 3-methyl-	109-08-0	1250 ± 25	0.426	0.460	0.000	0.225	0.000	0.077	0.435	0.207	0.150
Pyrazine, methyl-	116-09-6	1260 ± 25	1.472	1.130	1.578	0.693	0.766	0.398	1.228	0.796	1.462
2-Propanone, 1-hydroxy-	513-86-0	1290 ± 25	0.073	0.105	0.000	0.000	0.157	0.157	0.264	0.240	0.322
2-Butanone, 3-hydroxy-	629-50-5	1290 ± 15	1.598	1.661	0.000	0.000	0.097	0.095	0.000	0.229	0.000
Tridecane	123-32-0	-	0.000	0.000	0.279	0.000	0.000	0.000	0.000	0.000	0.000
Pyrazine, 2,5-dimethyl-	111-27-3	1320 ± 20	12.279	11.466	15.562	2.393	4.865	3.188	10.909	3.771	13.510
1-Hexanol	3658-80-8	1350 ± 20	0.000	0.000	0.000	0.000	0.157	0.058	0.000	0.000	0.000
Dimethyl trisulfide	821-55-6	1365 ± 25	0.175	0.000	0.043	0.000	0.378	0.000	0.000	0.000	0.142
2-Nonanone	629-59-4	1380 ± 20	0.000	0.000	0.224	0.000	1.593	0.000	0.000	0.000	0.000
Tetradecane	5704-20-1	-	0.000	0.000	0.000	0.000	0.079	0.020	0.000	0.000	0.000
2-Hydroxy-3-pentanone	13925-03-6	1380 ± 20	0.000	0.164	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Pyrazine, 2-ethyl-6-methyl-	13360-64-0	1385 ± 15	0.011	0.109	0.000	0.000	0.000	0.000	0.371	0.000	0.124
Pyrazine, 2-ethyl-5-methyl-	14667-55-1	1390 ± 15	0.641	0.318	0.237	0.000	0.000	0.000	0.000	0.000	0.285
Pyrazine, trimethyl-	106-32-1	1400 ± 20	0.621	0.616	0.632	0.181	0.174	0.108	0.233	0.202	0.604
Octanoic acid, ethyl ester	628-99-9	1425 ± 20	0.000	0.000	0.000	0.000	0.101	0.054	0.000	0.000	0.000
Pyrazine, 3-ethyl-2,5-dimethyl-	64-19-7	1455 ± 25	0.177	0.595	0.265	0.000	0.000	0.087	0.000	0.000	0.676
Acetic acid	112-31-2	1460 ± 30	3.687	1.481	0.000	0.000	2.703	0.470	5.811	6.569	0.000
Decanal	100-52-7	1470 ± 20	0.000	0.000	0.000	0.000	0.126	0.036	0.000	0.000	0.000
2-Decanone	79-31-2	1490 ± 25	0.000	0.000	0.000	0.000	1.322	0.248	0.000	0.000	0.000
2-Nonanol	13360-65-1	1505 ± 20	0.000	0.000	0.196	0.368	0.315	0.000	0.000	0.000	0.000
Pyrrole	93-58-3	1515 ± 25	0.000	0.000	0.279	0.415	0.000	0.000	0.000	0.000	0.000
Benzaldehyde	629-62-9	1520 ± 25	0.000	0.000	0.000	0.000	0.000	0.446	0.000	0.000	15.211
Pentadecane	111-87-5	-	0.000	0.000	0.421	0.000	0.000	0.000	0.000	0.000	0.000

139

1-Octanol	109-97-7	1550 ± 25	0.000	0.000	0.000	0.000	1.390	0.380	0.000	0.000	0.000
Benzoic acid, methyl ester	693-54-9	1615 ± 25	0.000	0.000	0.000	0.000	0.000	0.000	0.412	0.000	0.000
Propanoic acid, 2-methyl-	636-41-9	1560 ± 25	0.201	0.375	0.000	0.000	0.000	0.000	0.143	0.301	0.000
1H-Pyrrole, 2-methyl-	112-12-9	1555 ± 15	0.000	0.000	0.205	0.000	0.000	0.000	0.000	0.000	0.000
2-Undecanone	503-74-2	1590 ± 15	0.000	0.000	0.043	0.192	0.380	0.213	0.000	0.000	0.000
3-Methylbutyric acid	1653-30-1	1660 ± 25	2.460	7.043	0.000	0.000	0.000	0.000	4.277	7.039	0.000
2-Undecanol	112-30-1	1712 ± 15	0.000	0.000	0.000	0.241	0.562	0.067	0.000	0.000	0.000
1-Decanol	111-82-0	1750 ± 20	0.000	0.000	0.000	0.000	6.722	0.000	0.000	0.000	0.000
Dodecanoic acid, methyl ester	593-08-8	1800 ± 15	0.000	0.000	0.000	0.000	0.264	0.010	0.000	0.000	0.000
2-Tridecanone	2345-27-9	1805 ± 15	0.000	0.000	0.000	0.000	1.497	0.316	0.000	0.000	0.000
Benzyl Alcohol	100-51-6	1850 ± 30	0.132	0.000	0.000	0.000	0.144	0.138	0.000	0.000	0.000
Phenylethyl Alcohol	60-12-8	1900 ± 25	0.000	0.000	0.000	0.000	0.236	0.113	0.000	0.000	0.000
2-Tridecanol	1653-31-2	1904 ± 10	0.000	0.000	0.000	0.000	0.501	0.000	0.000	0.000	0.000
1-Dodecanol	112-53-8	1965 ± 20	0.000	0.000	0.000	0.000	3.642	1.159	0.000	0.000	0.057
Methyl tetradecanoate	124-10-7	2010 ± 25	0.000	0.000	0.000	0.000	1.496	0.014	0.000	0.000	0.000
2-Pentadecanone	2345-28-0	2020 ± 20	0.000	0.000	0.000	0.000	0.423	0.023	0.000	0.000	0.000
Phenol	108-95-2	2000 ± 30	0.042	0.060	0.060	0.000	0.000	0.000	0.000	0.000	0.055
1-Tetradecanol	112-72-1	2175 ± 30	0.000	0.000	0.000	0.000	0.275	0.513	0.000	0.000	0.000
Hexadecanoic acid, methyl ester	112-39-0	2250 ± 45	0.000	0.000	0.000	0.000	0.683	0.000	0.000	0.000	0.000
1-Hexadecanol	36653-82-4	2375 ± 25	0.000	0.000	0.000	0.000	0.000	0.460	0.000	0.000	0.000
Indole	120-72-9	2420 ± 25	0.000	0.000	0.000	0.000	88.617	41.254	0.000	0.000	0.000

140

Chapter 3: An investigation of stability and species and strain-level specificity in bacterial volatilomes – Supplementary Information
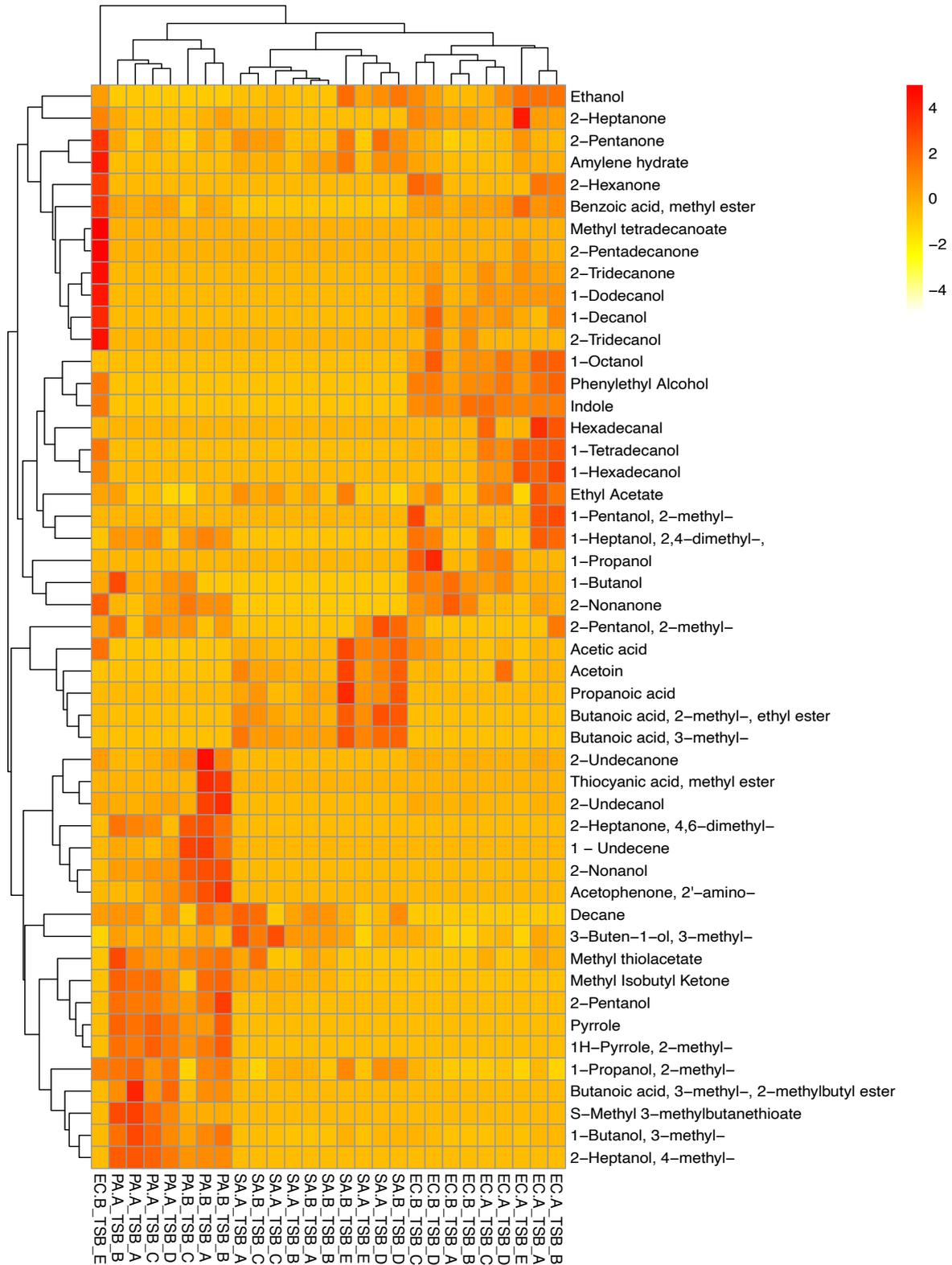
Shane Fitzgerald[1], Linda Holland[2], Aoife Morrin*[1]
School of Chemical Sciences, National Centre for Sensor Research, Insight SFI Research Centre for Data Analytics, Dublin City University, Ireland
School of Biotechnology, Dublin City University, Ireland

*aoife.morrin@dcu.ie

Figure S3.1: Heatmap showing the abundance of VOCs recovered (rows) from each bacterial strain (columns) cultured in BHI media. Values were scaled and centred by their respective rows, with highly abundant VOCs being coloured red, and less abundant VOCs being marked

orange - yellow . Dissimilarity between the samples was measured using Euclidean distance. The complete agglomeration method was used. The corresponding strain names to the abbreviated titles of the bacterial samples shown in this plot are as follows: **EC.A:** *E. coli DSM103372*, **EC.B:** *E. coli DSM30083*, **PA.A**: *P. aeruginosa DSM105372*, **PA.B**: *P. aeruginosa DSM25642*, **SA.A:** *S. aureus DSM2569*, **SA.B:** *S. aureus DSM799*

*Figure S3.2: Heatmap showing the abundance of VOCs recovered (rows) from each bacterial strain (columns) cultured in TSB media . Values were scaled and centred by their respective rows, with highly abundant VOCs being coloured red, and less abundant VOCs being marked orange - yellow . Dissimilarity between the samples was measured using Euclidean distance. The complete agglomeration method was used. The corresponding strain names to the abbreviated titles of the bacterial samples shown in this plot are as follows: **EC.A:** E. coli DSM103372, **EC.B:** E. coli DSM30083, **PA.A**: P. aeruginosa DSM105372, **PA.B**: P. aeruginosa DSM25642, **SA.A:** S. aureus DSM2569, **SA.B:** S. aureus DSM799*

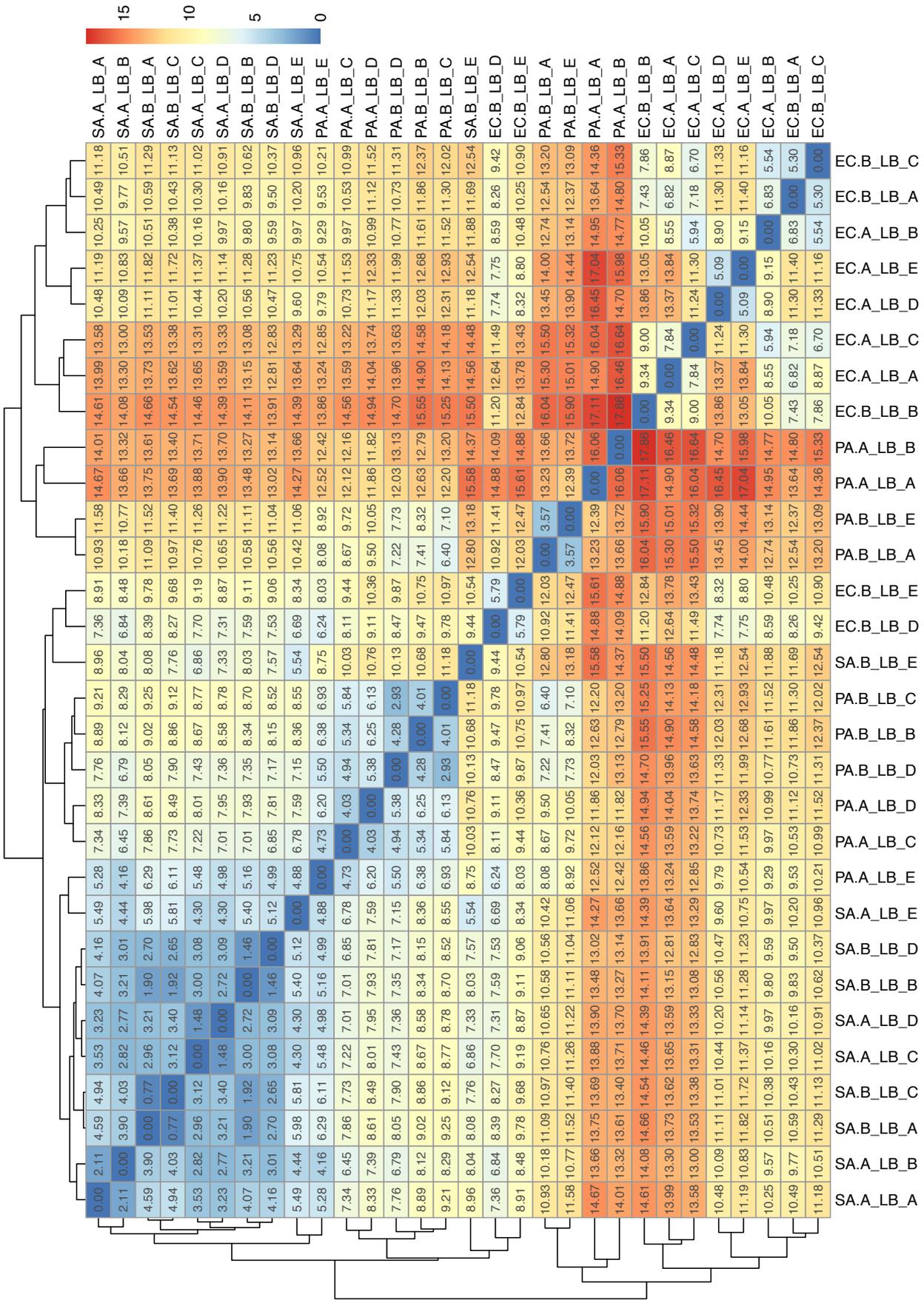Pre-examination copy
Submitted 27th May 2022

