PROFILES OF VOLATILE COMPOUNDS AS MICROBIAL MARKERS IN APPLICATIONS OF BIOSECURITY AND BIOENERGY

By

Kristen Leigh Reese

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Chemistry−Doctor of Philosophy

2020

ABSTRACT

PROFILES OF VOLATILE COMPOUNDS AS MICROBIAL MARKERS IN APPLICATIONS OF BIOSECURITY AND BIOENERGY

By

Kristen Leigh Reese

All organisms emit volatile organic compounds (VOCs) and profiling of volatiles (volatilomics) is finding diverse applications. Some VOCs are consistently present throughout the lifecycles of organisms, while other VOCs are biomarkers, quantifiable indicators of changes in physiological state or reflective of environmental stresses. This dissertation describes research into volatile biomarkers of different microorganisms in the context of biosecurity and bioenergy. Untargeted analyses of microbial biomarkers were accomplished using solid phase microextraction (SPME) coupled to gas chromatography-mass spectrometry (GC-MS).

In the context of biosecurity, pathogenic bacteria can be used as the basis for a bio-terrorism attack. There is a need for deeper understanding of the chemical signatures of organisms, in particular when they infect individuals, and a need for methods for detecting these pathogens in the context of infections of humans. Current research has performed metabolite profiling of VOCs emitted in culture by surrogates for potential bacterial bioterrorism agents, *Bacillus anthracis* Sterne and *Francisella tularensis* novicida in conjunction with measurements of VOCs released by their fully virulent counterparts, *F. tularensis* SCHU S4 and *B. anthracis* Ames, both on the CDC category A bioterrorism and disease agent list. Methyl ketones, alcohols, esters, carboxylic acids, and nitrogen- and sulfur-containing compounds were attributed to the bacteria. The two genera showed distinct VOC profiles whereas the taxa within each genus showed subtler differences in VOC profiles. Growth phase influenced absolute and relative VOC abundances, indicating the potential for markers to discriminate growth phases. This *in vitro*

determination of VOC profiles laid groundwork for non-invasive probing of bacterial metabolism.

Towards bioenergy efforts, microalgae present a renewable alternative to producing biofuels. However, biofuels are more costly per gallon compared to non-renewable fossil fuels due to production and harvesting costs. Therefore, research driving increases in biomass production are of interest, specifically (1) better early-warning tools to anticipate and/or diagnose the presence of predators and (2) understanding algae-bacteria interactions, as they are challenging to manage and may help or harm algal productivity. Research towards part (1) aimed to better define the physiological state of algae ponds. A biofuel-relevant alga, *Microchloropsis salina*, was infected with a predator, the rotifer *Brachionus plicatilis*. SPME-GC-MS aided discovery of seven putative culture crash biomarkers, including carotenoid degradation products trans-β-ionone and β-cyclocitral, over several timepoints during active crashing of algal ponds that were not observed in healthy controls. These biomarkers offer potential as diagnostic tools to signal the need for crash mitigation strategies, as signals were detected before observed losses in algal cell density. Research towards part (2) aimed to detect and identify VOC biomarkers related to the micro-scale interactions of a model system of alga *P. tricornutum* and bacterium *Marinobacter* spp. 3-2. The presence of *Marinobacter* spp. 3-2, either in the form of live bacterial cells or sterile exudates, caused modest inhibition in growth rates of *P. tricornutum*. Substantial differences in VOC biomarker profiles were observed between 1) co-cultures of both organisms, 2) *P. tricornutum* exposed to *Marinobacter* spp. 3-2 exudates, and 3) *Marinobacter* spp. 3-2 exposed to *P. tricornutum* exudates, all relative to the VOC biomarker profiles of corresponding monocultures. Increasing the knowledge base of algae-bacterial interactions will enable a deeper understanding of the basic science of microorganism signaling.

ACKNOWLEDGEMENTS

As this dissertation was the culmination of an unusual path, there are many people deserving of my deepest thanks for help along the way.

First, I would like to thank to my committee. To my Michigan State University Ph.D. advisor Dr. A. Daniel Jones, you have led by example with high standards and strong work ethic, admirable qualities which, through emulation, have allowed me to truly flourish. Your patience, understanding, and care have allowed me to finally reach the end of my graduate studies, and I will forever be grateful. To Dr. Matthias Frank, my LLNL mentor, thank you for your encouragement, support, and insight. You have helped me to envision the bigger picture of my research and provide context. Your invaluable mentorship over the past four years has truly been appreciated. Finally, to Dr. Dana Spence and Dr. Gary Blanchard, thank you both for serving on my committee and for guiding me through teaching assistant positions and academic courses.

I would like to thank collaborators at Lawrence Livermore National Laboratory, including Dr. Amy Rasley, Dr. Julia Avila, Dr. Matt Coleman, Dr. Xavier Mayali, Dr. Rhona Stuart, Dr. Vanessa Brisson, and Kristina Rolison. It has been a great privilege to work alongside these talented people and learn from them. I am extremely grateful to collaborators at Sandia National Laboratories, in particular Dr. Todd Lane, Pamela Lane, and Dr. Carolyn Fisher. Thank you for showing me the wonderful world of algae.

I want to extend my appreciation to other members of the MSU Community that have helped me to not only persist but thrive over the entire course of my graduate career. To Dr. Ruth Smith for supporting me through my Forensic Master's degree and beyond. In particular, to

members of Dr. Jones' research group and Dr. Smith's research group, thank you for sharing your knowledge and expertise and helping me to hone my research skills.

Thank you to the friends I've made during my time at MSU. A special thanks to Fanny Chu, a dear friend who has supported me throughout -- from listening to my presentations for the 1000th time or willing to grab early morning dim sum, I don't know how I could have managed this journey without you. To Michelle McDaniel, Kristin Shannon, Tim Shannon, Tanner McDaniel, Corey Jones, Travis Bethel, and Pengchao Hao, thank you for reminding me to always embrace a life both inside and outside of the lab, to daily laughs on the Chemistry Chat, and that good friends persist even while far apart. To Yu-Ling Lien, always a voice of reason and wit with a smile, you are gone too soon but never far from my thoughts.

Thank you to my family -- you are my reason to keep going. Mom, Dad, and Kelly, you've continuously encouraged my pursuit of higher education, even when you didn't fully understand my science-lingo. You are my biggest fans, inspiring me even when I felt like nothing was going right. I'm so lucky to have you in my life and can't wait to come home.

Finally, I will acknowledge several groups and funding sources that made my research possible. I thank the LLNL Forensic Science Center (FSC) and, in particular, Audrey Williams and Deon Anex, for use of laboratory equipment and space and Roald Leif for assisting with instrumental setup and maintenance. Partial funding for this work was provided by the LLNL Laboratory-Directed Research and Development (LDRD) projects 17-LW-021 and 19-FS-035 as well as by Sandia National Laboratories. This work was performed, in parts, under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52–07NA27344 and at Sandia National Laboratories for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

PREFACE

The Lawrence Livermore National Laboratory, Office of Scientific and Technical Information, Information Management (IM) number associated with this document is LLNL-TH-811155. This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1 VOLATILE MARKERS AS DIAGNOSTIC TOOLS TO ASSESS HEALTHY VS. UNHEALTHY STATES OF A VARIETY OF ORGANISMS

The generation of volatile organic compounds (VOCs) from natural or man-made processes is ubiquitous throughout the world. VOCs have been defined as low molecular mass carbon-containing compounds that have low boiling points and measurable vapor pressures at Standard Temperature and Pressure (20 °C, 1 atm) [1]. Human perception of VOCs is through the sensation of smell, whether fragrant or foul, arising from VOCs interacting with receptors in the nasal cavity. Applications of VOCs are widespread in society, from industry pursuits - cosmetics, pharmaceuticals and medicine, environmental science, and food science – to military pursuits - energy and national security. The study of volatiles in this plethora has been recently referred to as a new branch of the -omics world, volatilomics, where the VOC profile of an organism is referred to as its "volatilome" [2-4]. Research within volatilomics has rapidly expanded within the last decade [5], indicating this field provides complementary information to the existing "-omics" fields, in particular, metabolomics and exposomics.

The volatilomes of human, plant, and animal models are composed of VOCs that are consistently present throughout the lifecycle of the organism, while VOCs that vary in abundance are biomarkers, quantifiable indicator of a change in health state, such as an infection or disease or as a result of environmental exposure. The study of biomarkers in human and animal models has been accelerated by improvements in the analytical approaches (sample collection and pre-concentration) and development of novel instrumentation (chromatography and mass spectrometry), enabling the detection of qualitative and/or quantitative levels of VOCs in the headspace of different sample matrices.

Areas of national security where the study of the volatilome has gained momentum include biosecurity and bioenergy. The measures designed to lessen the transmission or intentional exposure of infectious diseases to the general public are characterized as biosecurity measures or biological countermeasures, while the restoration of a stable state to people, animals, or wildlife exposed to biological threats is characterized as remediation. Research presented in chapter 3 investigates biosecurity efforts into detection of potential volatile biomarkers for diagnosing the presence of pathogenic bacteria, specifically the species *Francisella tularensis* and *Bacillus anthracis*. For the area of bioenergy, increasing the productivity of commercial algae ponds offers potential to improve the economics of algal biofuels. In this light, research described in chapters 4 and 5 focuses on detection of volatile biomarkers indicative of biotic stress including the action of predators or symbiosis/competition with bacteria in culture, respectively.

1.2 OVERVIEW OF DISSERTATION CHAPTERS

1.2.1 VOC data acquisition and data processing – Chapter 2

Chapter 2 will present an overview of the data acquisition and data processing methods used in this dissertation. Solid-phase microextraction (SPME) is a popular approach to the collection and pre-concentration of VOCs in the gaseous headspace of samples and was used throughout multiple investigations. Within the context of this work, SPME offered non-invasive, non-destructive monitoring of the metabolite profiles for microorganisms relevant to biosecurity and bioenergy applications. Analytes adsorbed/absorbed onto the SPME fibers were separated and detected through a gas chromatography-mass spectrometry (GC-MS) analysis. An untargeted metabolomics approach was utilized in this work to broadly annotate the metabolites

detected across growth states. VOCs were identified as candidate microbial biomarkers through development of filtering criteria designed to eliminate background VOCs while allowing for biological variability observed in biological systems.

1.2.2 Health and Biodefense: Biomarkers for detection of pathogenic bacteria – Chapter 3

Over a century of research has been dedicated to studying the relationships between specific microorganisms and diseases caused by their interaction with humans and animals[6]. Some of the research has produced great benefits for human health and medicine, such as the development of vaccines and therapeutic drugs, which have increased the human lifespan. Responsible and safe handling of infectious pathogens, facilitated by greater protective measures to minimize risks, occurs in thousands of labs daily, with few incidents of release into the general public. However, there is a subset of pathogens that pose severe threats to human, animal, and plant health due to increased virulence, and some of them have been used deliberately to cause disease and instigate fear against military personnel and/or the general public. The U.S. Centers for Disease Control and Prevention (CDC) lists pathogens that pose a high risk to national security as Category A agents because of high person-to-person transmission rates, high mortality rates, a high likelihood of public disruption, and requires special preparations to ensure public health[7]. Examples of Category A agents and the resulting diseases include *Bacillus anthracis* (Anthrax), *Clostridium botulinum* (Botulism), *Yersinia pestis* (Plague), *Variola major* (Smallpox), and *Francisella tularensis* (Tularemia). The National Institute of Allergy and Infection Diseases (NIAID)[8] has stated “biological weapons in the possession of hostile states or terrorists… are among the greatest security challenges to the United States”. Therefore, it is of great importance to develop (1) a deeper understanding of the organisms, in particular their

actions when actively infecting individuals, and (2) methods capable of detecting these pathogens in the environment, in particular in the context of human/pathogen infections.

One of the most direct routes for human exposure for pathogens is through aerosolized microbes entering human airways via the mouth or nose during inhalation[9]. The pathogen can colonize and proliferate on human cells in the bronchial airways. Detection of pathogens at this early stage, before the individual displays relevant clinical symptoms, would facilitate early diagnosis and the work of healthcare providers by providing more time to treat an infection. Exhaled breath analysis has been explored as a diagnostic tool due to the breath's large chemical complexity and information content[4]. Breath is composed of thousands of VOCs and non-volatile trace compounds (e.g., proteins) embedded in aerosol droplets. Through the circulatory system, oxygen from the air is transferred via the lungs into blood, which is delivered to muscles and organs, and carbon dioxide is simultaneously removed. Besides oxygen and carbon dioxide exchanging at the blood-air interface, other compounds, both VOCs and non-volatiles, are incorporated into this system. The VOCs in this system are directly influenced by the immediate environment of the subject as well as the natural human metabolic processes. Thus, air exhaled from the lungs can be considered a headspace of circulating blood, where compounds are constantly partitioning. Changes in breath volatiles *in vivo* have been previously researched in relation to respiratory diseases, such as chronic obstructive pulmonary disease[10] and asthma[10], or in relation to non-respiratory diseases, such as diabetes[11] and breast cancer[12]. The use of exhaled breath holds additional advantages in the field of personal diagnostics, as the sampling is non-invasive and samplers have been developed that can be performed by subjects in a home environment.

Study of VOCs in breath research has also begun expansion beyond *in vivo* profiling of biomarkers associated with a specific disease to design *in vitro* models that could approximate a host body, and thus study the cellular origins of VOCs. Select prior studies have established that metabolites from proliferating bacteria detected in the headspace of *in vitro* cultures can be detected in the exhaled breath of infected animal models or humans, including *Mycobacterium tuberculosis* [13], *Heliobacter pylori* [14], and *Pseudomonas aeruginosa* [15]. However, replication of the complex cellular makeup of the human airways and respiration system in a cell culture is challenging. The generation of VOCs can occur from the mammalian cells and pathogen cells, which are likewise dependent on physiological state, and the host immune system may generate a specific response to the infection. Therefore, initial studies into the growth of pathogens can employ monocultures, allowing observations of species-specific and growth stage dependent VOC emissions on different nutrient-rich medias, all of which impact the volatile profile. Thereafter, increasing the complexity of *in vitro* models, and eventually transitioning into *in vivo* studies in animal/human models, would assist in establishing the potential origins of exhaled breath VOCs.

Chapter 3 describes efforts to conduct comprehensive (untargeted) metabolite profiling of volatile organic compounds (VOCs) emitted in culture by bacterial taxa *Francisella tularensis (F. tularensis)* novicida and *Bacillus anthracis (B. anthracis)* Sterne, surrogates for potential bacterial bioterrorism agents, as well as selective measurements of VOCs from their fully virulent counterparts, *F. tularensis* SCHU S4 and *B. anthracis* Ames that are both on the CDC category A bioterrorism and disease agent list. *F. tularensis* and *B. anthracis* were grown in liquid broth for time periods that covered logarithmic growth, stationary, and death phases. VOCs emitted from these cultures were collected from the headspace above the cultures as well

as from control samples without bacteria using solid phase microextraction (SPME) and were analyzed using gas chromatography-mass spectrometry (GC-MS). Criteria were developed for distinguishing VOCs originating from bacteria versus background VOCs (originating from growth media or sampling devices). Analyses of collected volatiles revealed methyl ketones, alcohols, esters, carboxylic acids, and nitrogen- and sulfur-containing compounds that were attributed to the bacteria. The two bacterial genera showed distinct VOC profiles whereas the taxa within each genus showed more subtle differences in VOC profiles, illustrating the potential for VOC profiles to distinguish biosecurity-relevant pathogens at the genus and species-level. Growth phase influenced VOC abundance, indicating the potential for VOC profiles to discriminate growth phases. This *in vitro* determination of VOC profiles lays the groundwork for non-invasive probes of bacterial metabolism. Such VOC profiles offer prospects for detection of taxa-specific VOC biomarkers from various potential biowarfare agents, with applications in bio-detection and in targeted breath-based diagnostics for detection of biological agents proliferating in the lungs of infected patients after a suspected biological attack.

1.2.3 Bioenergy: Evaluation of algal wounding signals in the presence or absence of predators – Chapter 4

There is high interest in the monitoring of VOC emissions from algal cultures due to the importance of algae as renewable energy resource and as a potential commercial feed or food source[16,17]. Microalgae are important autotrophs in ecological communities, serving as a sources of nutrition for higher organisms[18]. Microalgae are a diverse collection of photosynthetic, microscopic organisms, with an estimated several thousand species ranging in size from 1-100^{+} microns[18]. Select algae species produce and accumulate an abundance of

lipids and carbohydrates, which can be converted into biodiesel and ethanol, respectively, each of which can be used as alternative renewable energy source to compete with the production of non-renewable petroleum[18,19]. Additionally, algal proteins have a high nutritional value and can be utilized for human consumption. Biofuel production using algae as a precursor has several potential benefits. Algae fix atmospheric carbon dioxide and sunlight to create biomass for growth, producing approximately 50-66% of atmospheric oxygen as a by-product, and could be used to scrub industrial carbon dioxide emissions from the air[20]. High-production cultivation of algae is facilitated by rapid growth rates. Algae can be grown in freshwater, saltwater, and terrestrial areas (including some not suited for agricultural use)[19].

Algae can be cultured in the laboratory using controlled conditions, but relatively low volumes. Larger volumes of algae in commercial practice involve culture in open ponds or photobioreactors, where the environmental and exposure variables are more challenging to control, similar to a natural marine habitat. However, generation and maintenance of low-cost volumes of biomass production required for industrial pursuits is threatened by biological contaminations. In particular, algae ponds used in industrial biomass production are susceptible to pathogen or grazer infestation, resulting in pond crashes with high economic costs.

Current methods to monitor and mitigate unhealthy ponds are hindered by a lack of early indicators that precede culture crash[20]. It is known that algae produce VOCs as secondary metabolites during growth, and levels are affected by the algal species, growth phase, and environmental conditions (light, pH, water, nutrients, dissolved gasses, etc.)[20]. The research presented in chapter 4 describes efforts towards identifying early changes in the volatile biomarker metabolite profiles emitted from healthy and unhealthy cultures of *Microchloropsis salina* exposed to a known predator, the marine rotifer *Brachionus plicatilis*. An untargeted

analysis of VOCs, followed by preliminary identification, was accomplished using SPME-GC-MS. The addition of *B. plicatilis* to healthy cultures of *M. salina* resulted in decreased algal cell numbers, relative to uninfected controls, and generated *trans*-β-ionone and β-cyclocitral, which were attributed to carotenoid degradation. The abundances of the carotenoid-derived VOCs increased with rotifer consumption of algae and were detected prior to when conventional monitoring methods indicated distress. The results indicate that specific VOCs released by infected algae cultures may be early indicators for impending pond crashes, providing a useful tool to monitor algal biomass production and pond crash prevention.

1.2.4 Bioenergy: Introduction to algal-bacterial interactions – Chapter 6

Expanding on the discussion in Section 1.2.3, there are several routes to lowering the cost of biofuel production from microalgae to be competitive relative to petroleum. One avenue involves producing higher densities of algal cultures in order to increase the biomass that can be converted into biofuel. In order to increase the density of algal communities, a potential route involves exploiting interactions between algae and microorganisms, such as bacteria, that naturally inhabit the same space. Bacteria and microalgae have coexisted over millions of years of evolution. Exchange of nutrients, signaling molecules, and other dissolved materials between organisms happens in the area of closest contact, which has been termed the phycosphere (analogous to the rhizosphere around roots)[21]. Within the phycosphere, the exchange of metabolites, both volatile and non-volatile, can further affect the growth cycles of each organism. The detection, identification, and quantification of these molecules in complex communities would enable prediction and/or manipulation of these interactions for the purposes of basic science and commercial applications.

Both symbiotic and inhibitory interactions have been observed between select combinations of algae and bacteria, where the complexity of the systems is limited in order to more accurately determine the origins of identified metabolites. In combinations where bacteria promote algal growth, identification of exuded metabolites could identify targets that could be added to commercial cultures, and thus increase biomass production[22]. Several studies have concluded that algae can utilize vitamins and recycled minerals exuded by bacteria, while bacteria can utilize exuded ammonia from algae as a nitrogen source[23-25]. In combinations where bacteria inhibit algal growth, competition occurs between the organisms for nutrients and dissolved organic materials[26]. The effects can be either specific to a species or inhibit a broad range of algae. Studies have indicated that inhibitory or non-growth promoting interactions play an important role in organizing the structure of marine communities and ensuring survival[26,27]. Identification of inhibitory compounds may allow control of harmful phenomenon in algal systems, such as algae blooms, that produce toxic effects on marine wildlife and can contaminate sources of drinking water[28].

Chapter 5 presents efforts to detect, identify, and understand how algae and bacteria interact at the microscopic level using model algae-bacteria co-cultures. To this end, volatile metabolites detected in the headspace of co-cultures of model algae *Phaeodactylum tricornutum (P. tricornutum)* and model bacteria *Marinobacter spp. 3-2* were investigated. *P. tricornutum* is an alga used for biofuel production because of its high lipid content (estimated ~30% of dry weight[29]), and *Marinobacter* is a diverse genus of Gram-negative, aerobic bacteria found in most oceans[30]. This work demonstrated the feasibility of an untargeted SPME-GC-MS profiling of VOC marker detection and researched how VOC profiles changed between exponential and stationary growth states. Additionally, biomarkers from cultures of algae or

bacteria supplemented with exuded materials from the opposing organism were profiled to check if the same biomarkers from co-cultures could be replicated when the presence of one species is "simulated". If true, this would indicate soluble compounds produced by one species may trigger production of the VOCs by the other species.

The presence of *Marinobacter* spp. 3-2, either as live bacteria or as exudate containing bacterial metabolites, inhibited the growth of *P. tricornutum* to a small degree. Compared to monocultures of algae or bacteria, some differences were observed in biomarker profiles for 1) co-cultures of both organisms, 2) *P. tricornutum* exposed to *Marinobacter* spp. 3-2 exudates, and 3) *Marinobacter* spp. 3-2 exposed to *P. tricornutum* exudates. Increasing the knowledge base of algae-bacterial interactions at the phycosphere and alterations in microorganism physiology will enable better prediction and/or manipulation of these interactions for commercial purposes as well as a deeper understanding of the basic science of microorganism signaling.

REFERENCES

REFERENCES

1 U.S.E.P.A. *Technical Overview of Volatile Organic Compounds*, <https://www.epa.gov/indoor-air-quality-iaq/technical-overview-volatile-organic-compounds> (2017).

2 Amann, A. *et al.* The human volatilome: volatile organic compounds (VOCs) in exhaled breath, skin emanations, urine, feces and saliva. *J. Breath Res.* **8**, 034001, doi:10.1088/1752-7155/8/3/034001 (2014).

3 Rowan, D. D. Volatile metabolites. *Metabolites.* **1**, 41-63, doi:10.3390/metabo1010041 (2011).

4 *Breathborne Biomarkers and the Human Volatilome*. 2nd edn, (ed. Beauchamp, J., Davis, C, Pleil, J. D.) (Elsevier, 2020).

5 Broza, Y. Y., Mochalski, P., Ruzsanyi, V., Amann, A. & Haick, H. Hybrid volatolomics and disease detection. *Angew. Chem.* **54**, 11036-11048, doi:10.1002/anie.201500153 (2015).

6 Gest, H. The discovery of microorganisms by Robert Hooke and Antoni van Leeuwenhoek, Fellows of The Royal Society. *Notes and Records of the Royal Society of London* **58**, 187-201, doi:doi:10.1098/rsnr.2004.0055 (2004).

7 U.S. Centers for Disease Control and Prevention. CDC. *Bioterrorism Agents/Diseases*, <https://emergency.cdc.gov/agent/agentlist-category.asp> (2018).

8 U.S. Department of Heath and Human Services, U.S. National Institues of Health. NIAID Strategic Plan for Biodefense Research. (Bethesda, MD, 2007).

9 Tellier, R., Li, Y., Cowling, B. J. & Tang, J. W. Recognition of aerosol transmission of infectious agents: a commentary. *BMC Infectious Diseases* **19**, 101, doi:10.1186/s12879-019-3707-y (2019).

10 van de Kant, K. D. G., van der Sande, L. J. T. M., Jöbsis, Q., van Schayck, O. C. P. & Dompeling, E. Clinical use of exhaled volatile organic compounds in pulmonary diseases: a systematic review. *Respir Res* **13**, 117, doi:10.1186/1465-9921-13-117 (2012).

11 Minh, T. D. C., Blake, D. R. & Galassetti, P. R. The clinical potential of exhaled breath analysis for diabetes mellitus. *Diabetes Res Clin Pract* **97**, 195-205, doi:10.1016/j.diabres.2012.02.006 (2012).

12 Lavra, L. *et al.* Investigation of VOCs associated with different characteristics of breast cancer cells. *Scientific reports* **5**, 13246, doi:10.1038/srep13246 (2015).

13 Phillips, M. *et al.* Volatile biomarkers of pulmonary tuberculosis in the breath. *Tuberculosis* **87**, 44-52, doi:10.1016/j.tube.2006.03.004 (2007).

14 Lechner, M. *et al.* H. pylori infection increases levels of exhaled nitrate. *Helicobacter* **10**, 385-390 (2005).

15 Shestivska, V. *et al.* 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**, 2459-2467 (2011).

16 Odjadjare, E. C., Mutanda, T. & Olaniran, A. O. Potential biotechnological application of microalgae: a critical review. *Crit Rev Biotechnol* **37**, 37-52, doi:10.3109/07388551.2015.1108956 (2017).

17 Hamed, I. The Evolution and Versatility of Microalgal Biotechnology: A Review. *Comprehensive Reviews in Food Science and Food Safety* **15**, 1104-1123, doi:10.1111/1541-4337.12227 (2016).

18 Chisti, Y. Biodiesel from microalgae. *Biotechnology Advances* **25**, 294-306, doi:https://doi.org/10.1016/j.biotechadv.2007.02.001 (2007).

19 Chauton, M. S., Olsen, Y. & Vadstein, O. Biomass production from the microalga Phaeodactylum tricornutum: Nutrient stress and chemical composition in exponential fed-batch cultures. *Biomass and Bioenergy* **58**, 87-94, doi:10.1016/j.biombioe.2013.10.004 (2013).

20 Achyuthan, K. E., Harper, J. C., Manginell, R. P. & Moorman, M. W. Volatile metabolites emission by in vivo microalgae—An overlooked opportunity? *Metabolites* **7**, 39 (2017).

21 Seymour, J. R., Amin, S. A., Raina, J. B. & Stocker, R. Zooming in on the phycosphere: the ecological interface for phytoplankton-bacteria relationships. *Nature microbiology* **2**, 17065, doi:10.1038/nmicrobiol.2017.65 (2017).

22 Amin, S. A., Parker, M. S. & Armbrust, E. V. Interactions between diatoms and bacteria. *Microbiol Mol Biol Rev* **76**, 667-684, doi:10.1128/MMBR.00007-12 (2012).

23 Le Chevanton, M. *et al.* Screening and selection of growth-promoting bacteria for Dunaliella cultures. *Algal Research* **2**, 212-222, doi:10.1016/j.algal.2013.05.003 (2013).

24 Kim, B.-H., Ramanan, R., Cho, D.-H., Oh, H.-M. & Kim, H.-S. Role of Rhizobium, a plant growth promoting bacterium, in enhancing algal biomass through mutualistic interaction. *Biomass and Bioenergy* **69**, 95-105, doi:10.1016/j.biombioe.2014.07.015 (2014).

25 Wang, H., Hill, R. T., Zheng, T., Hu, X. & Wang, B. Effects of bacterial communities on biofuel-producing microalgae: stimulation, inhibition and harvesting. *Critical Reviews in Biotechnology* **36**, 341-352, doi:10.3109/07388551.2014.961402 (2016).

26 Mayali, X. & Azam, F. Algicidal bacteria in the sea and their impact on algal blooms1. *Journal of Eukaryotic Microbiology* **51**, 139-144, doi:10.1111/j.1550-7408.2004.tb00538.x (2004).

27 Paul, C., Mausz, M. A. & Pohnert, G. A co-culturing/metabolomics approach to investigate chemically mediated interactions of planktonic organisms reveals influence of bacteria on diatom metabolism. *Metabolomics* **9**, 349-359, doi:10.1007/s11306-012-0453-1 (2013).

28 Zerrifi, S. E. A., El Khalloufi, F., Oudra, B. & Vasconcelos, V. Seaweed bioactive compounds against pathogens and microalgae: Potential uses on pharmacology and harmful algae bloom control. *Mar Drugs* **16**, 55, doi:10.3390/md16020055 (2018).

29 Yang, Y. H. *et al.* Fatty acid and lipid class composition of the microalga phaeodactylum tricornutum. *Journal of oleo science* **66**, 363-368, doi:10.5650/jos.ess16205 (2017).

30 Gauthier, M. J. *et al.* Marinobacter hydrocarbonoclasticus gen. nov., sp. nov., a new, extremely halotolerant, hydrocarbon-degrading marine bacterium. *International Journal of Systematic and Evolutionary Microbiology* **42**, 568-576, doi:10.1099/00207713-42-4-568 (1992).

CHAPTER 2: DEVELOPMENT OF AN UNTARGETED SPME-GC-MS WORKFLOW TO IDENTIFY VOLATILE BIOMARKERS EMITTED FROM LIQUID CULTURES

2.1 INTRODUCTION

The research described in this dissertation explores ways in which detection of biomarkers unique to the volatilomes of a microorganism can assist in biodefense and bioenergy applications. The objectives of the investigations described in subsequent chapters are divergent, but the types of data generated and the methods for separating, detecting, and identifying biomarkers from a complex matrix are similar. Therefore, this chapter provides a brief overview of the pertinent data acquisition and data processing strategies.

2.2 STRATEGIES FOR DATA ACQUISITION OF VOLATILE METABOLITES

2.2.1 Overview of common VOC samplers

Several methods and/or devices have been developed for collection of VOCs from different environments, each with advantages and disadvantages. Solid phase microextraction (SPME) has grown to become one of the most popular approaches owing to its convenience and simplicity. SPME uses a sorbent material coated on the end of a fiber to passively extract volatile analytes from the vapor phase. SPME fibers can be introduced into heated gas chromatography (GC) injectors and the elevated temperatures release trapped VOC analytes into the GC column. A competing approach, known as thermal desorption (TD) sampling, involves active sampling of a gaseous headspace by drawing vapors over an adsorbent mix packed inside a glass or stainless steel tube. Adsorbed volatiles are removed by flowing carrier gas through the heated tube and into the GC column for further analysis. A growing alternative to the use of selective, sorbent materials involves the use of broad-range sensor arrays in electronic noses. Electronic noses have

been developed as mimics to the functionality of a human nose for rapid, real-time approaches to VOC detection. This approach couples sensor arrays and pattern recognition software to create a characteristic profile of sensor responses for a given odor, and has been used to a great extent in food [1] and medical/pharmaceutical applications [2]. However, as the sensors have limited response selectivity, different VOCs can trigger very similar responses. Additionally, the individual VOCs comprising a novel pattern are not identified in the electronic nose detection process, precluding the detection and identification of novel biomarkers. Finally, when analytical instruments such as GC and GC-mass spectrometry (MS) systems cannot be located where headspace gas is generated (such as an environmental collection or collection from human/animal subjects), air samples can be stored in plastic bags for transportation to a laboratory, sometimes made applicable for direct injection analyses in highly sensitive analytical detectors (see recent example with detection using proton transfer reaction (PTR) mass spectrometry [3]).

A variety of factors affect which sampling technique should be chosen for a given experiment, such as selectivity, sample recovery rates, the potential for contaminants in the collection vessel (called “outgassing”), storage conditions (e.g., time and temperature), and intervals between collection and analysis in a lab (ranging from direct analysis to storage for extended periods of time). Recovery rates between different VOC devices were shown to vary: yields of >95% recovery have been shown TD tubes in select conditions where compounds were well- suited to the given sorbent material [4,5], while recoveries for SPMEs [6] and storage bags were lower [5]. Compared to SPMEs and TD tubes, bags also have a higher potential for cross-contamination during storage and the polymer material may introduce outgassing contaminants [7,8]. A higher sample capacity is achievable for TD tubes (µg to low mg) [9] compared to

SPME fibers (ng to μg levels) due to a larger surface area of adsorbent material [9,10]. Conversely, TD tubes have a disadvantage of requiring the generation of a pressure gradient for active sampling (usually requiring active pumping) while SPME fibers involve passive sampling with minimal disruption of sample composition. This is important to consider for broader implications of VOC sampling. While the research for this dissertation was conducted by sampling from controlled environment, future applications could be conducted in the field, thus requiring evaluation of the portability and power requirements of external devices to the samplers. Ultimately, SPME samplers were chosen for the purposes of this work and will be further discussed.

2.2.2 Sample collection and concentration via solid-phase microextraction

The group of Janusz Pawliszyn developed SPME as a method to rapidly, cheaply, and easily achieve preconcentration of target analytes [11,12], and this approach has been applied to applications in environmental work [13], foods/fragrances [14], medicine [15], and forensics [16]. Thin polymer films are coated onto a support, typically a fused silica or stainless steel fiber that is a few hundreds of micrometer thick. Volatile or semi-volatile analytes can either absorb into the interior of the fiber coating, depending on film thickness, or adsorb on the exterior of the polymer film. The fiber can be exposed to analytes through immersion in a liquid sample or through exposure to the headspace of a solid, liquid, or gaseous sample.

SPME methods integrate a single step for sampling, extraction, and concentration prior to sample introduction into the heated injection port of a GC system. Conditions for VOC collection for all projects in this work involve a SPME fiber suspended in the headspace above actively outgassing samples, similar to the illustration in Figure 2.1. In order for analytes to be

detected via SPME, two processes must occur. First, volatile analytes must partition between the liquid media and the gaseous headspace above the sample. Second, analytes must partition between the gaseous phase above the sample and the SPME fiber coating. Once equilibrium has been established between the headspace and fiber coating, the amount of analyte deposited on the fiber is described by Equation 2.1, which depends on the volume of fiber coatings, volume of

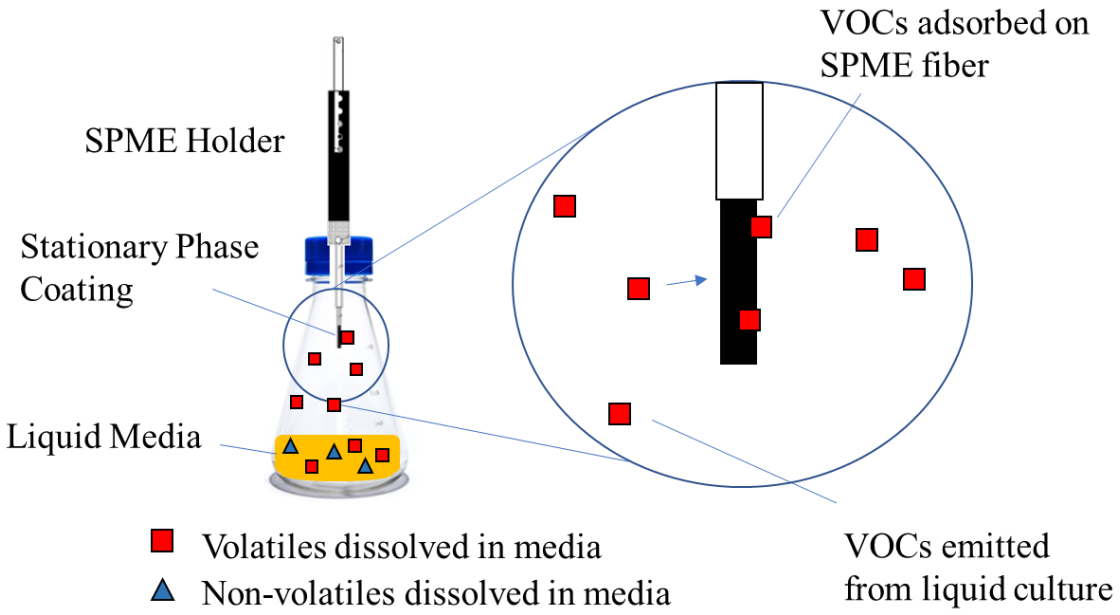


Figure 2.1 Schematic representing the partitioning of volatiles onto a SPME fiber (represented by black rectangle) suspended above a liquid media, translatable to the experiments to be covered in Chapters 3, 4, and 5. The volatility of compounds is a spectrum, and some compounds with low volatility may still be detected depending on experimental conditions (e.g., sampling time, temperature).

$$n = \frac{(K_{fs}V_fV_sC_o)}{(K_{fs}V_f+K_{hs}V_h+V_s)} \qquad (1)$$

Equation 2.1: n is the amount of analyte extracted into the SPME fiber, K_{fs} is the distribution constant of analyte partitioning between the fiber coating and sample (independent of the location of the fiber in the system), V_f is the fiber coating volume, V_s is the liquid sample volume, C_o is the initial analyte concentration in the sample, K_{hs} is the distribution constant of analyte partitioning between headspace and fiber coating, and V_h is the volume of the gas in headspace.

sample matrix (i.e. gaseous headspace, liquid media), the initial concentration of analyte in the sample, and equilibrium constants for each analyte between the sample/fiber or the headspace/fiber. The amount of analyte adsorbed on the fiber is proportional to its concentration in the headspace, which in turn is influenced by its concentration in the liquid.

Several different stationary phase coatings are commercially available, including polydimethylsiloxane (PDMS), divinylbenzene (DVB), carboxen (CAR), carbowax-polyethylene glycol (PEG), polyacrylate (PA), and C18 silica particles. Table 2.1 gives recommendations of which fiber types are advantageous to detect different analytes. Additionally, a mixture of sorbent materials can be coated on to the same fiber, thus broadening the capture performance to more substances.