Figure S3.3: Heatmap showing the abundance of VOCs recovered (rows) from each bacterial strain (columns) cultured in LB media. Values were scaled and centred by their respective rows, with highly abundant VOCs being coloured red, and less abundant VOCs being marked orange - yellow. Dissimilarity between the samples was measured using Euclidean distance. The complete agglomeration method was used. The corresponding strain names to the abbreviated titles of the bacterial samples shown in this plot are as follows: **EC.A:** *E. coli*

145

DSM103372, **EC.B:** *E. coli DSM30083,* **PA.A**: *P. aeruginosa DSM105372,* **PA.B**: *P. aeruginosa DSM25642,* **SA.A:** *S. aureus DSM2569,* **SA.B:** *S. aureus DSM799*

Figure S3.4: Heatmap dissimilarity matrix representation of bacterial samples cultured in BHI. Values are calculated Euclidean distances. Red represents highly dissimilar samples; blue represents highly similar samples. The corresponding strain names to the abbreviated titles of the bacterial samples shown in this plot are as follows: **EC.A:** *E. coli DSMI03372,* **EC.B:** *E. coli DSM30083,* **PA.A:** *P. aeruginosa DSMI05372,* **PA.B:** *P. aeruginosa DSM799,* **SA.A:** *S. aureus DSM25642,* **SA.B:** *S. aureus DSM25569,* **SA.B:** *S. aureus DSM799*

148

*Figure S3.5: Heatmap dissimilarity matrix representation of bacterial samples cultured in TSB. Values are calculated Euclidean distances. Red represents highly dissimilar samples; blue represents highly similar samples. The corresponding strain names to the abbreviated titles of the bacterial samples shown in this plot are as follows: **EC.A**: E. coli DSMI03372, **EC.B**: E. coli DSM30083, **PA.A**: P. aeruginosa DSMI05372, **PA.B**: P. aeruginosa DSM25642, **SA.A**: S. aureus DSM25569, **SA.B**: S. aureus DSM799*

149

Figure S3.6: Heatmap dissimilarity matrix representation of bacterial samples cultured in LB. Values are calculated Euclidean distances. Red represents highly dissimilar samples; blue represents highly similar samples. The corresponding strain names to the abbreviated titles of the bacterial samples shown in this plot are as follows: **EC.A**: E. coli DSM103372, **EC.B**: E. coli DSM30083, **PA.A**: P. aeruginosa DSM105372, **PA.B**: P. aeruginosa DSM25642, **SA.A**: S. aureus DSM2569, **SA.B**: S. aureus DSM799
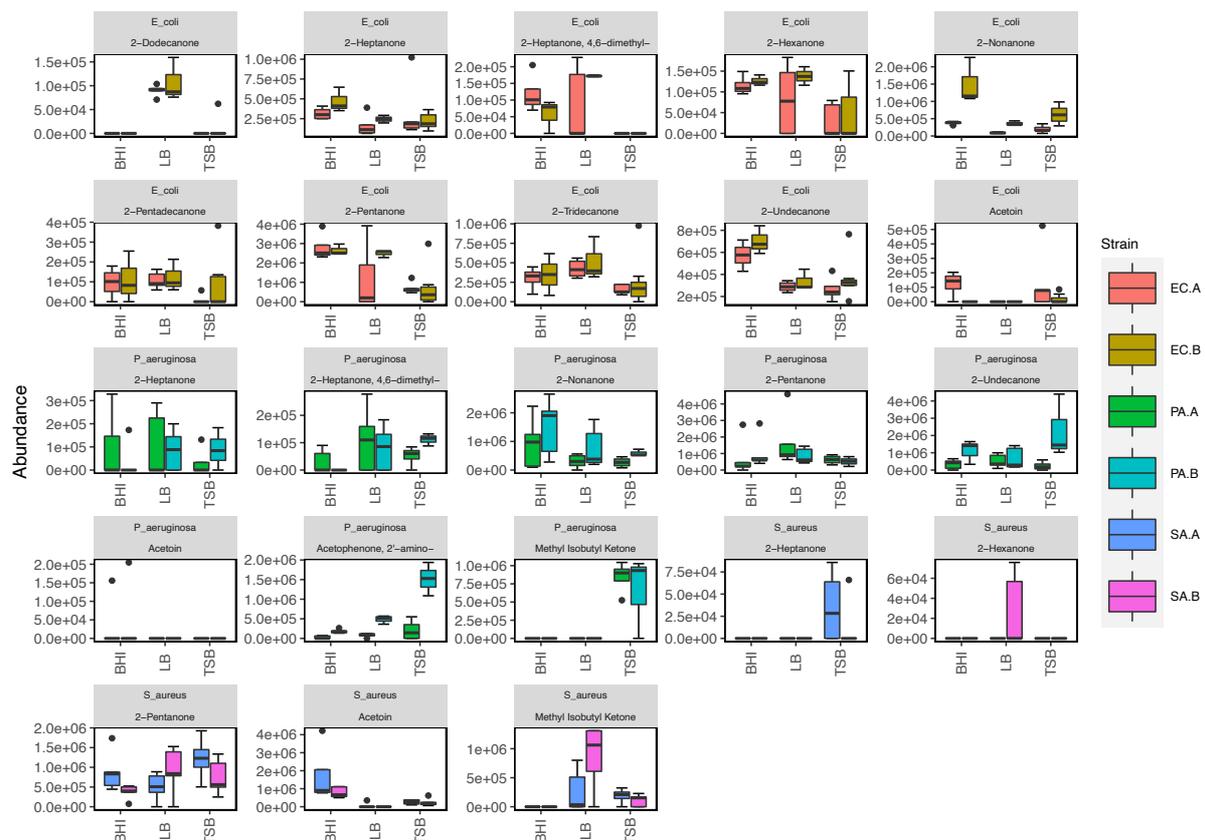
150

151
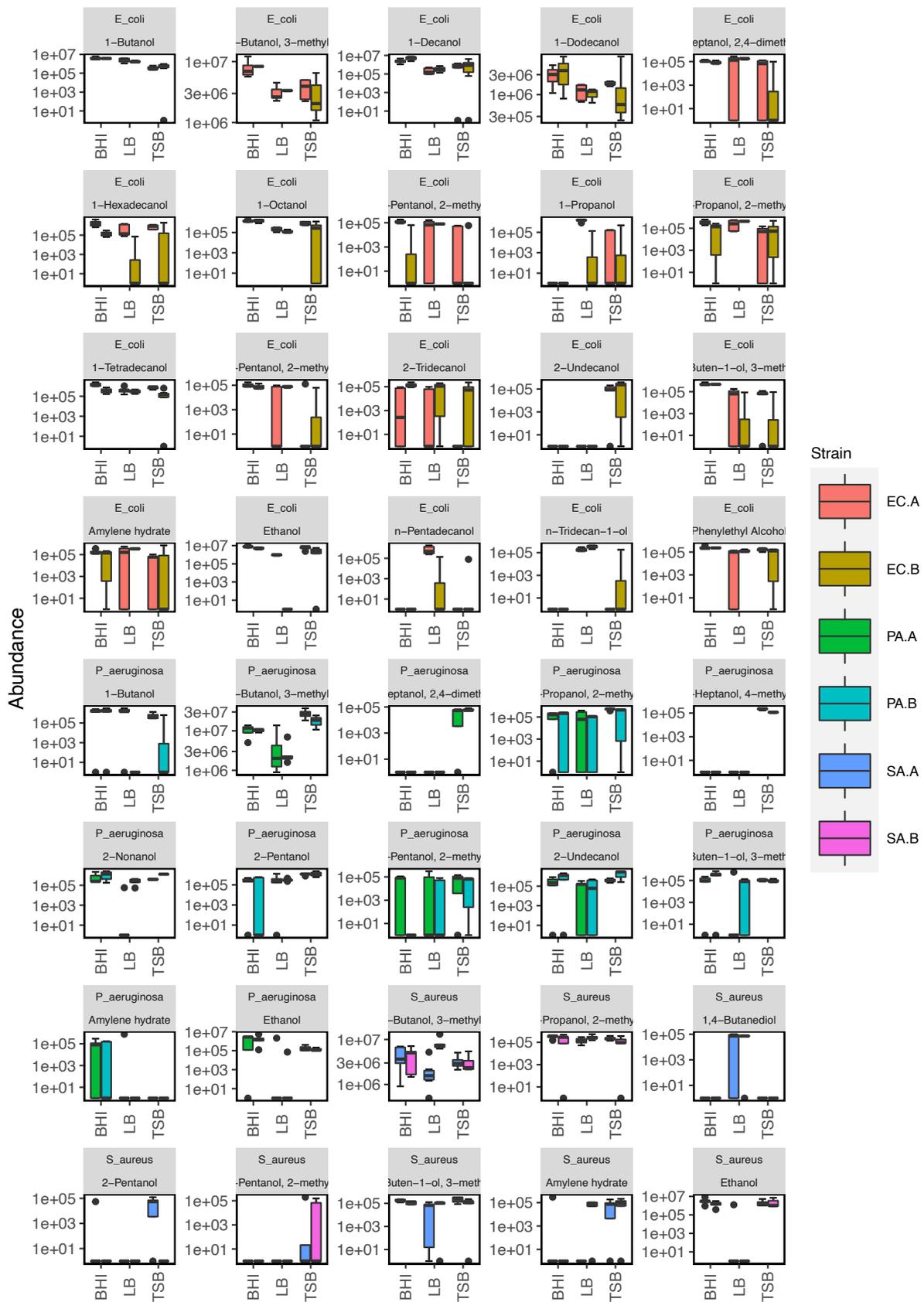
Figure S3.7 : Species- and alcohol-specific boxplots. Each plot illustrates the abundance of an individual alcoholic compound emitted by two strains of a species (*E. coli, P. aeruginosa*, and *S. aureus*) across three nutrient-rich media (Brain Heart Infusion – BHI; Lysogeny Broth – LB;

152

Tryptone Soy Broth – TSB). Each strain is colour-coded according to the legend at the left side of the plot. *For each examined strain in BHI, n=5;, TSB, n=5; and LB media, n = 5. *Five replicates were analysed for each strain in each media except for E. coli (EC.A and EC.B) in BHI (n = 4) and P. aeruginosa (PA.B) in TSB (n = 3). The following compound names are partially visible: row 1 , 2nd and 5th : 3-methyl-1-butanol and 1-Heptanol, 2,4-dimethyl-; row 2, 3rd and 5th: 1-Pentanol, 2-methyl- and 1-Propanol, 2-methyl-; row 3, 2nd and 5th : 2-Pentanol, 2-methyl- and 3-Buten-1-ol, 3-methyl-; row 5, 2nd, 3rd, and 5th : 3-methyl-1-butanol, 1-Heptanol, 2,4-dimethyl-, 1-Propanol, 2-methyl-, and 2-Heptanol, 4-methyl-; row 6, 3rd and 5th : 1-Pentanol, 2-methyl and 3-Buten-1-ol, 3-methyl-; row 7, 3rd and 4th : 3-methyl-1-butanol and 1-Propanol, 2-methyl-; row 8, 2nd and 3rd : 1-Pentanol, 2-methyl and 3-Buten-1-ol, 3-methyl-*.