This work utilized "portable field samplers" (manufactured by Sigma-Aldrich/Supelco), where the SPME coating is located at the end of an adjustable length needle and can be retracted

Table 2.1 Recommended SPME polymer coatings for different analyte types, recommended by the commercial supplier Supelco/Sigma-Aldrich (See SPME Applications Guide 925F)

Fiber Types	**Analytes**
100 µm PDMS	Volatiles (MW < 275)
30 µm PDMS	Non-polar semi-volatiles (MW 80-500)
7 µm PDMS	Non-polar high molecular weight compounds (MW 125-600)
85 µm / 75 µm PDMS/CAR	Gases and low molecular weight compounds (MW > 275)
65 µm PDMS/DVB	Volatiles, amines and nitro-aromatic compounds (MW 50-300)
85 µm PA	Polar semi-volatiles (MW 80-300)
60 µm PEG	Alcohols and polar compounds (MW 40-275)
50 µm / 30 µm DVB/CAR on PDMS	Trace compound analysis, C3-C20 (MW 40-275)

into an external casing. Only three polymer coatings were commercially available for the portable field samplers, limiting the extent to which the analysis could be tailored. The fiber coating utilized during all three projects for this dissertation was the biphasic, 65 μm PDMS/DVB coating, typically used for the broad analysis of VOCs and amines (nitro-aromatic compounds are rare in nature). The portable samplers are suitable for use in a field-setting, are lightweight and easily transported back to a laboratory, and provide minimal sample loss for up to 1 month when properly stored (4°C-20°C).

2.2.3 Separation and detection of VOCs

GC-MS is an easily adaptable and commonly used methodology for the identification and measurement of VOCs [17]. Figure 2.2 shows the most basic components of a GC-MS system. Volatiles introduced in the GC heated inlet are carried via a mobile gaseous phase (e.g., helium gas) through interactions with a stationary phase based upon their volatility and affinity for each phase. Compounds eluting from the end of a chromatography column exit the column at different times after injection (the "retention time") and are detected in the mass spectrometer. Detection of analytes at each retention time, where 1) ionization of compounds occurs via bombarding compounds with a 70 eV beam of electrons, 2) energy imparted during ionization causes fragmentation, with masses of fragment ions often characteristic of functional groups, 3) molecular and fragment ions are separated by their mass-to-charge (*m/z*) ratios in the mass analyzer. Extensive libraries of electron ionization (EI) mass spectra are available to facilitate identification of analytes separated and detected in complex samples by GC-MS, a common example being the National Institute for Standards and Technology (NIST) database that is updated every few years.

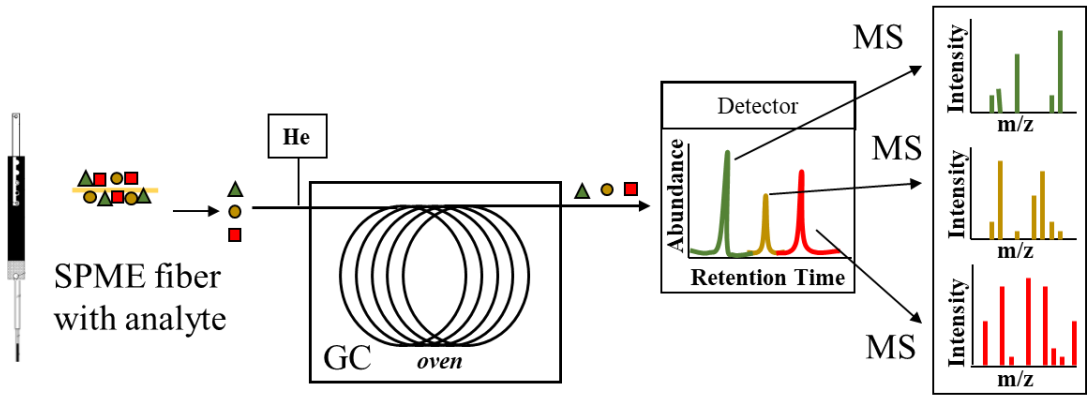


Figure 2.2 Schematic of a GC-MS system, where a sample collected on a SPME fiber is injected onto a capillary column. Flow of helium carrier gas drives transport through the column, where analytes separate depending on interactions between the stationary phase and mobile phase (accelerated by increasing oven temperatures). Compounds exiting the column are ionized, fragmented, and separated in the mass analyzer (quadrupole, nominal *m/z* reported). Detected ions are summed by repetitive scanning through the mass spectrum, creating a total ion current chromatogram. Abundances of ions at each *m/z* value are stored, and mass spectra can be constructed from ion abundances for each *m/z* value.

2.3 APPROACHES FOR DATA PROCESSING OF VOLATILE METABOLITES

2.3.1 Untargeted metabolomics vs. targeted metabolite profiling

Metabolomics is the qualitative and/or quantitative study of small molecules, called metabolites, produced by a biological system (organism, tissue, or cells). Changes in the metabolites result from altered biochemical activities and vary with the growth state of a biological system, thus representing the metabolic phenotype. Strategies for defining metabolic phenotypes are divided into two approaches, untargeted (or discovery) metabolomics and targeted metabolite profiling. Untargeted metabolomics aims to generate a broad, comprehensive analysis of both identified and unidentified analytes in samples without prior knowledge of metabolite identities. This approach attempts to detect as many compounds as possible (within the confines of the chosen sampling technique and detection system), presenting an avenue for novel biomarker discovery. Metabolomics commonly employs data reduction techniques to narrow an extensive list to a more manageable, smaller set of signals, where the smaller set

reflects metabolites relevant to changes in the biological system. These signals then are subjected to metabolite annotation using pre-existing *in silico* libraries or by experimental analysis of analytical standards on the same instrument.

Targeted metabolite profiling involves measurements of pre-determined groups of chemically annotated metabolites. When target compounds are known, the use of internal standards, whether similar in structure or isotopically labeled analogues, can be used to rigorously quantify changes of target compound levels across different sample types and has a higher sensitivity compared to untargeted approaches. Additionally, determining relationships in biochemical relationships between targeted metabolites, especially when changing between different physiological states, is enhanced due to having prior knowledge of their metabolic pathways. In the hope of guiding pathway discoveries, the initial stages of development of research projects for biosecurity and bioenergy utilized an untargeted metabolomics approach for identification of biomarkers within the datasets generated using SPME-GC-MS.

2.3.2 Development of data processing criteria to identify biomarker signatures

A multi-step data processing workflow facilitated conversion of GC-MS datasets into a list of volatile biomarkers for different matrices in this research. First, the information from "raw" data files (Agilent ChemStation files) was subjected to chromatographic deconvolution. Chromatographic deconvolution is the process of separating compounds in a chromatogram based on differences in elution time of various detected ions and can be used to separate compounds that elute at similar elution times leading to chromatographic peak overlap in the total chromatogram. All mass spectral ions that share the same chromatographic peak shape and reach apex at the same retention time were grouped to produce a "clean" mass spectrum for each

compound. Deconvoluted compounds were reported by Agilent MassHunter Qualitative software for each sample. Retention time alignment of compounds across samples was accomplished using Mass Profiler Professional for each experiment, which corrects for minor retention time drift. Project-specific parameters utilized for peak detection are further described in each chapter.

Detected VOCs were then annotated based upon comparisons to the NIST 2014 (NIST14) mass spectral database. Compounds with mass spectral matches ≥ 70% were subsequently identified by the name of the match with the highest score. The confidence in NIST14 annotation was furthered by calculation of chromatographic Retention Index (RI) for each analyte on a non-isothermal, temperature-programmed gradient, where the RI of each analyte is calculated based upon its retention compared to a series of n-alkane standards analyzed under the same temperature program. Calculated experimental RIs differing more than 5% total deviation from the theoretical RI value in the reference library resulted in rejection of a NIST ID. Compounds that did not exceed the mass spectral match or retention index threshold were annotated using the base peak *m/z* and retention index (e.g., *m/z* 177 at RI 1495).

Identification of VOCs as candidate compounds for potential microbial biomarkers was achieved through development of two filtering criteria, with the purpose to eliminate background VOCs contributed from the experimental setup, growth media, and other compounds that exhibited inadequate consistency of detection. The first criterion required detection of a potential biomarker in at least two replicate samples at a given timepoint. This criterion was developed to assess levels of reproducibility while also allowing for biological variability that leads to occasional non-detects that is often encountered in experiments involving live biological systems. Some of the detected compounds were present at low concentrations and some biological variability could have easily pushed a compound below the detection threshold in one

of the replicates. This criterion avoids the analysis missing some potentially interesting markers by applying a too stringent criterion. For different experimental conditions, the number of sample replicates occasionally varied depending on the experimental setup; i.e., an experiment with n=3 replicates would require detection in 66% of samples (2 out of 3 replicates) while an experiment with n=2 replicates required detection in 100% of samples (2 out of 2). Details on the sample types and sample replicates is included more within Chapters 3-5. The second filtering criterion concerned the presence or absence of a marker in a biological culture relative to the media blank controls appropriate for each organism. A compound was removed from consideration as a potential candidate if its relative abundance in the biological replicates was less than ten times the average relative abundance in the control, where the threshold was set high due to observed changes in chromatographic peak areas across experimental timepoints regularly changing 1-2 orders of magnitude.

2.4 Example of VOCs contributed from commercial sampling flasks.

The complexity of untargeted data analysis can be complicated when both pertinent and non-pertinent VOCs are present. One contributor of non-pertinent VOCs seen in this work that highlights that complexity includes outgassing from the experimental setup. Of note, Erlenmeyer flasks comprised of either glass or polymers were utilized in different projects. Polymer-based Erlenmeyer flasks are advantageous for growing pathogens due to their lighter weight and resistance to breakage. Their use minimizes potential safety hazards including exposure of researchers to pathogenic cultures. However, polymer-based materials contain higher levels of outgassing compounds that can be detected by VOC samplers, thus increasing endogenous background levels of VOCs.

Different polymer vessels were investigated for the purpose of growing pathogen cultures, comprised of polystyrene (PS), polyethylene terephthalate glycol-modified (PEG), or polycarbonate (PC). Figure 2.3 documents the VOC complexity observed for each flask when incubated at 37°C with 75-100 mL of liquid media present. Total ion currents were least for PS, and greatest for PEG, with PC in between. In particular, PEG released three VOCs of significantly higher abundance than anything observed in PS and PC. These intense signals matched the mass spectra of decamethyl-cyclopentasiloxane, dodecamethyl-cyclohexasiloxane, and tetradecamethyl-cycloheptasiloxane at 95%, 98% and 89% accuracy, respectively and all experimental RI values fell within 5% of theoretical values. The siloxane compounds are thought to originate during manufacturing, as siloxanes are used in proprietary additive mixtures to keep polymers malleable for manufacturing purposes as well as machinery lubricants [18]. As these vessels were adapted to VOC sampling and not the intended purpose of these vessels, future studies should evaluate release of VOCs from flasks at the beginning of experiments. The PC flasks were utilized for the work in Chapter 3 due to product availability.

2.5 SUMMARY

Rationales for choosing SPME-GC-MS as the combined VOC collection, data acquisition, and data processing methodology were outlined in this chapter. Minor modifications in the procedures, if any, will be addressed in each methods section of subsequent chapters. The developed workflow can be broadly employed to general VOC metabolomics datasets to enable detection and annotation of putative biomarkers and is not limited to the applications described in this work.

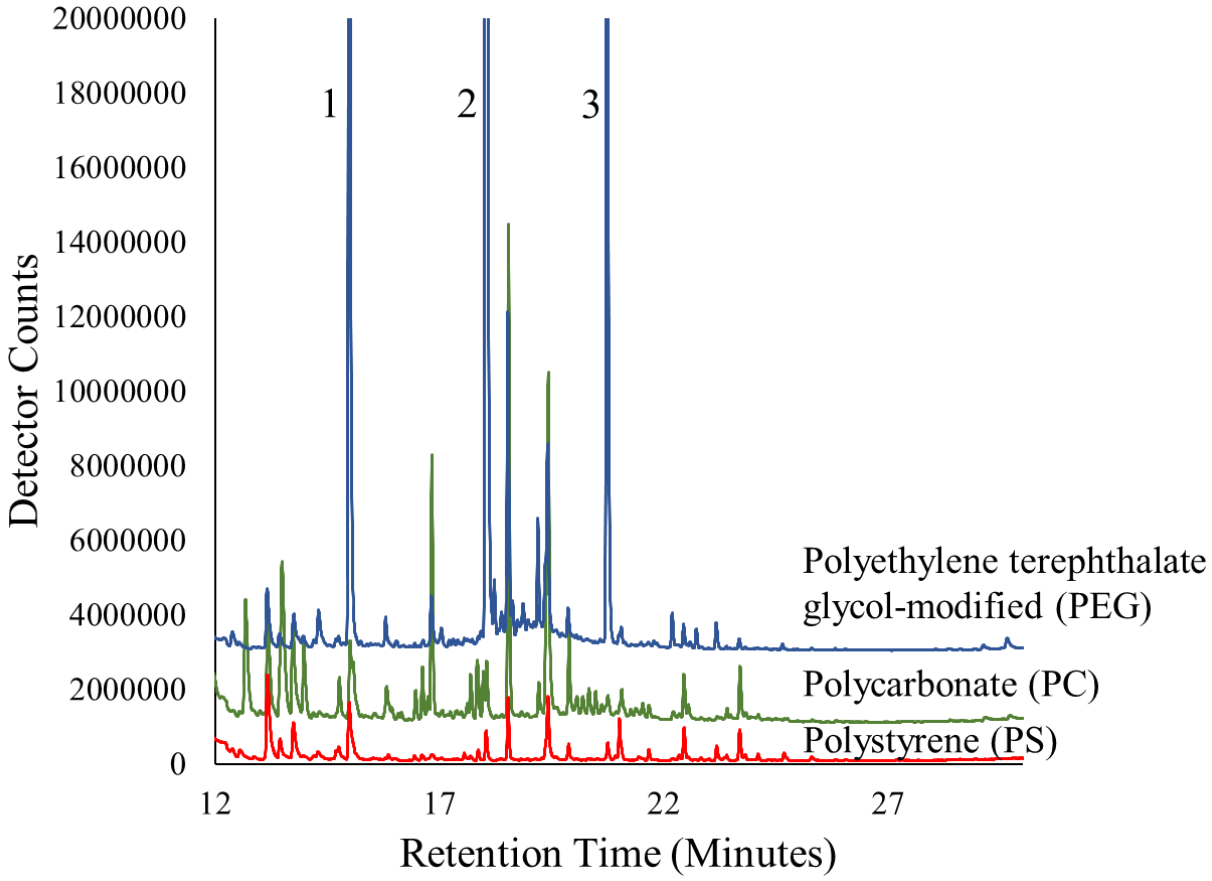


Figure 2.3 Examples of the endogenous VOCs exhibited by polymer-comprised sample flasks Polystyrene (PS), Polycarbonate (PC), and Polyethylene terephthalate glycol-modified (PEG) through comparison of the VOC total ion current chromatograms. Special annotation of peaks in PEG are given for 1) decamethyl-cyclopentasiloxane, 2) dodecamethyl-cyclohexasiloxane, and 3) tetradecamethyl-cycloheptasiloxane, high-outgassing VOCs contaminating the background VOC profiles when no cultures are present.

REFERENCES

REFERENCES

1 Sanaeifar, A., ZakiDizaji, H., Jafari, A. & Guardia, M. d. l. Early detection of contamination and defect in foodstuffs by electronic nose: A review. *TrAC Trends in Analytical Chemistry* **97**, 257-271, doi:10.1016/j.trac.2017.09.014 (2017).

2 Wasilewski, T., Migoń, D., Gębicki, J. & Kamysz, W. Critical review of electronic nose and tongue instruments prospects in pharmaceutical analysis. *Analytica Chimica Acta* **1077**, 14-29, doi:10.1016/j.aca.2019.05.024 (2019).

3 Papurello, D. Direct injection mass spectrometry technique for the odorant losses at ppb(v) level from nalophan™ sampling bags. *International Journal of Mass Spectrometry* **436**, 137-146, doi:10.1016/j.ijms.2018.12.008 (2019).

4 Kim, Y.-H. & Kim, K.-H. Extent of sample loss on the sampling device and the resulting experimental biases when collecting volatile fatty acids (VFAs) in air using sorbent tubes. *Analytical Chemistry* **85**, 7818-7825, doi:10.1021/ac401385m (2013).

5 Koziel, J. A. *et al.* Evaluation of sample recovery of malodorous livestock gases from air sampling bags, solid-phase microextraction fibers, tenax TA sorbent tubes, and sampling canisters. *Journal of the Air & Waste Management Association* **55**, 1147-1157, doi:10.1080/10473289.2005.10464711 (2005).

6 Merkle, S., Kleeberg, K. K. & Fritsche, J. Recent developments and applications of solid phase microextraction (SPME) in food and environmental analysis—A Review. *Chromatography* **2**, 293-381 (2015).

7 Harreveld, A. P. v. Odor concentration decay and stability in gas sampling bags. *Journal of the Air & Waste Management Association* **53**, 51-60, doi:10.1080/10473289.2003.10466121 (2003).

8 Gilchrist, F. J. *et al.* An investigation of suitable bag materials for the collection and storage of breath samples containing hydrogen cyanide. *Journal of Breath Research* **6**, 036004, doi:10.1088/1752-7155/6/3/036004 (2012).

9 Moein, M. M., Said, R., Bassyouni, F. & Abdel-Rehim, M. Solid phase microextraction and related techniques for drugs in biological samples. *Journal of Analytical Methods in Chemistry* **2014**, 921350, doi:10.1155/2014/921350 (2014).

10 Pawliszyn, J. Theory of Solid-Phase Microextraction. *Journal of Chromatographic Science* **38**, 270-278, doi:10.1093/chromsci/38.7.270 (2000).

11 Belardi, R. P. & Pawliszyn, J. B. The application of chemically modified fused silica fibers in the extraction of organics from water matrix samples and their rapid transfer to capillary columns. *Water Quality Research Journal* **24**, 179-191, doi:10.2166/wqrj.1989.010 (1989).

12 Arthur, C. L. & Pawliszyn, J. Solid phase microextraction with thermal desorption using fused silica optical fibers. *Analytical Chemistry* **62**, 2145-2148, doi:10.1021/ac00218a019 (1990).

13 Ouyang, G. & Pawliszyn, J. SPME in environmental analysis. *Analytical and bioanalytical chemistry* **386**, 1059-1073, doi:10.1007/s00216-006-0460-z (2006).

14 Xu, C.-H. *et al.* Applications of solid-phase microextraction in food analysis. *TrAC Trends in Analytical Chemistry* **80**, 12-29, doi:10.1016/j.trac.2016.02.022 (2016).

15 Looby, N. T. *et al.* Solid phase microextraction coupled to mass spectrometry via a microfluidic open interface for rapid therapeutic drug monitoring. *Analyst* **144**, 3721-3728, doi:10.1039/C9AN00041K (2019).

16 Gherghel, S. *et al.* Development of a HS-SPME/GC–MS method for the analysis of volatile organic compounds from fabrics for forensic reconstruction applications. *Forensic Science International* **290**, 207-218, doi:10.1016/j.forsciint.2018.07.015 (2018).

17 Makoś, P., Przyjazny, A. & Boczkaj, G. Methods of assaying volatile oxygenated organic compounds in effluent samples by gas chromatography—A review. *Journal of Chromatography A* **1592**, 143-160, doi:10.1016/j.chroma.2019.01.045 (2019).

18 Landon, S. J. *et al.* Room temperature-cured siloxane sealant compositions of reduced gas permeability. USA patent US7674857B2 (2005).

CHAPTER 3: METABOLIC PROFILING OF VOCs EMITTED BY THE PATHOGENS *Francisella tularensis* AND *Bacillus anthracis* IN LIQUID CULTURE

FOREWORD

The material presented in this chapter has been adapted from work first published in 2020 in the journal *Scientific Reports* [1]. Contributions from others to conduct the experiments described in this chapter are as follows: A. Rasley and J. R. Avila were integral in providing stock cultures of RG2 pathogens, experimental development, and culturing RG3 pathogens and conducting volatile organic compound sampling in Biosafety Level – 3 (BSL3) facilities at LLNL.

3.1 INTRODUCTION

The study and detection of volatile organic compounds (VOCs) originating from or interacting with organisms ranging from bacteria to humans have numerous applications in biology, environmental sciences, medicine, food industry, and national security. The measurement of VOCs and non-volatiles in exhaled breath is becoming an important rapid and non-invasive diagnostic tool to assess human physiology and health as well as a diagnostic tool for infections and systemic disease [2-5]. VOC markers in exhaled breath are also being explored to assess human chemical pharmacokinetics and environmental exposures to drugs, toxic materials, chemical or biological agents and other illicit materials. In this context, *in vitro* systems are often used to explore human exposome, microbiome and disease pathogenesis biomarkers [6,7]. One under-explored area of human volatile analysis is the analysis of exhaled breath to detect biomarkers indicative of an individual's exposure to biosecurity-relevant bacterial pathogens. However, human breath VOC profiles may be specific to an individual,

influenced by one's unique personal microbiome, external exposures, and immunological responses. VOCs derived from metabolic processes specific to bacterial taxa are hypothesized to occur in the breath of infected individuals, requiring differentiation of the chemical signatures from background substances in exhaled breath, those formed by the action of biological agents, and those resulting from human-microbe interactions. This study has focused on identification of VOCs emitted from actively growing bacterial agents *in vitro* (independent of host) and represents an important, albeit only a first step towards potential VOC-based detection of an infection by select agents.

Bacteria emit VOCs as major metabolic products during their growth cycles, many of which have important functions as signaling molecules to either neighboring bacteria or higher organisms [8]. Bacterial VOC profiles comprise complex mixtures containing diverse structural and chemical complexity. Monitoring of volatile emissions of bacteria has been facilitated by use of SPME-GC-MS, and extensive literature on bacteria related to food safety and hospital-acquired infections demonstrates how volatile chemical signatures can differentiate bacteria from various genera, species, and subspecies. Prior work into sampling volatiles of active bacterial growth has largely focused upon pathogens related to clinical settings, with common examples originating from the genera *Pseudomonas* [9], *Staphylococcus* [9], *Klebsiella [10]*, and *Mycobacterium* [11]. For example, Rees, et al. [10] reported aliphatic 2-ketones as the most abundant VOCs produced by *Klebsiella pneumoniae*, a common Gram-negative human pathogen, with less-abundant compound classes including esters, benzene derivatives, heterocycles, and nitrogen-containing compounds. Chen et al. [12] focused on the time-dependent VOC emissions of several foodborne pathogens, identifying long chain methyl ketones (2-heptanone, 2-nonanone, 2-undecanone) and alcohols (1-octanol, 1-decanol, 1-

dodecanol) as markers of three Gram-negative species (*E. coli, S. flexneri,* and *S. enteritidis*), while 3-hydroxy-2-butanone was identified as a marker of two Gram-positive bacteria (*S. aureus* and *L. monocytogenes*).

At present, there is less knowledge about volatiles released from bacterial threat agents, which would be useful for detecting the presence of and distinguishing such threat agents. Horsmon and Crouse [13] used thermal desorption tubes coupled to gas chromatography-mass spectrometry (GC-MS) to describe VOC profiles emitted by cultures of *Yersinia pestis (Y. pestis),* the causative agent of plague, and several strains from the genus *Bacillus*. They showed that VOC profiles and relative abundances of individual compounds distinguished bacterial genera as well as species within the same genus. However, their study did not provide a comprehensive analysis of detected VOCs and only qualitative differences determined by inspection of the profiles were used to distinguish species or genera. Lonsdale et al. [14] used colorimetric sensor arrays (CSAs) to differentiate volatiles in the headspace of *Y. pestis* and *B. anthracis* cultures, and reported high specificity, accuracy, and sensitivity to very low bacterial concentrations. However, individual biomarkers leading to the colorimetric changes were not identified, and signal response could have been influenced by the growth media utilized.

Two bacteria of concern to biosecurity and subjects of the work presented here are the aerobic, facultative intracellular pathogen *Francisella tularensis (F. tularensis)* and the obligate, endospore-forming pathogen *Bacillus anthracis* (*B. anthracis*). Both are classified by the Center for Disease Control (CDC) as Tier 1 Select Agents on the CDC category A Bioterrorism Agents list [15]. *F. tularensis*, the causative agent of the disease tularemia, has been isolated from more than 200 separate organisms, and several subspecies are known human pathogens. The bacterium is highly infectious and easily aerosolized, requiring as few as ten bacteria to cause infections

[16]. *B. anthracis*, the causative agent of the disease anthrax, forms resilient spores that survive chemical treatments, heat, lack of nutrients, and radiation, and has previously been developed into a bioweapon [17]. Detection of volatile biomarkers specific to the presence and growth of *F. tularensis* or *B. anthracis* through headspace sampling would be an important step towards developing a non-invasive metabolomics tool for rapid diagnosis of their presence in the lungs of subjects exposed to a biological attack.

Prior studies aimed at identification and/or differentiation of metabolites from *F. tularensis* or *B. anthracis* have largely focused on measuring profiles of pre-selected molecular targets in whole cell extracts. In particular, fatty acids have been profiled using GC-MS following esterification. The Voorhees group distinguished strains of *F. tularensis, B. anthracis, Brucella* spp. *abortus, melitensis,* and *neotomae,* and *Yersinia pestis* through analysis of fatty acid methyl ester (FAME) profiles using pyrolysis mass spectrometry in combination with an in-situ thermal transesterification [18,19]. Fatty acids of carbon chains ranging from 12:0 to 24:1 were identified, and principal components analysis (PCA) of the fatty acid profiles discriminated bacterial species. Li et al. distinguished *Francisella tularensis* subspecies *novicida, Escherichia coli,* and *Bacillus subtilis* by derivatizing fatty acids to form trimethylsilyl esters [20]. However, these studies required whole bacteria and sample preparation that was destructive to the bacteria, precluding analysis of metabolite changes over time in an unperturbed culture.

Our work presented here focused on the *in vitro*, non-invasive, untargeted profiling of VOCs from cultures of *F. tularensis* subspecies *novicida* (*Ft novicida*) and *B. anthracis* Sterne (*Ba* Sterne), both risk group 2 (RG2) surrogates for more virulent species, and from *F. tularensis* subspecies *tularensis* SCHU S4 (*Ft* SCHUS4) and *B. anthracis* Ames (*Ba* Ames), two fully virulent, risk group 3 (RG3) organisms. (For descriptions of risk group (RG) classifications of

infectious microorganisms and recommended biosafety level (BSL) for their handling see the U.S. Department of Health and Human Services guide on Biosafety in Microbiological and Biomedical Laboratories[21] or the World Health Organization Laboratory Biosafety Manual [22].) This work involved solid phase microextraction (SPME) sampling during multiple phases of pathogen growth of bacteria grown in biosafety level 2 (BSL-2) and biosafety level 3 (BSL-3) laboratories, respectively, and analysis by GC-MS. This *in vitro* determination of VOC profiles lays the groundwork for non-invasive investigation of bacterial metabolism of such organisms and represent the first steps towards potential VOC-based detection of an infection by such agents.

3.2 METHODS

3.2.1 Strains and Growth Media

F. tularensis subspecies *novicida* (strain U112; RG2) and subspecies *tularensis* (strain SCHU S4; RG3) were obtained from the CDC and Brigham Young University, respectively. *B. anthracis* (strain Sterne; RG2) and Ames (strain Ames; RG3) were obtained from a collaborator at Dugway Proving Grounds. The agar plates and liquid growth media for bacterial growth were prepared separately for each species. Different media were chosen to achieve optimal bacterial growth. *F. tularensis* was grown using a modified Mueller-Hinton (MH) growth media [23] (Becton Dickinson (BD) Difco, Franklin Lakes, NJ) supplemented with 0.1% glucose, 0.025% ferric pyrophosphate (Sigma-Aldrich, St. Louis, MO), and 0.02% IsoVitaleX (BD Difco); *B. anthracis* was grown with Brain-Heart Infusion (BHI) growth media (Becton Dickinson (BD) Difco, Franklin Lakes, NJ). RG2 strains and RG3 strains were grown, prepared and sampled in a biosafety level 2 (BSL-2) laboratory and biosafety level 3 (BSL-3) laboratory, respectively.

3.2.2 Preparation of Bacterial Headspace

Bacterial colonies were selected after overnight incubation on agar plates and transferred to 10 mL of liquid modified MH media or BHI media, respectively. Bacteria were cultured in media under aerobic conditions with overnight incubation at 37 °C and 170 rpm shaking. For each species and experiment, three 100-µL aliquots were inoculated into three separate 20 mL portions of fresh liquid media (1:200 dilutions) and incubated in three 250-mL disposable polycarbonate Erlenmeyer flasks with vented caps at 37 °C and 170 rpm shaking. The VOC profiles from the headspaces of each of the triplicate bacterial cultures (replicates) and the number of viable bacteria were sampled and assessed at multiple timepoints. In addition to the three replicates of each pathogen species, an uninoculated liquid media flask was simultaneously prepared and VOCs sampled from it as a negative (media-only) control.

3.2.3 Sampling VOCs from Bacterial Headspace (RG2 strains)

The VOC profiles of bacterial headspaces and media controls were sampled at different time intervals depending on bacterial growth rates and experimental setups using a protocol developed here that some of the authors also applied for headspace analysis of algal cultures in other work.[24] *Ft novicida* cultures were sampled at the following timepoints: 0, 2, 4, 8, 12, 16, 20, 24, 28, 32, 48, and 52 hours post-inoculation. *Ba* Sterne cultures were sampled at the following timepoints: 0, 4, 8, 12, 20, and 24 hours post-inoculation. At the time of sample collection, Erlenmeyer flasks were removed from the incubator-shaker and transferred to a biosafety cabinet. Headspace VOCs were immediately collected for 30 minutes on a field-portable 2 cm solid-phase microextraction (SPME) fiber with a 65 µm polydimethylsiloxane/divinylbenzene (PDMS/DVB) coating (Supelco, Bellefonte, PA) with no

agitation of the flask. At each timepoint, one unexposed SPME fiber (fiber remaining retracted behind the septum in the SPME housing) was placed within the biosafety cabinet where the SPME sampling of cultures was taking place. These fibers served as "travel blanks" to account for potential background volatiles leaking onto retracted fibers over time during storage or transportation to the GC-MS analysis laboratory. These "travel blanks" were analyzed concurrently with fibers exposed to cultures. After collection, all SPME fibers were stored in refrigerators at 2-4 °C until analysis. Data acquisition on the GC-MS occurred within 3 weeks of collection.

3.2.4 Sampling VOCs from Bacterial Headspace (RG3 strains) and Transfer of SPME Samples to BSL-2 Facility

The VOC profiles of RG3 *Ft* SCHUS4 and *Ba* Ames as well as corresponding media controls were sampled at the following timepoints: 0, 6, and 24 hours post-inoculation. Timepoints were chosen to capture the exponential and stationary growth phase in each species. At the time of sample collection, Erlenmeyer flasks were removed from the incubator-shaker and transferred to a biosafety cabinet within the BSL-3 facility. Flasks were allowed to sit in the BSC for 30 minutes prior to sampling in order to allow any aerosols to settle. Headspace VOCs were collected for 30 minutes on SPME fibers with no agitation of the flask. After collection, SPME fiber devices were decontaminated by bleach wiping the entirety of their external housing for 1 min apiece, and residual bleach was removed via wiping. The process of bleach wiping to prevent accidental transfer of pathogens out of the BSL-3 facilities was tested and validated. The overall protocol was approved by the Institutional Biosafety Committee (IBC) at LLNL (see Protocol in Appendix). SPME fibers were transferred from the BSL-3 to BSL-2 facilities and

stored in refrigerators at 2-4 °C until analysis, as previously described. In analyzing the samples collected in the BSL-3 we did not find any indication that bleach wiping may have altered the compounds detected e.g., by introducing chlorinated compounds.

3.2.5 Determination of Bacterial Concentrations

The growth phase (logarithmic, stationary, decline) of each organism was estimated by monitoring the concentration of viable bacteria over the course of the experiment for all biological replicates. Aliquots (1 mL) of all bacterial cultures were collected immediately following VOC sampling at each of the timepoints, and the Erlenmeyer flasks were subsequently placed back into the incubator-shaker. The aliquots were serially diluted between 10^{-2} to 10^{-7} depending on expected growth phase. A preliminary experiment was performed by plating in duplicate 10-fold dilutions to determine the appropriate serial dilution for each growth phase. The dilution factor was selected to achieve a target concentration of 30-300 cells per plate for counting. Dilutions were plated in duplicate (100-µL aliquots) on agar plates to determine the number of colony-forming units (CFU). Bacterial concentrations are reported as CFU counts per mL of liquid culture.

3.2.6 Data Acquisition Parameters

The data acquisition followed a procedure similar to the one previously described for algal VOCs and is briefly summarized here.[24] VOC analyses were performed on an Agilent 5975T GC-MSD (Agilent Technologies, Santa Clara, CA) using an Agilent HP-5ms column (30 m x 250 µm x 0.25 µm) coupled to a single quadrupole mass analyzer with helium carrier gas at

a constant flow rate of 1.2 mL/min. Volatiles absorbed by the SPME fiber were desorbed in the heated (250˚C) GC inlet for 60-seconds using splitless injection. The column temperature was programmed to start at 40 °C for 6 min, then heated at 8 °C/min from 40 to 280 °C and held for 4 min (total run time = 40 min). Ions were generated using electron ionization (EI) (70 eV) and acquired at 4 scans/s over *m/z* 35-450. Data acquisition was performed using ChemStation (version E.02.02). A commercial GC-MS reference standard (S-22329; AccuStandard, New Haven, CT) was used to evaluate day-to-day performance of the GC-MS system and to calculate retention indices.

3.2.7 Data Processing

After data acquisition, data processing procedures and criteria were applied to detect and identify taxa-specific biomarkers similar to the work previously described for algal VOCs [24]. All ChemStation data files (consisting of data from biological replicates, media controls, and travel fibers) were translated using MassHunter GC/MS Translator B.07.05 for compatibility with Agilent's Mass Hunter Qualitative software (version B.07.00 SP2) and Mass Profiler Professional (MPP) 12.6.1 software. These programs enabled sophisticated organization of individual MS files into complex datasets for chemometric analyses.

Chromatographic deconvolution and visualization were performed using MassHunter Qualitative using a Retention Time window size factor of 90.0, signal-to-noise ratio threshold of 2.00, and absolute ion height filter of 1000 counts. An arbitrary small value of 1 was assigned across all samples to the signal value for compounds that were not detected. Detected peaks were transferred into MPP and inter-aligned using a retention time tolerance of 0.15 minutes, mass spectral match factor of 0.6 (of maximum 1.0), and a delta *m/z* tolerance of 0.2 Da. Annotation of

the aligned compounds was performed by searching spectra against the NIST14 mass spectral database. Compounds with mass spectral matches ≥ 70% were subsequently identified by the name of the match with the highest score. Identifications with literature retention indices deviating more than 5% from the experimental retention indices were rejected. Compounds that did not exceed the mass spectral match or retention index threshold were annotated using the base peak *m/z* and retention index (e.g., "Unknown *m/z* 121_RI 1002").

The reported abundance values in this work are relative abundances of compounds, obtained by integrating the signal in their chromatographic peaks. Relative abundances are compared between different measurements (timepoints, species). Absolute quantification of VOCs in the headspace above bacterial cultures is challenging with our method. For example, our culture vessels were not fully enclosed due to the use of vented caps designed to facilitate gas exchange and avoid pressure buildups, and some loss of VOCs may have occurred. The retention of analytes is also affected by sorbent material, sampling time, and potential saturation, whereas the desorption is affected by extraction time and temperature. Some relative quantitation could be achieved using internal standards, whether pre-loaded or spiked into cultures, but also has a number of practical issues. Therefore, for the purposes of our work, absolute quantification was not attempted.

Two filtering criteria were used to identify relatively robust and reproducible VOCs as the most likely candidate compounds for potential taxa-specific biomarkers. The first criterion required detection of a potential biomarker in at least two of three culture replicates at a given sampling timepoint. This "2 out of 3 replicates" filter criterion was chosen as a compromise to require some level of reproducibility while also allowing for some biological variability that is often encountered in experiments involving live biological systems. Some of the detected

compounds were present at fairly low concentrations and some biological variability could have easily pushed a compound below the detection threshold in one of the replicates. The "2 out of 3" criterion was chosen in order to avoid missing some potentially interesting markers by applying too stringent a criterion. The second criterion concerned the presence or absence of a marker in a biological culture relative to the media blank controls appropriate for each organism. A compound was removed from consideration as a potential candidate if its relative abundance in the biological replicates was less than ten times the relative abundance in the control.

The VOCs identified as putative taxa-specific biomarkers were compared with regard to both individual markers and groups of markers encompassing a compound class. First, the presence or absence of these markers in each growth phase (logarithmic, stationary, and decline) was determined. Second, the calculated peak areas of markers, also referred to as relative abundances here, were compared amongst biological replicates to assess consistency of detection. Finally, principal component analysis (PCA) was used as a dimension-reduction strategy to visualize covariance in the dataset. Only markers remaining after the filtering criteria were applied were utilized. Using the MPP software, prior to PCA analysis, markers were individually mean-centered and variance-scaled. PCA was performed on the transformed dataset, and the results are presented as a scores plot of the first two principal components (PCs) and a loadings plot to elucidate the contribution of each marker to PC positioning. PCA was not performed on the RG3 taxa due to the limited number of acquired samples.

3.3 RESULTS

The complexity of chromatographic peaks detected through GC-MS analysis of each sample is illustrated by an observed profile of Ft *novicida* sampled 24 hours post-inoculation,

representing the early stationary phase (Figure 3.1a). However, many peaks originated from the Mueller-Hinton media and the SPME sampling device (Figure 3.1b) and were considered background. Peaks representative of the bacterial signature were of lower relative abundance, highlighted on a smaller chromatographic scale in Figure 3.1c. This complexity was similar for *Ft* SCHUS4 and both *B. anthracis* taxa (not pictured).

Detection of thousands of volatile compounds in the various cultures and timepoints for each of the taxa (Table 3.1) required specified data-filtering criteria (see Methods) to remove background compounds and those compounds not reproducibly detected. For example, more than 2000 VOCs were detected across all *Ft novicida* samples. Eliminating VOCs that did

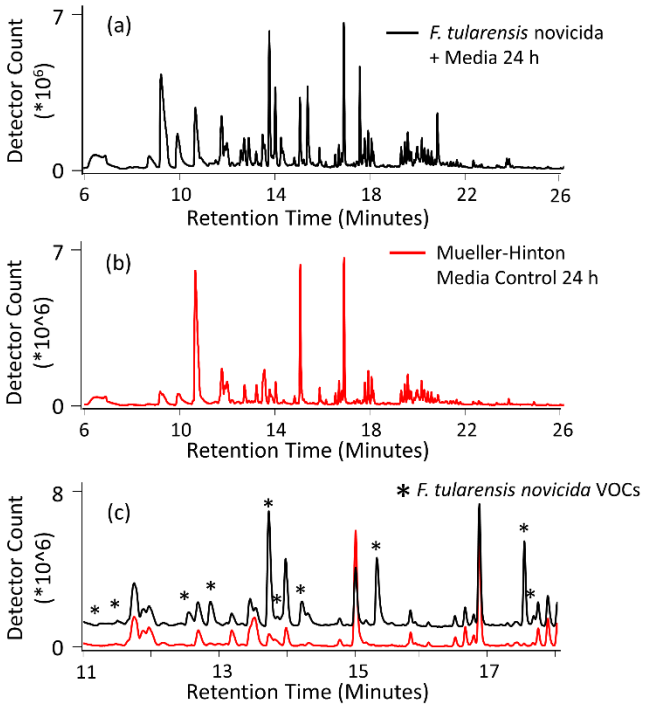


Figure 3.1 Examples of the chemical complexity exhibited by *F. tularensis novicida* cultures through comparison of the VOC total ion chromatograms at (a) 24 hours post inoculation, (b) corresponding Mueller-Hinton media control, and (c) enlarged overlay of both 1a and 1b, where stars indicate the bacteria-specific VOC emissions.

Table 3.1 Number of VOCs from *F. tularensis* and *B. anthracis* taxa detected by GC-MS and remaining after application of filtering criteria

Species	Total VOCs Detected	VOCs Fail Criterion 1	VOCs Pass Criterion 1	VOCs Fail Criterion 2	VOCs Pass Criterion 2 (putative biomarkers)
F. tularensis novicida	2360	2239	121	103	18
F. tularensis SCHU S4	999	754	245	207	38
B. anthracis Sterne	1031	912	119	89	30
B. anthracis Ames	1022	745	277	221	56

not appear in at least two of the triplicate measurements (Criterion 1) narrowed the dataset to 121 VOCs, a reduction of approximately 95 % (Figure 3.2). Further elimination of VOCs with relative abundances less than 10x the average relative abundance in the negative controls (Criterion 2) narrowed the dataset to 18 putative volatile biomarkers that were confidently attributed to *Ft novicida*. The same criteria were applied to the data from other bacterial species studied here, resulting in 38 putative VOC biomarkers for *Ft* SCHUS4, 30 biomarkers in *Ba* Sterne, and 56 biomarkers in *Ba* Ames (Table 3.1).

3.3.1 Results from RG2 species

Candidate bacterial VOC biomarkers from all timepoints were annotated through examination of both mass spectral library matching scores using the NIST14 database and experimental retention indices. Since all metabolite annotations in this report are based on comparisons to literature spectra and retention index values, they should be considered as satisfying confidence level 2 of the Metabolomics Standards Initiative recommendations for

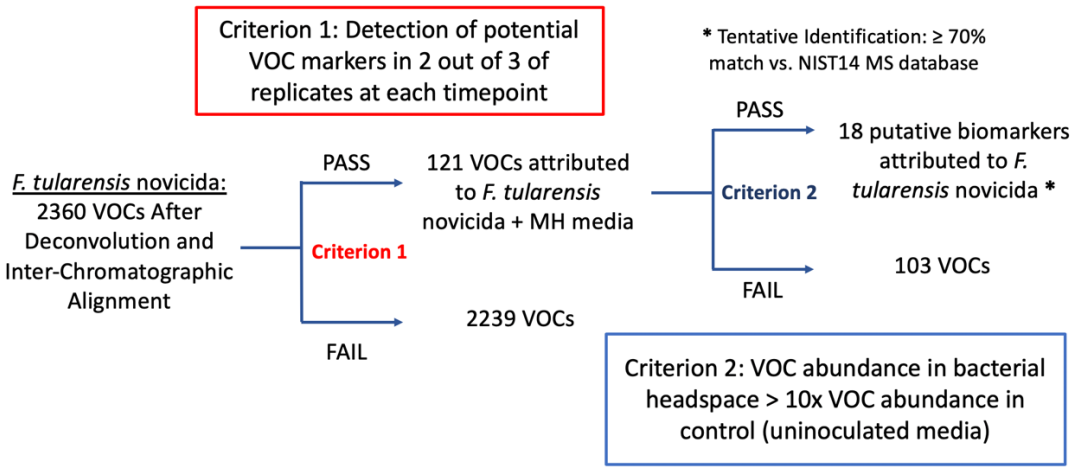


Figure 3.2 Example workflow of criteria utilized to filter list of detected VOCs to bacteria-specific biomarkers produced during growth, shown for *F. tularensis novicida*.

identification of compounds [25]. For *Ft novicida* (Table 3.2), 15 of the 18 biomarkers passed the set threshold of 70% match, while three were labeled as "unknowns" owing to poorer matches below that threshold. For *Ba* Sterne (Table 3.3), 18 of the 30 biomarkers passed the set threshold of 70 % match, while the remainder were labeled "unknowns". Inter-species diversity in emitted VOC biomarkers was observed. The *Ft novicida* profile contains odd-chain, aliphatic methyl ketones, alcohols, nitrogen-containing, and sulfur-containing volatiles. The *Ba* Sterne volatile profile is comprised of branched methyl ketones, followed by esters, carboxylic acids, alcohols, and sulfur-containing volatiles.

Evaluation of potential markers requires assessment of the growth phase at each sampled timepoint post-inoculation of the culture flask. The logarithmic, stationary, and decline phases were identified based upon CFU measurements taken alongside SPME-VOC sampling. The data

Table 3.2 Annotations of *F. tularensis* novicida-specific VOC markers through compound class, putative NIST ID, *m/z*, and retention index matching

Class	Compound	MS Base Peak (*m/z*)	RI (Lit)	RI (Exp)	NIST 14 Match Factor
Alcohol	1-Butanol, 2-methyl-	70	739	719	74
Alcohol	2-Nonanol	45	1101	1113	80
Alcohol	Phenylethyl Alcohol	91	1116	1127	70
Alcohol	1-Nonanol	56	1173	1183	83
Alcohol	2-Undecanol	45	1308	1306	75
Methyl ketone	2-Heptanone	43	891	890	71
Methyl ketone	2-Nonanone	58	1092	1104	72
Methyl ketone	2-Undecanone	58	1294	1298	73
Methyl ketone	2-Tridecanone	58	1497	1492	93
Methyl ketone	2-Pentadecanone	58	1698	1690	87
Methyl ketone	58.0@26.267332	58	1902	1892	83
Nitrogen-containing	Pyrazine, 2,5-dimethyl-	108	917	911	80
Nitrogen-containing	2-Methyl-3-isopropylpyrazine	121	1056	1064	83
Sulfur containing	Dimethyl trisulfide	126	970	974	96
Sulfur containing	1-Propanol, 3-(methylthio)-	106	981	993	70
Unknown	*m/z* 121 _ RI 1002	121		1002	
Unknown	*m/z* 108 _ RI 1049	108		1049	
Unknown	*m/z* 133 _ RI 1110	133		1110	

Table 3.3 Annotations of *B. anthracis* Sterne-specific VOC markers through compound class, putative NIST ID, *m/z*, and retention index matching

Class	Compound	MS Base Peak (*m/z*)	RI (Lit)	RI (Exp)	NIST 14 Match Factor
Alcohols	4-Heptanol	55	872	893	86
Carboxylic Acid	Propanoic acid, 2-methyl-	43	772	744	70
Carboxylic Acid	Butanoic acid, 2-methyl-	74	861	865	90
Carboxylic Acid	Butanoic acid, 3-methyl-	60	863	853	86
Ester	Propanoic acid, 2-methyl-, butyl ester	89	898	961	97

Table 3.3 (cont'd)

Ester	Butanoic acid, butyl ester	71	995	1006	97
Ester	Butyl 2-methylbutanoate	103	1043	1053	89
Ester	Butanoic acid, 3-methyl-, butyl ester	85	1047	1058	83
Methyl Ketone	Methyl Isobutyl Ketone	43	735	719	73
Methyl Ketone	2-Hexanone, 5-methyl-	43	862	848	80
Methyl Ketone	2-Heptanone	43	891	889	94
Methyl Ketone	2-Heptanone, 6-methyl-	43	956	962	97
Methyl Ketone	2-Heptanone, 5-methyl-	43	971	973	73
Methyl Ketone	5-Hepten-2-one, 6-methyl-	108	986	977	70
Methyl Ketone	2-Heptanone, 4,6-dimethyl-	58	1045	1067	76
Sulfur containing compound	Butanethioic acid, S-methyl ester	43	874	834	78
Sulfur containing compound	Thiopivalic acid	85	959	945	71
Unknown	*m/z* 80 _ RI 715	80		715	
Unknown	*m/z* 57 _ RI 769	57		769	
Unknown	*m/z* 43 _ RI 791	43		791	
Unknown	*m/z* 43 _ RI 873	43		873	
Unknown	*m/z* 45 _ RI 901	45		901	
Unknown	*m/z* 57 _ RI 912	57		912	
Unknown	*m/z* 43 _ RI 956	43		956	
(methyl ketone)[a]	*m/z* 58 _ RI 962	58		962	
Unknown	*m/z* 43 _ RI 997	43		997	
Unknown	*m/z* 90 _ RI 1005	90		1005	
(methyl ketone)[a]	*m/z* 58 _ RI 1104	58		1104	
Unknown	*m/z* 83 _ RI 1145	83		1145	
(methyl ketone)[a]	*m/z* 58 _ RI 1554	58		1554	

[a]: GC/MS fragmentation similar to observed methyl ketones
RI (Lit): Retention Index reported from NIST14
RI (Exp): Retention Index calculated from experiment

for both RG2 species, *Ft novicida* and *Ba* Sterne, are presented in Figure 3.3. Logarithmic or "Log" phase, characterized by exponential bacterial growth, was observed to last for 20 hours and 8 hours, respectively. The bacterial counts rose approximately 3 orders of magnitude for both species, peaking at 1-2*10^9 CFU/mL for *Ft novicida* and 5*10^8 CFU/mL for *Ba* Sterne. For *Ft novicida* (Figure 3.3 (a)) the observed growth during that phase appeared rather variable. *Ft* cultures are known to be difficult to grow. Sampling more replicates may improve statistical confidence in future experiments. Stationary phase, occurring when the bacteria exhibit no

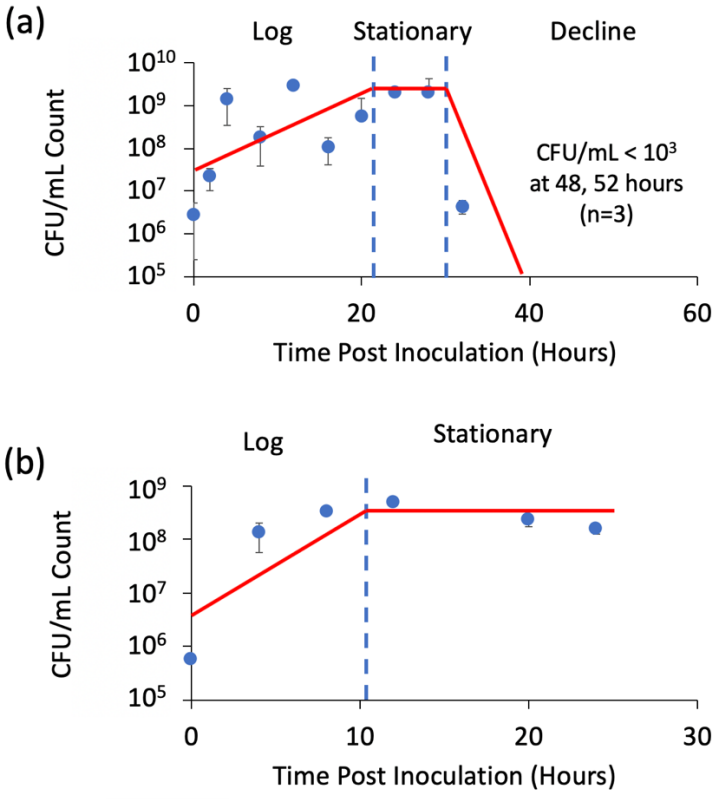


Figure 3.3 Growth curves of (a) *F. tularensis novicida* in modified Mueller-Hinton media over a 52-hour time period and (b) *B. anthracis* Sterne in Brain-Heart Infusion media over a 24-hour time period. Data points and error bars represent the means and standard deviations of CFU/mL determined from three culture replicates. Red lines represent visually determined trends in bacterial growth across each growth phase.

additional growth due to a depleted nutrient source, was observed in both species. *Ba* Sterne measurements were completed at 24 hours post-inoculation while still in stationary phase. For *Ft novicida*, further growth phase changes were observed through a decline in viable bacterial growth to ~$1*10^6$ CFU/mL at 32 h and no observable growth at the 48 h and 52 h post inoculation. The limit of detection for concentrations of viable bacteria was less than 1000 CFU/mL. Regardless, the CFU/mL counts were fairly consistent across triplicate measurements in both taxa, allowing assessments of growth phase.

The observed profiles of the VOC biomarkers varied with growth phase. The averaged relative abundances of biomarkers from *Ft novicida* and *Ba* Sterne are listed across all measured timepoints and grouped by compound class (Tables 3.4, 3.5; Appendix Tables A.3.1, A.3.2). The mean combined areas of each compound class are also stacked as a function of total area for each of the two RG2 species (Figure 3.4). Of the tentatively identified markers, chemical diversity was observed in the presence of ketones, aldehydes, alcohols, esters, carboxylic acids, nitrogen- or sulfur- containing markers, and alkanes. Although more biomarkers were detected for *Ba* Sterne (30) versus *Ft novicida* (18), the combined peak areas (total signal) of markers from *Ft novicida* at its peak growth (32 hours post-inoculation, stationary phase) were approximately 5x the total combined peak area of *Ba* Sterne VOCs at its peak growth (8 hours, logarithmic phase), attributed to the ~10x higher concentration of bacteria (compare Figure 3.4 and Figure 3.3). There was only a moderate correlation between combined marker peak areas and the bacteria concentration at any single timepoint. Peak areas and bacterial counts rose during the logarithmic phase for both species, but cumulative peak areas were stagnant or dropped during stationary phase despite bacterial concentration remaining steady.

Table 3.4 Average relative abundances (Log-10 Scale) of *F. tularensis novicida*-associated VOCs at all measured timepoints, separated by growth phase

	Time Post Inoculation (Hours)	**0**	**2**	**4**	**8**	**12**	**16**	**20**	**24**	**28**	**32**	**48**	**52**
	Growth Phase	**Log Phase**							**Stationary Phase**		**Decline Phase**		
Class	**Compound**	**Abundance (Log 10 Values)**											
Alcohols	1-Butanol, 2-methyl-	0.00	0.00	0.00	0.00	5.87 b	5.84 b	6.77	6.55 a	6.77	6.76	5.70	5.44 b
Alcohols	2-Nonanol	0.00	0.00	0.00	0.00	6.05	6.61	6.08	5.89	0.00	0.00	0.00	0.00
Alcohols	Phenylethyl Alcohol	0.00	0.00	5.66 a	6.07	6.29	6.36	6.54	6.75	6.74	6.81	6.67	6.67
Alcohols	1-Nonanol	0.00	0.00	5.80	6.35	6.70	7.26	7.46	6.98	6.43	6.24	5.90	5.85
Alcohols	2-Undecanol	0.00	0.00	0.00	0.00	6.14	6.35	6.27	6.13	5.96	0.00	0.00	0.00
Methyl Ketones	2-Heptanone	0.00	0.00	0.00	6.02	6.03 a	6.48	6.74	6.84	6.82	6.88	6.46	6.33
Methyl Ketones	2-Nonanone	0.00	0.00	0.00	6.44	6.60	6.84	7.11	7.31	7.41	7.48	6.96	6.85
Methyl Ketones	2-Undecanone	4.70 b,c	5.27 b,c	5.86	6.13	5.98	6.03	6.72	7.16	7.29	7.23	6.25	6.15
Methyl Ketones	2-Tridecanone	0.00	0.00	0.00	0.00	0.00	0.00	6.14	6.91	7.01	6.89	6.09	6.02
Methyl Ketones	2-Pentadecanone	0.00	0.00	0.00	0.00	0.00	0.00	4.96 b	6.23	6.50	6.51	6.03	6.04
Methyl Ketones	2-Heptadecanone	0.00	0.00	0.00	4.71 b	0.00	0.00	0.00	5.27	5.80	5.88	5.63	5.69
Nitrogen-Containing Compounds	Pyrazine, 2,5-dimethyl-	5.69 c	6.37 c	6.41 c	6.38 c	6.78 c	7.18 c	7.36	7.60	7.60	7.76	7.81	7.81

Table 3.4 (cont'd)

Nitrogen-Containing Compounds	2-Methyl-3-isopropylpyrazine	0.00	0.00	5.47 b	0.00	6.13	6.24	6.47	6.75	6.87	7.05	7.06	7.05
Sulfur-Containing Compounds	Dimethyl trisulfide	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6.90	6.84
Sulfur-Containing Compounds	1-Propanol, 3-(methylthio)-	0.00	0.00	0.00	0.00	5.72 a	5.76 a	5.98	6.05	6.00	6.07	6.08	6.13
Unknown	*m/z* 121 _ RI 1002	0.00	6.14	5.71 b	6.10	5.90	5.79	5.86	6.15	6.18	6.33	6.16	6.09
Unknown	*m/z* 108 _ RI 1049	0.00	5.80	5.79	6.37	6.51 a	6.50 a	6.54 a	6.62	6.60	6.69	6.56	6.58
Unknown	*m/z* 133 _ RI 1110	0.00	0.00	0.00	0.00	0.00	0.00	5.07 b	6.07	6.01	6.08	5.94	5.87

Notes:

[a]: VOC detected in 2/3 of triplicate measurements
[b]: VOC detected in 1/3 of triplicate measurements
[c]: VOC detected at levels less than 10x abundance in media blank

Table 3.5 Average relative abundances (Log-10 Scale) of *B. anthracis* Sterne-associated VOCs at all measured timepoints, separated by growth phase

	Time Post Inoculation (Hours)	**4**	**8**	**12**	**20**	**24**
	Growth Phase	**Log Phase**		**Stationary Phase**		
Class	**Compound**	**Abundance (Log 10 Values)**				
Alcohols	4-Heptanol	6.15	0.00	0.00	0.00	0.00
Carboxylic Acid	Propanoic acid, 2-methyl-	0.00	5.38 a	5.53 a	0.00	0.00
Carboxylic Acid	Butanoic acid, 2-methyl-	0.00	5.93	5.31 a	0.00	0.00
Carboxylic Acid	Butanoic acid, 3-methyl-	0.00	6.16	5.68	0.00	0.00
Ester	Propanoic acid, 2-methyl-, butyl ester	6.50	6.65	0.00	0.00	0.00
Ester	Butanoic acid, butyl ester	6.70	6.22	0.00	0.00	0.00
Ester	Butyl 2-methylbutanoate	5.82	6.11	0.00	0.00	0.00
Ester	Butanoic acid, 3-methyl-, butyl ester	5.44	6.23	4.81 a	0.00	0.00
Methyl Ketone	Methyl Isobutyl Ketone	0.00	0.00	5.71	6.22 a	6.14
Methyl Ketone	2-Hexanone, 5-methyl-	0.00	5.43 a	6.05	6.52	6.76
Methyl Ketone	2-Heptanone	5.12 a	6.58	6.65	6.69	6.89
Methyl Ketone	2-Heptanone, 6-methyl-	0.00	0.00	6.62	6.92	7.02
Methyl Ketone	2-Heptanone, 5-methyl-	0.00	0.00	5.54 a	6.03	6.25
Methyl Ketone	5-Hepten-2-one, 6-methyl-	0.00	0.00	0.00	4.34	5.49
Methyl Ketone	2-Heptanone, 4,6-dimethyl-	0.00	0.00	4.80 b	5.26 a	5.64
Sulfur containing compound	Butanethioic acid, S-methyl ester	0.00	5.35	5.57	5.89	0.00
Sulfur containing compound	Thiopivalic acid	4.60 b	5.46	5.47	5.86	4.34 b
Unknown	*m/z* 80 _ RI 715	5.08 b,c	5.80	5.68 a	5.82	5.20 b
Unknown	*m/z* 57 _ RI 769	0.00	0.00	4.54 b	5.19	0.00
Unknown	*m/z* 43 _ RI 791	6.93	7.23	6.83	6.15	5.64 a
Unknown	*m/z* 43 _ RI 873	0.00	0.00	0.00	4.34 b	4.86
Unknown	*m/z* 45 _ RI 901	0.00	0.00	5.16 a	5.30	0.00
Unknown	*m/z* 57 _ RI 912	5.64	5.82	5.06 b	0.00	0.00
Unknown	*m/z* 43 _ RI 956	5.18	0.00	0.00	0.00	0.00
(methyl ketone)[a]	*m/z* 58 _ RI 962	0.00	5.82	0.00	0.00	0.00
Unknown	*m/z* 43 _ RI 997	0.00	4.88 b	0.00	5.03 a	5.18 a
Unknown	*m/z* 90 _ RI 1005	0.00	0.00	5.22	5.00 a	0.00
(methyl ketone)[a]	*m/z* 58 _ RI 1104	5.62	5.79	5.67	5.13 b	5.79
Unknown	*m/z* 83 _ RI 1145	0.00	5.11	4.57 a	0.00	0.00
(methyl ketone)[a]	*m/z* 58 _ RI 1554	0.00	4.54 b	5.00	0.00	0.00

Table 3.5 (cont'd)

Notes:
[a]: VOC detected in 2/3 of triplicate measurements
[b]: VOC detected in 1/3 of triplicate measurements
[c]: VOC detected at levels less than 10x abundance in media blank
[d]: GC/MS fragmentation similar to observed methyl ketones

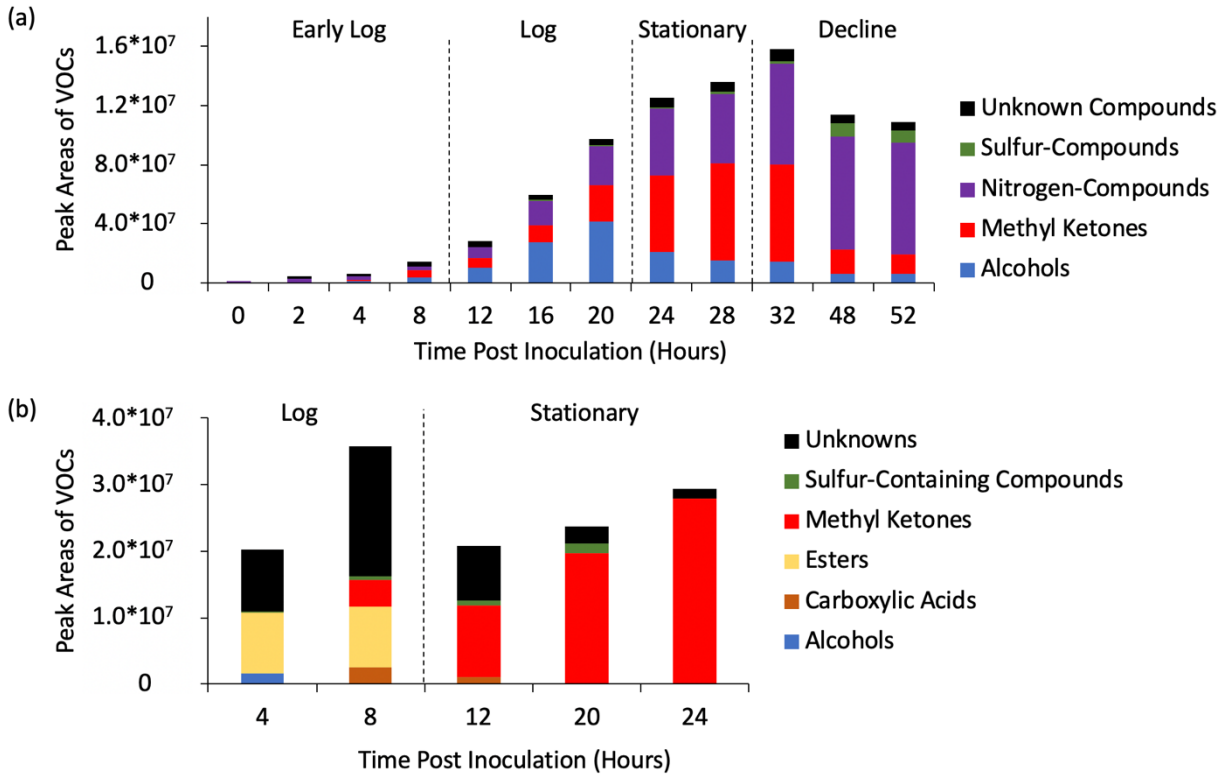


Figure 3.4 Mean combined peak areas (integrated detector counts) for compounds within individual classes at each timepoint post bacterial inoculation of cultures for (a) *F. tularensis novicida* and (b) *B. anthracis* Sterne.

The biomarker peak areas for *Ft novicida* steadily increased throughout the logarithmic and stationary phases before decreasing during the decline phase. Alcohols steadily rose in relative abundance throughout the log phase and were dominant in the early log and log phases. While some alcohols persisted throughout the entire study, several were fully depleted at the

longest timepoints measured (see 2-nonanol and 2-undecanol). Linear, odd-chain methyl ketones (or 2-ketones) were present throughout all growth phases, with ketones consisting of more than 13 carbons (longer than 2-tridecanone) being present only in the stationary phase and beyond. The contribution of methyl ketones peaked in stationary phase growth, and their decrease in the decline phase lowered total VOC relative abundances. Nitrogen-containing markers were present throughout the analysis of *F. tularensis* species due to the presence of 2,5-dimethylpyrazine, a marker that was a component of the growth media. However, the signal emitted from the bacterial cultures first exceeded 10x the signal in the media control at the 20-hour timepoint, prompting its inclusion as a potential *F. tularensis* marker. Combined with the signal from 2-methyl-3-isopropylpyrazine, nitrogen-containing markers comprised almost 70% of the chemical profile for the decline phase. Finally, *Ft novicida* noticeably displayed a large signal of dimethyltrisulfide as an abundant marker in the decline phase, comprising 7-8.5 % of the total VOC signal, formed as by-products of cysteine and methionine amino acid decomposition after the death of bacterial cells.

Biomarker areas for *Ba* Sterne also changed dependent on growth phase, though fewer timepoints were measured compared to *Ft novicida*. Esters, carboxylic acids, and alcohols comprised a significant portion of the logarithmic phase VOC marker profiles. Esters were based on butanoic or propanoic acids, with methyl groups at the 2 or 3-carbon positions. Two carboxylic acids were also based on butanoic and propanoic acids, both methylated at the 2-carbon position. Alcohols were only present in the early log phase. The non-detection of these markers during the stationary phases (with the exception of 2-methyl-propanoic acid in early stationary phase) suggests use as precursors for further synthesis. Relative abundances of methyl ketones significantly increased during stationary phase. Moreover, while *Ft novicida* was

dominated by straight-chain aliphatics, the methyl ketones in *Ba* Sterne contained methyl and aromatic substituents.

Levels of VOC biomarkers for *Ft novicida* and *Ba* Sterne were subjected to principal component analysis (PCA) to visualize VOC profiles observed at different growth phases. The scores plots in Figure 3.5(a) and 3.5(b) show clustering of all three culture replicates of the respective strains. The loadings plots, depicting the relative importance of individual markers towards sample positioning on PCs 1 and 2, are described in greater detail in Appendix Figures A.3.1a and A.3.1b. PCA groupings similarly exhibited distinct groupings of timepoints into

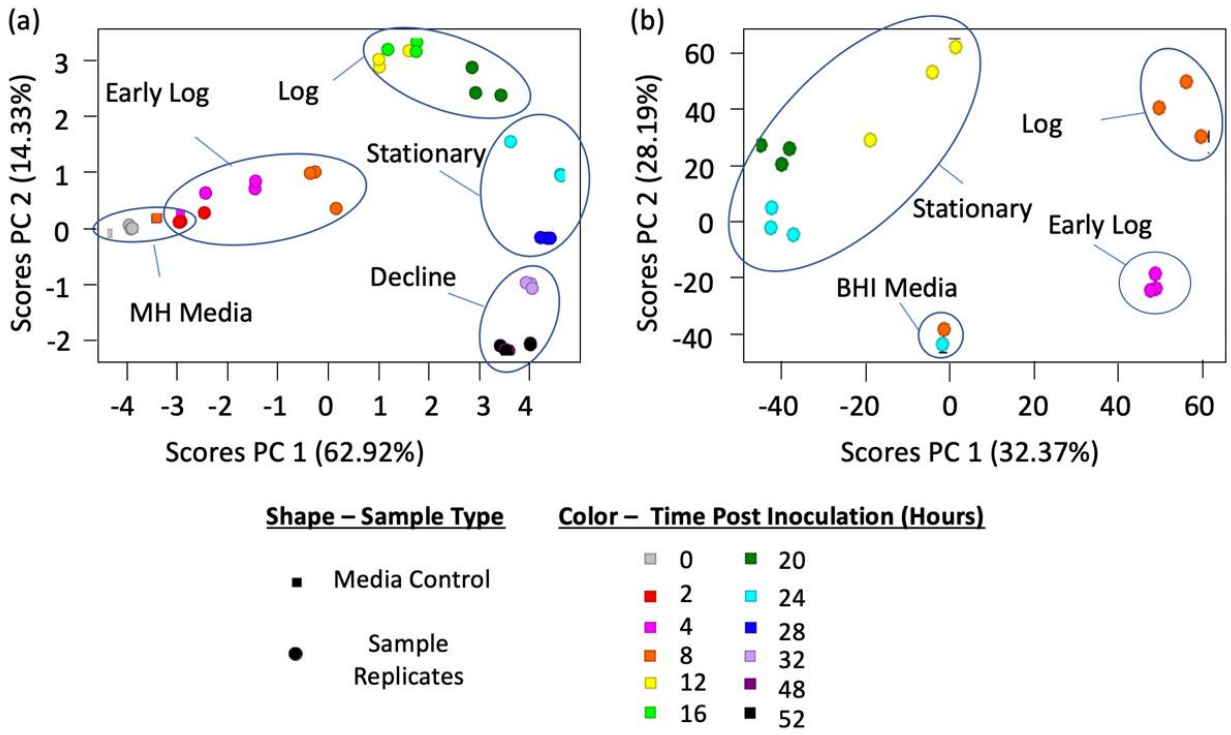


Figure 3.5 PCA scores plots for VOC marker profiles of (a) *F. tularensis novicida* and (b) *B. anthracis* Sterne generated using the peak areas of pathogen biomarkers across all timepoints. Each plot point represents one sample. Distinct chemical profiles were observed amongst labeled growth phases. Corresponding loadings plots to explain placement of samples are located in Appendix Figure A.3.1.

clusters as determined in Figure 3.4(a) and 3.4(b) for each species. The PCA scores plot provides additional verification of similarity of VOCs from culture replicates - profiles from each timepoint (same color) were positioned more closely to each other than to replicates of an adjacent timepoint, demonstrating fairly reproducible VOC marker profiles.

3.3.2 Results from RG3 species

Determination of bacterial concentrations and VOC sampling of the RG3 pathogens grown in our BSL-3 laboratory were performed at select time points, as shown in Appendix Table A.3.3 for both *Ft* SCHUS4 and *Ba* Ames. *Ba* Ames exhibited growth throughout 24 hours, with bacterial concentrations rising to $2.5*10^7$ CFU/mL at the 24 h time point. Meanwhile *Ft* SCHUS4 concentration remained stagnant around $6.7*10^5$ CFU/mL throughout 24 hours of culture, hypothesized to remain in a lag phase after inoculation.

The decontamination protocols developed for our work in the BSL-3 laboratory on both *Ft* SCHUS4 and *Ba* Ames included wiping the SPME fiber exterior casing using bleach (see Appendix Information). There was a potential for VOCs adsorbed on the internal fibers to be inadvertently oxidized. However, comparison of the VOC profiles from *B. anthracis* taxa obtained using the BSL-2 protocol without bleach wiping and the BSL-3 protocol that included the bleach wiping revealed a range of similar markers and/or compound classes, with no evidence of oxidized by-products for the profiles obtained using the BSL-3 protocol.

The VOC marker profile of *Ba* Ames displayed similarities to its RG2 counterpart *Ba* Sterne and is detailed in Table 3.7, where 18 of the 56 putative markers were identified. The VOC marker profile at the 6-hour timepoint of *Ba* Ames resembled the logarithmic VOC marker

profile of *Ba* Sterne. Esters were the most abundant identified markers, consisting of propanoic and butanoic acid esters. Five esters were shared between both *B. anthracis* taxa. Subsequent compound classes included methyl ketones and carboxylic acids. Conversely, the VOC marker profile at the 24-hour timepoint of *Ba* Ames more closely resembled the stationary VOC marker profile of *Ba* Sterne. Methyl ketones were the dominant markers, while all esters have been depleted. Four methyl ketones were shared between *B. anthracis* taxa. Principal component analysis (PCA) of the level of VOC biomarkers was also applied to *Ba* Ames to visualize VOC profiles at different growth phases. Similar to *Ba* Sterne, it appears that different chemical profiles can be associated with different growth phases of *Ba* Ames (not shown here), however, more data points would be needed to draw stronger conclusions.

Conversely, the profile of VOCs from *Ft SCHUS4* (Table 3.6) had fewer similarities with *Ft novicida*. The majority of putative markers for *Ft SCHUS4* were classified as unknowns, with only 5 markers passing the conservative identification criteria. While none of the observed compounds passing the filtering criteria were shared between either species, the 6 and 24-hour timepoints both contained alcohols such as 4-methyl-3-heptanol and 1-dodecanol. Alcohols were also the dominant class of the logarithmic phase for *Ft novicida*. The lack of compound class similarities for determined markers could result from genetic differences between *Ft* SCHUS4 and *Ft novicida*. However, in agreement with the bacterial concentration data shown in Appendix Table A.3.3, the CFU counts suggests that *Ft SCHUS4* remained in a lag phase or a very early logarithmic phase throughout the first 24 h after inoculation. Additional measurements of the growth phases of *Ft* SCHUS4 over longer time periods are needed for a more comprehensive comparison of *Ft* SCHUS4 VOC markers against those of *Ft novicida*.