Figure S3.8 : Species- and ketone-specific boxplots. Each plot illustrates the abundance of an individual ketone compound emitted by two strains of a species (*E. coli, P. aeruginosa*, and *S. aureus*) across three nutrient-rich media (Brain Heart Infusion – BHI; Lysogeny Broth – LB; Tryptone Soy Broth – TSB). Each strain is colour-coded according to the legend at the left side of the plot. *For each examined strain in BHI, n=5;, TSB, n=5; and LB media, n = 5. *Five replicates were analysed for each strain in each media except for E. coli (EC.A and EC.B) in BHI (n = 4) and P. aeruginosa (PA.B) in TSB (n = 3).*
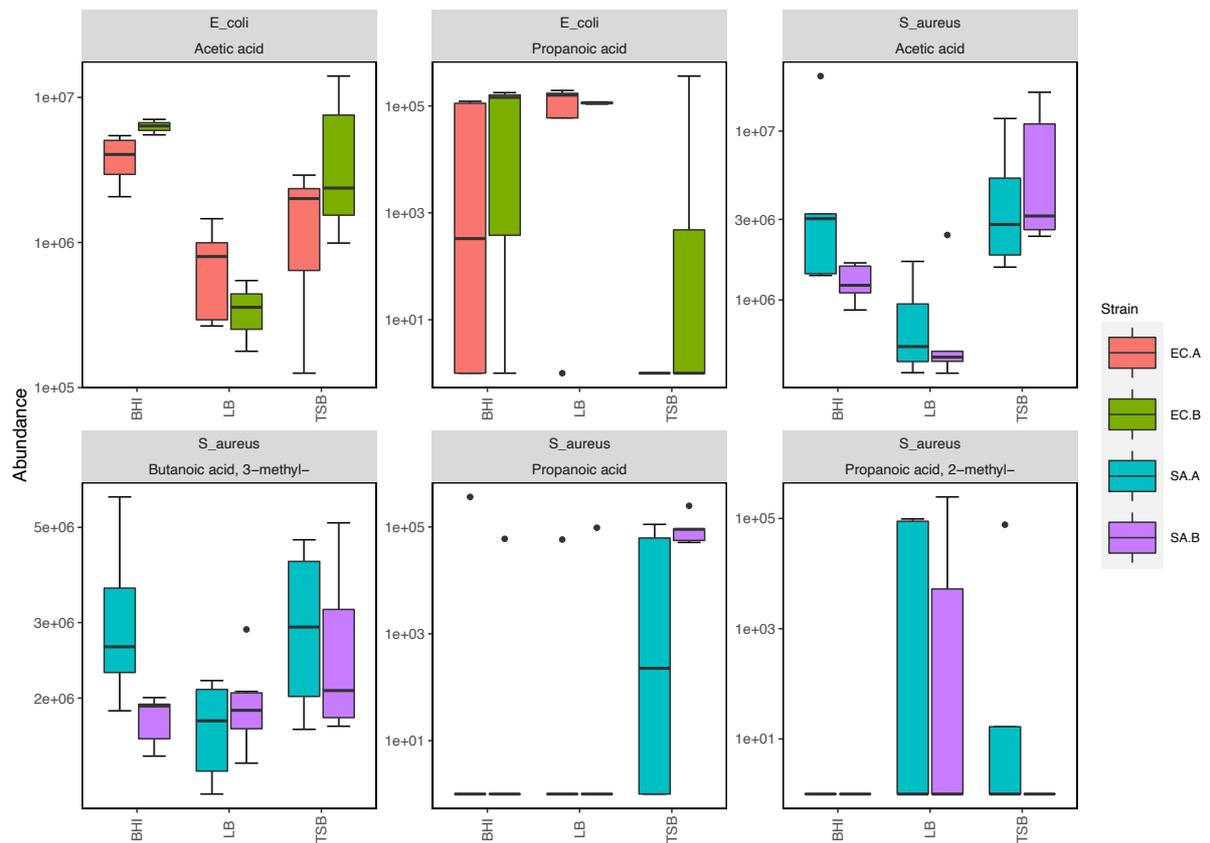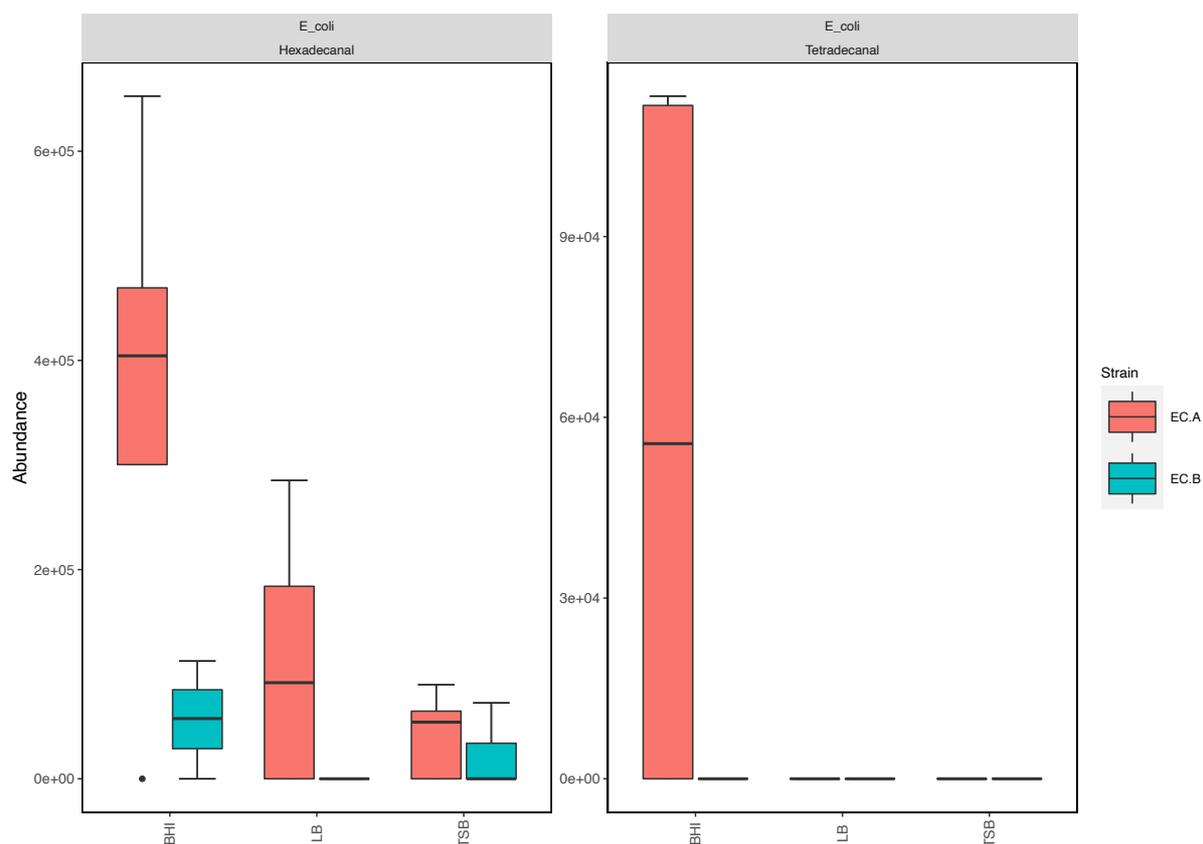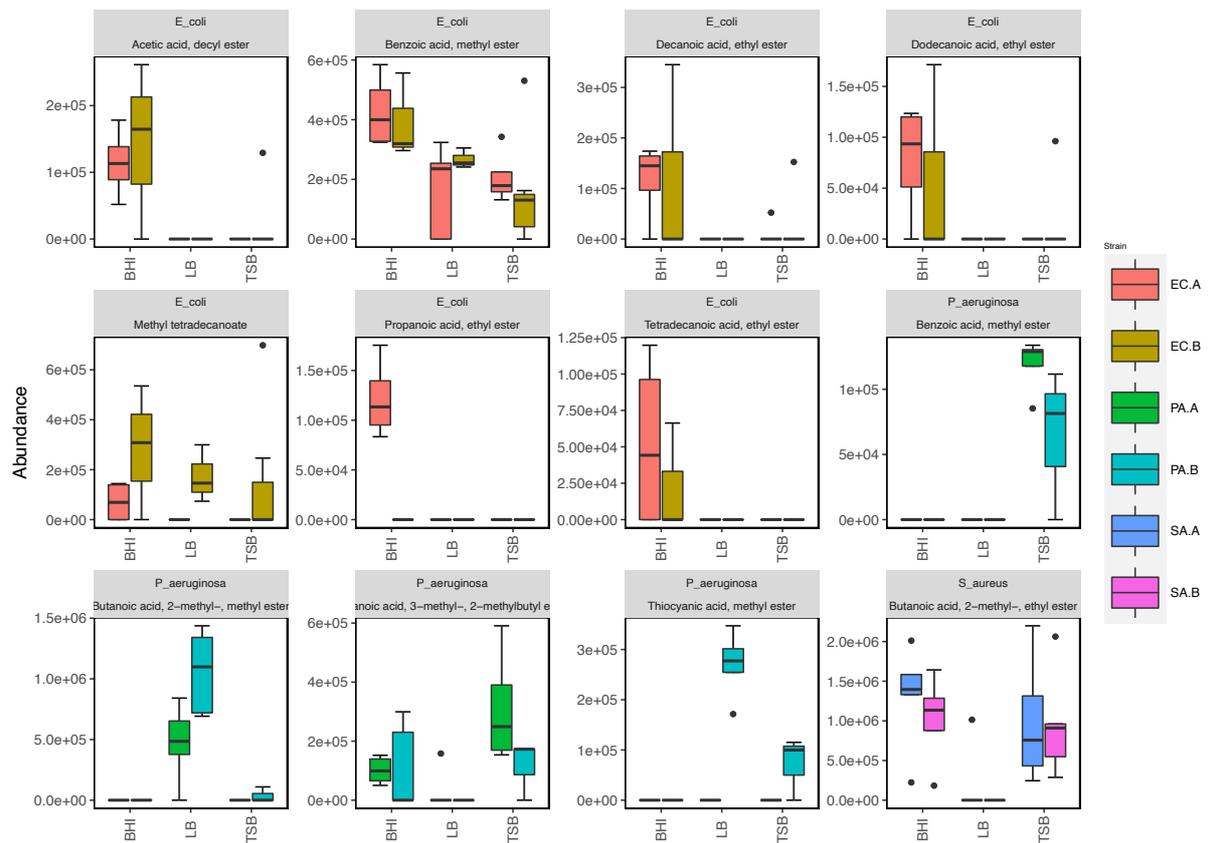
153

Figure S3.9 : Species- and acid-specific boxplots. Each plot illustrates the abundance of an individual acidic compound emitted by two strains of a species (*E. coli* and *S. aureus*) across three nutrient-rich media (Brain Heart Infusion – BHI; Lysogeny Broth – LB; Tryptone Soy Broth – TSB). Each strain is colour-coded according to the legend at the left side of the plot. *For each examined strain in BHI, n=5;, TSB, n=5; and LB media, n = 5. *Five replicates were analysed for each strain in each media except for E. coli (EC.A and EC.B) in BHI (n = 4) and P. aeruginosa (PA.B) in TSB (n = 3).*

154

Figure S3.10 : Species- and aldehyde-specific boxplots. Each plot illustrates the abundance of an individual aldehydic compound emitted by two strains of a species (*E. coli, P. aeruginosa*, and *S. aureus*) across three nutrient-rich media (Brain Heart Infusion – BHI; Lysogeny Broth – LB; Tryptone Soy Broth – TSB). Each strain is colour-coded according to the legend at the left side of the plot. *For each examined strain in BHI, n=5;, TSB, n=5; and LB media, n = 5. *Five replicates were analysed for each strain in each media except for E. coli (EC.A and EC.B) in BHI (n = 4) and P. aeruginosa (PA.B) in TSB (n = 3).*
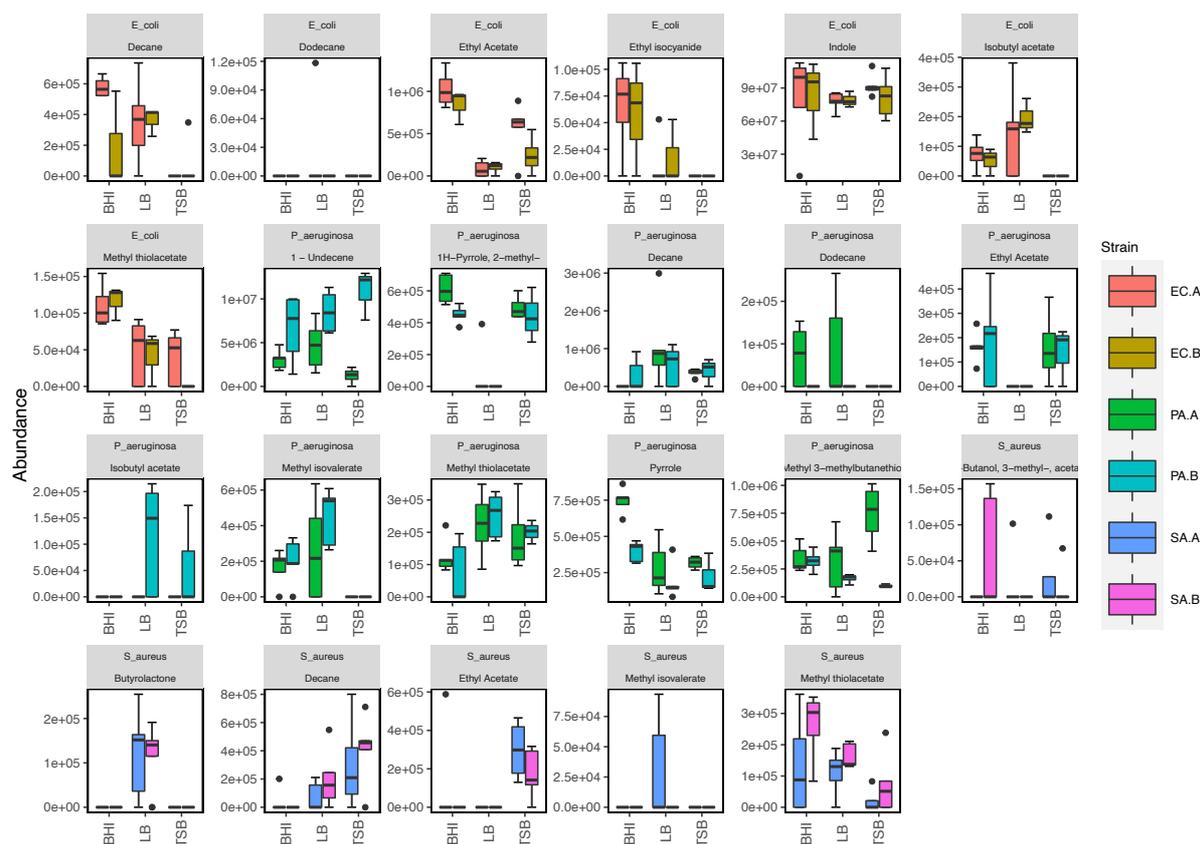
Figure S3.11 : Species- and ester-specific boxplots. Each plot illustrates the abundance of an individual fatty acid ester compound emitted by two strains of a species (*E. coli, P. aeruginosa*, and *S. aureus*) across three nutrient-rich media (Brain Heart Infusion – BHI; Lysogeny Broth – LB; Tryptone Soy Broth – TSB). Each strain is colour-coded according to the legend at the left side of the plot. *For each examined strain in BHI, n=5;, TSB, n=5; and LB media, n = 5. *Five replicates were analysed for each strain in each media except for E. coli (EC.A and EC.B) in BHI (n = 4) and P. aeruginosa (PA.B) in TSB (n = 3). *The name of the second compound in the third row is 'Butanoic acid, 3-methyl-, 2-methylbutyl ester'.*