Table 3.6 Annotations of *F. tularensis* SCHU S4-specific VOC markers and average relative abundances (n=3) at 6 and 24 hours (Hr) post inoculation

Class	Compound	MS Base Peak (*m/z*)	RI (Lit)	RI (Exp)	NIST14 Match Factor	Abundance (Log10 Values) 6 Hr	24 Hr
Alcohol	3-Heptanol, 4-methyl-	43	915	831	75	5.49 b	5.85
Alcohol	1-Dodecanol	55	1473	1469	82	0.00	5.77 a
Alcohol	Phenol, 2,5-bis(1,1-dimethylethyl)-	191	1514	1508	76	4.59 b	4.87 a
Aldehyde	Furfural	96	833	818	73	4.77 b	4.94 a
Nitrogen-containing compound	Pyrazine, methyl-	94	801	802	76	0.00	5.43 a
Unknown	*m/z* 41 _ RI 672	41		672		0.00	5.65 a
Unknown	*m/z* 133 _ RI 913	133		913		0.00	7.42
Unknown	*m/z* 57 _ RI 941	57		941		5.30	0.00
Unknown	*m/z* 43 _ RI 986	43		986		4.81 a	0.00
Unknown	*m/z* 105 _ RI 990	105		990		4.81 b	5.23
Unknown	*m/z* 207 _ RI 998	207		998		0.00	5.46 a
Unknown	*m/z* 121 _ RI 1010	121		1010		5.30	0.00
Unknown	*m/z* 69 _ RI 1034	69		1034		5.25 a	4.89 b
Unknown	*m/z* 43 _ RI 1063	43		1063		4.79 b	4.93 a
Unknown	*m/z* 71 _ RI 1185	71		1185		0.00	5.02 a
Unknown	*m/z* 57 _ RI 1199	57		1199		0.00	4.96 a
Unknown	*m/z* 57 _ RI 1303	57		1303		5.00 a	4.74 b
Unknown	*m/z* 119 _ RI 1365	119		1365		0.00	4.91 a
Unknown	*m/z* 57 _ RI 1370	57		1370		4.90	0.00
Unknown	*m/z* 57 _ RI 1398	57		1398		5.58 a	5.55 a
Unknown	*m/z* 69 _ RI 1527	69		1527		5.36	0.00
Unknown	*m/z* 40 _ RI 1536	40		1536		4.74 a	0.00
Unknown	*m/z* 163 _ RI 1667	163		1667		5.02 a	0.00
Unknown	*m/z* 71 _ RI 1810	71		1810		4.23 b	4.61 a
Unknown	*m/z* 40 _ RI 1972	40		1972		4.50 a	0.00
Unknown	*m/z* 73 _ RI 2363	73		2363		5.17 a	5.56
Unknown	*m/z* 73 _ RI 2520	73		2520		5.00 a	4.56 b
Unknown	*m/z* 73 _ RI 2521	73		2521		4.80 b	5.31 a
Unknown	*m/z* 73 _ RI 2679	73		2679		4.58 b	5.09 a
Unknown	*m/z* 96 _ RI 2777	96		2777		4.34 a	0.00
Unknown	*m/z* 208 _ RI 3117	208		3117		5.52 a	0.00
Unknown	*m/z* 207 _ RI 3154	207		3154		4.85 b	5.63 a

Table 3.6 (cont'd)

Unknown	*m/z* 97 _ RI 3306	97	3306	4.94 a	5.11
Unknown	*m/z* 207 _ RI 3356	207	3356	5.16 a	0.00
Unknown	*m/z* 97 _ RI 3393	97	3393	0.00	4.88 a
Unknown	*m/z* 97 _ RI 3425	97	3425	0.00	4.88 a
Unknown	*m/z* 97 _ RI 3548	97	3548	4.63 b	4.95 a
Unknown	*m/z* 208 _ RI 3572	208	3572	0.00	4.84 a

Notes:
[a]: VOC detected in 2/3 of triplicate measurements
[b]: VOC detected in 1/3 of triplicate measurements
RI (Lit): Retention Index reported from NIST14
RI (Exp): Retention Index calculated from experiment

Table 3.7 Annotations of *B. anthracis* Ames-specific VOC markers and average relative abundances (n=3) at 6 and 24 hours (Hr) post inoculation

Class	Compound	MS Base Peak (*m/z*)	RI (Lit)	RI (Exp)	NIST14 Match Factor	Abundance (Log10 Values)	
						6 Hr	24 Hr
Alcohol	3-Octanol, 3,6-dimethyl-	73	1043	1110	83	5.96 b	6.03 a
Carboxylic Acid	Butanoic acid, 2-methyl-	74	861	844	89	5.73	0.00
Ester	Acetic acid, butyl ester	43	812	794	96	7.29	0.00
Ester	Propanoic acid, 2-methyl-, butyl ester	89	898	964	97	7.03	0.00
Ester	Butanoic acid, butyl ester	71	995	1009	80	5.95	0.00
Ester	Butyl 2-methylbutanoate	103	1043	1055	97	6.48	0.00
Ester	Butanoic acid, 3-methyl-, butyl ester	85	1047	1060	95	6.40	0.00
Ester	2-Butenoic acid, 2-methyl-, 2-methylpropyl ester, (E)-	101	1112	1146	74	5.36	0.00
Ketone	3-Octanone	43	986	999	71	5.28	5.31
Methyl Ketone	Acetoin	45	713	691	85	6.86	0.00
Methyl Ketone	Methyl Isobutyl Ketone	43	735	712	89	0.00	6.39
Methyl Ketone	2-Pentanone, 3-methyl-	43	752	722	77	5.29 a	5.84
Methyl Ketone	2-Hexanone, 5-methyl-	43	862	851	94	0.00	6.64
Methyl Ketone	2-Heptanone	43	891	892	96	6.28	6.67
Methyl Ketone	2-Heptanone, 6-methyl-	43	956	965	93	0.00	7.16
Methyl Ketone	2-Nonanone	58	1092	1069	86	0.00	5.97
Methyl Ketone	Benzyl methyl ketone	91	1110	1141	85	0.00	5.67
Nitrogen-Containing Compound	Pyrazine, tetramethyl-	136	1089	1099	77	4.85 b	5.80
Unknown	*m/z* 42 _ RI 748	42		748		0.00	5.29
Unknown	*m/z* 40 _ RI 874	40		874		5.01 a	0.00
Unknown	*m/z* 42 _ RI 912	42		912		7.18 b	7.67
Unknown	*m/z* 57 _ RI 915	57		915		5.68 a	0.00
Unknown	*m/z* 43 _ RI 938	43		938		5.15 a	0.00
Unknown	*m/z* 93 _ RI 939	93		939		4.63 b	5.04 a
Unknown	*m/z* 43 _ RI 976	43		976		4.95 a	6.26

Table 3.7 (cont'd)

Unknown	*m/z* 57 _ RI 1012	57	1012	5.89 a	0.00
Unknown	*m/z* 69 _ RI 1034	69	1034	5.41 a	0.00
Unknown	*m/z* 43 _ RI 1054	43	1054	0.00	5.23
Unknown	*m/z* 55 _ RI 1057	55	1057	0.00	5.36
Unknown	*m/z* 43 _ RI 1063	43	1063	0.00	5.00 a
Unknown	*m/z* 57 _ RI 1071	57	1071	4.60 b	4.99 a
Unknown	*m/z* 130 _ RI 1126	130	1126	0.00	5.40 a
Unknown	*m/z* 149 _ RI 1169	149	1169	5.41 a	0.00
Unknown	*m/z* 57 _ RI 1180	57	1180	5.21	5.31
Unknown	*m/z* 91 _ RI 1189	91	1189	0.00	5.29
Unknown	*m/z* 55 _ RI 1201	55	1201	5.26	0.00
Unknown	*m/z* 108 _ RI 1205	108	1205	4.42 b	4.83 a
Unknown	*m/z* 339 _ RI 1218	339	1218	4.32 b	4.65 a
Unknown	*m/z* 71 _ RI 1285	71	1285	4.56 b	4.72 a
Unknown	*m/z* 119 _ RI 1365	119	1365	0.00	4.85 a
Unknown	*m/z* 401 _ RI 1564	401	1564	0.00	4.96 a
Unknown	*m/z* 405 _ RI 1685	405	1685	0.00	4.66 a
Unknown	*m/z* 40 _ RI 1747	40	1747	0.00	4.36 a
Unknown	*m/z* 40 _ RI 2271	40	2271	4.46 a	0.00
Unknown	*m/z* 73 _ RI 2521	73	2521	0.00	5.09 a
Unknown	*m/z* 281 _ RI 2674	281	2674	0.00	5.07 a
Unknown	*m/z* 281 _ RI 2983	281	2983	5.39 a	4.66 b
Unknown	*m/z* 281 _ RI 3008	281	3008	5.54 a	0.00
Unknown	*m/z* 208 _ RI 3069	208	3069	5.86 a	5.43 b
Unknown	*m/z* 209 _ RI 3135	209	3135	5.31 a	0.00
Unknown	*m/z* 207 _ RI 3154	207	3154	5.04 b	5.72
Unknown	*m/z* 281 _ RI 3173	281	3173	5.71 a	0.00

Table 3.7 (cont'd)

Unknown	*m/z* 281 _ RI 3190	281	3190	5.64 a	0.00
Unknown	*m/z* 208 _ RI 3439	208	3494	0.00	5.04 a
Unknown	*m/z* 97 _ RI 3521	97	3521	0.00	4.87 a
Unknown	*m/z* 97 _ RI 3533	97	3533	0.00	4.91 a

Notes:
[a]: VOC detected in 2/3 of triplicate measurements
[b]: VOC detected in 1/3 of triplicate measurements
RI (Lit): Retention Index reported from NIST14
RI (Exp): Retention Index calculated from experiment

3.4 DISCUSSION

The methodology and results described here provide initial groundwork for detection and identification of volatile biomarkers from bacterial pathogens including fully virulent RG3 strains. The application of this non-invasive methodology for VOC profiling applied to actively growing *F. tularensis* and *B. anthracis* bacterial cultures revealed dynamic profiles, influenced by both the bacterial growth phase and bacterial concentration. At any given timepoint, isolation of the bacterial biomarkers was complicated by background volatiles, and data processing was applied uniformly across all sample types to identify bacterial biomarkers.

As the VOC profiles observed here are discussed, one should keep in mind that measured VOC profiles were influenced by the sampling and detection methods used. For example, the type of sorbent material could have introduced a sampling bias (sampling efficiency is dependent on partition behavior of each compound) and the sampling time and mass spectral analysis method could have influenced the sensitivity with which compounds can be detected. Generally, detection limits for the basic type of SPME-GC-quadrupole MS used here for untargeted analysis

(scan, not select ion mode) are on the order 1 ng of a compound injected into a column. It is conceivable that only the most prevalent VOCs were detected. A larger number of relevant may be found if more efficient sampling techniques (e.g., thermal desorption tubes) and more sensitive mass spectrometry protocols such as selected ion monitoring are used. It should be noted here that absolute VOC quantification was not attempted here. For various practical reasons, absolute quantification of VOCs in entire cultures is challenging. The ideal standards, stable isotope-labeled internal standards, are susceptible to metabolic degradation. Also, cultures used here were not fully enclosed (flasks with vented caps) because gas exchange (oxygen) was required to sustain growth and needed to be vented to avoid buildup of pressure. However, relative abundances of compounds were compared among cultures by integrating chromatographic peaks across species and timepoints.

The cumulative VOC profiles of *Ft novicida*, *Ft* SCHUS4, *Ba* Sterne, and *Ba* Ames determined here included representatives of different compound classes such as methyl ketones, alcohols, nitrogen-containing compounds, sulfur-containing compounds, carboxylic acids, esters, and various unidentified biomarkers. Exhaustive identification of every pathway that produces these volatiles is beyond the scope of this discussion, but several likely routes of biosynthesis are enumerated below.

3.4.1 Ketones

Ketones were abundant markers, present in all pathogens except *Ft* SCHUS4, and largely as methyl ketones. The methyl ketones are likely formed by modifying products of the fatty acid biosynthesis pathway, specifically the β-oxidation of fatty acids [26]. Odd-chain methyl ketones

can be formed through the decarboxylation of even-carbon β-keto acids. Conversely, even-carbon methyl ketones arise from odd-carbon fatty acids and occur with lower frequency [27].

Interestingly, methyl ketones with straight-chain alkane branches were abundant in *Ft novicida*, while primarily branched and aromatic methyl ketones were prevalent in *Ba* Sterne and Ames. This difference may stem from *B. anthracis* being Gram-positive, whereas *F. tularensis* is Gram-negative. Synthesis of fatty acids in Gram-positive and Gram-negative bacteria is controlled by enzymes with different preferred substrates. For example, comparison of the enzyme β-ketoacyl-acyl carrier protein synthase III from Gram-negative *E. coli* and Gram-positive bacteria *S. aureus* demonstrated a larger binding pocket in the Gram-positive bacteria, thus having a higher likelihood for synthesis of branched-chain alkyl substrates [28]. A smaller binding pocket for the Gram-positive bacteria would limit the use of branched-chain alkyl substrates.

3.4.2 Alcohols

While alcohols were present for both *F. tularensis* and *B. anthracis*, their number and relative abundances were greater in *F. tularensis*. Alcohols may be synthesized from the breakdown products of β-oxidation of fatty acids, for example after enzymatic reduction of carboxylic acids [27,29]. The fatty acid chains observed for the alcohols class exhibited diversity, including straight-chain, branched chain, and aromatic substituents. 1-nonanol was likely formed by reduction of fatty acids. The 2-alkanols (2-nonanol and 2-undecanol) are postulated to be derived from corresponding methyl ketones as reduced intermediates, since the corresponding methyl ketones were also detected at all timepoints where the 2-alkanols were detected, usually at a higher relative abundance. 2-nonanol was not observed after 24 hours post-

inoculation, indicating it was utilized by *F. tularensis* as a precursor for stationary phase metabolism. The aromatic alcohol phenylethyl alcohol is a widely occurring VOC produced by several bacterial species. Volatile alcohols have been shown to play a role in growth inhibition of several bacteria and fungi [30,31].

3.4.3 Sulfur-containing compounds

Dimethyltrisulfide was an abundant VOC uniquely present in *Ft novicida* during the decline phase, when no viable bacteria were detected. Sulfur-containing VOCs are attributed to breakdown of the amino acids cysteine and methionine [27]. Dimethyltrisulfide has previously been observed as product of human decomposition caused by bacteria.

3.4.4 Nitrogen-containing compounds

The detection of nitrogen-containing pyrazine markers produced by bacteria is complicated by endogenous pyrazine VOCs present in the growth media [27]. The sterilization of growth media through autoclaving heats amino acids and reducing sugars, producing pyrazines via the Maillard Reaction [32]. The Mueller-Hinton growth media controls consistently produced 2,5-dimethylpyrazine, and a similar relative abundance was observed in the *Ft novicida* cultures through the first 16 h of growth. At 20 h of growth and beyond, the relative abundance of 2,5-dimethylpyrazine rose more than 10x the relative abundance of the controls, suggesting the growing bacteria have active involvement in biosynthesis of pyrazines. An additional pyrazine, 2-methyl-3-isopropylpyrazine, was also observed. Isopropyl substituents to pyrazine compounds are not common constituents in bacterial volatiles [27]. Therefore, we

hypothesize both pyrazines originate from *Ft novicida* under the chosen growth conditions. Pyrazines have also been observed as volatile byproducts of bacterial metabolism, for example, in the genera *Streptomyces* [33] and *Bacillus* [34]. Only one nitrogen-containing compound, tetramethylpyrazine, was observed in *Ba* Ames during the last observed timepoint, estimated to be in the logarithmic phase, but was not as abundant in as in *Ft novicida*. This is compounded by the different growth media utilized, which emphasizes the need for careful evaluation when comparing biomarkers across different growth conditions and species.

3.4.5 Esters and carboxylic acid compounds

Esters and carboxylic acids were detected exclusively in both *Ba* Sterne and *Ba* Ames, but not in the *F. tularensis* strains. The identified *B. anthracis* markers contained either propanoic or butanoic acid as the backbone for methylated esters or the side-chain for carboxylic acids. The formation of esters and carboxylic acids can be derived from shared metabolic pathways occurring during normal bacterial growth, such as oxidation of fatty acids or amino acid metabolism. As the ester and carboxylic acid compounds were only observed during the logarithmic growth stage, this demonstrates a shift in *B. anthracis* metabolism once bacteria reach the stationary phase.

3.4.6 Evidence of dynamic metabolic processes

For all four bacterial taxa studied here, their VOC marker profiles varied as function of time after inoculation/culture start and, as observed for *Ft novicida*, *Ba* Sterne and *Ba* Ames, varied distinctly across their growth phases. For *Ft novicida* and *Ba* Sterne, this was also shown

through application of PCA, which produced distinct groupings for the VOC markers of different growth phases. For example, in the *Ft novicida* cultures, the methyl ketones, once produced, were present throughout the remainder of the experiment. However, select alcohols (2-nonanol and 2-undecanol) were not detected after a mid-stationary (28-hour) timepoint. This suggests alcohols were depleted in the liquid culture, potentially as precursors in ongoing bacterial metabolism. Once the basic metabolism of isolated pathogens is determined, changes in the marker profiles when additional variables are added (e.g., different substrates) can help drive inferences on metabolic activity of complex systems.

3.4.7 Comparison of VOC markers for RG3 vs. RG2 strains

In comparing the putative VOC biomarkers identified for *Ba* Ames (RG3) to those for *Ba* Sterne (RG2) some similarities were found, but also distinct differences. In contrast, the profile of VOCs from *Ft* SCHU S4 (RG3) had fewer similarities with *Ft novicida* (RG2). The relative similarities between *Ba* Ames and *Ba* Sterne may stem from the close genetic relationship of these two strains (*Ba* Sterne is missing one of the two plasmids that Ba Ames has but is otherwise genetically very similar to *Ba* Ames[35]). In contrast, *Ft* SCHU S4 and *Ft novicida* are genetically more distinct[36]. If further confirmed in future studies, this may have implications for the use of RG2 "simulants" to develop sensors and algorithms for detecting exposures to the related RG3 pathogens. In the biodefense community *Ba* Sterne is generally considered a good simulant for *Ba* Ames. For *Ft* RG2 simulants other than *Ft novicida* may be considered.

Future VOC sampling should be performed on additional subspecies of *F. tularensis* (e.g., spp. holarctica) and *B. anthracis* (e.g., spp. Vollum or H9401) to investigate whether these profiles are unique to a subspecies, species, or bacterial pathogens in general. Several markers

identified in this study have been previously reported as emissions of other bacterial types. For example, Chen et al. [12] reported 2-heptanone, 2-nonanone, and 2-undecanone in *E. coli* but did not detect higher carbon methyl ketones. Rees et al. [10,11] reported both even and odd-chain methyl ketones, including 2-hexanone, 2-heptanone, 2-nonanone, and 2-decanone products from *Klebsiella pneumoniae*, where the presence of even-chain methyl ketones suggests a different or complementary synthesis pathway for volatile production.

One of the long-term goals of this project seeks to use VOCs as breath-based diagnostic markers towards the detection of biowarfare agents in patients after a suspected biological attack. During a hypothetical pathogen infection in humans, the VOCs in breath may be derived from 1) the invading pathogen, 2) the human breath volatilome, or 3) interactions between the human host and pathogen. This study represents a first step in non-invasive methodology and data analysis optimization, extensively profiling two attenuated pathogen or RG2 species and screening their RG3 virulent counterparts in optimized growth media. The number of compounds identified in the human breath volatilome continues to grow through targeted and untargeted studies. A searchable database of breath-specific compounds in the "human volatilome" has been curated by the U.S. Environmental Protection Agency (EPA) and is continuously updated [37,38]. A survey of this list against the *F. tularensis* RG2 and RG3 profiles found here revealed 2-heptanone and 2-nonanone have been detected in human breath, while a comparison against the *B. anthracis* RG2 and RG3 profiles revealed 2-methylpropanoic acid, 2-heptanone, 6-methyl-2-heptanone, and 5-methyl-5-hepten-2-one that have been reported in human breath. Also, it is important to note that the volatiles in this database may not be commonly shared among all people, as human breath has been shown to be influenced by one's unique personal microbiome, external exposures, and immunological responses. The effects of shared volatiles between

pathogens and a human host must be evaluated in further studies that better simulate an *in vivo* infection, as well as identifying markers unique to that interaction.

Future work into "baseline" human breath signatures, a pathogen-specific volatilome, and host-microbe interactions are required for evaluation of VOCs as diagnostic tools for human health. Towards a pathogen-specific volatilome, further efforts include expanding both the number of bacterial species and evaluating the effects of chosen growth media on VOCs produced. Additionally, as animal model studies have established a low bacterial count can establish infections (e.g., 10 bacterial counts for *F. tularensis* in primate models), optimization of signal detection will also be investigated, as the conditions employed here used relatively high bacterial counts. Finally, future work should make the transition from *in vitro* studies into experiments more closely aligned with *in vivo* studies, such as initiating bacterial infection of human lung cell cultures and analyzing the resultant profiles for discovery of overlapping volatile compounds that may serve as diagnostic markers of human exposures to biosecurity-relevant pathogens.

3.5 CONCLUSIONS

This study adapted a SPME-GC-MS methodology for noninvasive profiling of VOCs emitted from actively growing pathogens, specifically potential biowarfare bacterial agents and their surrogates, in both BSL-2 and BSL-3 settings. The devised methodology detected volatile biomarkers that were reflective of both the presence and physiological growth phase of pathogens. The data processing employed distinguished signals from the pathogens against a complex chemical background, in this case aided by the use of powerful software (MassHunter, MPP) for compound annotation and visualization of GC-MS data. Although the devised

methodology based on SPME-GC-quadrupole MS does not represent the pinnacle of sensitivity, a number of relatively robust and reproducible putative volatile biomarkers could be detected. Further confirmation of these markers should be pursued in more repeat experiments across a wider range of growth conditions. More efficient VOC collection methods and more sensitive mass spectral analysis techniques may also uncover additional markers in the future.

Detection and identification of metabolites specific to taxa or species provides the first steps to understanding their formation via various metabolic pathways and the genetic basis for these pathways. It should be acknowledged that the work presented here constitutes only initial scoping experiments. While this work demonstrates the applicability of this method and found a number of interesting volatile biomarkers, this work needs to be expanded to determine the influence of various experimental factors on markers. Future research should include determining the dependence of pathogen-produced volatiles on environmental conditions (e.g., chosen growth media) and use of different VOC collection methods (e.g., thermal desorption tubes) to achieve lower detection limits. Elucidation of comprehensive bacterial profiles is expected to provide clues about bacterial metabolism in controlled environments, which can further inform research into metabolic processes when pathogens are in other settings (i.e,. a host). Ultimately, such biomarkers may yield useful information about metabolism in bacterial taxa and may facilitate new applications in biodetection. Distinct volatile profiles have potential to be used for the detection of pathogens in the context of biosecurity-relevant exposures of humans during a biological attack.

Results from this work have implications in the larger volatilomics community, both within the field of pathogens study and beyond. While volatile compounds from *B. anthracis* have been previously studied, this is the first study, to our knowledge, to profile volatile

emissions of *F. tularensis*. Future databases can incorporate biomarker signatures from various pathogen species for means of relevant comparisons.

APPENDICES

APPENDIX A: Figures

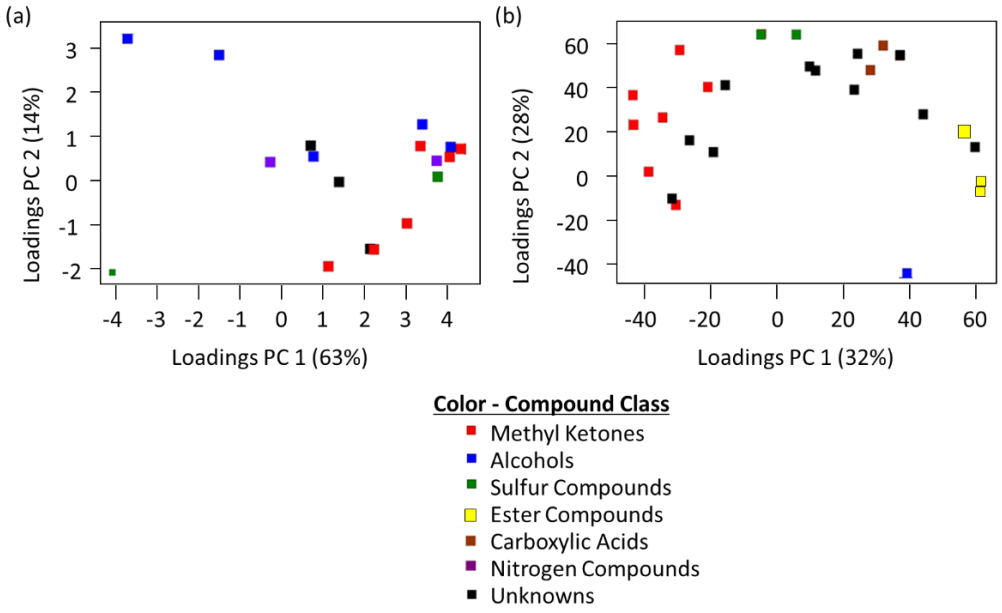


Figure A.3.1 PCA loadings plots for VOC profiles of (a) *F. tularensis novicida* and (b) *B. anthracis* Sterne generated using the relative abundances of pathogen VOCs measured using GC-MS across all timepoints (For PCA score plots see Figure 3.5). Points represent individual VOC markers (colored by compound class) explaining placement of samples on scores plot.

The loadings plots shown in Appendix Figure A.3.1a and A.3.1b depicts the relative importance of each volatile towards the positioning of individual samples in the scores plots (Figures 3.5a and 3.5b, main manuscript). For *F. tularensis* spp. novicida, the first two principal components (PCs) described 63 % and 14 ± % of the variance in data, respectively. Positioning on PC1 is dependent on the abundances of 2,5-dimethylpyrazine and 2-undecanone. Both compounds are present at all timepoints and generally increase with higher timepoints, creating the left-right distribution seen in the scores plot. Positive loading on PC2 indicates a higher percentage of alcohol-containing compounds, indicative of the early log phase and early stationary phase. A progressive shift towards a negative PC2 loading indicates an increasing percentage of methyl ketones, which rise to prominence in the late stationary phase, culminating in the appearance of dimethyltrisulfide. Notably, the volatiles that contributed the most to growth

phase separation, those of PC 2, were not the most abundant volatile classes, demonstrating the power of this analysis when characterizing trace compounds in a complex chemical profile. For *B. anthracis* Sterne, the first two PCs described 32 % and 28 % of the variance in data, respectively. Positioning on PC1 in the scores plots exhibits strong dependence on the abundances of 4-heptanol (negative loading) and the sulfur-containing compounds (S-methyl ester butanethioic acid and thiopivalic acid) and carboxylic acids (methylated butanoic acids and 2-ethyl-propanoic acid). A positive positioning on PC2 separates the logarithmic phase, containing a high percentage of esters, from the negatively positioned stationary phase, containing a high percentage of methyl ketones.

APPENDIX B: Tables

Table A.3.1 Relative abundances of *F. tularensis novicida*-associated VOCs for each replicate at all measured timepoints, separated by growth phase

	Time Post Inoculation (Hours)		0	2	4	8	12	16	20	24	28	32	48	52
	Growth Phase		Log Phase							Stationary Phase		Decline Phase		
Compounds	Compound Class	Replicate	Abundance (Log 10 Values)											
1-Butanol, 2-methyl-	Alcohols	1	0.00	0.00	0.00	0.00	0.00	0.00	6.72	6.65	6.82	6.83	0.00	0.00
		2	0.00	0.00	0.00	0.00	6.35	6.31	6.66	0.00	6.33	6.51	6.18	5.92
		3	0.00	0.00	0.00	0.00	0.00	0.00	6.89	6.79	6.96	6.85	0.00	0.00
2-Nonanol	Alcohols	1	0.00	0.00	0.00	0.00	6.02	6.61	6.07	5.88	0.00	0.00	0.00	0.00
		2	0.00	0.00	0.00	0.00	6.06	6.60	6.10	5.87	0.00	0.00	0.00	0.00
		3	0.00	0.00	0.00	0.00	6.07	6.63	6.09	5.90	0.00	0.00	0.00	0.00
Phenylethyl Alcohol	Alcohols	1	0.00	0.00	5.88	6.12	6.31	6.34	6.52	6.79	6.81	6.80	6.72	6.64
		2	0.00	0.00	5.79	6.03	6.27	6.35	6.47	6.70	6.68	6.83	6.66	6.72
		3	0.00	0.00	0.00	6.04	6.28	6.40	6.61	6.75	6.73	6.79	6.64	6.64
1-Nonanol	Alcohols	1	0.00	0.00	5.88	6.44	6.72	7.25	7.46	7.02	6.48	6.34	5.94	5.83
		2	0.00	0.00	5.75	6.26	6.72	7.28	7.54	7.11	6.52	6.26	5.96	5.95
		3	0.00	0.00	5.76	6.33	6.67	7.24	7.35	6.69	6.26	6.09	5.76	5.76
2-Undecanol	Alcohols	1	0.00	0.00	0.00	0.00	6.13	6.34	6.26	6.16	5.97	0.00	0.00	0.00
		2	0.00	0.00	0.00	0.00	6.15	6.33	6.25	6.17	5.90	0.00	0.00	0.00
		3	0.00	0.00	0.00	0.00	6.15	6.37	6.30	6.03	6.00	0.00	0.00	0.00
2-Heptanone	Methyl Ketones	1	0.00	0.00	0.00	6.05	0.00	6.52	6.71	6.82	6.84	6.87	6.51	6.21
		2	0.00	0.00	0.00	5.74	6.18	6.46	6.73	6.82	6.68	6.88	6.46	6.38
		3	0.00	0.00	0.00	6.16	6.24	6.47	6.79	6.87	6.90	6.90	6.41	6.38
2-Nonanone	Methyl Ketones	1	0.00	0.00	0.00	6.42	6.53	6.85	7.04	7.26	7.37	7.48	7.01	6.76
		2	0.00	0.00	0.00	6.37	6.68	6.81	7.12	7.25	7.32	7.44	6.97	6.87
		3	0.00	0.00	0.00	6.52	6.58	6.85	7.17	7.41	7.52	7.51	6.89	6.91

Table A.3.1 (cont'd)

2-Undecanone	Methyl Ketones	1	5.17	0.00	5.90	6.15	5.96	6.04	6.64	7.09	7.20	7.26	6.28	6.06
		2	0.00	0.00	5.87	6.12	6.05	6.05	6.60	7.07	7.23	7.17	6.25	6.19
		3	0.00	5.75	5.80	6.12	5.92	5.99	6.87	7.30	7.41	7.26	6.21	6.20
2-Tridecanone	Methyl Ketones	1	0.00	0.00	0.00	0.00	0.00	0.00	5.96	6.82	6.94	6.97	6.09	5.92
		2	0.00	0.00	0.00	0.00	0.00	0.00	5.99	6.84	7.00	6.80	6.06	6.07
		3	0.00	0.00	0.00	0.00	0.00	0.00	6.34	7.03	7.09	6.88	6.13	6.06
2-Pentadecanone	Methyl Ketones	1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6.20	6.43	6.55	6.03	5.92
		2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6.12	6.43	6.42	5.96	6.09
		3	0.00	0.00	0.00	0.00	0.00	0.00	5.43	6.34	6.62	6.55	6.10	6.08
2-Heptadecanone	Methyl Ketones	1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.42	5.90	5.65	5.68	5.65
		2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.69	5.53	5.57	5.78
		3	0.00	0.00	0.00	5.19	0.00	0.00	0.00	5.47	5.78	6.17	5.62	5.64
Pyrazine, 2,5-dimethyl-	Nitrogen-Containing Compounds	1	0.00	6.35	6.45	6.46	6.72	7.10	7.36	7.57	7.56	7.78	7.71	7.76
		2	5.72	6.39	6.36	6.30	6.86	7.24	7.44	7.71	7.44	7.66	7.82	7.78
		3	5.97	6.38	6.40	6.36	6.77	7.19	7.27	7.48	7.74	7.81	7.89	7.88
2-Methyl-3-isopropylpyrazine	Nitrogen-Containing Compounds	1	0.00	0.00	0.00	0.00	6.11	6.21	6.45	6.74	6.90	7.04	7.08	7.00
		2	0.00	0.00	5.94	0.00	6.15	6.23	6.45	6.71	6.77	7.04	7.03	7.08
		3	0.00	0.00	0.00	0.00	6.12	6.27	6.51	6.79	6.92	7.08	7.07	7.07
Dimethyl trisulfide	Sulfur-Containing Compounds	1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	7.00	6.55
		2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6.91	6.86
		3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6.74	7.01
1-Propanol, 3-(methylthio)-	Sulfur-Containing Compounds	1	0.00	0.00	0.00	0.00	5.87	0.00	5.88	6.04	5.93	6.11	6.09	6.05
		2	0.00	0.00	0.00	0.00	5.93	5.86	6.01	6.01	5.93	6.07	6.07	6.19
		3	0.00	0.00	0.00	0.00	0.00	6.00	6.03	6.09	6.11	6.03	6.08	6.13
m/z 121 _ RI 1002	Unknown	1	0.00	6.19	6.19	6.29	6.07	5.98	5.99	6.15	6.23	6.30	6.22	5.99
		2	0.00	6.13	0.00	5.96	0.00	0.00	0.00	6.15	6.06	6.34	6.19	6.16
		3	0.00	6.10	0.00	5.96	6.08	5.95	6.08	6.16	6.24	6.35	6.08	6.11

Table A.3.1 (cont'd)

m/z 108 _ RI 1049	Unknown	1	0.00	5.91	5.67	6.39	6.53	6.48	6.54	6.67	6.62	6.70	6.62	6.55
		2	0.00	5.54	5.83	6.38	6.49	6.51	6.55	6.62	6.56	6.70	6.59	6.63
		3	0.00	5.87	5.85	6.34	6.52	6.51	6.53	6.57	6.61	6.67	6.44	6.57
m/z 133 _ RI 1110	Unknown	1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6.22	6.09	6.03	5.92	5.83
		2	0.00	0.00	0.00	0.00	0.00	0.00	5.55	6.02	5.92	6.05	5.98	5.91
		3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.90	6.00	6.14	5.91	5.87

Table A.3.2 Relative abundances of *B. anthracis Sterne*-associated VOCs for each replicate at all measured timepoints, separated by growth phase

	Time Post Inoculation (Hours)		**4**	**8**	**12**	**20**	**24**
	Growth Phase		**Log Phase**		**Stationary Phase**		
Compound Class	**Compound**	**Replicate**	**Abundance (Log 10 Values)**				
Alcohols	4-Heptanol	1	6.10	0.00	0.00	0.00	0.00
		2	6.21	0.00	0.00	0.00	0.00
		3	6.12	0.00	0.00	0.00	0.00
Carboxylic Acid	Propanoic acid, 2-methyl-	1	0.00	5.49	0.00	0.00	0.00
		2	0.00	5.62	5.93	0.00	0.00
		3	0.00	0.00	5.20	0.00	0.00
Carboxylic Acid	Butanoic acid, 2-methyl-	1	0.00	5.82	0.00	0.00	0.00
		2	0.00	5.78	5.62	0.00	0.00
		3	0.00	6.11	5.29	0.00	0.00
Carboxylic Acid	Butanoic acid, 3-methyl-	1	0.00	6.09	5.47	0.00	0.00
		2	0.00	6.08	5.87	0.00	0.00
		3	0.00	6.27	5.58	0.00	0.00
Ester	Propanoic acid, 2-methyl-, butyl ester	1	6.50	6.48	0.00	0.00	0.00
		2	6.53	6.59	0.00	0.00	0.00
		3	6.46	6.81	0.00	0.00	0.00
Ester	Butanoic acid, butyl ester	1	6.70	6.12	0.00	0.00	0.00
		2	6.74	6.15	0.00	0.00	0.00
		3	6.66	6.35	0.00	0.00	0.00
Ester	Butyl 2-methylbutanoate	1	5.83	5.95	0.00	0.00	0.00
		2	5.86	6.05	0.00	0.00	0.00
		3	5.75	6.26	0.00	0.00	0.00

Table A.3.2 (cont'd)

Ester	Butanoic acid, 3-methyl-, butyl ester	1	5.42	6.10	4.94	0.00	0.00
		2	5.48	6.17	0.00	0.00	0.00
		3	5.40	6.37	5.03	0.00	0.00
Methyl Ketone	Methyl Isobutyl Ketone	1	0.00	0.00	0.00	6.30	6.10
		2	0.00	0.00	5.83	6.47	6.23
		3	0.00	0.00	5.93	0.00	6.07
Methyl Ketone	2-Hexanone, 5-methyl-	1	0.00	0.00	6.04	6.55	6.63
		2	0.00	5.40	5.90	6.63	6.96
		3	0.00	5.75	6.18	6.30	6.59
Methyl Ketone	2-Heptanone	1	0.00	6.50	6.71	6.68	6.68
		2	5.30	6.60	6.54	6.82	7.12
		3	5.30	6.64	6.68	6.52	6.72
Methyl Ketone	2-Heptanone, 6-methyl-	1	0.00	0.00	6.65	6.88	6.79
		2	0.00	0.00	6.56	7.03	7.28
		3	0.00	0.00	6.63	6.83	6.81
Methyl Ketone	2-Heptanone, 5-methyl-	1	0.00	0.00	5.66	5.99	6.04
		2	0.00	0.00	0.00	6.12	6.51
		3	0.00	0.00	5.76	5.96	5.98
Methyl Ketone	5-Hepten-2-one, 6-methyl-	1	0.00	0.00	0.00	0.00	5.39
		2	0.00	0.00	0.00	0.00	5.72
		3	0.00	0.00	0.00	4.81	5.21
Methyl Ketone	2-Heptanone, 4,6-dimethyl-	1	0.00	0.00	5.27	0.00	5.46
		2	0.00	0.00	0.00	5.59	5.88
		3	0.00	0.00	0.00	5.21	5.44
Sulfur containing compound	Butanethioic acid, S-methyl ester	1	0.00	0.00	5.56	5.97	0.00
		2	0.00	5.48	5.38	5.88	0.00
		3	0.00	5.57	5.70	5.82	0.00

Table A.3.2 (cont'd)

Sulfur containing compound	Thiopivalic acid	1	5.08	5.41	5.49	5.99	0.00
		2	0.00	5.50	5.37	5.71	0.00
		3	0.00	5.48	5.54	5.84	4.82
Unknown	*m/z* 80 _ RI 715	1	0.00	5.82	0.00	5.76	0.00
		2	5.55	5.80	5.70	5.88	5.68
		3	0.00	5.78	5.98	5.79	0.00
Unknown	*m/z* 57 _ RI 769	1	0.00	0.00	5.02	5.05	0.00
		2	0.00	0.00	0.00	5.28	0.00
		3	0.00	0.00	0.00	5.21	0.00
Unknown	*m/z* 43 _ RI 791	1	6.95	7.17	6.89	6.29	5.92
		2	6.97	7.10	6.67	6.09	5.69
		3	6.84	7.37	6.91	6.03	0.00
Unknown	*m/z* 43 _ RI 873	1	0.00	0.00	0.00	0.00	4.92
		2	0.00	0.00	0.00	4.82	4.89
		3	0.00	0.00	0.00	0.00	4.74
Unknown	*m/z* 45 _ RI 901	1	0.00	0.00	0.00	5.41	0.00
		2	0.00	0.00	5.21	5.07	0.00
		3	0.00	0.00	5.43	5.35	0.00
Unknown	*m/z* 57 _ RI 912	1	5.63	5.67	0.00	0.00	0.00
		2	5.70	5.81	0.00	0.00	0.00
		3	5.58	5.94	5.54	0.00	0.00
Unknown	*m/z* 43 _ RI 956	1	5.11	0.00	0.00	0.00	0.00
		2	5.20	0.00	0.00	0.00	0.00
		3	5.23	0.00	0.00	0.00	0.00
Unknown	*m/z* 58 _ RI 962	1	0.00	5.82	0.00	0.00	0.00
		2	0.00	5.81	0.00	0.00	0.00
		3	0.00	5.84	0.00	0.00	0.00

Table A.3.2 (cont'd)

Unknown	*m/z* 43 _ RI 997	1	0.00	0.00	0.00	5.18	0.00
		2	0.00	0.00	0.00	5.23	5.39
		3	0.00	5.35	0.00	0.00	5.32
Unknown	*m/z* 90 _ RI 1005	1	0.00	0.00	5.11	5.16	0.00
		2	0.00	0.00	5.22	0.00	0.00
		3	0.00	0.00	5.31	5.18	0.00
Unknown	*m/z* 58 _ RI 1104	1	5.69	5.67	5.73	0.00	5.76
		2	5.65	5.87	5.60	5.61	5.94
		3	5.49	5.81	5.66	0.00	5.60
Unknown	*m/z* 83 _ RI 1145	1	0.00	5.05	0.00	0.00	0.00
		2	0.00	5.10	4.51	0.00	0.00
		3	0.00	5.16	4.90	0.00	0.00
Unknown	*m/z* 58 _ RI 1554	1	0.00	0.00	4.92	0.00	0.00
		2	0.00	5.02	5.10	0.00	0.00
		3	0.00	0.00	4.97	0.00	0.00

Table A.3.3 Bacterial concentrations of *F. tularensis* SCHU S4 in modified Mueller-Hinton media and *B. anthracis* Ames in Brain-Heart Infusion media measured at the VOC sampling time points 6 h and 24 h post-inoculation of cultures. The numbers represent the mean of CFU/mL determined from 2 plate counts for each of the 3 culture replicates (total number of plate counts for each time point = 6), the errors represent standard deviation.