Figure S3.12 : Species- and compound-specific boxplots. Each plot illustrates the abundance of an individual compound emitted by two strains of a species (*E. coli, P. aeruginosa*, and *S. aureus*) across three nutrient-rich media (Brain Heart Infusion – BHI; Lysogeny Broth – LB; Tryptone Soy Broth – TSB). Each strain is colour-coded according to the legend at the left side of the plot. Compounds shown in this plot belong to a variety of chemical classes. *For each examined strain in BHI, n=5;, TSB, n=5; and LB media, n = 5. *Five replicates were analysed for each strain in each media except for E. coli (EC.A and EC.B) in BHI (n = 4) and P. aeruginosa (PA.B) in TSB (n = 3).*

157
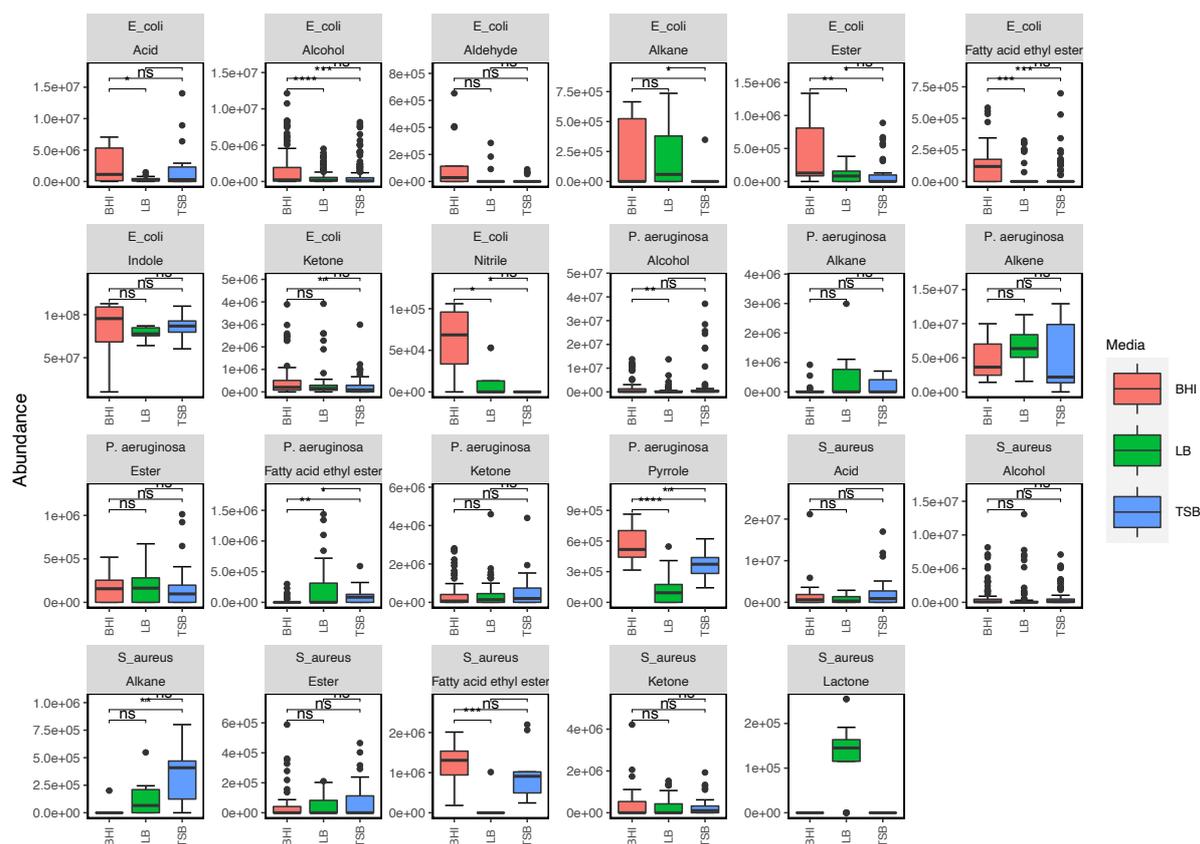
Figure S3.13 : Grouped boxplot representation *illustrating the differences in emission of individual chemical classes in BHI, LB and TSB growth media by E. coli, P. aeruginosa and S. aureus. This bar plot was obtained by summing the mean abundance of each chemical class detected in each of the examined bacteria. The following symbols were used to indicate statistical significance (*ns: p > 0.05; *: p <= 0.05; **: p <= 0.01; ***: p <= 0.001; ****: p <= 0.0001). In row 1, compound 1, '*' between LB and TSB is not visible. In row 3, compound 2, '*' between LB and TSB is not visible. In row 5, compound 3, '**' between LB and TSB is not visible.

158
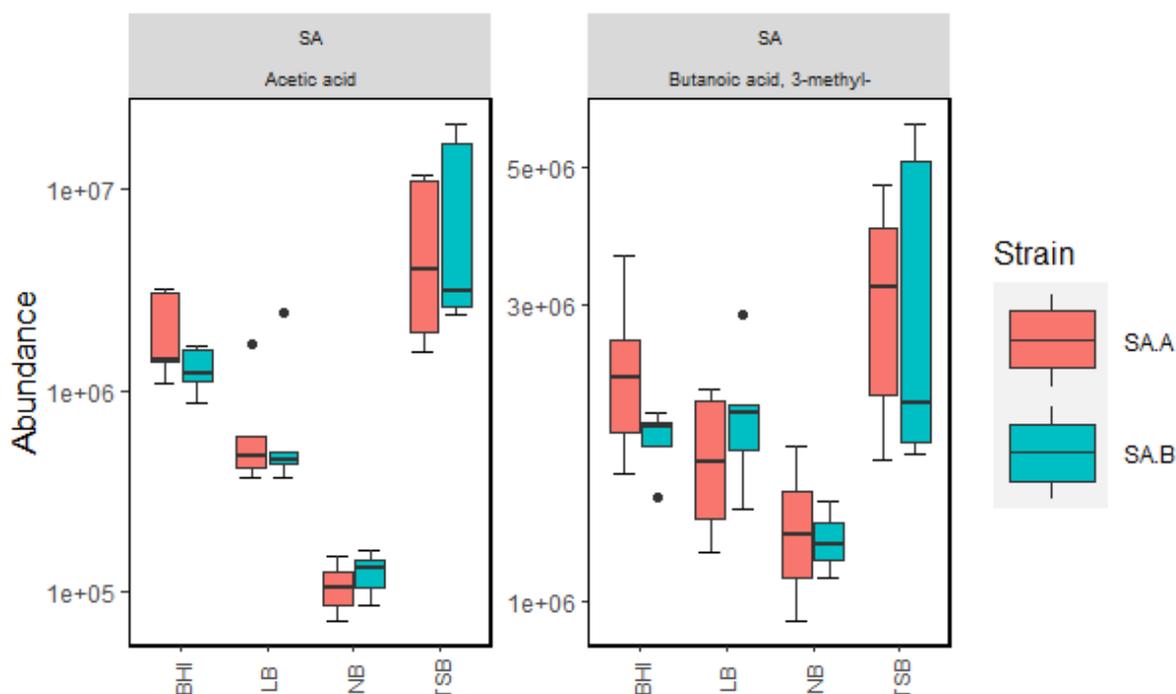
Figure S3.14 : S. aureus - specific box plots illustrating the difference in acid abundances between glucose-containing media (BHI and TSB) and glucose-free media (LB and NB). Nutrient broth – NB

Table S3.1: List of contaminant peaks present in background media, fiber and column controls. 'X' marks presence of compound in chromatograms

Compound	Base peak	BHI	LB	TSB	Fiber blank
Acetone	43	X	X	X	
Cyclotrisiloxane, hexamethyl-	207	-	-	-	X
Cyclotetrasiloxane, octamethyl	281	-	-	-	X
Furan, 3-methyl-	82	X	-	X	-
2,4-Dimethyl-1-heptene	70	-	-	X	-
2-Butanone	43	X	X	X	-
Butanal, 2-methyl-	57	X	X	X	-
Butanal, 3-methyl-	41	X	X	X	-
Benzene	78	X	-	-	-
2,3-Butanedione	43	X	X	X	-
Methyl Isobutyl Ketone	43	X	-	-	-
2-Butanol	45	X	-	X	-
Trichloromethane	83	X	-	-	-
Toluene	91	X	X	X	-
Disulfide, dimethyl	94	X	X	X	-
2-Butenal, 2-methyl-	84	X	X	X	-
1-Butanol	56	X	X	X	-

159

Pyrazine	80	X	-	X	-
Styrene	104	X	-	X	-
Pyrazine, methyl-	94	X	-	X	-
2-Propanone, 1-hydroxy-	43	X	-	X	-
Pyrazine, ethyl-	107	X	-	X	-
Pyrazine, 2,5-dimethyl	108	X	X	X	-
Pyrazine, 2,6-dimethyl-	108	-	-	X	-
Dimethyl trisulfide	126	X	X	X	-
Pyrazine, 2-ethyl-5-methyl-	121	X	X	X	-
Pyrazine, trimethyl-	122	X	-	X	-
Pyrazine, 3-ethyl-2,5-dimethyl-	135	X	-	X	-
Acetic acid	43	-	-	X	-
3-Furaldehyde	95	X	-	-	-
Nonanal	57	X	X	X	X
1-Hexanol, 2-ethyl-	57	X	X	X	X
Decanal	57	X	X	X	X
Benzaldehyde	106	X	X	X	X
Silanediol, dimethyl-	77	X	X	X	X
Oxime-, methoxy-phenyl-_	133	X	X	X	X
2-Furanmethanol	98	X	-	X	-
Benzaldehyde, 4-methyl-	119	X	-	-	-
2-Acetylthiazole	43	X	-	-	-
Furan, 3-phenyl	115	-	-	-	-

Table S3.2: Agilent MassHunter parameters used for chromatographic data analysis

Chromatographic parameters	
Peak finding method	*Chromatogram deconvolution*
Peak filters	*>= 50000 counts*
Peak area calculation	*Chromatogram deconvolution*
Compound identification	*NIST Mass Spectral Library 2017*
Signal smoothing	*No signal smoothing was performed*
Baseline correction	*No baseline correction was performed*

Chapter 4: Multi-Strain and -Species Investigation of Volatile Metabolites emitted from Planktonic and Biofilm *Candida* cultures - Supplementary Information

Shane Fitzgerald[1], Ciara Furlong[2], Linda Holland[2], Aoife Morrin*[1]
School of Chemical Sciences, National Centre for Sensor Research, Insight SFI Research
Centre for Data Analytics, Dublin City University, Ireland
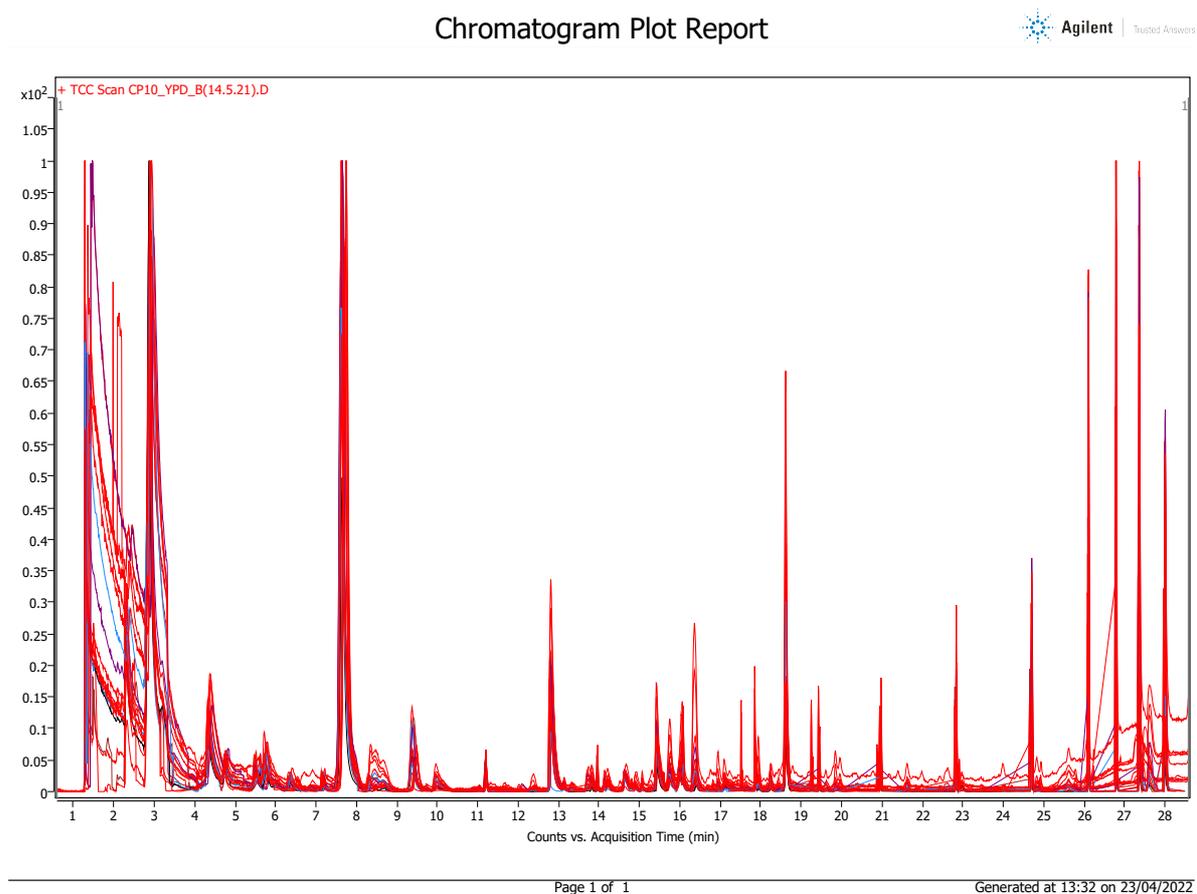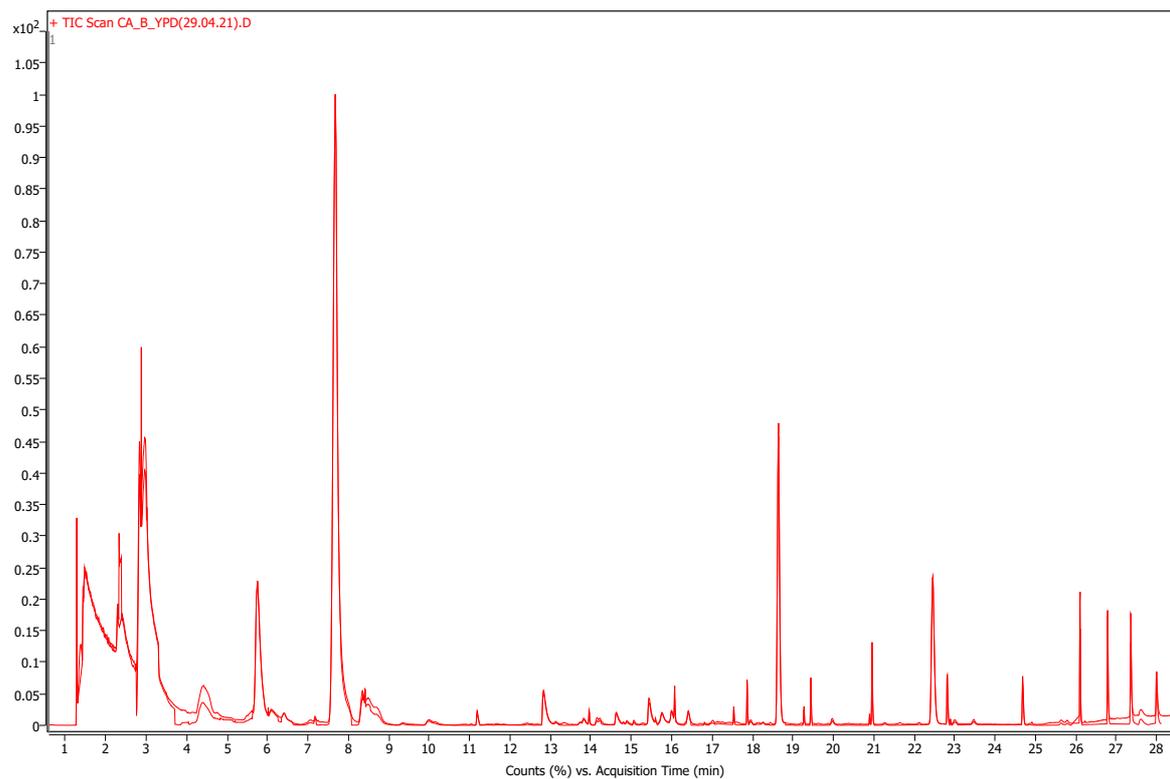School of Biotechnology, Dublin City University, Ireland