	6 h	24 h
Ft SCHU S4	$(6.7 \pm 1.6) * 10^5$	$(6.1 \pm 2.9) * 10^5$
Ba Ames	$(6.7 \pm 1.4) * 10^6$	$(2.5 \pm 0.4) * 10^7$

APPENDIX C: Protocol

Solid-Phase Microextraction (SPME) Sampling for Volatile Organic Compounds (VOCs) of Liquid Bacterial Cultures

Lawrence Livermore National Laboratory

Protocol: Rasley-2018-001

Facility/Program: SCA_BSL-2_BSL-3 Global Security

Effective Date: 09/26/2018

Level of Use: Continuous Use

LLNL-MI-791181

Table A.3.P.1 Protocol Document Revision History

Revision #	Date	Author	Summary of Changes
Initial	07-03-2018	Amy Rasley	This SOP was developed to establish the procedure for SPME fiber head-space sampling of liquid bacterial cultures without contamination of the SPME fibers with viable bacteria.
Rev 1	9/17/2018	Amy Rasley	Revised to include non-contamination test results when SPME fibers are exposed to fully virulent *Bacillus anthracis*.
Rev 2	9/26/2018	Amy Rasley	Revised to include non-contamination test results when SPME fibers are exposed to fully virulent *Francisella tularensis*.

1. Purpose

The purpose of this procedure is to demonstrate that solid-phase microextraction (SPME) fibers are not contaminated with viable agent when exposed to the head space of liquid bacterial cultures during sampling for volatile organic compounds (VOCs) prior to removal of SPME fiber sampling devices to a lower physical containment level [e.g., lower biosafety level (BSL), non-BSL laboratory] for GS/MS analysis.

2. Scope

This procedure applies to and has been validated for head space sampling for VOCs using SPME fiber VOC collection devices with the following agents:

- *Bacillus anthracis* Sterne (BSL-2)
- *Francisella tularensis* subsp. *novicida* (BSL-2)
- *Bacillus anthracis* Ames (BSL-3)
- *Francisella tularensis* SCHU S4 (BSL-3)

Note: This protocol has been developed and validated using attenuated, select agent-exempt strains in a non-registered BSL-2 laboratory for the purpose of demonstrating that SPME fibers used for head space sampling of VOCs from liquid bacterial cultures are not contaminated with viable agent during the sampling process.

Prior to removing SPME fibers from any registered BSL-2 or BSL-3 BSAT laboratory, this procedure shall be validated on the fully virulent agents and the SOP updated and approved by the Biosafety Officer (BSO) and Responsible Official (RO).

As of the current revision, this protocol has been validated for use on the fully virulent *Bacillus anthracis* and *Francisella tularensis*. Validation results indicated that there is no potential for SPME fiber contamination during the sampling process. The current SOP has been updated and approved by the Biosafety Officer (BSO) and Responsible Official (RO).

3. Responsibility

Research Personnel

Research personnel who will be performing head space sampling using SPME fiber collection devices with any of the above listed agents for which this procedure has been validated must be familiar with and adhere to the exact procedures outlined in this SOP. Any deviation from the written procedure will require re-review and re-validation of the procedure.

All research personnel who will be performing this procedure are required to read, understand and review this procedure with the appropriate PI/RI who developed the procedure and must sign the Training Record in Appendix A prior to performing this procedure.

Principal Investigator (PI)/Responsible Individual (RI)

The PI/RI of the protocol is responsible for submitting the written procedure to the RO and BSO for review and approval prior to initiating validation of the procedure. In addition, the PI/RI is responsible for submitting validation data to the RO for review and final approval of the procedure prior to initiating experimental work utilizing this procedure.

The PI/RI is also responsible for training and reviewing this procedure with all personnel who will be performing the validated inactivation procedure.

4. Procedure

A. Preparation of Bacterial Cultures for SPME fiber head-space sampling of VOCs

i. Using a sterile, disposable loop, transfer a loopful of frozen glycerol stock of the bacterium onto an appropriate agar medium.

ii. Streak for isolation and incubate (static incubator) at 37°C for 24-28 hours until viable growth of isolated colonies is observed on the plate.

iii. Inoculate 10 ml of broth media with 1-5 colonies of bacteria and incubate with shaking (170 rpm) overnight at 37°C.

iv. Subculture bacteria into 25-50 mL of broth media based on desired optical density (OD) in disposable 250 mL Erlenmeyer flasks with vented caps.

1. Control Flasks:

- Empty Flask (no media)
- Media Only Flask (no bacteria)

v. Incubate flasks with shaking (170 rpm) at 37°C for various time points (varies by experiment). Maximum bacteria concentrations estimated to not exceed 10^9 colony-forming units (CFUs)/mL.

B. Head Space Sampling using SPME fibers

i. At each specific time point, remove the culture flasks from the incubator and transfer to the biological safety cabinet (BSC).

1. Allow the culture flasks to sit for a minimum of 30 minutes prior to head space sampling to allow for any potential aerosols generated during the incubation phase to settle.

ii. After 30 minutes, insert the SPME fiber (Supelco Portable Field Sampler, Product #57359-U) through the vent in the cap into the head space above the liquid bacterial culture (see Protocol Figure A.3.P.1). Sampling times vary by experiment. Distance between the fiber tip and liquid culture is approximately 3 inches, minimizing the risk of bacteria transfer.

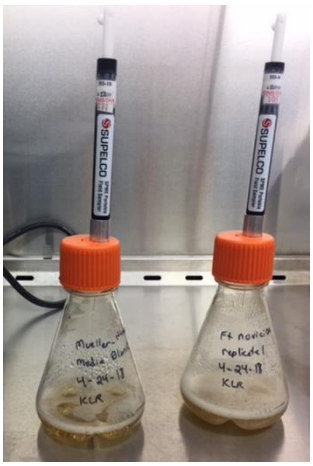


Figure A.3.P.1 SPME fiber sampling of bacterial culture headspace

iii. Once head space sampling is complete, retract the SPME fiber back into the sampling device.

1. Surface decontaminate the exterior of the sampling device prior to removing the device from the BSC.
2. Place the device in a secondary container for transport out of the laboratory.
    - Surface decontaminate the exterior of the secondary container prior to removal from the laboratory.

Note: Refer to Protocol Appendix A.3.P.B for procedures and data used to validate that this head space sampling procedure does not result in contamination of the SPME fiber with viable agent.

5. Change Control

Revisions to this protocol will be made as necessary. This procedure will be reviewed, and revised as necessary, by the PI/RI, LLNL Responsible Official (RO) and BSO, at least annually or after any change in PI/RI, after any change in the validated procedure and after any failure of

the procedure. The PI/RI is responsible for communicating revisions to the appropriate research staff and training the staff on the changes to the procedure.

6. Responsible Individual

The PI/RI who authored the procedure is responsible for change control.

Protocol Appendix A.3.P.A. Training Record

All research personnel who will be performing this procedure are required to read, understand and review this procedure with the appropriate PI/RI who developed the procedure and must sign this Training Record prior to performing this procedure and again after any revision to this procedure.

The PI/RI must submit an electronic copy of the completed training record(s) to the RO who will maintain file copies in the Laboratory Select Agent Program Folder on UCM.

I have read, understood and agree to fully comply with and adhere to the exact procedure, as written in SOP #Rasley-2018-001, Revision 2, *Solid-Phase Microextraction (SPME) Sampling for Volatile Organic Compounds (VOCs) of Liquid Bacterial Cultures* .

Name (print)	**Signature**	**Date**

Protocol Appendix A.3.P.B. Method Validation

To demonstrate that SPME fibers are not contaminated with viable agent during headspace sampling of liquid bacterial cultures, viability testing was performed on the SPME fibers post-headspace sampling. Viability testing was performed one time using 3 biological replicates for each organism.

A. **Viability Testing Method**

1. SPME fibers exposed to sampling headspace above liquid bacterial cultures were submerged and agitated in 1 mL of an appropriate growth media for 30 seconds (see Protocol Figure A.3.P.2) and then removed.
    - Brain Heart Infusion (BHI) – *B. anthracis* Sterne
    - Modified Mueller Hinton – *F. tularensis* subsp. novicida
2. Media exposed to SPME fibers was incubated (static) for 48 hours at 37°C ±2°C.
3. Controls:
    i. Broth only (no bacteria) – negative control
    ii. 100 µL of inoculum from Replicate #1 in 1 mL media – positive control
4. After 48 hours, media was observed for growth (turbidity) and results were recorded.
5. 100% of the media exposed to the SPME fiber was then transferred to the appropriate agar plates (100 µL x 10 plates).
6. Agar plates were incubated (static) for a minimum of 48 hours at 37°C ±2°C.
7. After 48 hours, plates were observed for growth.

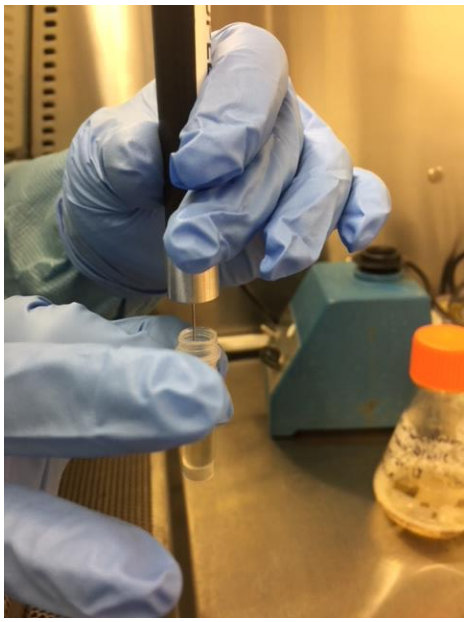


Figure A.3.P.2 Submerging SPME fiber exposed to headspace above liquid bacterial culture for viability testing.

B. **Results:**

No growth was observed in either the liquid culture exposed to the SPME fibers used to sample the headspace of the tested bacterial cultures for VOCs or when 100% of the exposed media was plated onto agar plates. The negative control (media only, no bacteria) also showed no growth, whereas the positive control (inoculum from replicate #1) showed expected growth indicating that the media uses was capable of supporting the growth of the tested bacteria. These results demonstrate that there is no potential for contamination of the SPME fibers when exposed to the head sampling space of a bacterial liquid culture when following this protocol, as written.

REFERENCES

REFERENCES

1 Reese, K. L., Avila, J. R., Rasley, A., Jones, A. D. & Frank, M. Metabolic profiling of volatile organic compounds (VOCs) emitted by the pathogens *Francisella tularensis* and *Bacillus anthracis* in liquid culture. *Sci. Rep.*, doi:10.1038/s41598-020-66136-0 (2020).

2 Amann, A. *et al.* The human volatilome: volatile organic compounds (VOCs) in exhaled breath, skin emanations, urine, feces and saliva. *J. Breath Res.* **8**, 034001, doi:10.1088/1752-7155/8/3/034001 (2014).

3 Lawal, O., Ahmed, W. M., Nijsen, T. M. E., Goodacre, R. & Fowler, S. J. Exhaled breath analysis: a review of 'breath-taking' methods for off-line analysis. *Metabolomics* **13**, 110, doi:10.1007/s11306-017-1241-8 (2017).

4 Bruderer, T. *et al.* On-line analysis of exhaled breath focus review. *Chem. Rev.* **119**, 10803-10828, doi:10.1021/acs.chemrev.9b00005 (2019).

5 Das, S. & Pal, M. Review—Non-invasive monitoring of human health by exhaled breath analysis: A comprehensive review. *J. Electrochem. Soc.* **167**, 037562, doi:10.1149/1945-7111/ab67a6 (2020).

6 Pleil, J. D. Cellular respiration: replicating in vivo systems biology for in vitro exploration of human exposome, microbiome, and disease pathogenesis biomarkers. *J. Breath Res.* **10**, 010201, doi:10.1088/1752-7155/10/1/010201 (2016).

7 Winters, B. R. *et al.* Review: Endogenously produced volatiles for in vitro toxicity testing using cell lines. *Appl. In Vitro Toxicol.* **4**, 129-138, doi:10.1089/aivt.2017.0038 (2018).

8 Audrain, B., Farag, M. A., Ryu, C. M. & Ghigo, J. M. Role of bacterial volatile compounds in bacterial biology. *FEMS Microbiol. Rev.* **39**, 222-233, doi:10.1093/femsre/fuu013 (2015).

9 Filipiak, W. *et al.* Molecular analysis of volatile metabolites released specifically by staphylococcus aureus and pseudomonas aeruginosa. *BMC Microbiol.* **12**, 113, doi:10.1186/1471-2180-12-113 (2012).

10 Rees, C. A., Smolinska, A. & Hill, J. E. Volatile metabolic diversity of Klebsiella pneumoniae in nutrient-replete conditions. *Metabolomics* **13** (2017).

11 Rees, C. A., Burklund, A., Stefanuto, P.-H., Schwartzman, J. D. & Hill, J. E. Comprehensive volatile metabolic fingerprinting of bacterial and fungal pathogen groups. *J. Breath Res.* **12**, 026001 (2018).

12 Chen, J., Tang, J., Shi, H., Tang, C. & 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**, 228-237, doi:10.1002/jobm.201600505 (2017).

13 Horsmon, J. R. & Crouse, K. Characterization of volatile organic compound profiles of bacterial threat agents. *US Army Edgewood Chemical Biological Center* (2008).

14 Lonsdale, C. L. *et al.* The use of colorimetric sensor arrays to discriminate between pathogenic bacteria. *PLOS ONE* **8**, e62726, doi:10.1371/journal.pone.0062726 (2013).

15 U.S. Centers for Disease Control and Prevention. *Select Agents and Toxins List*, <https://www.selectagents.gov/selectagentsandtoxinslist.html> (2018).

16 McLendon, M. K., Apicella, M. A. & Allen, L.-A. H. Francisella tularensis: taxonomy, genetics, and Immunopathogenesis of a potential agent of biowarfare. *Annu. Rev. Microbiol.* **60**, 167-185, doi:10.1146/annurev.micro.60.080805.142126 (2006).

17 Goel, A. K. Anthrax: A disease of biowarfare and public health importance. *World J. Clin. Cases.* **3**, 20-33, doi:10.12998/wjcc.v3.i1.20 (2015).

18 Hendricker, A. D., Abbas-Hawks, C., Basile, F., Voorhees, K. J. & Hadfield, T. L. Rapid chemotaxonomy of pathogenic bacteria using in situ thermal hydrolysis and methylation as a sample preparation step coupled with a field–portable membrane-inlet quadrupole ion trap mass spectrometer. *Int. J. Mass Spectrom.* **190-191**, 331-342, doi:10.1016/S1387-3806(98)14266-5 (1999).

19 Basile, F., Beverly, M. B., Voorhees, K. J. & Hadfield, T. L. Pathogenic bacteria: their detection and differentiation by rapid lipid profiling with pyrolysis mass spectrometry. *Trends Anal. Chem.* **17**, 95-109, doi:10.1016/S0165-9936(97)00103-9 (1998).

20 Li, Y. *et al.* Differentiation of bacteria using fatty acid profiles from gas chromatography–tandem mass spectrometry. *J. Sci. Food Agric.* **90**, 1380-1383, doi:doi:10.1002/jsfa.3931 (2010).

21 *Biosafety in Microbiological and Biomedical Laboratories*. 5th edn, (U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institutes of Health, HHS Publication No. (CDC) 21-1112 Revised December 2009; https://www.cdc.gov/labs/BMBL.html).

22 WHO. *Laboratory Biosafety Manual*. 3rd edn, (World Health Organization, 2004; https://www.who.int/csr/resources/publications/biosafety/Biosafety7.pdf).

23 El-Etr, S. H. *et al.* Francisella tularensis Type A strains cause the rapid encystment of Acanthamoeba castellanii and survive in amoebal cysts for three weeks postinfection. *Appl. Environ. Microbiol.* **75**, 7488-7500, doi:10.1128/AEM.01829-09 (2009).

24 Reese, K. L. *et al.* Chemical profiling of volatile organic compounds in the headspace of algal cultures as early biomarkers of algal pond crashes. *Sci. Rep.* **9**, 13866, doi:10.1038/s41598-019-50125-z (2019).

25 Sumner, L. W. *et al.* Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics.* **3**, 211-221, doi:10.1007/s11306-007-0082-2 (2007).

26 Forney, F. W. & Markovetz, A. J. The biology of methyl ketones. *J. Lipid Res.* **12**, 383-395 (1971).

27 Schulz, S. & Dickschat, J. S. Bacterial volatiles: the smell of small organisms. *Nat. Prod. Rep.* **24**, 814-842, doi:10.1039/b507392h (2007).

28 Qiu, X. *et al.* Crystal structure and substrate specificity of the β-ketoacyl-acyl carrier protein synthase III (FabH) from Staphylococcus aureus. *Protein Sci.* **14**, 2087-2094, doi:10.1110/ps.051501605 (2005).

29 Schmidt, R., Cordovez, V., de Boer, W., Raaijmakers, J. & Garbeva, P. Volatile affairs in microbial interactions. *ISME J.* **9**, 2329-2335, doi:10.1038/ismej.2015.42 (2015).

30 Elgaali, H. *et al.* 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**, 373-380, doi:10.1002/1521-4028(200212)42:6<373::AID-JOBM373>3.0.CO;2-4 (2002).

31 Togashi, N. *et al.* Antibacterial activity of long-chain fatty alcohols against Staphylococcus aureus. *Molecules.* **12**, 139 (2007).

32 Hwang, H.-I., Hartman, T. G., Rosen, R. T., Lech, J. & Ho, C.-T. Formation of pyrazines from the Maillard reaction of glucose and lysine-.alpha.-amine-15N. *J. Agric. Food Chem.* **42**, 1000-1004, doi:10.1021/jf00040a031 (1994).

33 Brana, A. F. *et al.* Activation and silencing of secondary metabolites in Streptomyces albus and Streptomyces lividans after transformation with cosmids containing the thienamycin gene cluster from Streptomyces cattleya. *Arch. Microbiol.* **196**, 345-355, doi:10.1007/s00203-014-0977-z (2014).

34 Larroche, C., Besson, I. & Gros, J.-B. High pyrazine production by Bacillus subtilis in solid substrate fermentation on ground soybeans. *Process Biochem.* **34**, 667-674, doi:10.1016/S0032-9592(98)00141-1 (1999).

35 Turnbull, P. C. Definitive identification of Bacillus anthracis--a review. *J. Appl. Microbiol.* **87**, 237-240, doi:10.1046/j.1365-2672.1999.00876.x (1999).

36 Rohmer, L. *et al.* Comparison of Francisella tularensis genomes reveals evolutionary events associated with the emergence of human pathogenic strains. *Genome Biol.* **8**, R102, doi:10.1186/gb-2007-8-6-r102 (2007).

37 Pleil, J. D. & Williams, A. Centralized resource for chemicals from the human volatilome in an interactive open-sourced database. *J. Breath Res.* **13**, 040201, doi:10.1088/1752-7163/ab2fa2 (2019).

38 U.S. Environmental Protection Agency. Volatilome: Human breath. < https://comptox.epa.gov/dashboard/chemical_lists/VOLATILOME> (2019).

CHAPTER 4. METABOLIC PROFILING OF VOCs IN THE HEADSPACE OF ALGAL CULTURES AS EARLY BIOMARKERS OF ALGAL POND CRASHES

FOREWORD

The material presented in this chapter has been adapted from work first published in 2019 in the journal *Scientific Reports*[1]. Contributions from others to conduct the experiments described in this chapter are as follows: C. L. Fisher, P. D. Lane, and J. D. Jaryenneh were integral in setting up and maintaining the experimental cultures and collecting algal density data.

4.1 INTRODUCTION

As the energy needs of the world increase, dependence on non-renewable sources of energy remains a concern. Increased production of corn starch or sugarcane-based ethanol has resulted in increased atmospheric carbon dioxide levels, diversion of arable land from food production, and increased consumer cost for sugar and corn[2]. For these reasons, microalgae production systems are considered a promising avenue for biofuel production. Microalgal strains are capable of growth in a range of environments (e.g., freshwater, marine, hypersaline, highly acidic) including high-nutrient municipal wastewater systems[3], allowing for simultaneous bioremediation and biofuel production. Microalgae's ability for rapid growth in non-potable (brackish or marine) water sources using non-arable land, combined with their high capacity for fixation of atmospheric carbon dioxide and high lipid-to-biomass ratios are significant advantages toward its use as a biofuel feedstock. The development of optimized systems for sustainable and dependable biofuel production through algal pond systems are necessary as global energy strategies continue to evolve (for review, see Kayitar 2017[4]).

A major challenge faced in algal production is ‘pond crashes’, which are devastating, often unpredictable losses of entire crops due to parasitism, grazing, weather, or many other factors. Closed photobioreactors are less likely to be susceptible to contamination with deleterious species (e.g., viruses, fungi, protozoans, detrimental microbes), but involve higher capital costs[5] and, once contaminated, can be challenging to disinfect. Open algal ponds are less expensive to set up but are more likely to succumb to crashes (Figure 4.1A) caused by grazing or parasitism[6,7]. Notably, a single adult marine rotifer, *Brachionus plicatilis* (Figure 4.1B), can consume 200 microalgal cells per minute and double in population within 1-2 days[8]. It has been estimated that pond crashes account for 30 % loss of annualized algal production[5] and represent a significant economic barrier to biofuel production[9].

However, algae are currently more costly to produce per unit due to insufficient rates of biomass production and harvesting costs [10]. Basic research that drives increases in algal cell culture density, and thus total lipid accumulation, is an initial step to address these issues. To produce higher yields of algae crops, research is needed to (1) have better early-warning tools to anticipate and/or diagnose the presence of predators and (2) understand algae-bacteria interactions, as they frequently happen in nature. The study of potential VOC biomarkers for (1) is covered in this chapter and for (2) in chapter 5.

Current pond crash mitigation strategies, both prophylactic and interdictive, are largely focused on chemical treatments, such as hypochlorite [11], copper [12,13], quinine sulfate [14], rotenone [15], additives that lower pH (to less than 3.0)[16], and biocides, such as tossendanin [17]. The use of chemical countermeasures as a prophylactic strategy to prevent pond crashes is prohibitively expensive for most algal industry business models. Chemical additives can degrade

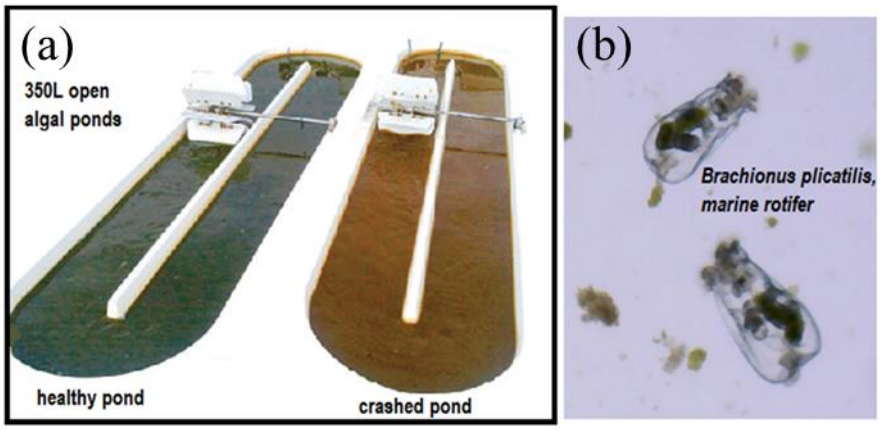


Figure 4.1 (a) Adapted from McBride et al, showing 350 L open algal production ponds with a healthy algal pond on the left compared with a crashed algal pond on the right. (b) *Brachionus plicatilis* (average length 160 µm), marine rotifer, in a field of algae, *Microchloropsis salina*.

from sun exposure and may need to be repeatedly added to cultures to maintain protection. In addition, frequent and repeated chemical application can be environmentally detrimental through the development of resistant pest species or through unacceptable off-target effects (e.g., pisicidal effects of rotenone). However, when applied early and in a targeted fashion after the detection of a deleterious species, chemical additives can be highly effective at saving algal cultures [for review of crop protection strategies, see Fisher et al. 2019 [18]].

In order to attain the production levels of 25 g m^{-2} d^{-1} ash free dry weight (or 2,500 gallons of biofuel per acre per year)[19] deemed necessary for economic algal biofuel production, cost-effective pond monitoring strategies are necessary to reduce culture loss and increase annualized production. Currently, algal production facilities utilize light microscopy to identify contaminants, pathogens, and competing algal strains that could lead to the demise of the desired algal strain[20]. However, microscopy is slow, labor-intensive, and requires advanced operator training for differentiating various microbiota. Alternative methods involving automated and semi-automated technologies, such as FlowCAM imaging flow-cytometry[21,22],

polymerase chain reaction, and hybridization-based assays[23,24] are under development to increase sensitivity and expedite analysis for daily algal culture monitoring.

All of the strategies described above require sampling of liquid from the algal culture, but an alternative approach involves sampling volatile molecular indicators from the air above algal cultures. Volatile organic compounds (VOCs) are carbon-containing molecules with high vapor pressures at ambient temperatures[25] that often occur in rich mixtures. Within the field of chemical ecology, VOCs have been identified as secondary metabolites and include, but are not limited to, pheromones, semiochemicals, odorants, and phytohormones[26,27] . Algal VOC production has been associated with intra- and inter-species communication, allelopathy, semiochemical production, and predator deterrence [for review, see Zuo 2019 [28]]. A well characterized example of an algal volatile involves conversion of nonvolatile dimethylsulfoniopropionate (DMSP) to volatile dimethyl sulfide (DMS). In intact *Emiliania huxleyi* cells, conversion of DMSP to DMS by the enzyme DMSP lyase is minimal, theorized to be segregated within the intact cell. However, during algal grazing, such as by the dinoflagellate *Oxyrrhis marina*, *E. huxleyi* is disrupted, releasing DMSP. Once released, conversion of DMSP to DMS[29,30] is catalyzed by DMSP lyases including those from bacteria. DMS acts not only as a deterrent against herbivory by *Oxyrrhis marina[31],* but additionally as an attractant for other species such as birds and fish[32].

The aims of this study were to 1) develop a methodology to detect VOCs from healthy algal cultures (*Microchloropsis salina)* as well as algal cultures in the presence of a grazer (*M. salina* cultures with marine rotifer *Brachionus plicatilis)* and 2) evaluate whether specific VOCs could serve as early indicators of an imminent culture crash. A setup based upon solid-phase microextraction (SPME) fibers coupled with gas chromatography-mass spectrometry (GC-MS)

allowed for non-invasive monitoring of volatile emissions. Compounds present during the active grazing period of rotifers on algal cultures, but not produced in healthy controls, were deemed potential biomarkers of high stress conditions. We propose that these biomarker compounds are potential diagnostic tools for chemical monitoring systems in algae cultivation systems to enabling the early detection of culture stress for improved algal crop production.

4.2 METHODS

The experimental setup (Figure 4.2) facilitated headspace volatile monitoring of *M. salina* with and without the algal grazer and rotifer, *B. plicatilis*. In total, three replicate experiments were performed using this setup, labeled Experiment 1, Experiment 2, and Experiment 3. A total of 6 culture vessels ("carboys") were utilized in each experiment, duplicates of every culture type. Algal cell concentrations and VOC headspace samples were collected at various time points for *M. salina* alone (abbreviated **Algae** or **A)**, *M. salina* and *B. plicatilis* (abbreviated **Algae + Rotifer** or **A+R**), and ESAW media blanks (abbreviated **Media Blank** or **MB**).

4.2.1 Axenic algae cultures

Microchloropsis salina (CCMP 1776) was obtained as an axenic stock culture (as determined by the supplier) from the National Center for Marine Algae and Microbiota (NCMA at Bigelow Laboratory, ME, USA). *M. salina* cultures were grown as previously described in Fisher et al, 2019[33]. For volatilomics experiments, Enriched Seawater, Artificial Water (ESAW) media was modified to contain 7.5 mM $NaNO_3$ and 0.5 mM Na_3PO_4. Cultures were

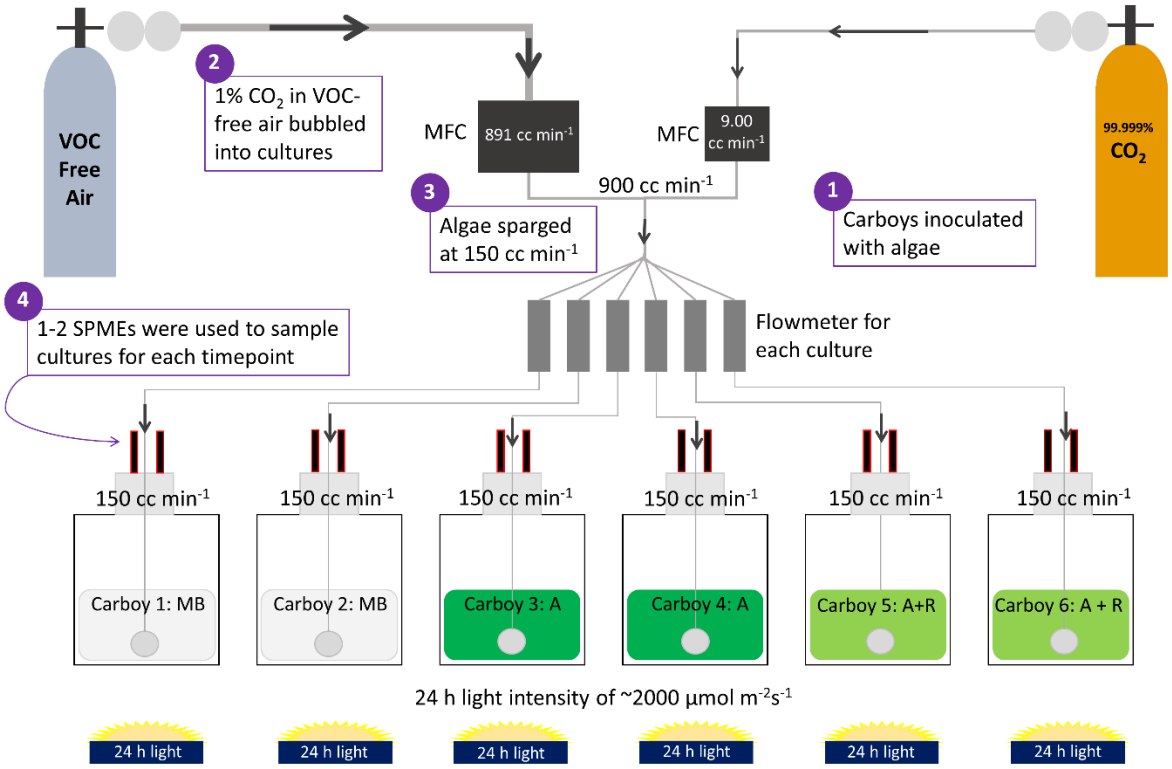


Figure 4.2 Schematic of experimental setup for growth *M. salina* (Algae, A) in the presence of *B. plicatilis* (Rotifer, R) for 5 days. Mass flow controllers (MFCs) mixed 1% CO2 with VOC-free air to sparge 15 L cultures at 150 cc min-1. In total, three replicate experiments (Experiment 1, Experiment 2, Experiment 3) were performed using this setup. SPME fibers (Experiment 1 used 1 fiber; Experiment 2,3 used 2 fibers) were used to sample the headspace of media blank (MB), Algae only (A), and Algae + Rotifer (A+R) carboys for 30-60 min each at various timepoints over 2-4 days.

grown in 15L of media in 20-L polycarbonate carboys at room temperature (RT) of ~ 22°C with 24-h light intensity of ~2000 µmol $m^{-2}s^{-1}$ for 5 days (d) Carbon dioxide gas, research purity 99.999 % (Matheson Tri-Gas, NJ, USA), and research grade air (70:30 N_2/O_2), VOC Free (Matheson Tri-Gas, NJ, USA), were supplied to all carboys via two mass flow controllers (one for CO_2 and one for air). The two mass flow controllers (Alicat, AZ, USA) were set to deliver 1% CO_2 (9.00 cc min^{-1}) and 99 % air (891 cc min^{-1}), for a total mass flow of 900 cc min^{-1} split

equally across six carboys(150 cc min^{-1} sparging rate for each sample through an air stone bubbler).

4.2.2 Xenic marine rotifers

Lots of 10-15 x 10^6 live, xenic, marine rotifers, *Brachionus plicatilis*, were obtained from Reed Mariculture, CA, USA 1-2 days before each inoculation and were shipped overnight on ice. Upon arrival, *B. plicatilis* were kept at 4 °C until concentrated and inoculated into algal cultures for each experiment. On the day of inoculation, *B. plicatilis* were allowed to warm for 1-3 h to room temperature (22 °C), gently concentrated using a 53 µm screen filter (Florida Aqua Farms, FL, USA) down to 100 mL of culture, and rinsed twice with 200 mL of ESAW media. Rotifers were enumerated by direct counting using a Rafter counter.

4.2.3 Preparation of cultures

Experimental cultures were grown in in 20-L polycarbonate carboys (ThermoFisher Scientific, MA USA) containing 15 L nutrient enriched ESAW medium. At the start of an experiment, 4 of the 6 carboys were inoculated with *M. salina* culture to a concentration of 4 - 5 x 10^6 cells mL^{-1} in 15 L. After 48 hours (h) of *M. salina* growth and acclimation to culturing conditions in the four carboys, 1.32 x 10^6 live rotifers (final concentration of 88 rotifers mL^{-1}) were added to two of the four algal cultures.

4.2.4 Monitoring algae growth and rotifer cultures

Algal culture density was determined daily by chlorophyll fluorescence (430 nm excitation, 685 nm emission) using a Tecan i-control infinite 200 PRO, version 1.11.1.0 (Tecan

Group, Switzerland). Monitoring growth of algal cultures was derived by calculation via a standard curve correlating chlorophyll fluorescence with algal density. Duplicate fluorescence measurements for each sample were averaged for each timepoint and then normalized to the final concentration measurements for the *M. salina* control in the absence of rotifers. Health and viability of rotifers within algal cultures was monitored daily via light microscopy. Significant differences between means of healthy or infected algal cultures were compared using two-way ANOVAs with Tukey's HSD test.

4.2.5 SPME headspace sampling and GC-MS data acquisition

VOCs were sampled from the headspaces of each culture and media control vessel using portable field sampler SPME fibers, with 65 µm polydimethylsiloxane/divinyl-benzene (PDMS/DVB) coatings (Supelco, Bellefonte, PA). One fiber was used per carboy in Experiment 1, totaling 6 measurements per timepoint, while two fibers were used per carboy in Experiments 2 and 3, totaling 12 replicate measurements for each time point. For this work, we required a "portable field sampler" fiber design to facilitate sample collection, transport, and storage over the time course of the experiments. The bi-phasic coating (one of three commercially available field-portable options) was chosen for sampling a wide range of compounds, including polar analytes, semi-volatiles, and larger weight volatiles. SPME samples were obtained within 1-2 h of the fluorescence measurements that were used to determine algae concentrations. SPME exposure times were shorter for Experiment 1 (30 min) compared to Experiments 2 and 3 (60 min). SPME fibers were stored in sealed containers at 2-4 °C after sampling. Unexposed SPME fibers served as "travel blanks" to account for potential extraneous volatiles arising from storage conditions. Samples were analyzed by GC-MS within 2 weeks of collection.

An untargeted GC-MS approach was used to analyze the collected VOCs with an Agilent 5975T GC-MSD (Agilent Technologies, Santa Clara, CA) using an Agilent HP-5ms column (30 m x 250 µm x 0.25 µm) coupled to a single quadrupole mass analyzer with helium carrier gas at a constant flow rate of 1.2 mL/min. VOCs absorbed in the SPME fiber were desorbed in the heated GC inlet (280 °C) for 15 seconds using splitless injection. The column temperature was programmed, starting at 40 °C for 3 min, ramped at 5 °C/minute from 40 to 150 °C, ramped at 15 °C/min from 150 to 280 °C and then held for 2 min. The total run time was 35.67 min. Ions were generated using electron ionization (70 eV) and spectra were acquired at 4 scans/s over *m/z* 35-450. Data acquisition was performed under control of ChemStation software (Agilent Technologies, version E.02.02). A commercial reference of 18 standard compounds (S-22329; AccuStandard, New Haven, CT) was used to evaluate day-to-day performance and to calculate retention indices.