Figure S4.1: Overlayed chromatogram of planktonic *C. parapsilosis* strains 1-10 in YPD media.

Figure S4.2: Chromatogram of planktonic *C. albicans* YPD culture
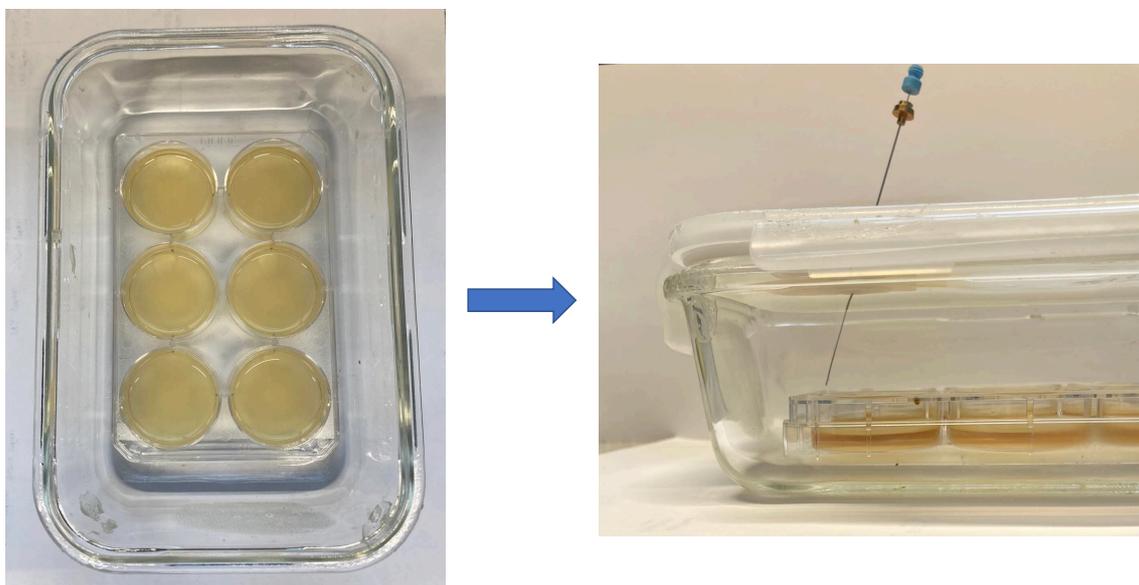
Figure S4.3: Experimental set-up for the SPME sampling of biofilm cultures. (Left) Top view of biofilm sample in container; (Right) Side view of SPME sampling of biofilm metabolites.

162
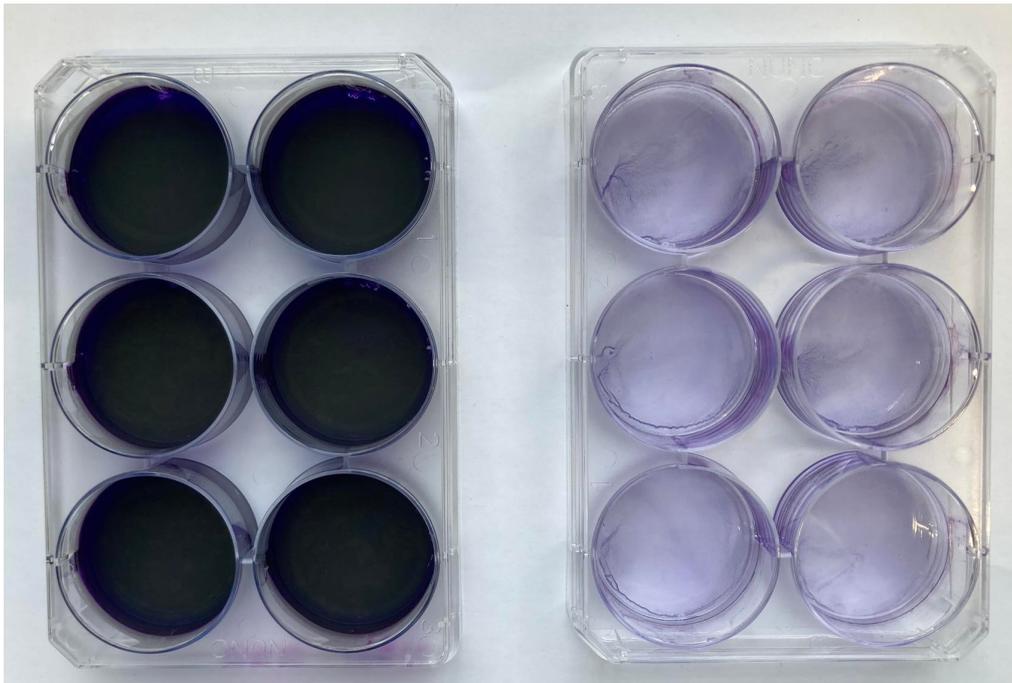
Figure S4.4 – Crystal violet stain of biofilm-positive C. parapsilosis (CP6) samples (left) to verify biofilm formation; and biofilm-negative C. albicans samples (right).
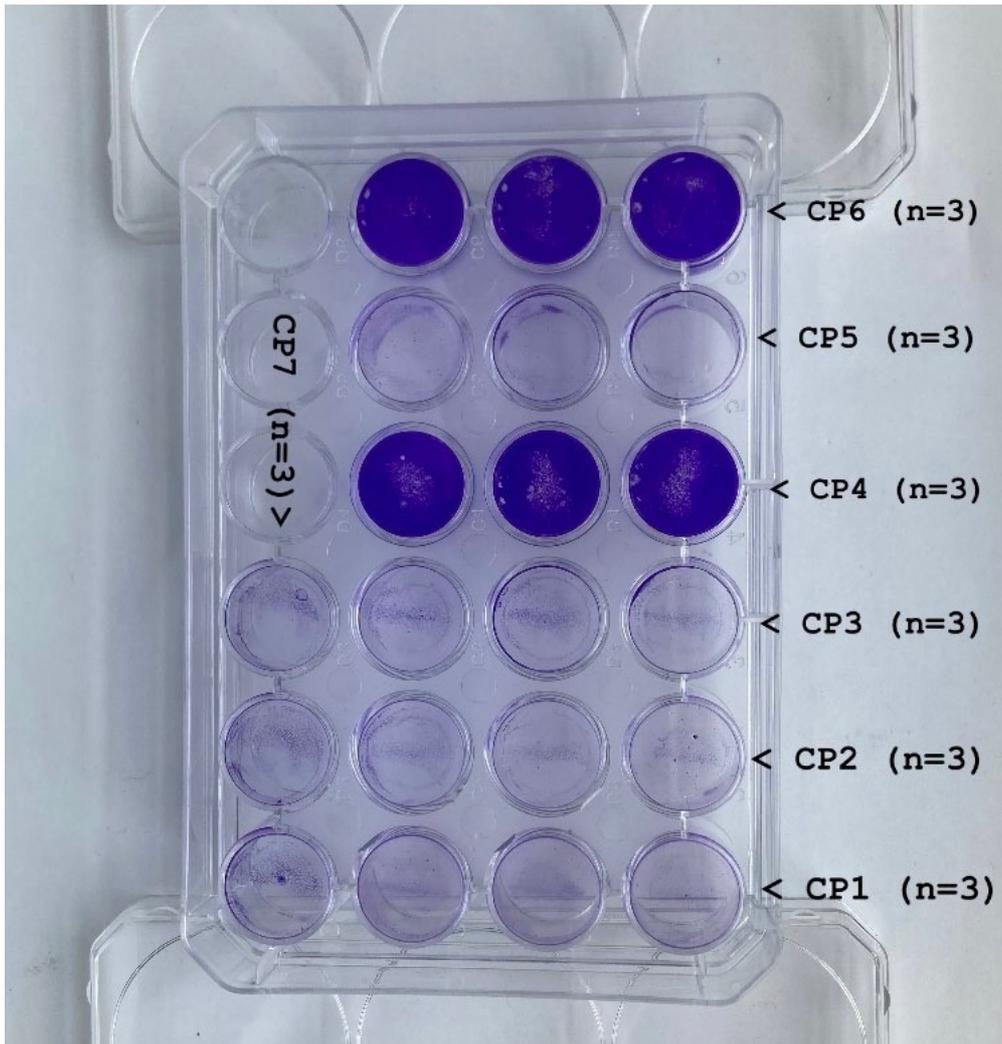
Figure S4.5: Crystal violet – based biofilm viability assay on CP1 – CP7. Out of the 7 CP strains tested, CP4 and CP6 were biofilm-positive.
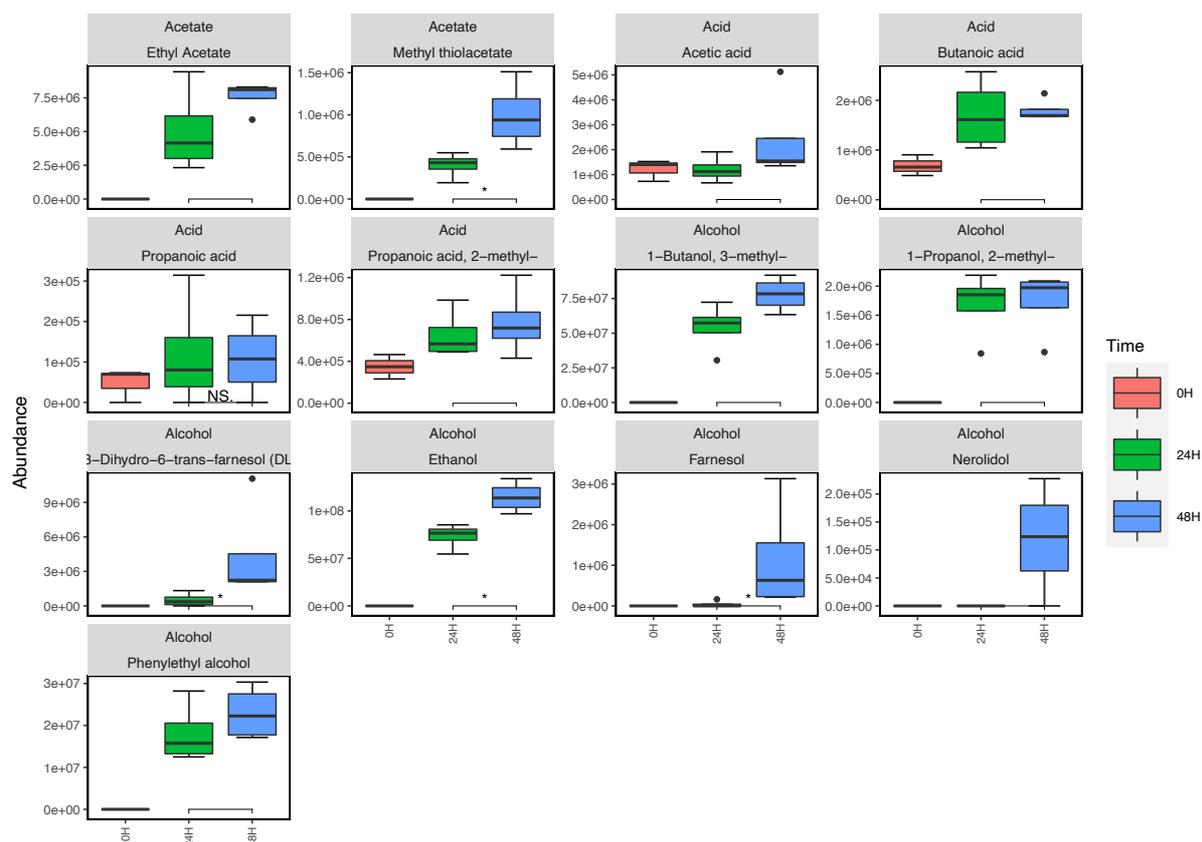
*Figure S4.6: Compound boxplots illustrating the abundances of selected compounds detected in the headspace of biofilm-negative C. albicans cultures at 0, 24, and 48 h growth. Statistically significant differences in the abundance of compounds at each time point are illustrated through the star system where * = p < 0.05, ** = p < 0.01.*
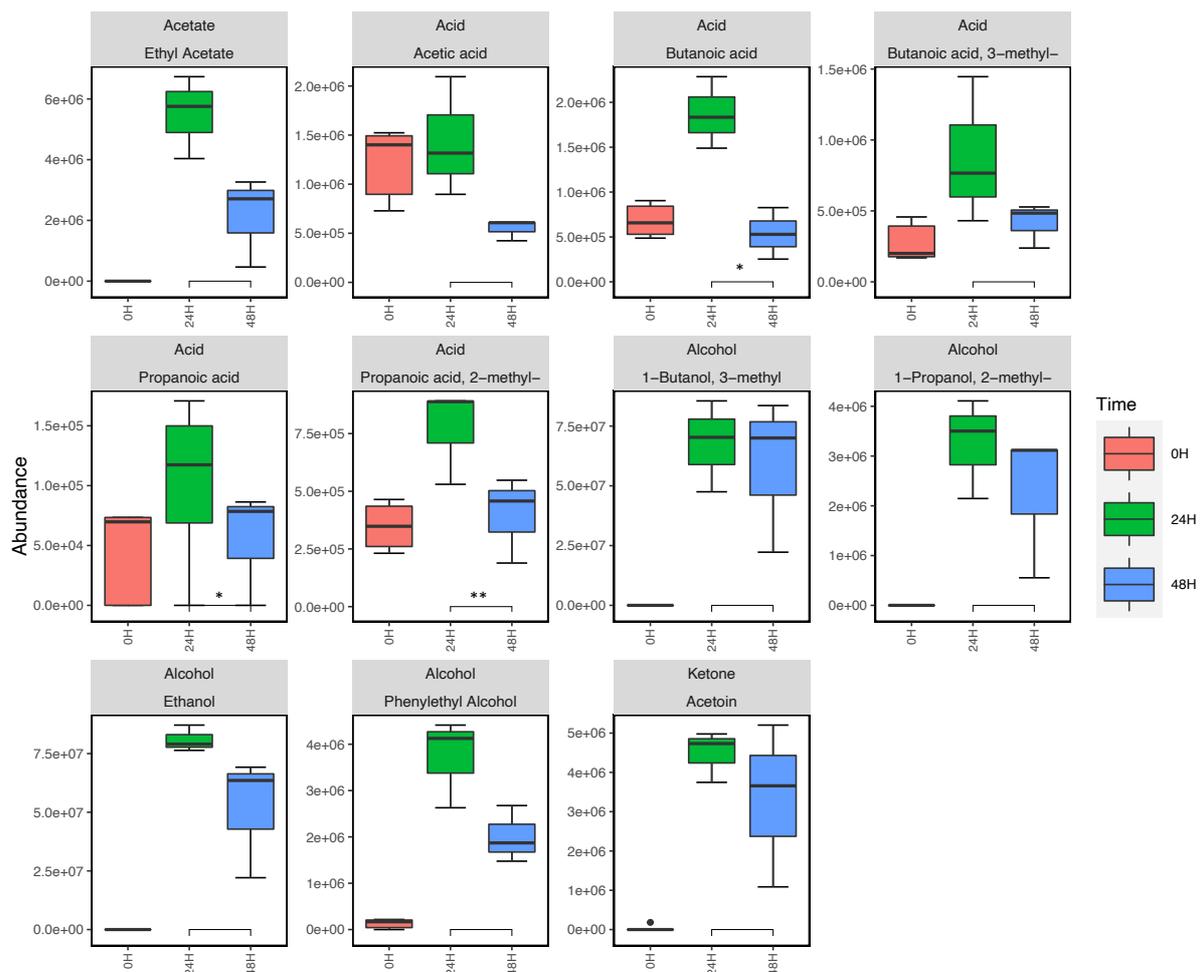
Figure S4.7: Compound boxplots illustrating the abundances of various compounds detected in the headspace of biofilm-negative CP1 cultures at 0, 24, and 48 h growth. Statistically significant differences in the abundance of compounds at each time point are illustrated through the star system where * = p < 0.05, ** = p < 0.01
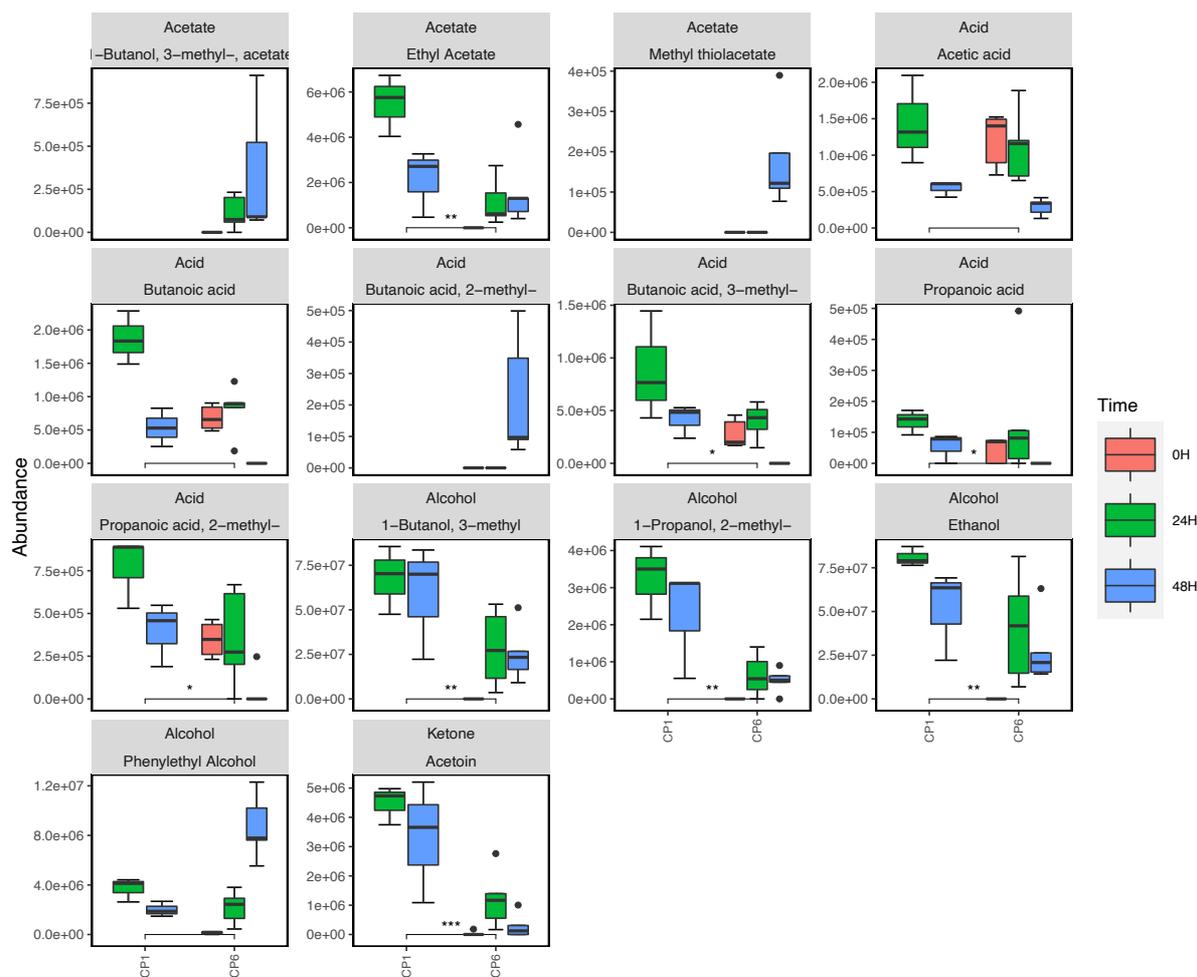
166