4.2.6 GC-MS data processing

After GC-MS data acquisition, data processing procedures and criteria were applied to detect and identify individual biomarkers in each condition. All ChemStation data files (consisting of biological duplicates, media controls, and unexposed fibers) were translated for compatibility with Agilent's MassHunter Software (MassHunter GC/MS Translator B.07.05). Chromatographic deconvolution and visualization were performed using MassHunter Qualitative (version B.07.00 SP2) using a Retention Time window size factor of 90.0, signal-to-noise ratio threshold of 2.00, absolute ion height filter of 1000 counts, and ≥ 5 ions required for compound detection (threshold of detection 5 x 10^3 counts per peak). An arbitrary small value of 1 was assigned to the signal value for compounds that were not detected.

Detected peaks were transferred into Mass Profiler Professional (MPP) 12.6.1 software and aligned across all samples in the data set using a retention time tolerance of 0.15 minutes, mass spectral match factor of 0.6 (of maximum 1.0), and a delta *m/z* tolerance of 0.2 Da. Putative identification of the aligned compounds was performed by searching spectra against the National Institute of Standards and Technology (NIST) mass spectral database, NIST14. Compounds with mass spectral matches ≥ 70% were subsequently annotated as the best match. Compounds that did not exceed the mass spectral match threshold were annotated using the base peak *m/z* and retention index (e.g., "Unknown *m/z* ##_RI ####").

Two criteria were used to identify volatile biomarkers unique to the Algae or Algae + Rotifer conditions: (1) detection of the biomarker in at least three of the four replicates at each sampled timepoint and (2) a) the biomarker was present in the Algae or Algae + Rotifer condition and not detected in the media blank or travel blank conditions; OR b) the biomarker was present in the Algae or Algae + Rotifer conditions at an abundance greater than 10x the abundance in the media blank or travel blank.

The peak areas of potential biomarkers passing the filter criteria were compared across the three performed experiments, with regards to both individual biomarkers and groups of biomarkers belonging to the same compound class. The presence or absence of these biomarkers in each experiment was determined, and the calculated peak areas were compared to algal density measurements.

4.3 RESULTS

4.3.1 Cell counts of infected and control cultures

Measured algae concentrations as a function of time for all three experiments is presented in Figure 4.3. At 48 h after inoculation, algal concentrations across all cultures were similar,

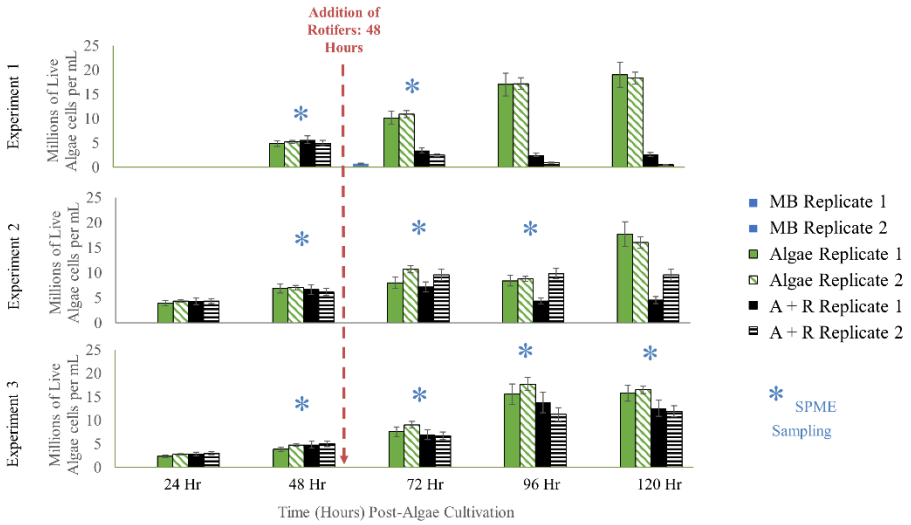


Figure 4.3 Algae concentration as determined by fluorescence measurements collected for three experiments. Similar coloring and patterns represent biological replicates of each condition: media blanks (MB), Algae (A), and Algae + Rotifer (A+R) cultures. Small fluorescence signals were observed in MB controls, most of which are not discernable on this scale. Error bars represent standard deviation derived from duplicate measurements for each sample. Significance levels for conditions that exhibited $p<0.05$ are in Appendix Table A.4.1. Blue asterisks (*) indicate the time points for headspace VOC sampling by SPME fibers.

approaching the mid to late stages of logarithmic growth. At this time, *B. plicatilis* were added to two of the four *M. salina* cultures, resulting in time-dependent decreases in algal density relative to the axenic cultures (Figure 4.3). Despite consistent growth conditions, 96 h after the initial cultures were started and 48 h after rotifers were added, the Algae + Rotifer cultures displayed different extents of algal biomass loss attributed to rotifer grazing (see Figure 4.3). This variation in rates of biomass loss may arise from differences in rotifer lots.

4.3.2 Headspace VOC results

Qualitative and quantitative differences were observed in the VOC profiles of Algae + Rotifer cultures compared to the Algae cultures. Example total ion chromatograms for Algae and Algae + Rotifer cultures taken approximately 24 h after addition of rotifers (Experiment #3) are

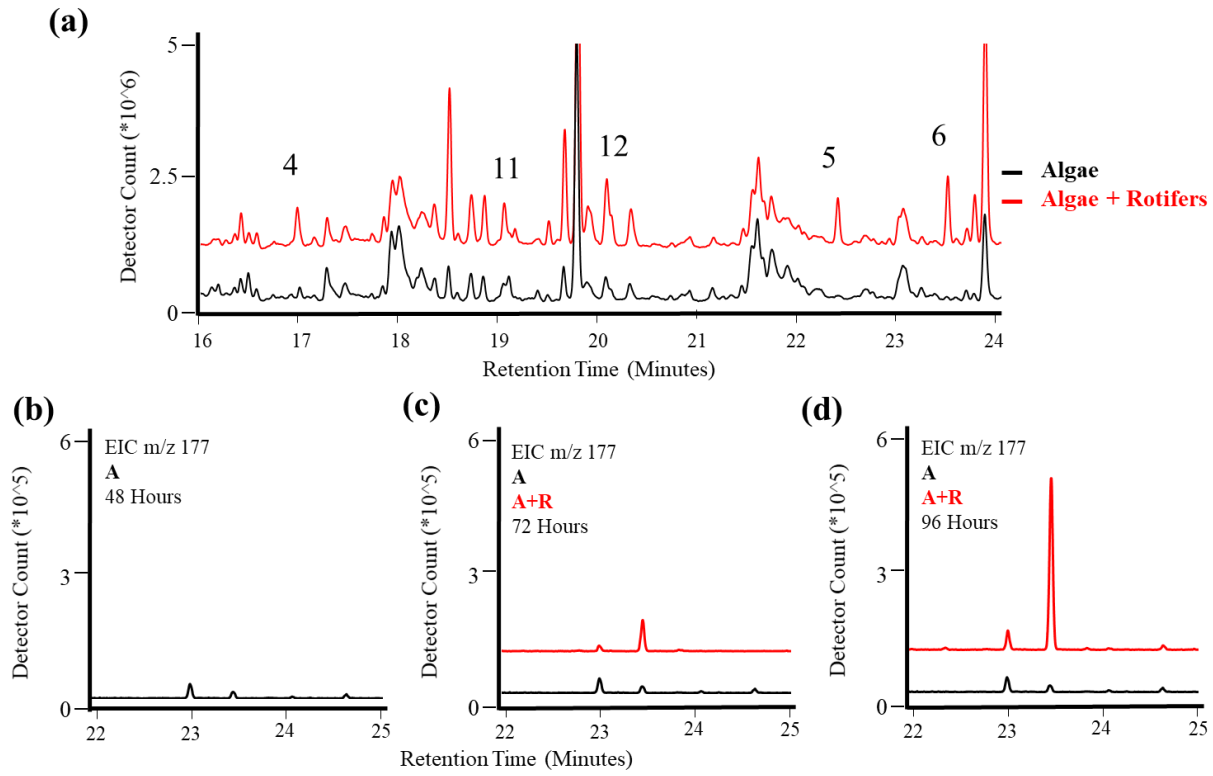


Figure 4.4 Example GC-MS chromatograms for observed VOCs sampled from Algae (A) and Algae + Rotifer (A+R) cultures between 16 and 24 min, (a) Total ion chromatogram with indicated VOCs (Annotations See Table 4.1), (b-d) extracted ion chromatograms monitoring increase in compound 6 over time (*m/z* 177, RT 23.46 min, RI 1495).

shown in Figure 4.4a. Several VOCs that differentiate the two culture conditions are indicated (Figure 4.4a, annotations in Table 4.1) and are potential early indicators of algal grazing or death. Extracted ion chromatograms were utilized to improve visualization of individual VOCs, as shown in Figures 4.4b-d. These demonstrate the increase in an VOC displaying a base peak *m/z* 177 and retention index 1495, observed over the time course of the experiment in Algae + Rotifer cultures. Although the Algae chromatogram for *m/z* 177 also displays a small peak at the same retention time, this VOC was not detected using the given filtering criteria for data processing.

Table 4.1 VOCs robustly and repeatedly detected from Algae and Algae + Rotifer experiments

	Compound #	Mass	Tentative Compound Class*	NIST14 ID	NIST % Match	Experimental Retention Index	Theoretical Retention Index	Experiment #		
								1	2	3
VOCs detected in A+R cultures	1	82	Carotenoid	2,2,6-trimethylcyclohexanone	79	1021	1036		X	X
	2	107				1181		X	X	X
	3	121	Phenol			1191			X	X
	4	137	Carotenoid	2,6,6-trimethyl-1-cyclohexene-1-carboxaldehyde	81	1209	1220	X	X	X
	5	121	Carotenoid	4-(2,6,6-trimethyl-1-cyclohexen-1-yl)-2-butanone	76	1419	1433	X	X	X
	6	177	Carotenoid	*trans*-beta-ionone	94	1495	1486	X	X	X
	7	57	Alkane			1691		X	X	X
VOCs detected in A+R and A cultures	8	71				1039			X	X
	9	96	Methyl Ester	3-Nonenoic acid, methyl ester	74	1134	1191	X	X	X
	10	341				1139			X	X
	11	71				1293			X	X
	12	138	Terpene/Carotenoid			1338			X	X
	13	73	Fatty acid (Hexadecanoic acid)			1983			X	X
	14	192				2197			X	X

*Tentative compound class for unknown compounds is based on fragmentation in averaged mass spectra determined via chromatographic deconvolution and alignment.

The number of compounds detected from deconvolution of chromatographic peaks varied with each sample. The analysis of a single sample typically detected 100-200 chemical compounds, many of which were attributed to background (present also in control measurements) or not found reproducibly. Application of chromatographic peak alignment across the data from all samples and at every timepoint generated a list of more than 1800 compounds, consisting of both algal VOCs and extraneous signals from the experimental setup. Application of the filtering criteria based upon algal abundance and detection frequency across experimental replicates identified the most robust compounds as potential VOC biomarkers from either Algae or Algae + Rotifer cultures, removed irreproducible compounds, and narrowed the extensive list to ~ 50 compounds in any single experiment. Table 4.1 shows only biomarkers that were observed across multiple experiments. For a detailed list of the volatile biomarkers detected in each individual experiment, refer to Appendix Table A.4.2.

Analysis of VOC data from Experiments 1-3 revealed several VOCs that were reproducibly observed in 1) Algae + Rotifer cultures and 2) both Algae and Algae + Rotifer cultures, as shown in Table 4.1, despite the different rates in algal biomass loss. For example, Compound 6 monitored in Figure 4.4B-D was identified with a 94% confidence score as *trans*-β-ionone using the NIST14 library. Confidence in this identification increases when considering the calculated experimental retention index (RI) of 1495 was within 5% of the literature theoretical value (1486, NIST 14 database). Within the Algae + Rotifer cultures, all of the discriminating VOCs were structurally-related ketones or aldehydes: (a) Compound 6: *trans*-β-ionone [IUPAC name: (*E*)-4-(2,6,6-trimethyl-1-cyclohexen-1-yl)-3-buten-2-one], (b) Compound 4: β-cyclocitral [IUPAC name: 2,6,6-trimethyl-1-cyclohexene-1- carboxaldehyde], (c) Compound 1: 2,2,6-trimethyl-cyclohexanone, and (d) Compound 5: 4-(2,6,6-trimethyl-1-

cyclohexene-1-yl)-2-butanone. Within the Algae and Algae + Rotifer cultures, 3-nonenoic acid methyl ester has an adequate confidence for identification (74% spectral match, < 5% RI deviation from theoretical value).

For those VOCs that could not be identified using the relatively conservative filtering criteria, the observed spectra and experimental retention indices provide suggestions for their identification. The suggested compound classes for unknown compounds are provided in Table 4.1, and their experimental mass spectra are included in Appendix Figure A.4.1. For example, the mass spectrum of Compound 12 has similar peaks and peak ratios to those of the identified carotenoids, suggesting a terpenoid structure with a molecular weight of 208 Da. The mass spectrum of Compound 13 contains ions characteristic of hexadecenoic acid including *m/z* 43, 60, 73, 129, 213, and 256 (M^+), and its experimental RI is within 1% of the literature RI of hexadecanoic acid (1968).

4.3.3 Abundance of VOCs

In addition to the qualitative analysis, the relationship between rotifer duration of grazing and the abundance of Algae + Rotifer distinguishing VOCs was examined. Levels of Compounds 4 and 6 were compared for individual Algae + Rotifer cultures (Carboys 5 and 6) across individual experiments (Figure 4.5). While there was no detected signal from either compound before the addition of rotifers (48 h), all detected signals after rotifer addition exceeded 2.0×10^5 counts, 2-3 orders of magnitude above detection threshold. Another carotenoid oxidation product, β-cyclocitral, appeared in Experiment 2 after 24 h of rotifer feeding, with the signal increasing to more than 6.0×10^5 counts after 48 h of rotifer feeding. Similar comparisons for Compounds 1-7 are included in Appendix Figure A.4.2. Of note,

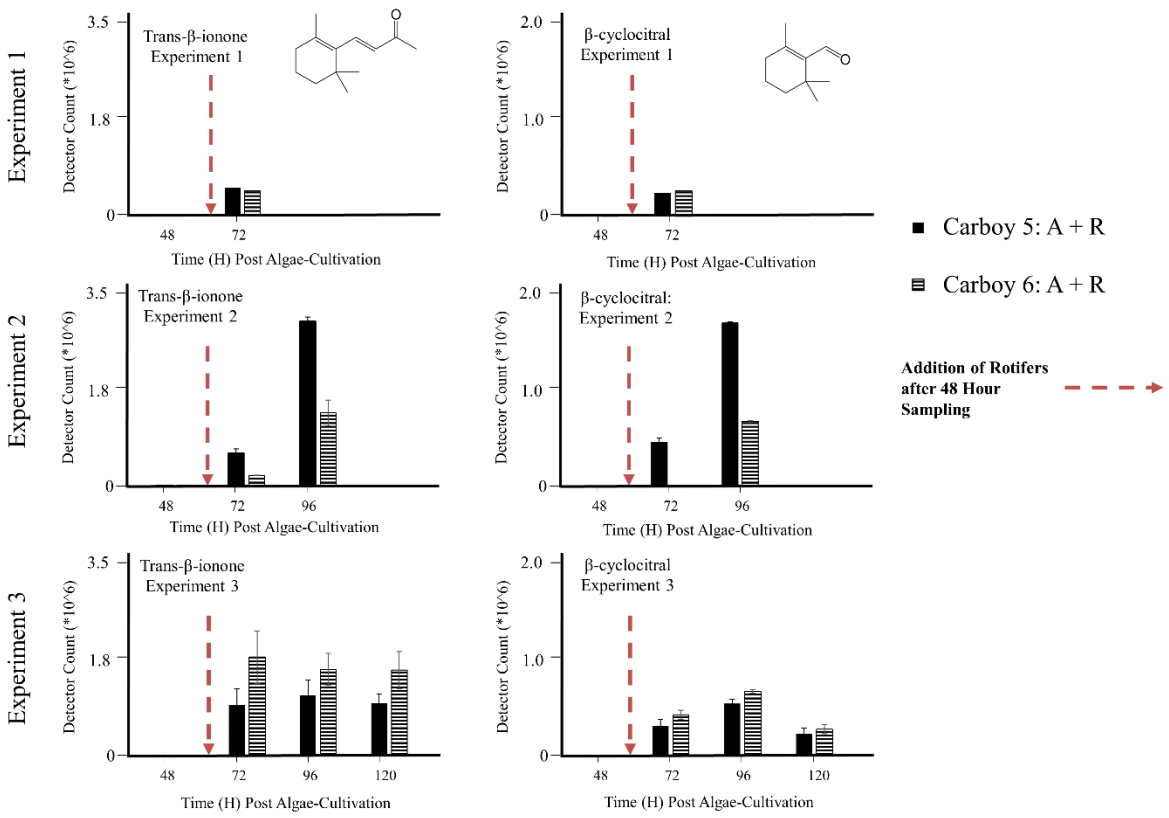


Figure 4.5 Peak areas of extracted compound chromatograms for trans-β-ionone and β-cyclocitral across Experiments 1, 2, and 3, separated by biological replicates. Error bars represent standard deviation derived from duplicate measurements for each sample. The exposure time for SPME fibers was 30 minutes in Experiment 1 and 60 minutes in Experiments 2 and 3.

comparison of Figure 4.5 to Figure 4.3 reveals several instances where these VOCs were detected in Algae + Rotifer cultures before biomass loss was apparent as compared to algae controls. For example, the second biological replicate of Algae + Rotifers in Experiment 2 did not differ in algal density from the healthy controls at the 72 h and 96 h timepoints. However, the signals for Compounds 4 and 6 were already large (6.0×10^5 and 1.0×10^6 counts, respectively).

4.4 DISCUSSION

This SPME-GC-MS analysis has identified seven discriminating VOCs in *M. salina* cultures in the presence of actively-grazing *B. plicatilis* (Algae + Rotifer). The absence of these volatiles in the time-matched Algae control cultures (Figure 4.3) suggests these chemicals are specific signals of active algal grazing, algal physiological stress, or algal death. Many of these chemicals were detected within 24 h after rotifer addition and before algal cell densities changed substantially relative to controls (Figure 4.2, 4.4). Specifically, Compounds 4, 5, 6, and 7 were identified as early and robust grazing signals observed in *M. salina* cultures containing rotifers. Several identified biomarkers – Compounds 1, 4, 5, 6 and 7 – were detected only during rotifer grazing and contained structural similarities, hinting at a shared metabolic pathway. Many of these compounds (Table 4.1) are known products of carotenoid oxidation[34,35]. Carotenoid-derived substances have been previously observed in algae volatile research, largely associated with investigations of flavor or smell components in food production. Carotenoids have important physiological functions as a component the light-harvesting complexes that transfer light energy to chlorophyll in photosystems[36]. Oxidative cleavage of the carotenoid backbone can occur through enzymatic (carotenoid cleavage dioxygenases) or non-enzymatic (light, oxygen, temperature) mechanisms[36]. Potential sites of oxidative cleavage of the carotenoid β-carotene are shown in Figure 4.6. In this work, the carotenoid-derived VOCs could be generated from the oxidation of β-carotene released upon lysis of *M. salina* cells during the digestive process of *B. plicatilis*. This would be in agreement with studies of *Arabidopsis* plants exposed to reactive oxygen species resulting in the subsequent release of β-ionone and β-cyclocitral[37]. The results from vascular plants suggest that carotenoid degradation products may be more general indicators of stressed or wounded algae cultures, not solely limited to the interaction of algae with rotifers.

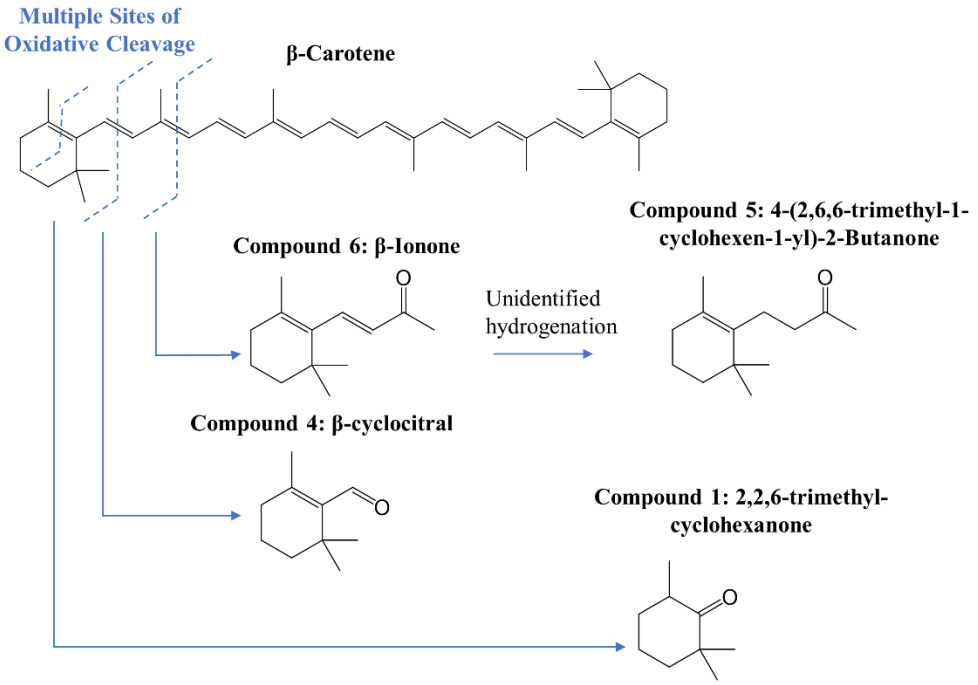


Figure 4.6 VOCs identified from the headspace of Algae + Rotifer cultures formed from the oxidative cleavage of the carotenoid β-carotene. Only those oxidation cleavages relevant to this study are pictured, but all double bonds across the β-carotene backbone are cleaved.

Although there are no reports of such analyses of *M. salina*, a small number of studies have examined algae from the genus *Nannochloropsis*, of which *M. salina* is a close relative[38]. Van Durme *et al* [39] investigated the volatile composition of five microalgae species (*Botryococcus braunii*, *Rhodomonas*, *Tetraselmis* sp., *Nannochloropsis oculata* and *Chlorella vulgaris*) by heating samples (40 °C) to enhance volatile signatures under heat stress conditions. Products of carotenoid oxidation, including α- or β- ionone and β-cyclocitral, were identified in all species tested. Interestingly, *N. oculata* contained a large abundance of ethanol, 2-hydroxy-2-butanone, and benzaldehyde, while only small amounts of β-cyclocitral and ionones were detected. Hosoglu [40] likewise characterized the volatilomes of several microalgae species using SPME-GC-MS and GC-olfactometry for both chemical profiles and olfactory properties to benefit incorporation into food products and to minimize unpleasant smells. The species *C.*

vulgaris, C. protothecoides, and *T. chuii* reportedly contain distinguishing amounts of the carotenoid degradation products α- and β-ionone and 6-methyl-5-hepten-2-one, while expressing a woody smell.

While VOCs have been observed in analyses of chemical compositions of algae in destructive manners (e.g., heating, sonication, solvent extraction etc.), there are fewer reports of volatiles emitted from live, actively-growing cultures. A variety of live algae-derived volatiles (terpenoids, aldehydes, halogenated compounds, etc.) have been shown to influence the odor quality of water[28]. Zhou et al [41] investigated changes in the volatilome of intact algae over different growth phases (logarithmic, stationary, and decline phase) for six microalgae (*Thalassiosira weissflogii*, *Nitzschia closterium*, *Chaetoceros calcitrans, Platymonas helgolandica*, *Nannochloropsis* spp. *(NMBluh014-1),* and *Dicrateria inornate)*. The *Nannochloropsis* volatilome was largely dominated by alkanes and alkenes and 8-heptadecene, but no carotenoid by-products were reported. Several functions for actively-released VOCs, including carotenoids, have been postulated, such as tolerance of light and oxidative stressors, signaling the presence of predators[36], and transfer of information throughout algal colonies [28]. β-cyclocitral has previously been reported as a volatile emitted by the bloom-forming cyanobacterium *Microcystis* as a defense mechanism against grazing by *Daphnia magna* [42]. Similarly, this work focused on the volatiles generated from active grazing of algae, which will generate more rapid algal death compared to natural growth cycles. Here, we confirmed the importance of β-cyclocitral as an indication of algal cell damage due to grazing. Future work will determine if VOCs produced by *M. salina* in the presence of *B. plicatilis* have similar roles in algal-defense as observed previously.

Using a non-invasive, non-destructive sampling and analysis technique, we have demonstrated that VOCs from the headspace of algae cultures can distinguish between algae cultures with grazing rotifers present and uninfected algal cultures, and may serve as general indicators of algal cell stress or death. It is worth noting that the list of biomarkers reported in this analysis may be considered conservative owing to the stringency of our filtering criteria. Additional biomarkers may be observed with more experimental repeats, more efficient VOC sampling and higher-sensitivity analyses. In order to discover and validate additional diagnostic biomarkers of grazer infection or other incipient crashes, more extensive studies of emitted volatiles from algae species are required. For example, low levels of grazer-associated VOCs in healthy algal cultures may result from background rates of algal death. Such background signals are likely modulated by physiological state of the culture (e.g., exponential growth or stationary phase) or by nutrient limitation. Thus, an improved understanding of the threshold biomarker concentrations that indicate the need for interdictive treatment is paramount.

Additionally, our SPME-GC-MS methodology is expected to have broader applications. Complex systems-level dynamics between algae, commensal bacteria, and various grazers will require more sophisticated sampling procedures alongside volatilomics data to include biological data sets, such as transcriptomics, metagenomics, and metabolomics. These systems-level analyses and bioinformatics analysis would be more likely to elucidate biological interactions or implications for the chemicals observed in the volatilome. Non-invasive and non-destructive VOC sampling is an attractive, analytical way to better understand and predict the health of microbial cultures.

4.5 CONCLUSIONS

The work presented in this chapter has aimed to increase the breadth and depth of reported algal and rotifer-specific VOCs, providing a tool to better define the physiological state of algae ponds and facilitate greater algal biomass production. A SPME-GC-MS methodology for non-invasive and non-destructive sampling of *M. salina* infected by *B. plicatilis* aided our discovery of seven putative culture crash biomarkers, including trans-β-ionone and β-cyclocitral, over several timepoints during active crashing of algal ponds. These biomarkers were not detected in cultures displaying natural background levels of cell death, suggesting that these signals are produced by high stress conditions, such as rotifer grazing. Finally, these biomarkers offer potential as diagnostic tools to signal the need for crash mitigation strategies, as several signals were detectable before cell death was evident from changes in cell density. Both VOC baselines and signatures from multiple healthy and infected cultures will be compiled in a data base. Early use of this technique would then include surveilling for the emergence of targeted VOC biomarkers of algal distress or injury above healthy baseline thresholds that are indicative of imminent culture failure.

The future of this work could see VOC based monitoring in open ponds. In such environments, one must take into account that the volatile "headspace" of open ponds may be influenced by external sources (e.g., VOCs from the environment, wind effects, particulates, etc.), creating a variable background that would require correction for the levels of biomarker compounds. The creation of temporarily closed "headspaces" above an area of open pond (perhaps using a large funnel), during sample collection, would serve to limit outside "background ".

The SPME fibers used in this experiment are field deployable and can easily be adapted to an algal pond production system. Although SPME-GC-MS has proven powerful for

untargeted discovery of algal volatile chemical signatures from healthy or grazed cultures, the cost of state-of-the art laboratory-based GC-MS systems and analyses efforts is prohibitive for using this method for continuous monitoring of industrial scale, open algal ponds. Knowledge gained and biomarkers annotated from our untargeted discovery efforts may guide development of targeted, lower-cost, field-deployable detectors capable of monitoring for changes in diagnostic chemical signatures and detecting volatile signals of infection in real-time to facilitate the timely deployment interdictive strategies to prevent pond crashes. Miniaturized GC-MS systems[43,44] for field deployable detector systems is one such technology currently under development and optimization[45] for this type of application.

APPENDICES

APPENDIX A: Tables

Table A.4.1 Significant difference determination between mean levels of algal cell densities across replicates of Algae (*M. salina*), Algae + Rotifer (*M. salina* and *B. plicatilis*) and Media Blank (MB, ESAW) calculated by ANOVA with Tukey's HSD test

	Experiment 1					Experiment 2					Experiment 3				
	24 H	48 H	72 H	96 H	120 H	24 H	48 H	72 H	96 H	120 H	24 H	48 H	72 H	96 H	120 H
MB 1 vs MB 2	-	>0.9999	0.9942	>0.9999	>0.9999	>0.9999	>0.9999	>0.9999	>0.9999	>0.9999	>0.9999	>0.9999	>0.9999	>0.9999	>0.9999
MB 1 vs Algae 1	-	0.0024	<0.0001	<0.0001	<0.0001	0.0202	<0.0001	<0.0001	<0.0001	<0.0001	0.4867	0.0696	<0.0001	<0.0001	<0.0001
MB 1 vs Algae 2	-	0.0011	<0.0001	<0.0001	<0.0001	0.0075	<0.0001	<0.0001	<0.0001	<0.0001	0.3291	0.0136	<0.0001	<0.0001	<0.0001
MB 1 vs Algae + Rotifer 1	-	0.0004	0.2087	0.2671	0.2335	0.0088	<0.0001	<0.0001	0.0079	0.0051	0.2848	0.0115	0.0001	<0.0001	<0.0001
MB 1 vs Algae + Rotifer 2	-	0.002	0.6522	0.9602	0.9978	0.0081	0.0001	<0.0001	<0.0001	<0.0001	0.2356	0.0086	0.0002	<0.0001	<0.0001
MB 2 vs Algae 1	-	0.0024	<0.0001	<0.0001	<0.0001	0.0202	<0.0001	<0.0001	<0.0001	<0.0001	0.4861	0.069	<0.0001	<0.0001	<0.0001
MB 2 vs Algae 2	-	0.0011	<0.0001	<0.0001	<0.0001	0.0075	<0.0001	<0.0001	<0.0001	<0.0001	0.3286	0.0135	<0.0001	<0.0001	<0.0001
MB 2 vs Algae + Rotifer 1	-	0.0004	0.0719	0.2662	0.2335	0.0088	<0.0001	<0.0001	0.008	0.0051	0.2844	0.0114	0.0001	<0.0001	<0.0001
MB 2 vs Algae + Rotifer 2	-	0.002	0.3351	0.9598	0.9978	0.0081	0.0001	<0.0001	<0.0001	<0.0001	0.2352	0.0085	0.0002	<0.0001	<0.0001
Algae 1 vs Algae 2	-	0.9996	0.9827	>0.9999	0.9898	0.9987	>0.9999	0.2066	0.9996	0.6884	0.9997	0.9814	0.8709	0.6052	0.9915
Algae 1 vs Algae + Rotifer 1	-	0.9818	<0.0001	<0.0001	<0.0001	0.9994	0.9999	0.9711	0.0148	<0.0001	0.999	0.9722	0.9966	0.7581	0.1778
Algae 1 vs Algae + Rotifer 2	-	>0.9999	<0.0001	<0.0001	<0.0001	0.9991	0.9868	0.7331	0.8134	<0.0001	0.9964	0.9503	0.9866	0.0302	0.064
Algae 2 vs Algae + Rotifer 1	-	0.9987	<0.0001	<0.0001	<0.0001	>0.9999	0.9992	0.0422	0.0069	<0.0001	>0.9999	>0.9999	0.6082	0.0622	0.0533
Algae 2 vs Algae + Rotifer 2	-	0.9999	<0.0001	<0.0001	<0.0001	>0.9999	0.9723	0.9268	0.9322	<0.0001	>0.9999	>0.9999	0.5053	0.0005	0.0163
Algae + Rotifer 1 vs Algae + Rotifer 2	-	0.9887	0.9615	0.7415	0.455	>0.9999	0.9984	0.2879	0.0006	0.0016	>0.9999	>0.9999	>0.9999	0.4138	0.9958

Timepoints are reported relative to the addition of algae to the growth media. Rotifers were added to each condition after the 48 hour timepoint.

"-" indicates no measurements were taken

Table A.4.2 List of VOCs in Individual Experiments

Experiment 1

	Compound Number	Mass	NIST ID	NIST % Match	Experimental Retention Index	Theoretical Retention Index
Algae + Rotifer Cultures	1	208			1074	
	2	152			1214	
	3	83	2-Butanone, 4-(2,6,6-trimethyl-1-cyclohexen-1-yl)-	70	1453	1433
	4	177	trans-.beta.-Ionone	87	1506	1486
	5	83	8-Heptadecene	75	1691	1719
Algae + Rotifer AND Algae Cultures	1	57			1134	
	2	148			1415	

Experiment 2

	Compound Number	Mass	NIST ID	NIST % Match	Experimental Retention Index	Theoretical Retention Index
Algae + Rotifer Cultures	1	59			987	
	2	82			1021	
	3	71			1101	
	4	107			1181	
	5	121	Phenol, 2,3,5-trimethyl-	71	1190	1235

Table A.4.2 (cont'd)

	6	137	1-Cyclohexene-1-carboxaldehyde, 2,6,6-trimethyl-	89	1209	1220
	7	121	2-Butanone, 4-(2,6,6-trimethyl-1-cyclohexen-1-yl)-	85	1443	1433
	8	177	trans-.beta.-Ionone	93	1495	1486
	9	57			1692	
	10	143			1881	
	11	135			2036	
	12	71			2115	
	1	55	Hexanoic acid, 2-ethyl-, methyl ester	68	1031	1043
	2	96	3-Nonenoic acid, methyl ester	73	1134	1191
	3	341			1138	
Algae + Rotifer AND Algae Cultures	4	71			1264	
	5	138			1338	
	6	73			1774	
	7	154			1854	
	8	73			1983	
	9	73			2079	
	10	192			2183	
	11	73			2210	
	1	71			757	
	2	56			1003	
	3	94			1207	
Algae Cultures	4	91			1530	
	5	109			1619	
	6	119	2,4-Diphenyl-4-methyl-1-pentene	83	1803	1846
	7	70			1955	
	8	70			2252	

Table A.4.2 (cont'd)

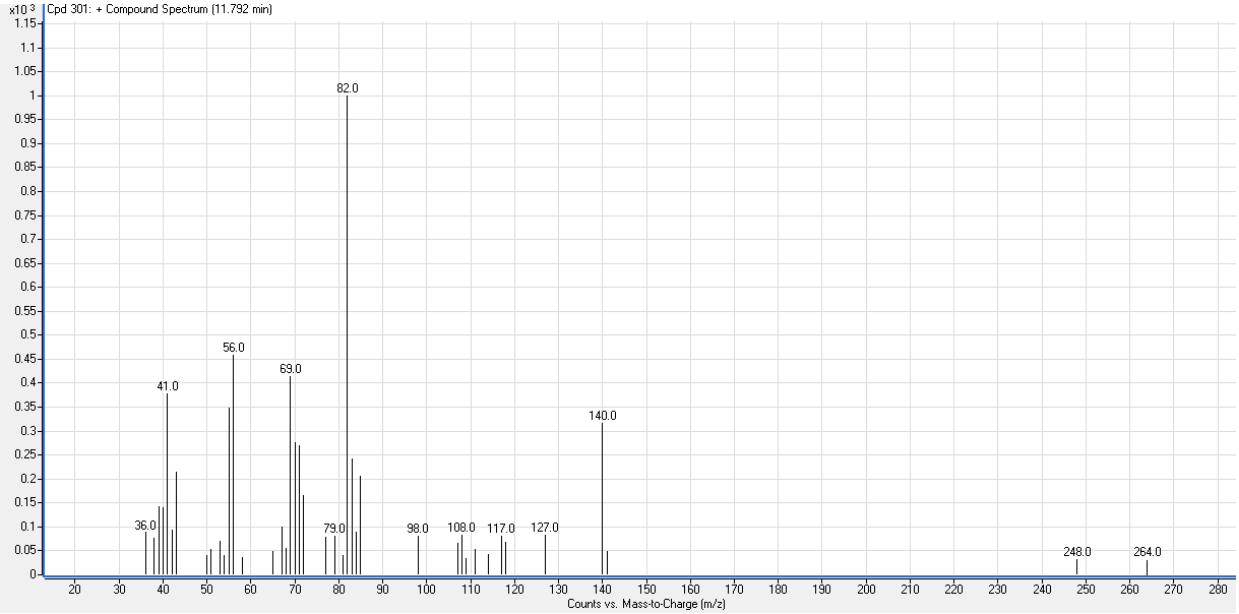
		Compound Number	Mass	NIST ID	NIST % Match	Experimental Retention Index	Theoretical Retention Index
Experiment 3	**Algae + Rotifer Cultures**	1	118			971	
		2	82	Cyclohexanone, 2,2,6-trimethyl-	79	1021	1036
		3	55			1054	
		4	107			1181	
		5	121			1191	
		6	137	1-Cyclohexene-1-carboxaldehyde, 2,6,6-trimethyl-	81	1209	1220
		7	122			1379	
		8	121	2-Butanone, 4-(2,6,6-trimethyl-1-cyclohexen-1-yl)-	76	1419	1433
		9	327			1429	
		10	401			1458	
		11	177	trans-.beta.-Ionone	94	1495	1486
		12	179			1522	
		13	158			1675	
		14	57			1691	
		15	172			1744	
		16	73			1744	
		17	150			1776	
		18	73			1907	
		19	251			2406	
		20	73			2422	
		21	149			2496	

Table A.4.2 (cont'd)

Algae + Rotifer AND Algae Cultures	1	57			978	
	2	71			1039	
	3	96	3-Nonenoic acid, methyl ester	74	1134	1191
	4	341			1139	
	5	71			1293	
	6	138			1338	
	7	71			1370	
	8	57			1507	
	9	73			1983	
	10	192			2112	
	11	192			2197	
	12	73			2345	

APPENDIX B: Figures

1a) Compound 1: Cyclohexanone, 2, 2, 6-trimethyl-



1b) Compound 2: *m/z* 107, RI 1181

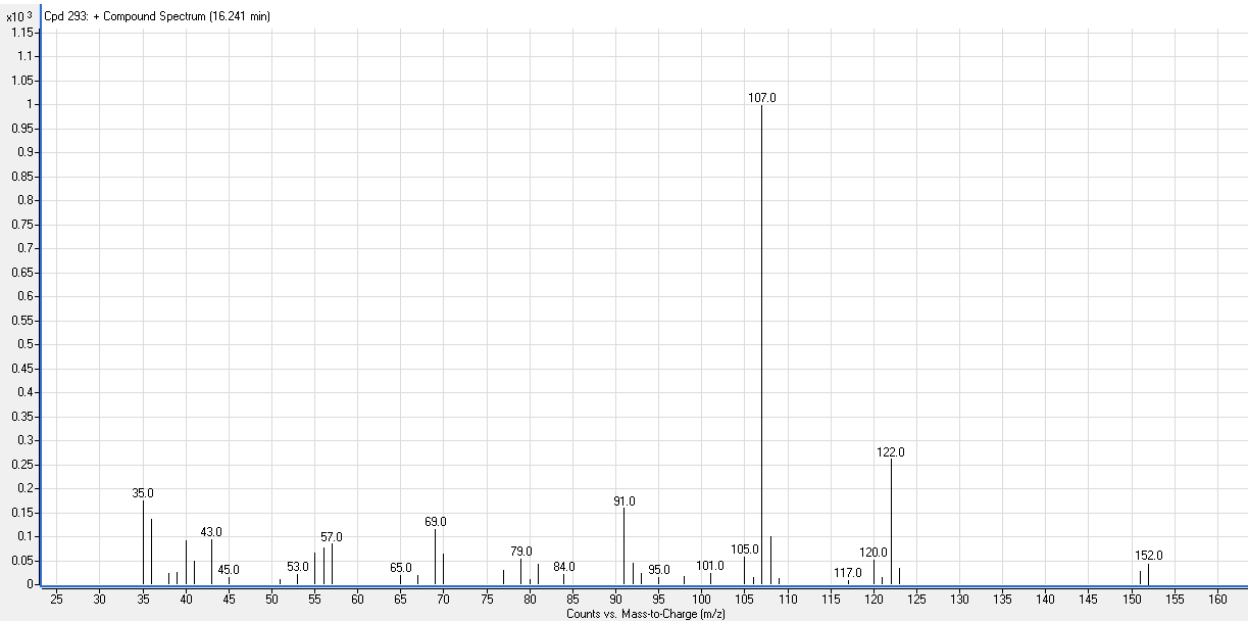
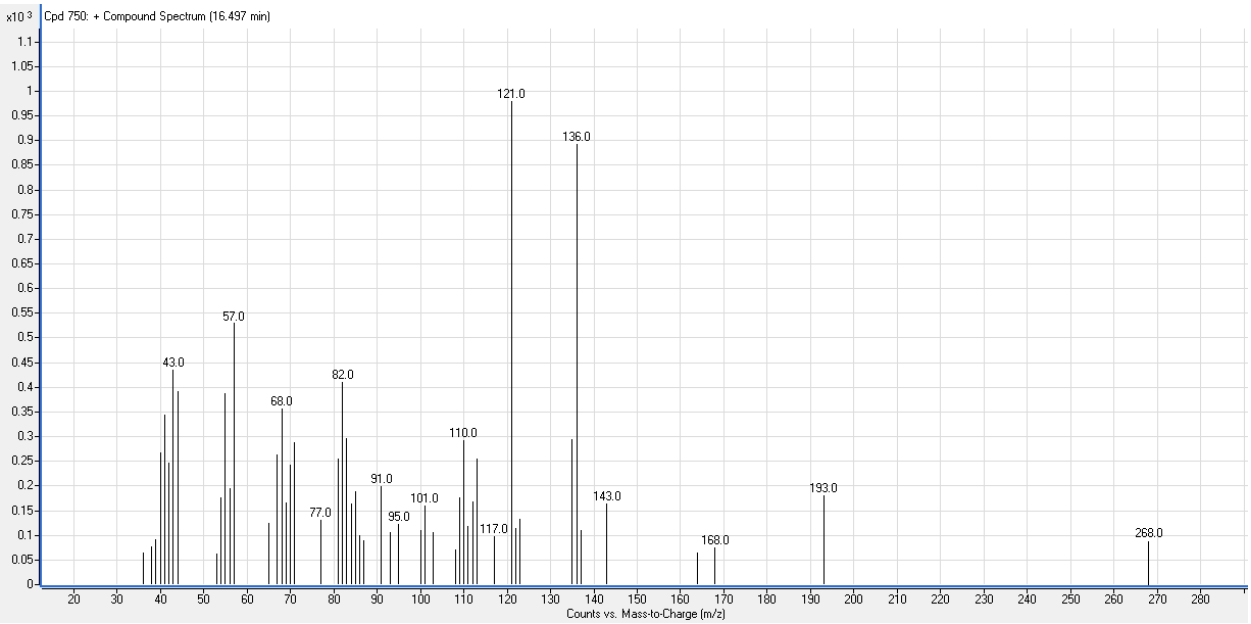


Figure A.4.1 Experimental mass spectra for all VOCs listed in Table 4.1, annotated either by NIST ID match or base peak and retention index pair.

Figure A.4.1 (cont'd)

1c) Compound 3: *m/z* 121, RI 1191



1d) Compound 4: β-cyclocitral

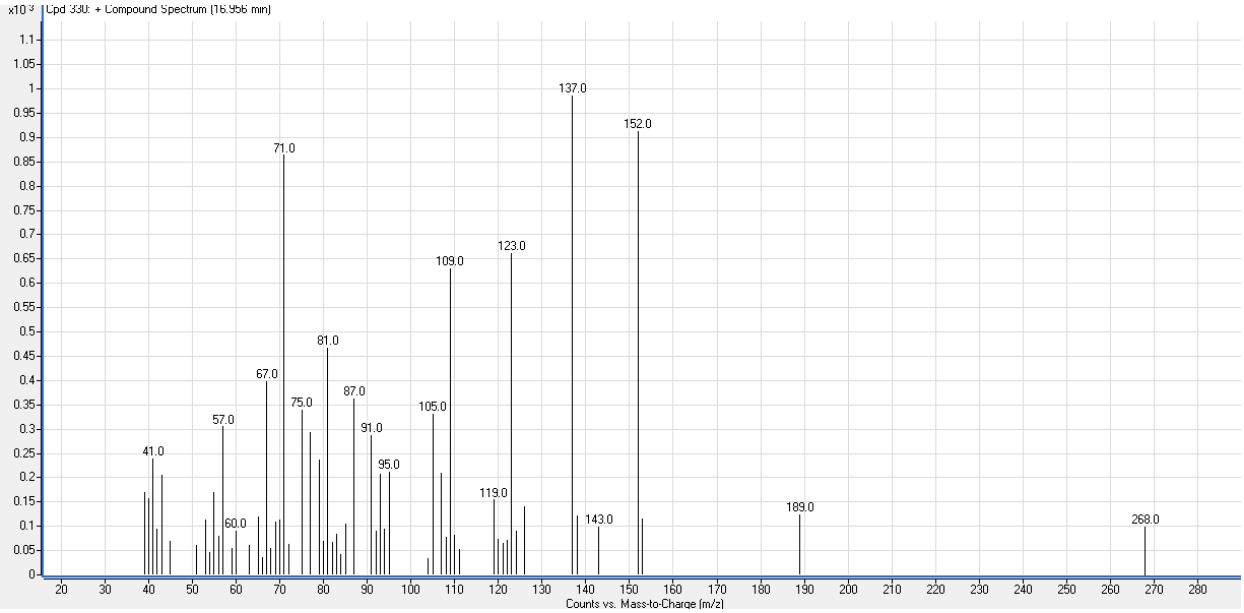
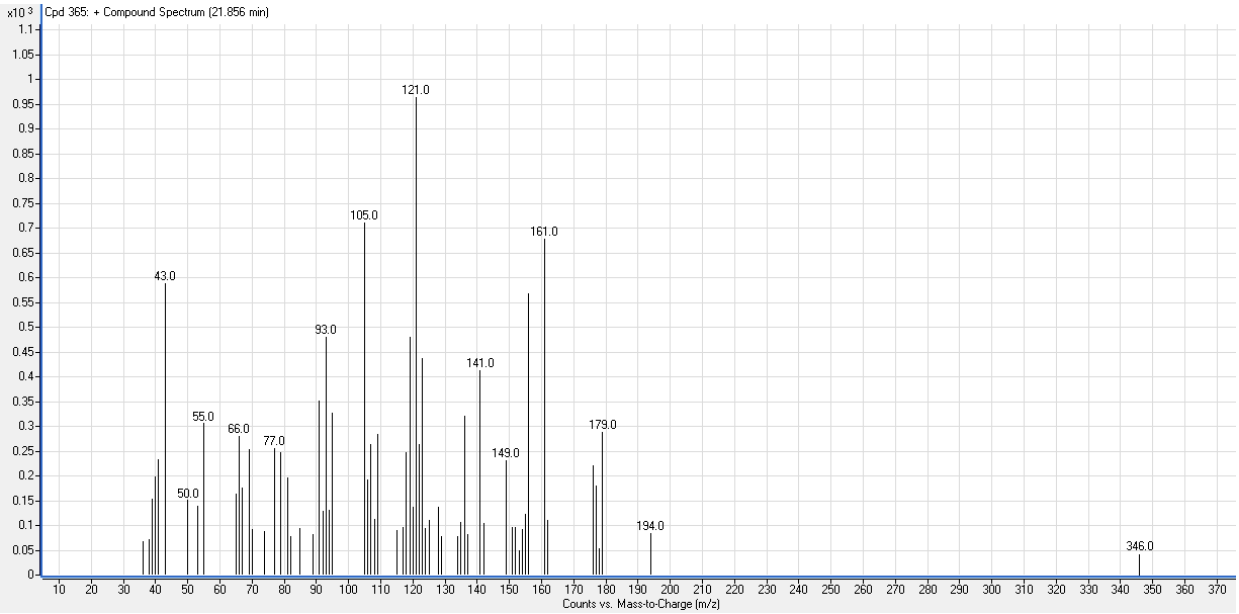


Figure A.4.1 (cont'd)

1e) Compound 5: 2-Butanone, 4-(2,6,6-trimethyl-1-cyclohexen-1-yl)-



1f) Compound 6: trans-β-ionone

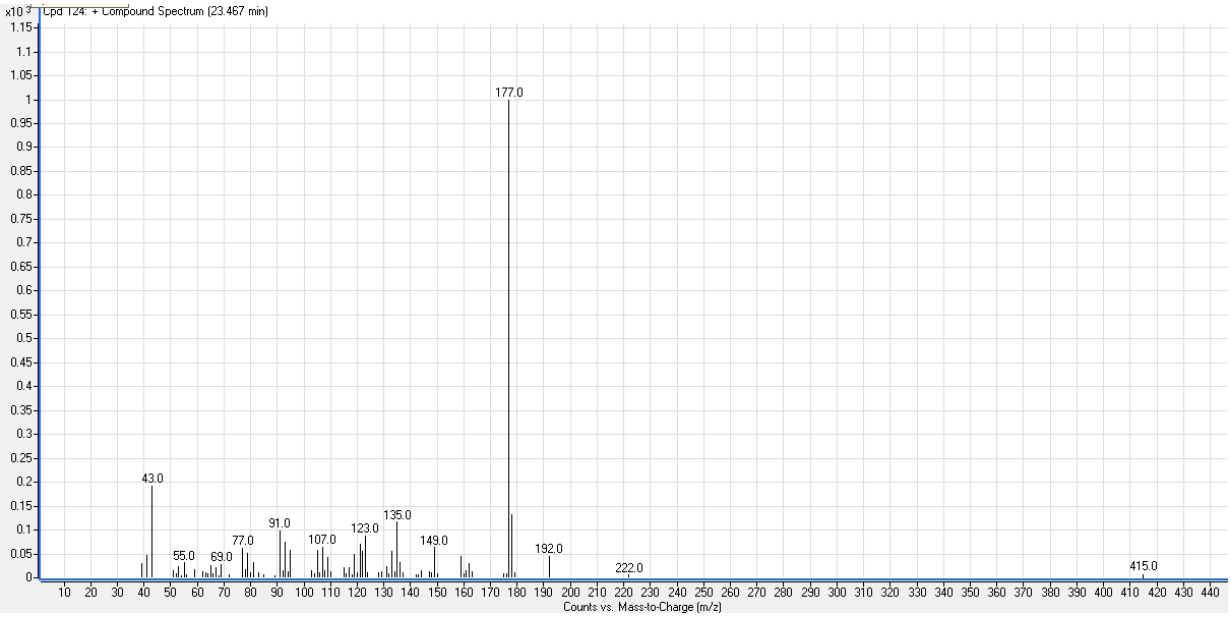
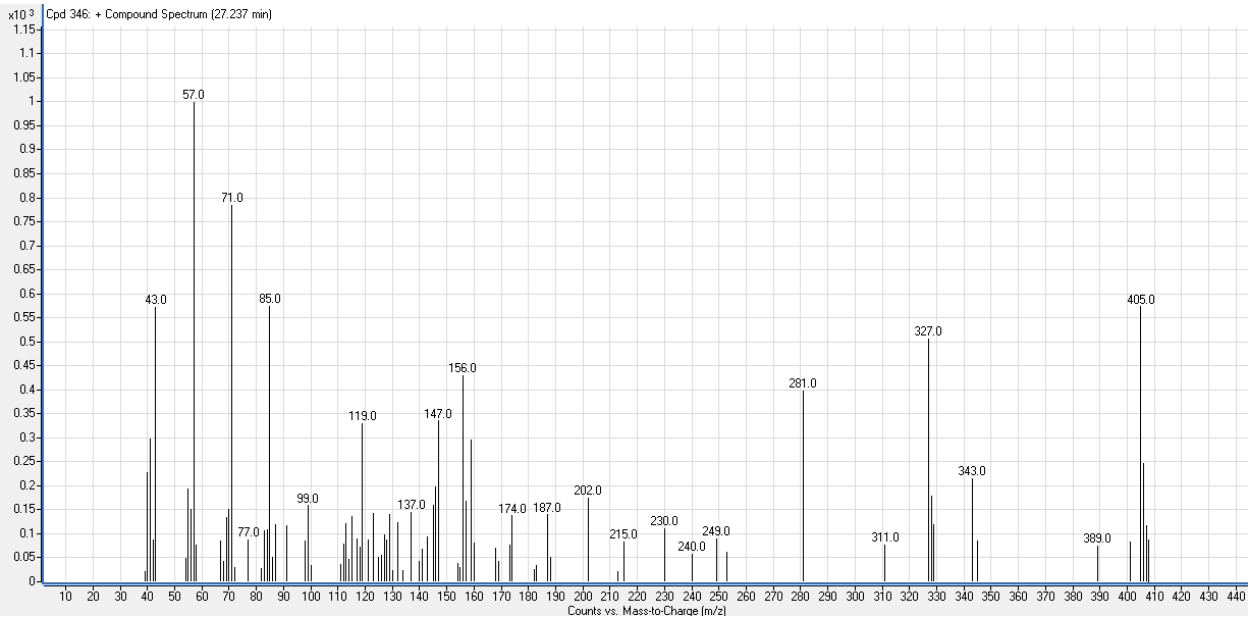


Figure A.4.1 (cont'd)

1g) Compound 7: *m/z* 57, RI 1691



1h) Compound 8: *m/z* 71, RI 1039

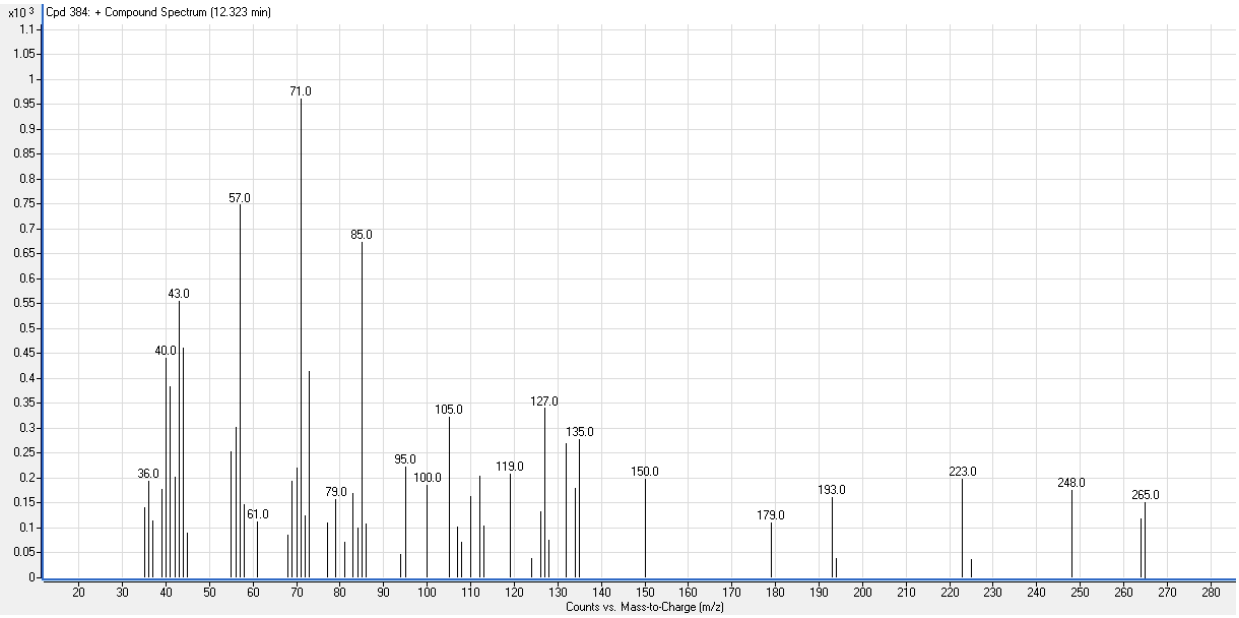
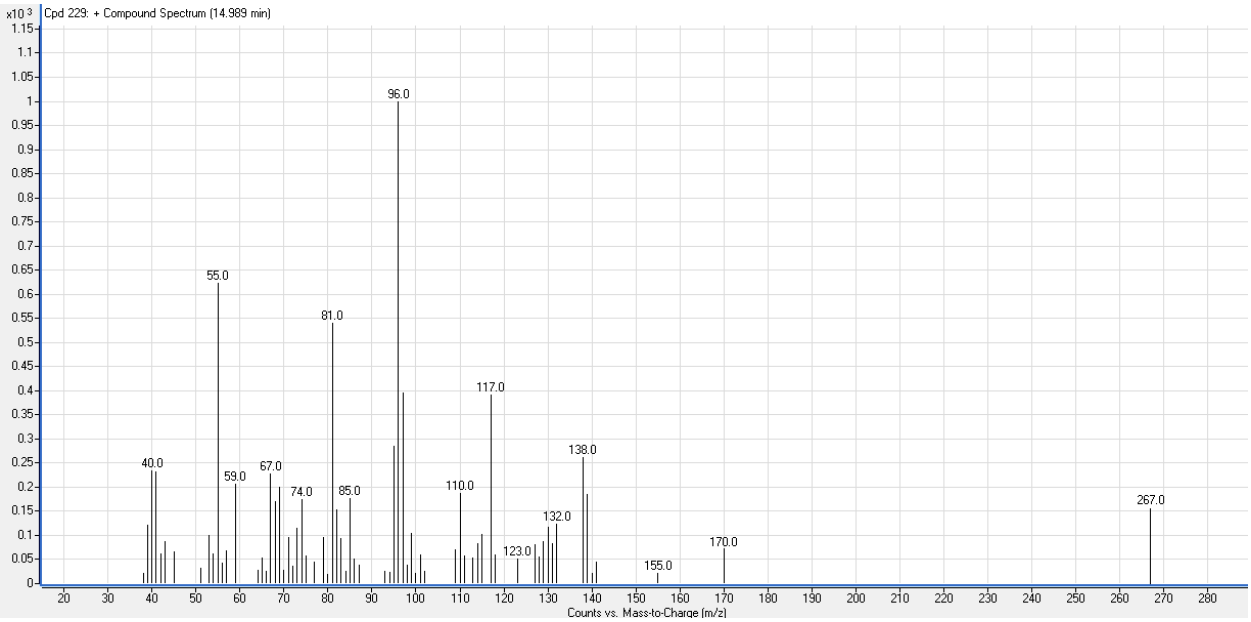


Figure A.4.1 (cont'd)

1i) Compound 9: 3-nonenoic acid, methyl ester



1j) Compound 10: *m/z* 341, RI 1139

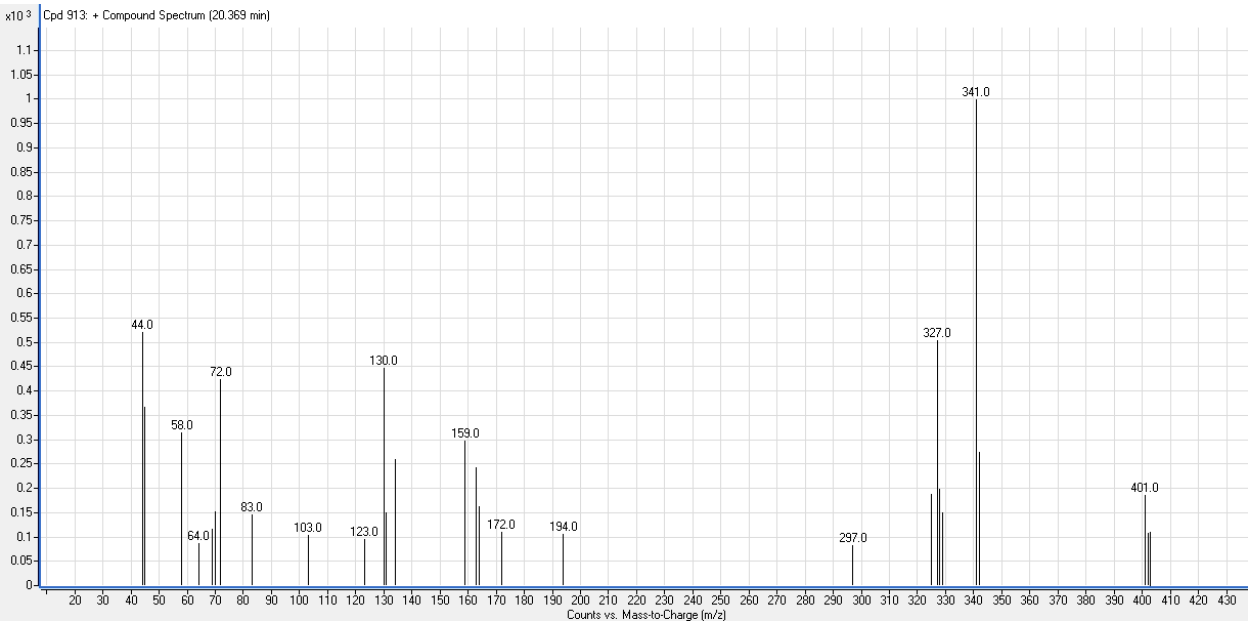
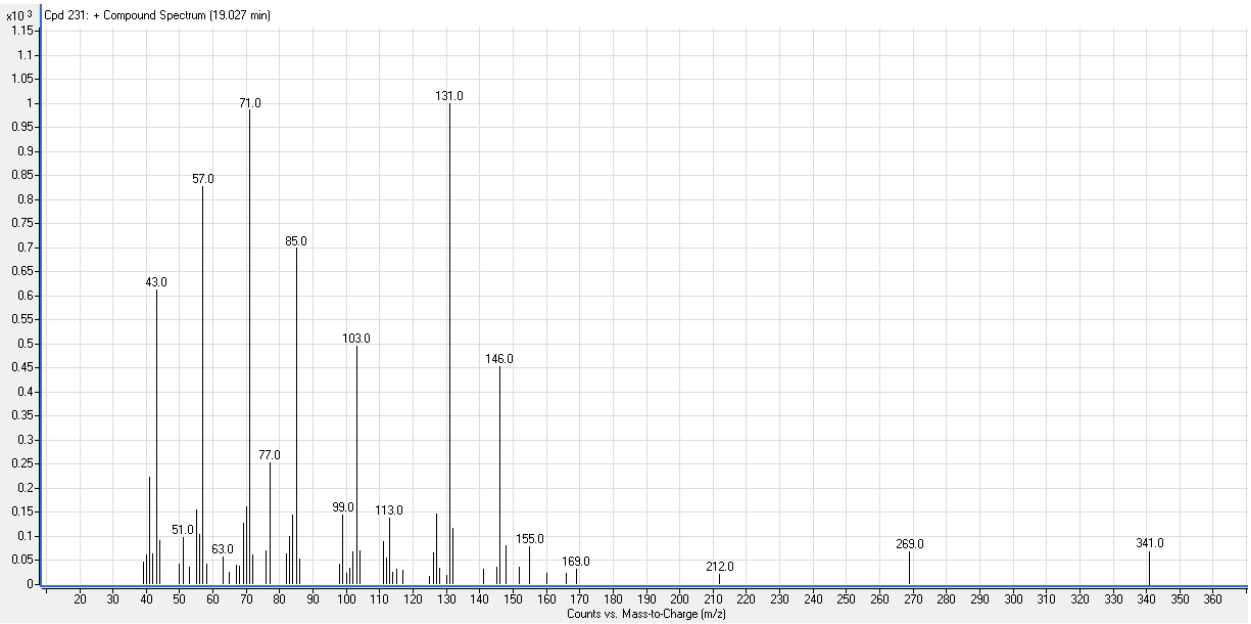


Figure A.4.1 (cont'd)

1k) Compound 11: *m/z* 71, RI 1293



1l) Compound 12: *m/z* 138, RI 1338

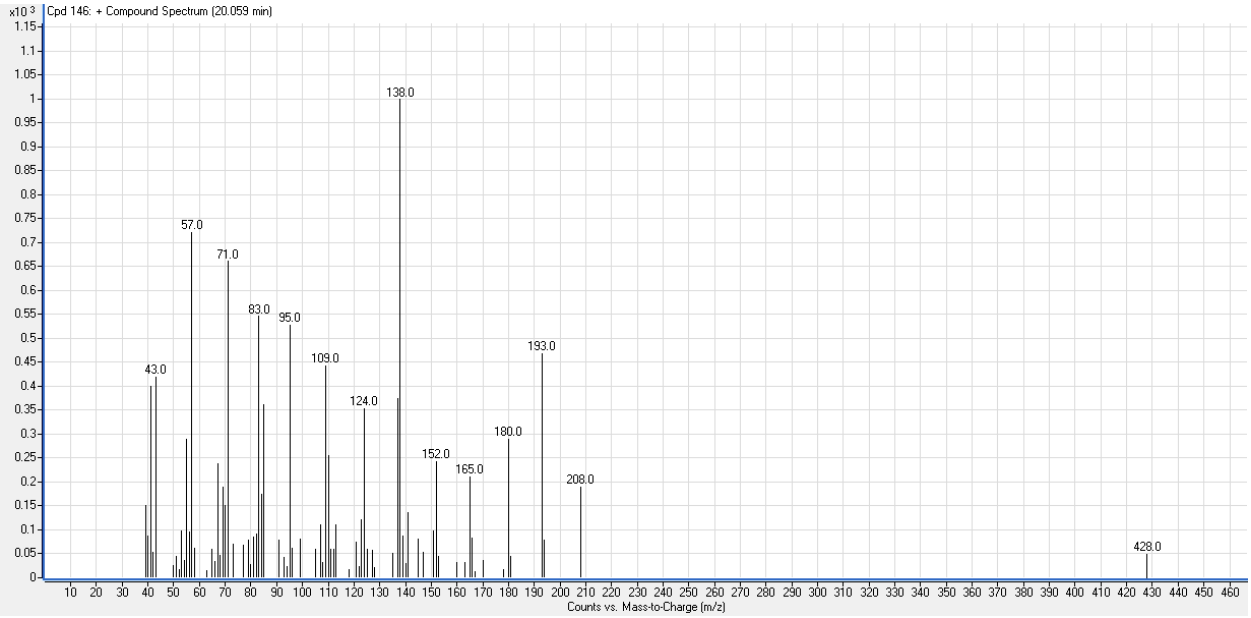
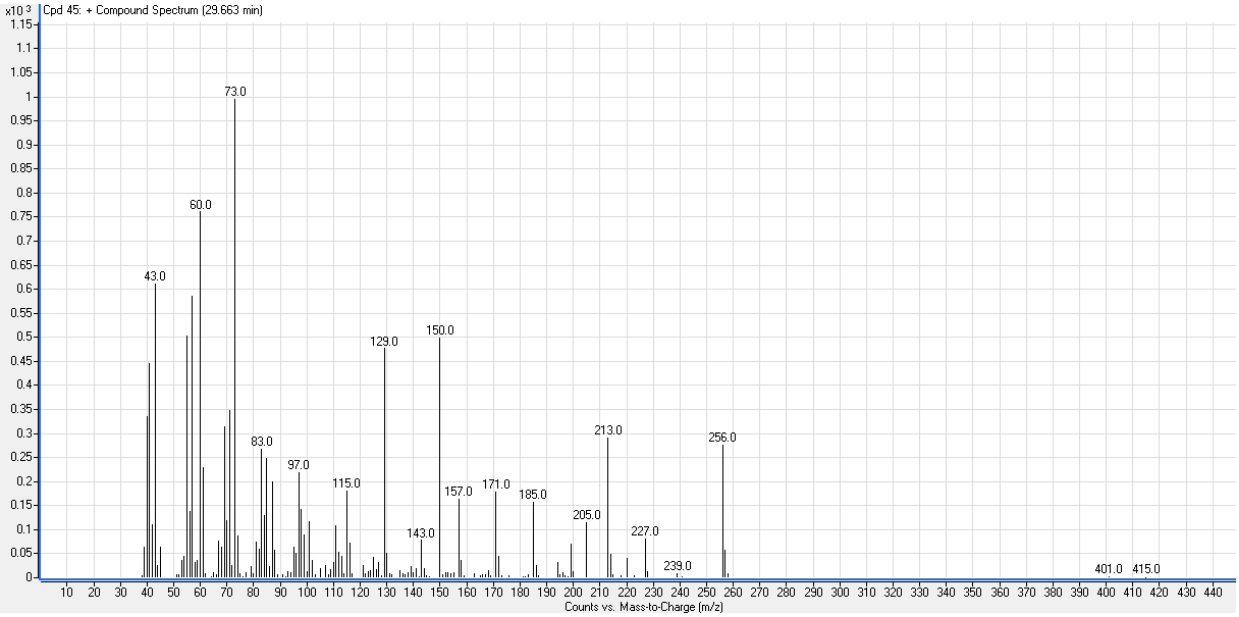
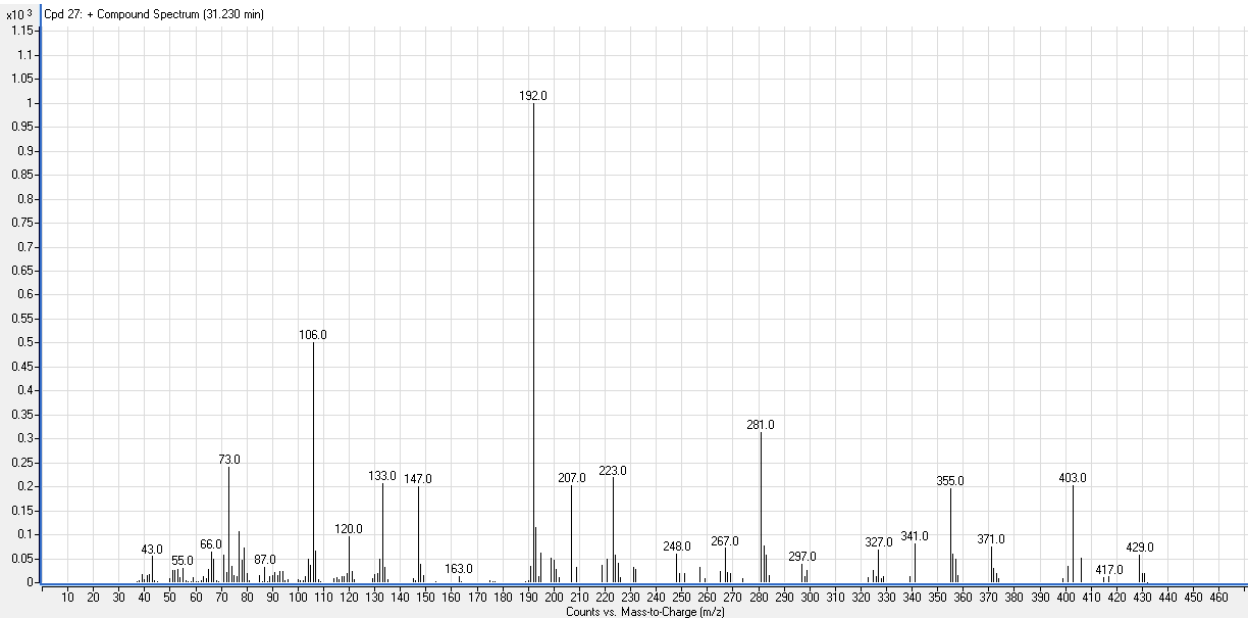


Figure A.4.1 (cont'd)

1m) Compound 13: *m/z* 73, RI 1983 : Hexadecanoic acid



1n) Compound 14: *m/z* 192, RI 2197



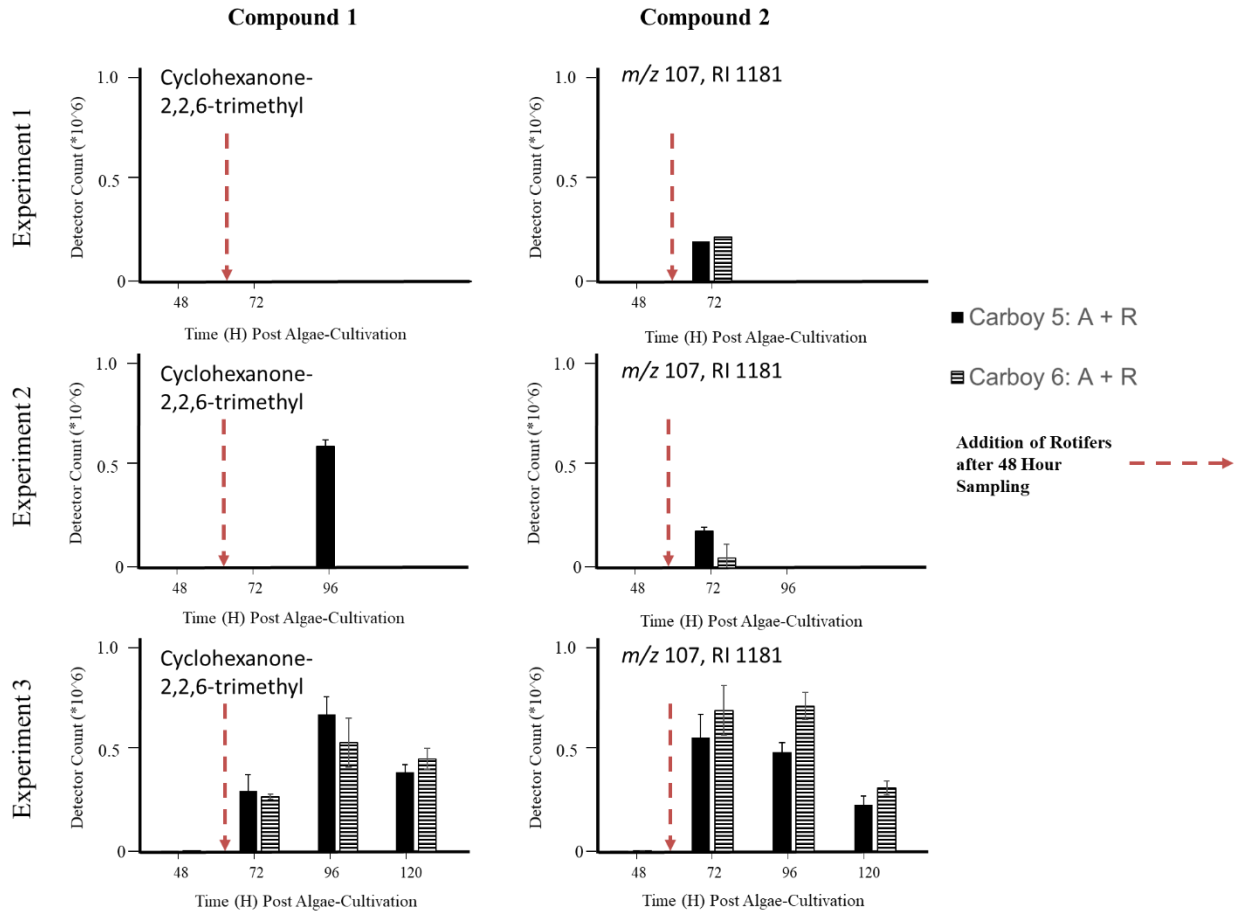


Figure A.4.2 Peak areas of extracted compound chromatograms for Compounds 1, 2, 3, 5, and 7 across Experiments 1, 2, and 3; Compounds 4 and 6 are displayed in Figure 4.5

Figure A.4.2 (cont'd)