Figure S4.8: Compound boxplots illustrating the differences in abundances of various compounds detected in the headspace of non-biofilm forming CP1 and biofilm-forming CP6 cultures at 0, 24, and 48 h growth. Statistically significant differences in the abundances of each compound detected between the two strains are illustrated through the star system where * = p < 0.05, ** = p < 0.01 and *** = p < 0.001.

Table S4.1: Strain names and origin information table for *C. parapsilosis* strains.

Number	Name	Origin	Isolated from	Reference
CP1	CLIB214	Puerto Rico	Faeces	Type strain
CP2	CDC317	USA	Healthcare workers hand	Clark et al 2004 doi: 10.1128/JCM.42.10.4468-4472.2004
CP3	CDC173	USA	Blood or catheter cultures	Kuhn et al 2004 doi: 10.3201/eid1006.030873
CP4	711701	Aberdeen, UK	Unknown	Tavanti et al doi 10.1128/JCM.43.1.284-292.2005
CP5	CDC167	USA	Blood or catheter cultures	Kuhn et al 2004 doi: 10.3201/eid1006.030873
CP6	J961250	Lisbon, Portugal	Nail	Tavanti et al doi 10.1128/JCM.43.1.284-292.2005
CP7	CDC179	USA	Blood or catheter cultures	Kuhn et al 2004 doi: 10.3201/eid1006.030873
CP8	J930733	Beerse, Belgium	Cat hair	Tavanti et al doi 10.1128/JCM.43.1.284-292.2005
CP9	103	London, UK	Anus	Tavanti et al doi 10.1128/JCM.43.1.284-292.2005
CP10	J930631/1	Africa	Cat hair	Tavanti et al doi 10.1128/JCM.43.1.284-292.2005

Table S4.2: Compound table with chromatographic and mass spectral validation

Retention time	Compounds	Base Peak	NIST match score	Chemical formula	Molecular weight	Retention index
2.354	Ethyl acetate	43	929	C4H8O2	88	888±8 (234)
2.865	Ethanol	45	946	C2H6O	46	932±8 (181
2.98	Furan, 2,5-dimethyl-	96	950	C6H8O	96	939±9 (40)
3.124	Propanoic acid, ethyl ester	57	952	C5H10O2	102	953±7 (87)
3.273	Propanoic acid, 2-methyl-, ethyl ester	43	651	C6H12O2	116	961±6 (99)
4.354	Butanoic acid, ethyl ester	71	761	C6H12O2	116	1035±8 (251)
4.464	Methyl thiolacetate	43	895	C3H6OS	90	1052±5 (15)
5.577	1-Propanol, 2-methyl	43	776	C4H10O	74	1092±9 (269)
5.774	1-Butanol, 3-methyl-, acetate	70	945	C7H14O2	130	1122±7 (168)
7.008	Propanoic acid, pentyl ester	57	765	C8H16O2	144	1239±13 (12)
7.649	1-Butanol, 3-methyl-	55	909	C5H12O	88	1209±9 (375)
8.493	Styrene	104	788	C8H8	104	1261±10 (102)

8.525	3-Buten-1-ol, 3-methyl-	56	795	C5H10O	86	1248±8 (72)
9.406	Acetoin	45	895	C4H8O2	88	1284±12 (241)
14.687	2,3-Butandiol	45	956	C4H10O2	90	1565±18 (4)
16.406	1-Propanol, 3-(methylthio)-	106	891	C4H10OS	106	1719±9 (91)
18.27	Benzyl alcohol	108	836	C7H8O	108	1870±14 (323)
18.649	Phenylethyl alcohol	91	840	C8H10O	122	1906±15 (423
19.885	2-Pyrrolidinone	85	944	C4H7NO	85	2020±17 (7)
19.976	Nerolidol	69	918	C15H26O	222	2042±10 (172)
21.674	1H-Pyrrole-2,5-dione	97	821	C4H3NO2	97	
22.455	2,3-Dihydrofarnesol	69	797	C15H28O	224	2262±10
22.941	Farnesol	69	876	C15H26O	222	2323±19 (16)
23.001	trans-Farnesol	69	903	C15H26O	222	2356±10 (61)
24.829	Succinimide	99	866	C4H5NO2	99	2438±21 (2)

170

Chapter 5: Non-Invasive Volatilomic Analysis of Infected and Non-infected Diabetic Foot Ulcers – Supplementary Information

Shane Fitzgerald[1], Linda Holland[2], Tommy Kyaw Tun[3], Eoghan O'Neill[3], Aoife Morrin*[1]
1. School of Chemical Sciences, National Centre for Sensor Research, Insight SFI Research Centre for Data Analytics, Dublin City University, Ireland
2. School of Biotechnology, Dublin City University, Ireland
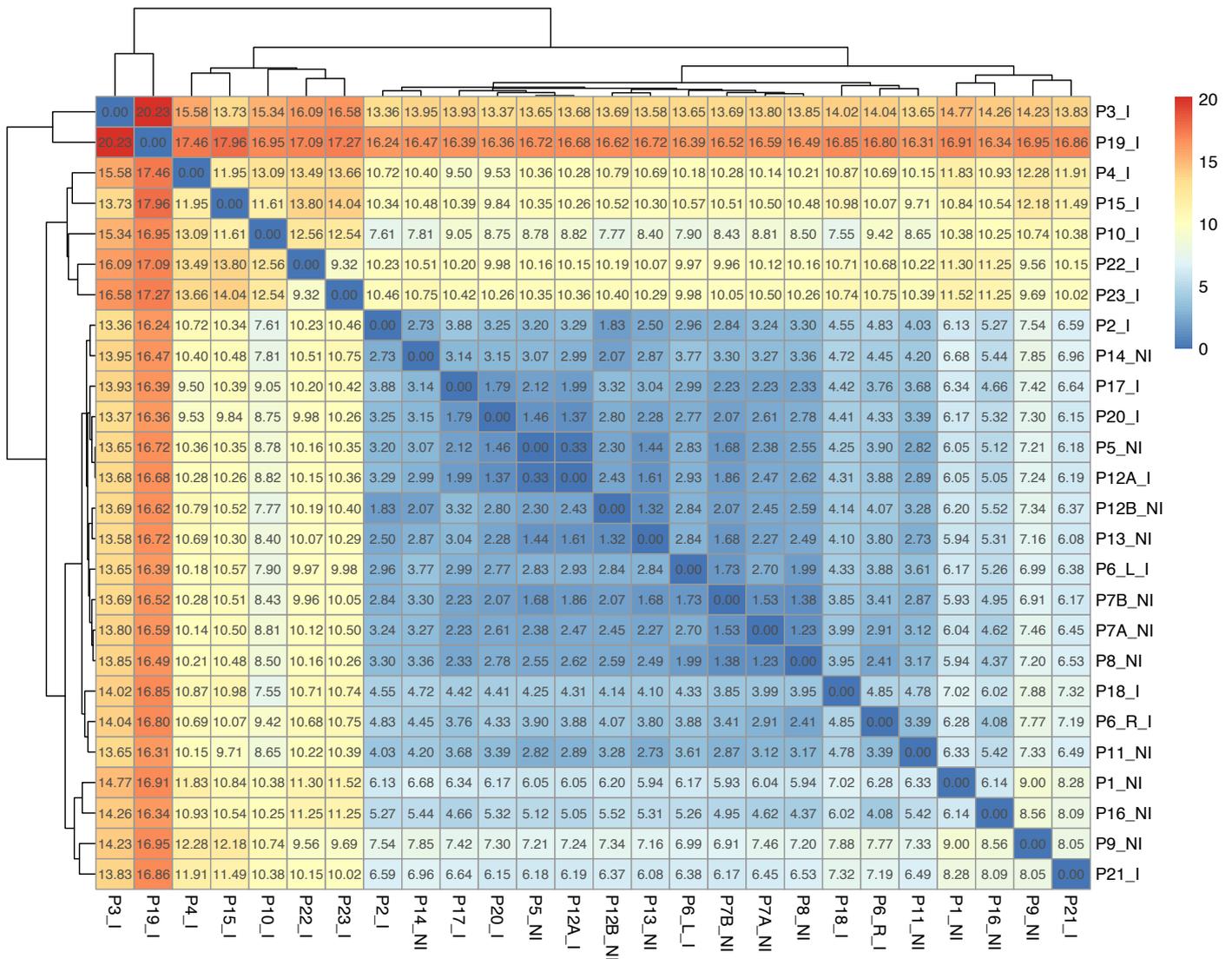3. Royal College of Surgeons, Connolly Hospital, Dublin, Ireland

Figure S5.1. Heatmap dissimilarity matrix representation of 26 patient samples samples. Values are calculated Euclidean distances. Red represents highly dissimilar samples; blue represents highly similar samples.
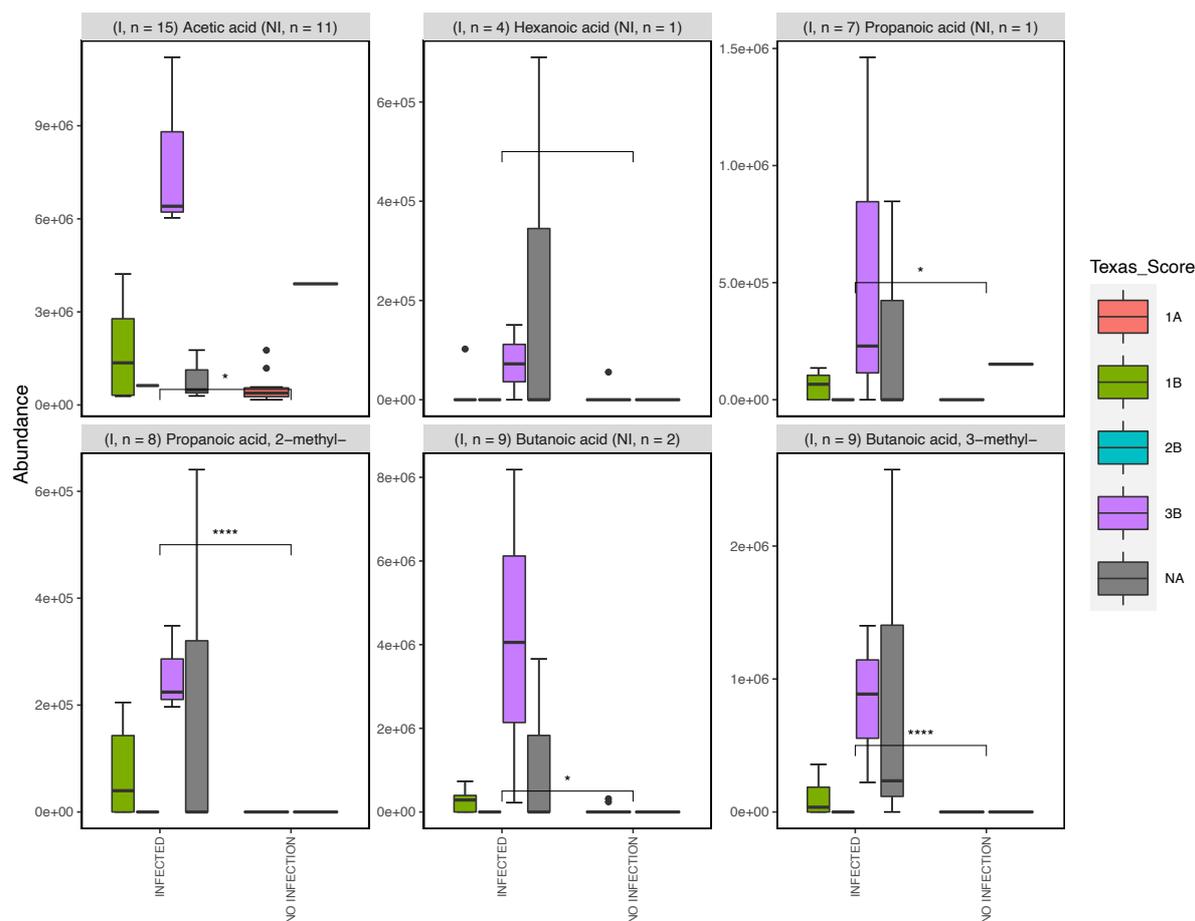
*Figure S5.2. : Comparative boxplots illustrating differences in the abundance of ketones across infected and non-infected wound samples. Each plot illustrates the abundance of an individual ketone compound recovered from varying wounds classifications (Texas score). The Texas score is a common classification system used for diabetic ulcers and is visualised through the various colours of the boxes: 1A (non-infected superficial wound; RED); 1B (infected superficial wound; GREEN); 2B (infected wound penetrating tendon; BLUE); 3B (infected wound penetrating to bone or joint; PINK); NA (infected and non-infected, non-diabetic wounds; GREY). The detection frequency for each compound across the samples is shown above each plot alongside the name of each compound where I = infected and NI = non infected. Statistically significant differences are illustrated through the star system where * = p < 0.05; ** = p < 0.01; *** = p < 0.001; **** = p < 0.0001.*
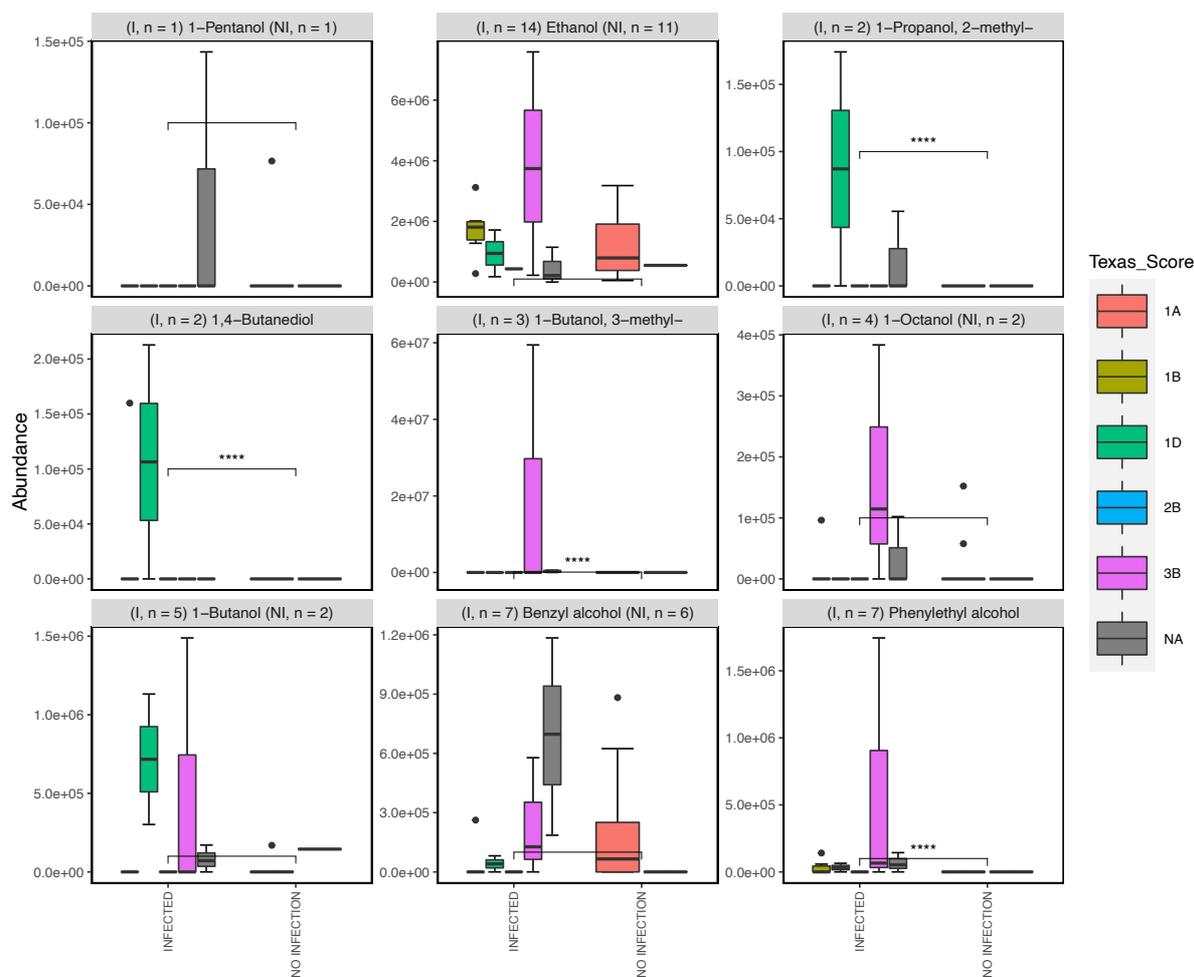
173

*Figure S5.3 : Comparative boxplots illustrating differences in the abundance of alcohols detected across infected and non-infected wound samples. Each plot illustrates the abundance of an individual alcohol compound recovered from varying wounds classifications (Texas score). The Texas score is a common classification system used for diabetic ulcers and is visualised through the various colours of the boxes: 1A (non-infected superficial wound; RED); 1B (infected superficial wound; GREEN); 2B (infected wound penetrating tendon; BLUE); 3B (infected wound penetrating to bone or joint; PINK); NA (infected and non-infected, non-diabetic wounds; GREY). The detection frequency for each compound across the samples is shown above each plot alongside the name of each compound where I = infected and NI = non infected. Statistically significant differences are illustrated through the star system where * = p < 0.05; ** = p < 0.01; *** = p < 0.001; **** = p < 0.0001.*
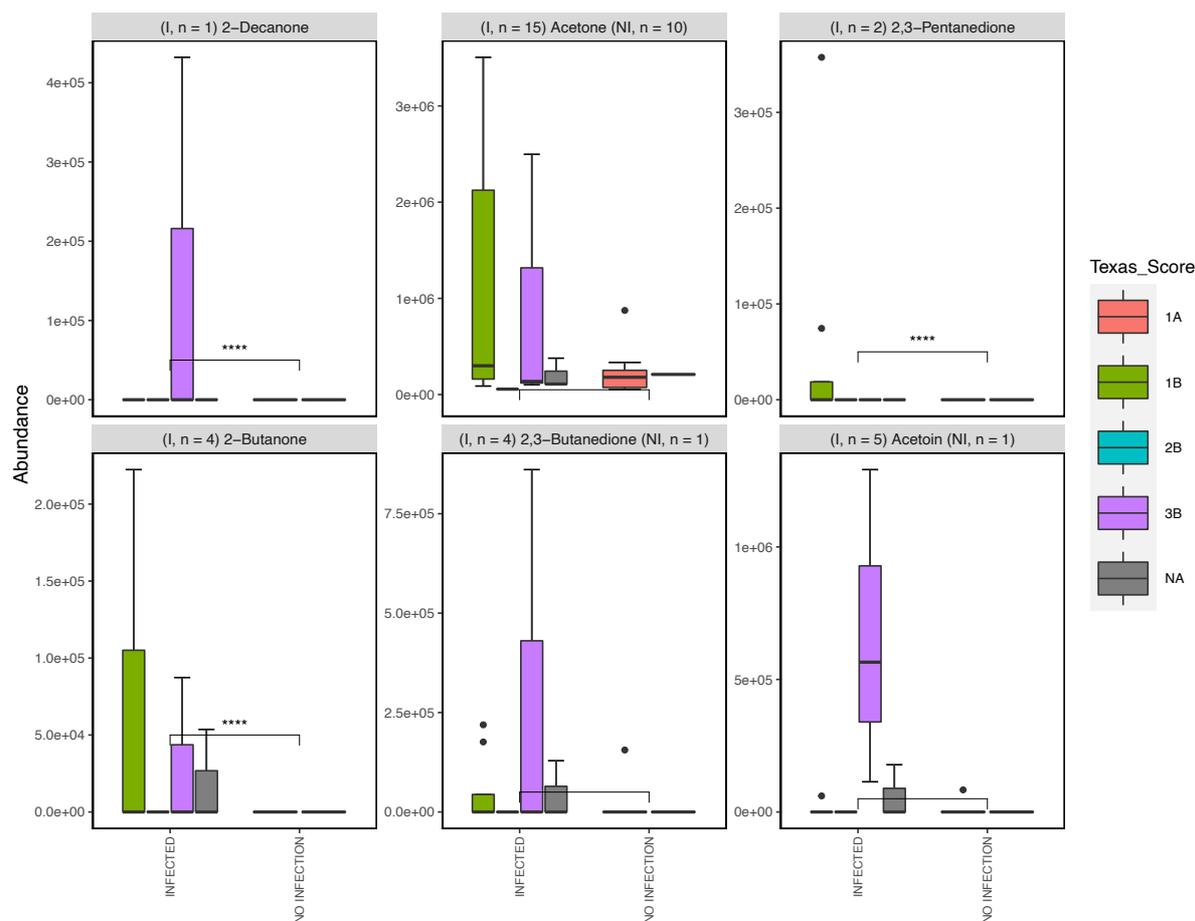
174

*Figure S5.4 : Comparative boxplots illustrating differences in the abundance of ketones detected across infected and non-infected wound samples. Each plot illustrates the abundance of an individual ketone compound recovered from varying wounds classifications (Texas score). The Texas score is a common classification system used for diabetic ulcers and is visualised through the various colours of the boxes: 1A (non-infected superficial wound; RED); 1B (infected superficial wound; GREEN); 2B (infected wound penetrating tendon; BLUE); 3B (infected wound penetrating to bone or joint; PINK); NA (infected and non-infected, non-diabetic wounds; GREY). The detection frequency for each compound across the samples is shown above each plot alongside the name of each compound where I = infected and NI = non infected. Statistically significant differences are illustrated through the star system where * = p < 0.05; ** = p < 0.01; *** = p < 0.001; **** = p < 0.0001.*
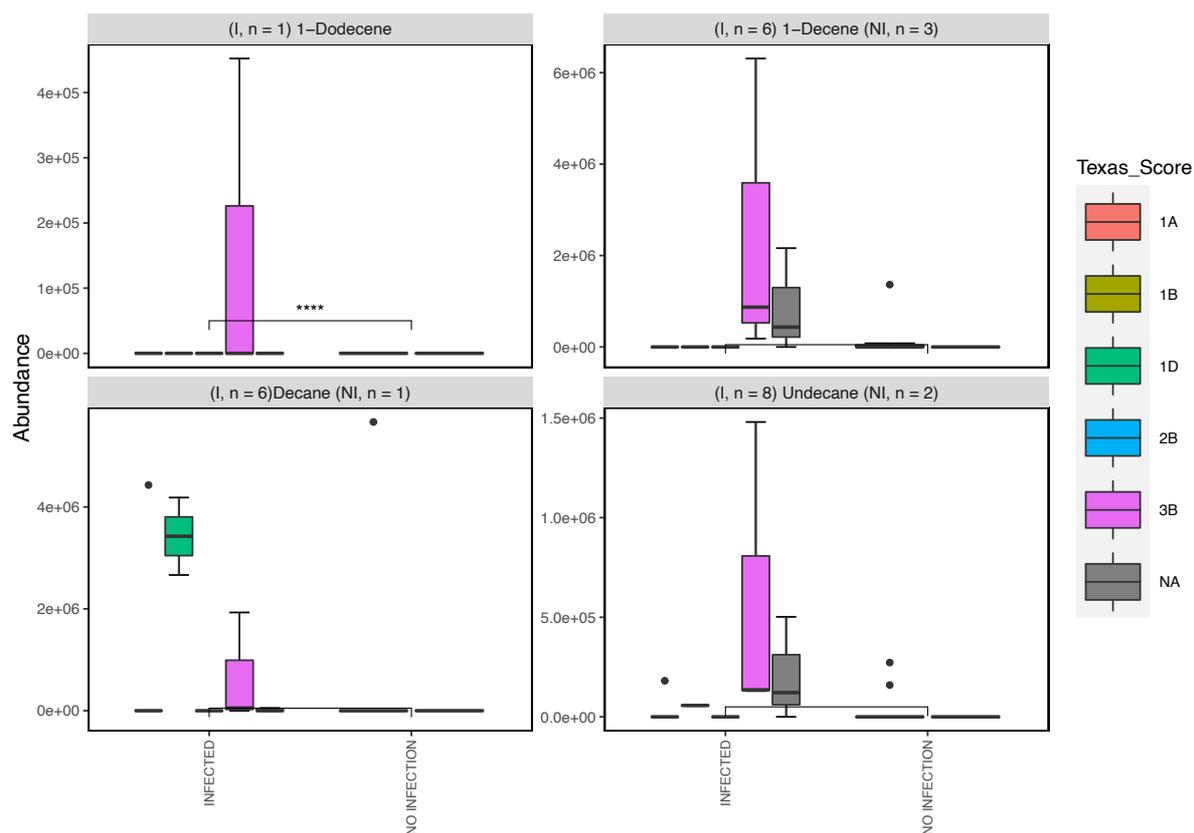
*Figure S5.5 : Comparative boxplots illustrating differences in the abundance of hydrocarbons detected across infected and non-infected wound samples. Each plot illustrates the abundance of an individual hydrocarbon compound recovered from varying wounds classifications (Texas score). The Texas score is a common classification system used for diabetic ulcers and is visualised through the various colours of the boxes: 1A (non-infected superficial wound; RED); 1B (infected superficial wound; GREEN); 2B (infected wound penetrating tendon; BLUE); 3B (infected wound penetrating to bone or joint; PINK); NA (infected and non-infected, non-diabetic wounds; GREY). The detection frequency for each compound across the samples is shown above each plot alongside the name of each compound where I = infected and NI = non infected. Statistically significant differences are illustrated through the star system where * = p < 0.05; ** = p < 0.01; *** = p < 0.001; **** = p < 0.0001.*
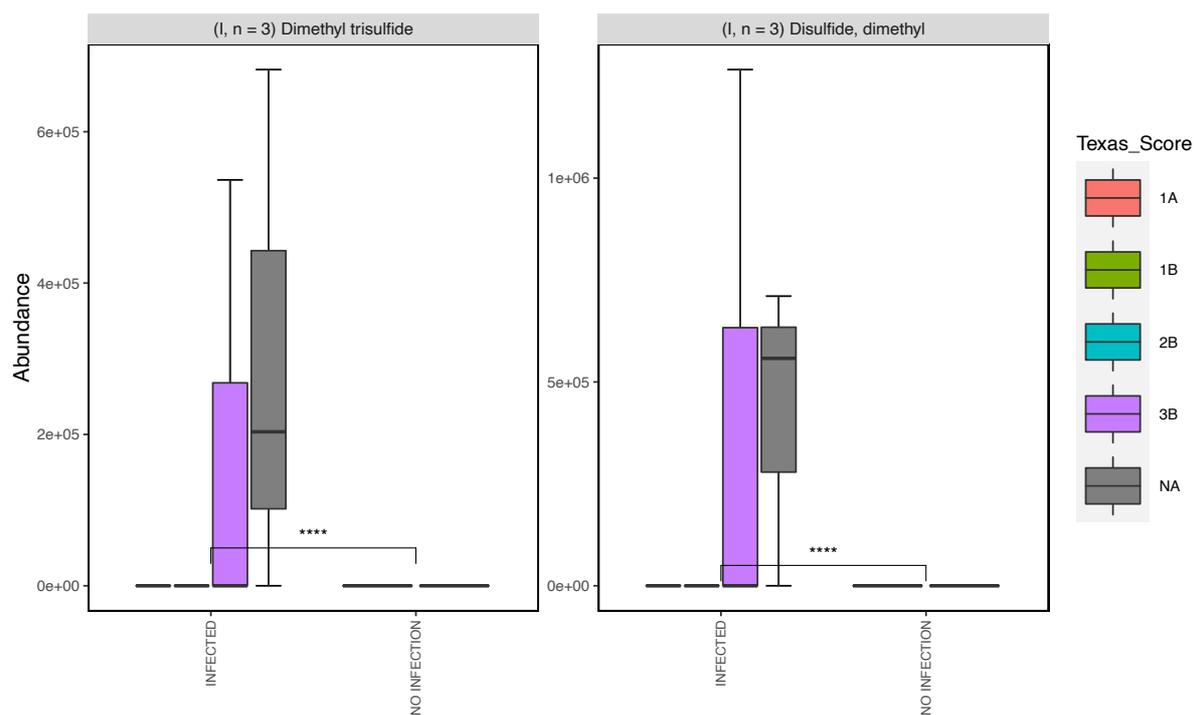
176

Figure S5.6 : Comparative boxplots illustrating differences in the abundance of sulfide detected across infected and non-infected wound samples. Each plot illustrates the abundance of an individual sulfide compound recovered from varying wounds classifications (Texas score). The Texas score is a common classification system used for diabetic ulcers and is visualised through the various colours of the boxes: 1A (non-infected superficial wound; RED); 1B (infected superficial wound; GREEN); 2B (infected wound penetrating tendon; BLUE); 3B (infected wound penetrating to bone or joint; PINK); NA (infected and non-infected, non-diabetic wounds; GREY). The detection frequency for each compound across the samples is shown above each plot alongside the name of each compound where I = infected and NI = non infected. Statistically significant differences are illustrated through the star system where * = p < 0.05; ** = p < 0.01; *** = p < 0.001; **** = p < 0.0001.
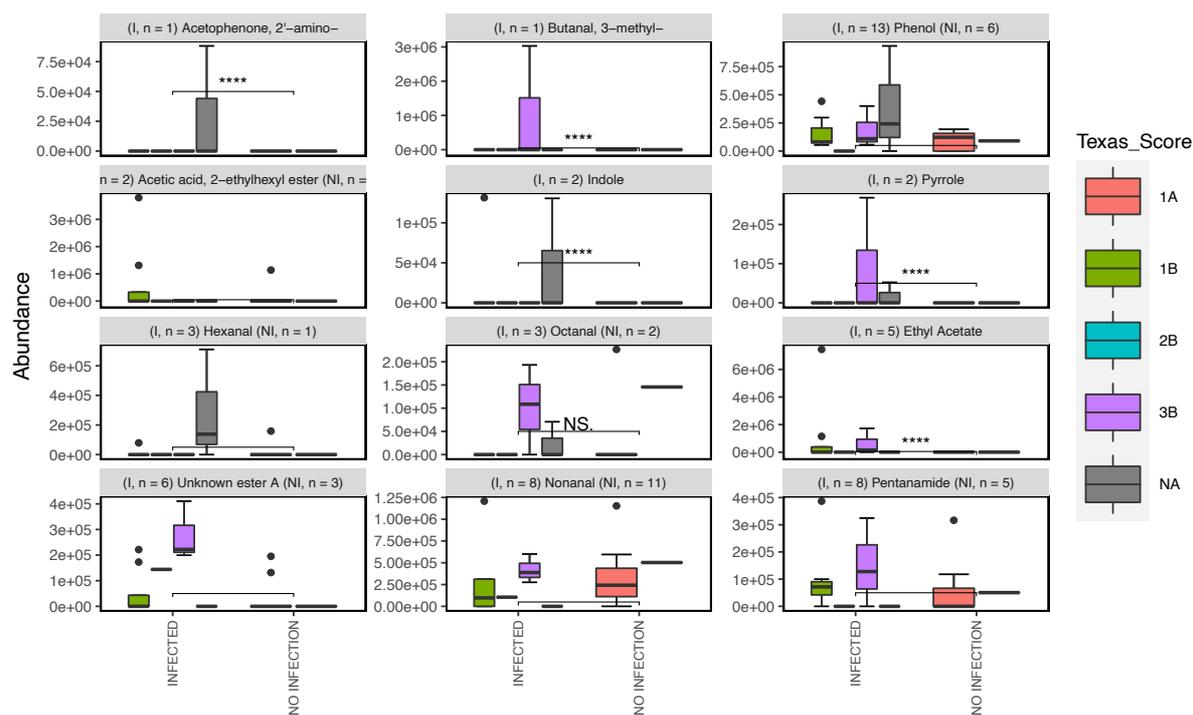
177

Figure S5.7 : Comparative boxplots illustrating differences in the abundance of various amine, amide, aldehyde, pyrrole, indole, acetate, and ester compounds detected across infected and non-infected wound samples. Each plot illustrates the abundance of an individual ketone compound recovered from varying wounds classifications (Texas score). The Texas score is a common classification system used for diabetic ulcers and is visualised through the various colours of the boxes: 1A (non-infected superficial wound; RED); 1B (infected superficial wound; GREEN); 2B (infected wound penetrating tendon; BLUE); 3B (infected wound penetrating to bone or joint; PINK); NA (infected and non-infected, non-diabetic wounds; GREY). The detection frequency for each compound across the samples is shown above each plot alongside the name of each compound where I = infected and NI = non infected. Statistically significant differences are illustrated through the star system where * = p < 0.05; ** = p < 0.01; *** = p < 0.001; **** = p < 0.0001.

Table S5.1: Patient clinical data

Wound ID	Date	Texas score	Plating results	Wound duration	DM status	DM type
1	08/06/2021	NA	scanty commensals	4 weeks	Not DM	
2	24/06/2021	NA	scanty commensals	5 years+	DM	Type 2
3	06/08/2021	NA	mixed anO2	3 years+	Not DM	
4	22/09/2021	B3	ent faecalis & acinetobacter p		DM	
5	30/09/2021	1A	mixed anaerobes, p mirabilis, Sau	6 Months	DM	Type 2
6L	04/10/2021	1B	P vulgarirs, ent faecalis	6 years +	DM	Type 2
6R	04/10/2021	1B	P vulgarirs, BHS Gp G	6 years +	DM	Type 2
7	05/10/2021	1A	P mirabilis	2 years +	DM	Type 2
8	06/10/2021	1A	S aureus, scanty e coli	2 years +	DM	Type 2
9	07/10/2021	1A		3 months	DM	TYPE 1
10	08/10/2021	NA				
11	13/10/2021	1A	BHS Group B S		DM	Type 2
12	13/10/2021	2B	Scanty Staph.aureus S	5 months +	DM	Type 2
13	18/10/2021	1A	Ent.faecalis , Serratia Liquef	6 months +	DM	Type 2
14	19/10/2021	1A		3 months	DM	Type 2
15	26/10/2021	3B	staph aureus	months	DM	Type 2
16	05/11/2021	1A	Proteus mirabil		DM	
17	05/11/2021	1B	Ps.aeruginosa, Staph aureus, Ent Faecalis		DM	
18	29/11/2021	1B	Proteus mirabil	4 months	DM	Type 2

179

19	08/12/2021	1A	Scanty Commensals	2 years +	DM	Type 2
20		1B			DM	
21	04/02/2022	1D	Proteus mirabil S		DM	
22	09/02/2022	1D	BHS Group B		DM	
23	10/02/2022	1B	Staph.aureus S Finegoldina mag		DM	