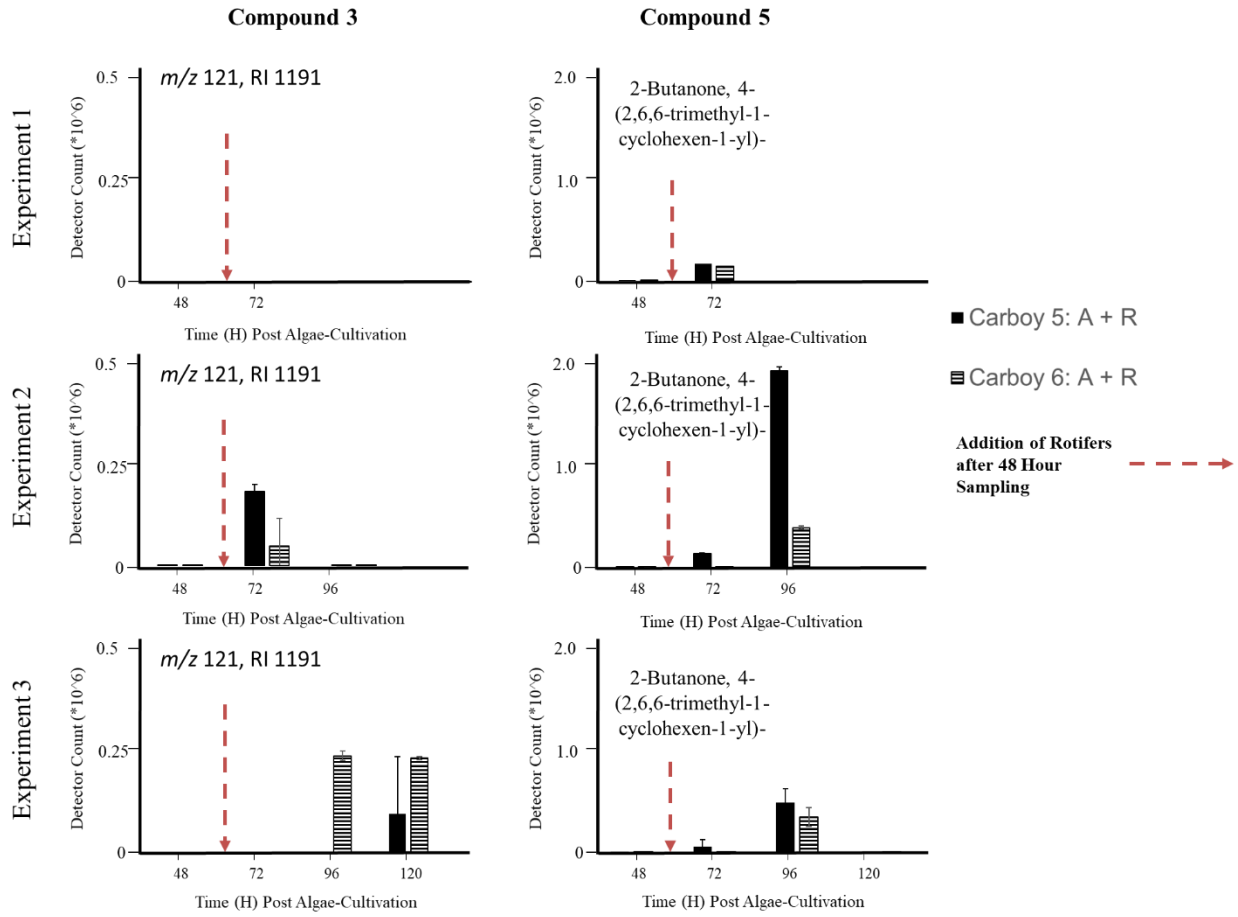
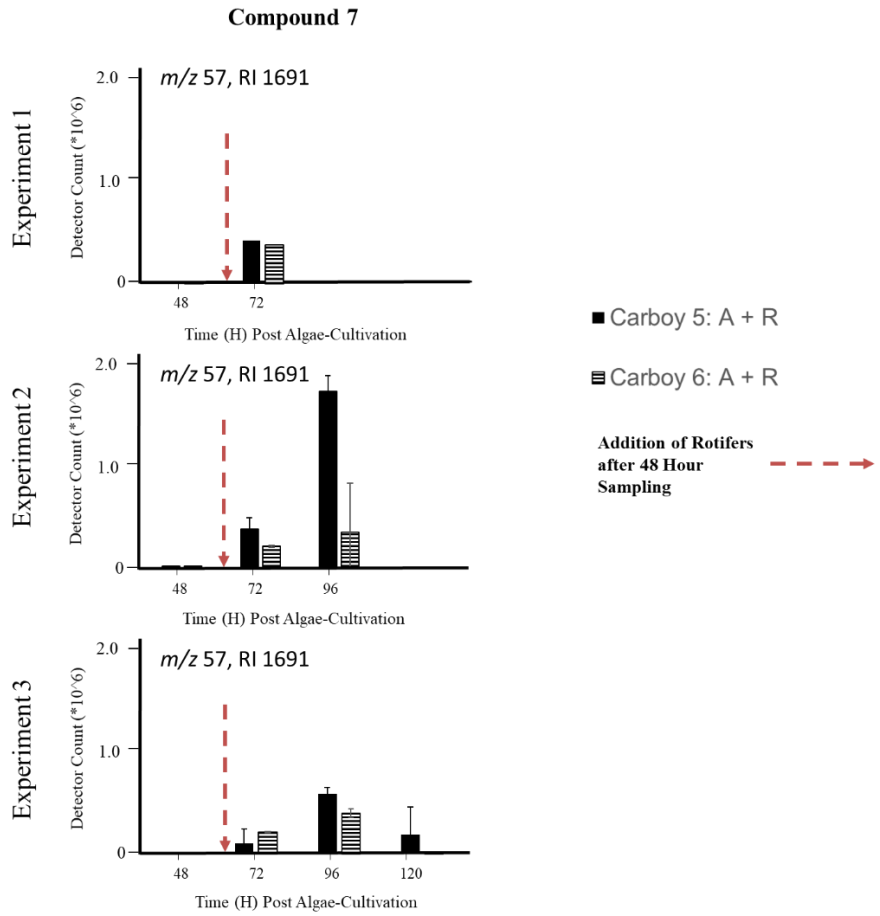


Figure A.4.2 (cont'd)



REFERENCES

REFERENCES

1 Reese, K. L. *et al.* Chemical profiling of volatile organic compounds in the headspace of algal cultures as early biomarkers of algal pond crashes. *Scientific Reports* **9**, 13866, doi:10.1038/s41598-019-50125-z (2019).

2 Hannon, M., Gimpel, J., Tran, M., Rasala, B. & Mayfield, S. Biofuels from algae: challenges and potential. *Biofuels* **1**, 763-784 (2010).

3 Leite, G. B., Abdelaziz, A. E. M. & Hallenbeck, P. C. Algal biofuels: Challenges and opportunities. *Bioresource technology* **145**, 134-141, doi:10.1016/j.biortech.2013.02.007 (2013).

4 Katiyar, R., Kumar, A. & Gurjar, B. R. in *Biofuels: Technology, Challenges and Prospects* (eds Avinash Kumar Agarwal, Rashmi Avinash Agarwal, Tarun Gupta, & Bhola Ram Gurjar) 157-175 (Springer Singapore, 2017).

5 Richardson, J. W. *et al.* A financial assessment of two alternative cultivation systems and their contributions to algae biofuel economic viability. *Algal Research* **4**, 96-104, doi:doi.org/10.1016/j.algal.2013.12.003 (2014).

6 Day, J. G., Gong, Y. & Hu, Q. Microzooplanktonic grazers – A potentially devastating threat to the commercial success of microalgal mass culture. *Algal Research* **27**, 356-365, doi:10.1016/j.algal.2017.08.024 (2017).

7 Carney, L. T. & Lane, T. W. Parasites in algae mass culture. *Front Microbiol* **5**, 278, doi:10.3389/fmicb.2014.00278 (2014).

8 Hirayama, K. & Ogawa, S. Fundamental studies on physiology of rotifer for its mass culture-I: Filter feeding of rotifer. *NIPPON SUISAN GAKKAISHI* **38**, 1207-1214, doi:10.2331/suisan.38.1207 (1972).

9 McBride, R. C. *et al.* Contamination management in low cost open algae ponds for biofuels production. *Industrial Biotechnology* **10**, 221-227, doi:10.1089/ind.2013.0036 (2014).

10 Zabed, H. M. *et al.* Biogas from microalgae: Technologies, challenges and opportunities. *Renewable and Sustainable Energy Reviews* **117**, doi:10.1016/j.rser.2019.109503 (2020).

11 Park, S. *et al.* The selective use of hypochlorite to prevent pond crashes for algae-biofuel production. *Water environment research : a research publication of the Water Environment Federation* **88**, 70-78, doi:10.2175/106143015x14362865227670 (2016).

12 Pradeep, V. *et al.* Use of copper to selectively inhibit Brachionus calyciflorus (predator) growth in Chlorella kessleri (prey) mass cultures for algae biodiesel production.

International journal of molecular sciences **16**, 20674-20684, doi:10.3390/ijms160920674 (2015).

13 Fott, B. Phlyctidium scenedesmi spec. nova, a new chytrid destroying mass cultures of algae. *Zeitschrift für allgemeine Mikrobiologie* **7**, 97-102, doi:10.1002/jobm.19670070203 (1967).

14 Xu, C. *et al.* The use of the schizonticidal agent quinine sulfate to prevent pond crashes for algal-biofuel production. *International journal of molecular sciences* **16**, 27450-27456, doi:10.3390/ijms161126035 (2015).

15 Van Ginkel, S. W. *et al.* Taking advantage of rotifer sensitivity to rotenone to prevent pond crashes for algal-biofuel production. *Algal Research* **10**, 100-103, doi:10.1016/j.algal.2015.03.013 (2015).

16 Liu, Z. & Lu, G. The sterilizing studies of flagellate and cilitate in marine unicellular algae liquid. *Zhanjiang Aquacult. Coll.* **6**, 36-41 (1990).

17 Wang, H., Zhang, W., Chen, L., Wang, J. & Liu, T. The contamination and control of biological pollutants in mass cultivation of microalgae. *Bioresource technology* **128**, 745-750, doi:10.1016/j.biortech.2012.10.158 (2013).

18 Fisher, C. L. *et al.* Bacterial communities protect the alga Microchloropsis salina from grazing by the rotifer Brachionus plicatilis. doi:10.1016/j.algal.2019.101500 (2019).

19 U.S. Department of Energy. Bioenergy Technologies office Multi-Year Program Plan: May, 2013. Washington, D.C., 2013.

20 Borowitzka M. A. in *Algal Culturing Techniques* (ed Andersen R. A.) 205-218 (Elsevier Academic Press, 2005).

21 Day, J. G., Thomas, N. J., Achilles-Day, U. E. & Leakey, R. J. Early detection of protozoan grazers in algal biofuel cultures. *Bioresource technology* **114**, 715-719, doi:10.1016/j.biortech.2012.03.015 (2012).

22 Wang, Y., Castillo-Keller, M., Eustance, E. & Sommerfeld, M. Early detection and quantification of zooplankton grazers in algal cultures by FlowCAM. *Algal Research* **21**, 98-102, doi:10.1016/j.algal.2016.11.012 (2017).

23 Carney, L., McBride, R., Smith, V. & Lane, T. Molecular Diagnostic Solutions in Algal Cultivation Systems. 183-204, doi:10.1201/b19464-9 (2016).

24 Carney, L. T. *et al.* Pond Crash Forensics: Presumptive identification of pond crash agents by next generation sequencing in replicate raceway mass cultures of Nannochloropsis salina. *Algal Research* **17**, 341-347, doi:org/10.1016/j.algal.2016.05.011 (2016).

25 Achyuthan, K. E., Harper, J. C., Manginell, R. P. & Moorman, M. W. Volatile metabolites emission by in vivo microalgae-An overlooked opportunity? *Metabolites* **7**, 39, doi:10.3390/metabo7030039 (2017).

26 Rowan, D. D. Volatile metabolites. *Metabolites.* **1**, 41-63, doi:10.3390/metabo1010041 (2011).

27 Leach, J. E., Triplett, L. R., Argueso, C. T. & Trivedi, P. Communication in the phytobiome. *Cell* **169**, 587-596, doi:10.1016/j.cell.2017.04.025 (2017).

28 Zuo, Z. Why Algae Release Volatile organic compounds-The emission and roles. *Front Microbiol* **10**, 491, doi:10.3389/fmicb.2019.00491 (2019).

29 Wolfe, G. & Steinke, M. *Grazing-Activated Production of Dimethyl Sulfide (DMS) by two clones of Emiliania huxleyi*. Vol. 41 (1996).

30 Wolfe, G. V., Sherr, E. B. & Sherr, B. F. Release and consumption of DMSP from *Emiliania huxleyi* during grazing by *Oxyrrhis marina*. *Mar. Ecol. Prog. Ser.* **111**, 111-119 (1994).

31 Wolfe, G. V., Steinke, M. & Kirst, G. O. Grazing-activated chemical defence in a unicellular marine alga. *Nature* **387**, 894-897, doi:10.1038/43168 (1997).

32 Hay, M. E. Marine chemical ecology: chemical signals and cues structure marine populations, communities, and ecosystems. *Annual review of marine science* **1**, 193-212, doi:10.1146/annurev.marine.010908.163708 (2009).

33 Fisher, C., Lane TW. Operational, prophylactic, and interdictive technologies in algal crop protection. *Grand Challenges in Algae Biotechnology* (2019).

34 C. M. de Jesus Benevides., M.C. da Cunha Veloso, P.A. de Paula Pereira, J.B. de Andrade. A chemical study of β-carotene oxidation by ozone in an organic model system and the identification of the resulting products. *Food Chemistry* **126**, 927-934, doi:10.1016/j.foodchem.2010.11.082 (2011).

35 Christaki, E., Bonos, E., Giannenas, I. & Florou-Paneri, P. Functional properties of carotenoids originating from algae. *Journal of the science of food and agriculture* **93**, 5-11, doi:10.1002/jsfa.5902 (2013).

36 Havaux, M. Carotenoid oxidation products as stress signals in plants. *The Plant journal : for cell and molecular biology* **79**, 597-606, doi:10.1111/tpj.12386 (2014).

37 Ramel, F. *et al.* Carotenoid oxidation products are stress signals that mediate gene responses to singlet oxygen in plants. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 5535-5540, doi:10.1073/pnas.1115982109 (2012).

38 Fawley, M. W., Jameson, I. & Fawley, K. P. The phylogeny of the genus Nannochloropsis (Monodopsidaceae, Eustigmatophyceae), with descriptions of N.

australis sp. nov. and Microchloropsis gen. nov. *Phycologia* **54**, 545-552, doi:10.2216/15-60.1 (2015).

39 Van Durme, J., Goiris, K., De Winne, A., De Cooman, L. & Muylaert, K. Evaluation of the volatile composition and sensory properties of five species of microalgae. *Journal of agricultural and food chemistry* **61**, 10881-10890, doi:10.1021/jf403112k (2013).

40 Isleten Hosoglu, M. Aroma characterization of five microalgae species using solid-phase microextraction and gas chromatography–mass spectrometry/olfactometry. *Food Chemistry* **240**, 1210-1218, doi:10.1016/j.foodchem.2017.08.052 (2018).

41 Zhou, L. *et al.* Change of volatile components in six microalgae with different growth phases. *Journal of the science of food and agriculture* **97**, 761-769, doi:10.1002/jsfa.7794 (2017).

42 Juttner, F., Watson, S. B., von Elert, E. & Koster, O. beta-cyclocitral, a grazer defence signal unique to the cyanobacterium Microcystis. *Journal of chemical ecology* **36**, 1387-1397, doi:10.1007/s10886-010-9877-0 (2010).

43 Whiting, J. J. *et al.* A high-speed, high-performance, microfabricated comprehensive two-dimensional gas chromatograph. *Lab on a chip* **19**, 1633-1643, doi:10.1039/c9lc00027e (2019).

44 Lewis, P. R. *et al.* Recent advancements in the gas-phase MicroChemLab. *IEEE Sensors Journal* **6**, 784-795, doi:10.1109/JSEN.2006.874495 (2006).

45 Snyder, D. T., Pulliam, C. J., Ouyang, Z. & Cooks, R. G. Miniature and fieldable mass spectrometers: Recent Advances. *Analytical chemistry* **88**, 2-29, doi:10.1021/acs.analchem.5b03070 (2016).

Chapter 5. EVALUATION OF DETECTION OF ALGAL-BACTERIAL INTERACTIONS BY TRACKING VOLATILE BIOMARKERS

FOREWORD

Contributions from others to conduct the experiments described in this chapter are as follows: X. Mayali and R. K. Stuart provided stock cultures of *Phaeodactylum tricornutum* and *Marinobacter* spp. 3-2. K. Rolison assisted with experimental set up, maintenance of experimental cultures, and discussions of data analysis.

5.1 INTRODUCTION

Algae have evolved in close relationships with other microbes, and interactions with other microbes influence their physiology and growth in natural systems [1] and likely in engineered systems as well. Our ability to assess and predict algal physiology and activity *in situ* is limited due to a lack of understanding about dynamic interactions with other microbes that naturally inhabit the same space. In contrast and complementary to the work of Chapter 4, where we investigated effect of predator gazing on algae on emitted algal volatile organic compound (VOC) markers, the research presented here in Chapter 5 aimed to lay further groundwork for understanding how microorganisms interact with each other at the molecular level using a model algal-bacterial co-culture and measuring volatile markers indicative of healthy (or stressed) algae in co-culture with bacteria.

Organic matter is exuded from all microorganisms, including bacteria and algae, throughout their lifecycles. Contained within these complex mixtures are volatile, semi-volatile, and non-volatile compounds. Such substances contribute to the biochemistry of aquatic systems by providing nutrition to organisms, facilitating organisms' ability to find food sources,

providing chemical defense from predators, or serving as an intra-species indicator of danger. However, the presence and roles of these compounds have not been fully understood, nor their effects on other microorganisms.

Algae and bacteria have coexisted for millions of years. Exchange of nutrients, signaling molecules, and other dissolved materials between organisms [2,3] happens in the area of closest contact, the microscale chemical environment surrounding algal cells, which has been termed the phycosphere (analogous to the rhizosphere around roots) [4]. Within the phycosphere, the exchange of volatile and non-volatile exuded metabolites can further affect growth cycles of each organism. Changes in the metabolic processes that affect levels of volatile compound accumulation might be detected in the headspace above actively growing algae ponds. The detection, identification, and quantification of these molecules in complex communities may enable diagnostics and/or manipulation of these interactions for a number of applications, including bioenergy, algal bioproducts and agriculture, and carbon capture.

Interactions of algae and bacteria, as well as their exuded organic matter, have been investigated, where the complexity of the systems was limited in order to determine the origins of identified metabolites. Several studies have concluded that algae can utilize vitamins and minerals exuded by bacteria, while bacteria can utilize ammonia exuded from algae [5-7]. Some exchanges are beneficial, where algae and bacterial growth is promoted through exchange of nutrients. Identification of exuded metabolites could guide addition of target microbes to commercial cultures to increase biomass production [1]. In contrast, competition between the co-cultured organisms for nutrients and dissolved organic material may inhibit algal growth [8]. The effects can be either specific to a species or inhibit a broad range of algae. Studies have indicated that inhibitory or non-growth promoting interactions play an important role in organizing the

structure of marine communities and ensuring survival [8,9]. Identification of inhibitory compounds may allow control of harmful phenomena in algal systems, such as algae blooms, that produce toxic effects on marine wildlife and can contaminate sources of drinking water [10].

The research presented in this chapter explored the feasibility of detecting and identifying algae-bacteria interactions at the molecular level by tracking volatile metabolites in the complex gaseous headspace of the model alga *Phaeodactylum tricornutum (P. tricornutum)* and model bacterium *Marinobacter spp. 3-2* at multiple timepoints over different algal growth states (exponential and stationary). *P. tricornutum* is a diatom with a sequenced genome used for biofuel production because of its high lipid content (estimated ~30% of dry weight [11] including abundant polyunsaturated fatty acids), and *Marinobacter* is a diverse genus of Gram-negative, aerobic bacteria found in most oceans [12]. The focus of this work was to detect and identify volatiles characteristic for each species as function of their growth states, and potentially, volatiles characteristic for the interaction of both species. To accomplish this, untargeted VOC biomarker profiles were obtained using solid phase microextraction (SPME) coupled with gas chromatography-mass spectrometry (GC-MS) on co-cultures, simulated co-cultures (i.e. algae exposed to exudates of bacteria, and vice versa), and controls. We hypothesized that monitoring VOC profiles of liquid cultures of algae and bacteria would reveal changes in the metabolism, specifically 1) co-culturing *P. tricornutum* and *Marinobacter* spp. 3-2 would produce a more complex VOC biomarker profile compared to monocultures of each organism, and 2) cultures of *P. tricornutum* exposed to exuded materials from *Marinobacter* spp. 3-2 cultures (and vice versa) would produce a more complex VOC biomarker profile (possibly more similar to that of the co-cultures) compared to monocultures due to the organisms potentially sensing volatile and non-volatile infochemicals emitted by their counterpart.

5.2 METHODS

5.2.1 Sample preparation

In order to evaluate the feasibility of collecting and detecting VOCs emitted from the headspace of algae and bacteria we set up single organism cultures of alga *Phaeodactylum tricornutum* (*P. tricornutum*) and bacterium *Marinobacter* subspecies 3-2 as well as co-cultures containing both organisms. The following nomenclature (Table 5.1) will be used throughout this chapter to designate mono- and co-cultures from *P. tricornutum* and *Marinobacter* spp. 3-2 microorganisms and samples comprised of exudates from monocultures of these microorganisms.

All samples prepared for VOC collection were prepared using Enriched Seawater/Artificial Water (ESAW) media. All samples used for VOC sampling were prepared in 250-mL glass Erlenmeyer flasks with stainless steel closures (Bellco Glass, Inc.) capable of allowing gas exchange and modified to accommodate insertion of solid-phase microextraction (SPME) fibers. The target working volume of each sample was 100 mL.

Table 5.1 Description of experimental sample types, abbreviations, and number of sample replicates

Sample Type	**Nomenclature**	**Number of Replicates**
P. tricornutum	**Algae**	2
Marinobacter spp. 3-2	**Bacteria**	2
P. tricornutum Exudates	**AlgEx**	2
Marinobacter spp. 3-2 Exudates	**BacEx**	2
P. tricornutum + *Marinobacter* spp. 3-2	**Co-Cultures**	3
P. tricornutum + *Marinobacter* spp. 3-2 Exudates	**Alg+BacEx**	3
Marinobacter spp. 3-2 + *P. tricornutum* Exudates	**Bac+AlgEx**	3
Enhanced Seawater/Artificial Water (ESAW) Growth Media	**ESAW**	1

ESAW media: ESAW media was prepared according to the National Center for Marine Algae and Microbiota (NCMA) guidelines and modified to contain elevated nutrient levels of 1.4 mM $NaNO_3$ and 58 μM Na_3PO_4. The ESAW media control sample was created by adding only 100 mL of liquid media to one of the flasks.

Algae: *P. tricornutum* (strain CCMP2561) was purchased from the National Center for Marine Algae (NCMA, West Boothbay Harbor, Maine, USA) and grown in ESAW media axenically (referred to in this work as monocultures), meaning no additional organisms were present at the start of the experiment. Stocks were incubated at 20°C under artificial sunlight on a 16:8 light: dark cycle. Stock cultures were regularly checked for contamination by collecting culture biomass on a filter, applying a blue-fluorescent DNA stain called DAPI (4′,6-diamidino-2-phenylindole), which targets adenine-thymine rich regions of double-stranded DNA, and visually checking for the presence of bacteria using fluorescence microscopy. Aliquots of stock cultures were transferred to sterile media 1-2x per month to maintain cultures, then again immediately prior to the experiment. At the start of the experiment, stock cultures were measured (fluorescence readings at 680 nm) to determine appropriate dilutions to obtain a consistent algal cell density (target of 100-150 relative counts). **Algae** replicates were created by inoculating sterile ESAW media with aliquots of the stock solution immediately prior to the start of the experiment.

Bacteria: *Marinobacter* spp. 3-2 (Genbank accession number ASM375135v1), a bacterium that was isolated and genetically characterized at LLNL, was grown axenically in Zobell media. Stocks were incubated at 20°C under artificial sunlight on a 16:8 light: dark cycle. Aliquots of stock cultures were transferred to sterile Zobell media 1-2x per month to maintain cultures, then again immediately prior to the experiment. Replicate monocultures of **Bacteria**

were created by diluting 5 mL of culture (estimated concentration 10^7 cells/mL) in 95 mL of ESAW.

Co-Cultures: Stock co-cultures of *P. tricornutum* and *Marinobacter* spp. 3-2 were maintained and provided by LLNL researchers. Stocks were incubated and maintained under the same conditions as **Algae** stock. At the start of the experiment, stock cultures were measured (excitation wavelength of 440 nm, fluorescence readings at 680 nm) to determine appropriate dilutions to obtain a similar algal cell density to **Algae**. (Distortions in the fluorescence readings due to the presence of bacteria are expected to be negligible.) **Co-cultures** replicates were created by inoculating sterile ESAW media with aliquots of the stock culture immediately prior to the start of the experiment.

AlgEx / BacEx: To generate **AlgEx** and **BacEx**, separate axenic algae and bacteria cultures, each with total volumes of 600 mL, were grown in 1 L baffled glass Erlenmeyer flasks with vented caps for air exchange. Algae cultures were grown until early stationary phase (approximately 144 hours post-inoculation, average fluorescence maximum 3000 counts) with incubation on a rotating platform at 20°C. Bacteria cultures were consistently grown for the same amount of time and displayed fluorescence counts comparable to an **ESAW** control. Subsequent vacuum filtration using a 0.2 μm filter tower (VWR) in a sterile environment removed algal or bacterial biomass from the filtrate containing dissolved organic matter (hereafter referred to as **AlgEx** and **BacEx**). Samples were prepared up to one month in advance of the experiment, sealed and stored in 1L Pyrex bottles at 20°C. For the control cultures, 100 mL of each exudate were added undiluted directly into the Erlenmeyer flasks.

Bac+AlgEx / Alg+BacEx: Samples containing cultures of one species mixed with exudates form the other species were created using the exudates **AlgEx** and **BacEx** as a base

media. Levels of nutrients (1.4 mM $NaNO_3$ and 58 μM Na_3PO_4) were replenished to support the growth of the new cultures. Subsequently, the same volumes of stock solution used to create the **Algae** and **Bacteria** stocks were added to exudate samples of the respective other species. To illustrate, one replicate sample of **Bac+AlgEx** was created by adding 95 mL of **AlgEx** and 5 mL of the same stock *Marinobacter* spp. 3-2 culture used to generate replicates of **Bacteria**.

5.2.2 Estimation of algal cell densities

Indirect measurements of live algal biomass were performed using fluorescence measurements of chlorophyll content on all sample types using a commercial plate reader (Cytation 5 Cell Imaging Multi-Mode Reader, BioTek Instruments, Inc./Agilent Technologies). Algal growth of all samples was monitored using 200 μL aliquots, read in the plate reader at the excitation and emission wavelengths 440 nm/680 nm approximately every 24 hours after sample inoculation. Additionally, at 72, 120, 168, and 240 hours post-inoculation, a 1 mL aliquot of each sample was placed in separate Eppendorf tubes and fixed with formaldehyde to prevent further biological growth – these archival samples were saved for future measurements of bacterial concentrations by microscopy.

A second planned method of measuring culture growth involved monitoring bacterial concentrations of bacteria in monocultures or in co-cultures using fluorescence microscopy and DAPI staining. Unpublished work at LLNL by researchers (X. Mayali, R.K. Stuart) on cultures of *P. tricornutum* and *Marinobacter* spp. 3-2 using DAPI staining and fluorescence microscopy to determine an estimated bacteria concentration of $1\text{-}6*10^6$ cells/mL over 8 days of growth in co-cultures. During each timepoint of VOC sampling (72, 120, 168, and 240 hours post-inoculation), aliquots of all sample types were taken in anticipation of these measurements.

However, I was unable to acquire the fluorescence microscopy data owing to limited laboratory access resulting from COVID-19. The bacterial concentration of all samples is anticipated to be measured at a later time. For the purposes of this dissertation, it is assumed the bacterial concentration was comparable to the 1-6*10^6 cells/mL levels in **Bacteria** and **Co-cultures** samples measured previously in similar mono- and and co-cultures.

5.2.3 VOC sampling

VOCs were collected from the headspaces of samples using PDMS-DVB SPME fibers, similar to the methods employed in Chapter 4 (Section 4.2.5). One fiber was used per flask. As demonstrated in Figure 5.1, within a sterile environment, SPME fibers were inserted into samples and exposed to the headspace for an average of 2.5 hours. VOCs were sampled (approximately) after 72, 120, 168, and 240 hours post-inoculation.

5.2.4 VOC data acquisition, processing, and biomarker identification

The VOC analysis and putative biomarker identification followed procedures similar to those described in Chapter 3 and 4 and in published works [13]. Briefly, an untargeted GC-MS approach was performed using a Agilent 5975T GC-single quadrupole-MSD system using 70 eV electron ionization and an Agilent HP-5ms column (the same system used for work described in chapters 3 and 4). Data acquisition was performed using ChemStation software (version E.02.02).

Figure 5.1 Experimental setup during passive VOC sampling of algal and bacterial samples - using SPME fibers (n=1 fiber per 250 mL flask); average exposure time ~ 2.5 hours); replicate cultures indicated by similar-colored labels.

Data processing was performed using MassHunter Qualitative and Mass Profiler Professional. Two filtering criteria similar to the ones used in Chapters 3 and 4 were applied to detect putative biomarkers from the large VOC dataset, but required modification accounting for smaller replicate numbers. For the first criterion, samples of n=3 replicates required a detected VOC to be present in two of three replicates for at least one timepoint to be further considered a biomarker candidate. For samples of n=2 replicates, a detected VOC was required to appear in both replicates for at least one timepoint to be further considered a biomarker. If a compound meeting either of the previous two requirements was also present in one replicate at another timepoint for the same sample type, its presence was noted (see Results). The second criterion for considering a VOC as a biomarker was unchanged compared to the work in chapters 3 and 4: the compound was not detected in the ESAW media blank or travel blank conditions; OR b) the compound was present in the sample type at an abundance greater than 10x the abundance in the ESAW media blank or travel blank.

Additional data processing was performed to identify the taxa-specific algal and bacterial biomarkers across all samples and timepoints. Retention indices were calculated with reference to a commercial reference standard, and identifications were performed by searching mass spectra against the NIST14 mass spectral database. Biomarker identification required a mass spectral library match greater than 70% and an experimental retention index within 3% match of the reference compound. Biomarkers unable to be identified were presented as the *m/z* value of the base peak at a given retention index (e.g., *m/z* 71 at RI 1271). Compounds identified with values of base peak *m/z* 207, 262, and 281 were removed from biomarker consideration, as these are common contaminant ions from polysiloxanes originating from either the chromatography column or the SPME stationary phase.

5.3 RESULTS AND DISCUSSION

5.3.1 Cell counts of microorganisms in cultures

Figure 5.2 presents algal growth (as measured by chlorophyll fluorescence) as a function of time after inoculation for the sample types with live algae. A higher biological variation in growth was observed amongst replicates of **Alg+BacEx** compared to the replicates of **Algae** and **Co-cultures**, as indicated by the error bars in this figure. For all sample types, logarithmic growth is maintained throughout approximately 120 hours before beginning to level off, indicating the beginning of the stationary phase. The average algal growth in **Co-cultures** (n=3 replicates) and **Alg+BacEx** (n=3) samples was consistently lower than that of **Algae** cultures (n=2). As algal growth is inhibitedin both **Co-cultures** and **Alg+BacEx** cultures, the reduction is possibly due to the actions of a growth inhibitor secreted by the bacteria rather than competition between live algae and bacteria for nutrients in the growth media.

5.3.2 Headspace VOC results for Cultures of *P. tricornutum* and *Marinobacter* spp. 3-2

This work demonstrated that our methodology 1) detected differences in the metabolite VOC biomarker profiles from the growth of *P. tricornutum* and *Marinobacter* subspecies 3-2, either as mono-cultures or co-cultures, using unique VOC biomarkers and 2) the biomarker profile varies within sample types across subsequent timepoints, indicating temporal metabolite changes. Data processing and analysis resulted in 150-250 compounds detected per sample. Peak alignment of all samples and timepoints generated a list of more than 3200 compounds. Application of filtering criteria based on detection frequency and compound abundance (described in the methods) removed over 90% of this list, thus identifying the most robust

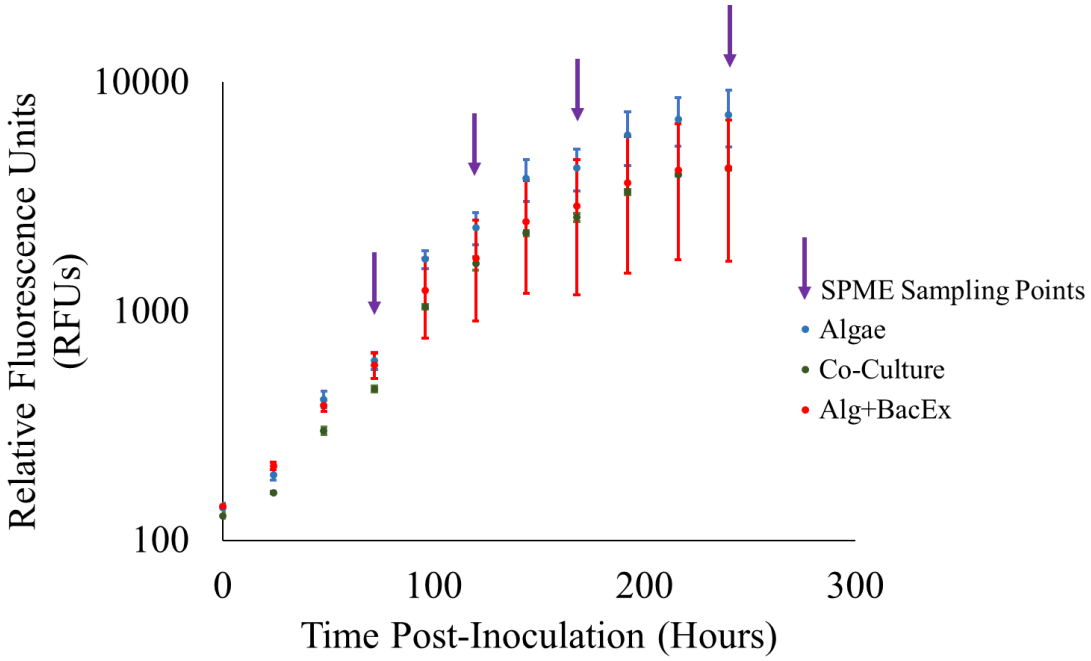


Figure 5.2 Growth curves of algae in samples containing *P. tricornutum*. **Algae**, **Co-Culture**, and **Alg+BacEx**, as determined by measurements of relative chlorophyll fluorescence units (RFUs) at increasing timepoints post-inoculation of cultures. Averages RFUs (averaged over replicates) are plotted with error bars showing +/- standard deviations. Colors of the error bars correspond to sample type in the legend. Counts of RFUs in **ESAW, Bacteria, and BacEx** samples were simultaneously acquired but did not exceed 20 RFUs, hence considered to be negligible signal and omitted from this figure. VOCs were collected via SPME sampling at 72, 120, 168, and 240 hours post-inoculation as indicated by arrows.

compounds as putative VOC biomarkers. In this section, the metabolite profiles for **Algae**, **Bacteria**, and **Co-cultures** samples will be described, followed (in section 5.3.3) by the results from the study of metabolite changes in a primary organism when exposed to exudates from a secondary organism (**Alg+BEx**, **Bac+AEx**).

Algae: A total of 19 putative VOC biomarkers were detected in **Algae** cultures and are presented in Appendix Table A.5.1. The following compound classes were represented: amines, aliphatics (cyclic olefins, alkanes), carotenoids, and methyl ketones. Three biomarkers were given specific identifications based on NIST14 mass spectral matching, specifically a) 6-[(*Z*)-1-Butenyl]-1,4-cycloheptadiene (common name ectocarpene), b) 6, 10, 14-trimethyl-2-pentadecanone (common name phytone or hexahydrofarnesyl acetone), and c) N-ethyl-N-methyl-2-propen-1-amine. Both 6-[(*Z*)-1-Butenyl]-1,4-cycloheptadiene and 6, 10, 14-trimethyl-2-pentadecanone were present at all timepoints measured. Additionally, 4 non-identified VOC biomarkers could be given tentative annotations to compound classes, such as alkanes, through use of compound class characteristic MS fragmentation patterns, base peak (*m/z*) values, and retention index values. One biomarker of note, annotated as *m/z* 177 at RI 1495, exhibited mass spectral fragments at *m/z* 43 and 135, similar to those of the carotenoid degradation product β-ionone that was observed as a signal of algal wounding in Chapter 4. Therefore, this was annotated as a carotenoid degradation product.

Bacteria: A total of 11 putative VOC biomarkers were detected in the **Bacteria** cultures (Appendix Table A.5.2). One biomarker with *m/z* 68 at RI 1015 was given a specific identification, the monoterpene D-limonene. An additional 2 markers, *m/z* 57 at RI 10 and *m/z* 57 at RI 1816, were assigned to the chemical functional groups of alkanes. Interestingly, biomarkers with base peak *m/z* ratios greater than 100 were observed more frequently at the 72 and 120

sampling points, while the 168 and 240 hour sampling points both contained biomarkers with base peak *m/z* ratios less than 100.

Co-cultures: An example of the chemical complexity of *P. tricornutum* + *Marinobacter* spp. 3-2 VOCs versus the ESAW growth media at 240 hours post-inoculation is shown for the total ion current chromatogram in Figure 5.3, where the putative VOC biomarkers for **Co-cultures** are indicated. A total of 31 putative VOC biomarkers were detected across **Co-cultures** (Appendix Table A.5.3). The following compound classes were represented: aliphatics (cyclic olefins, alkanes), carboxylic acids, carotenoids, methyl ketones, and terpenes. Six biomarkers were given specific identifications based on NIST14 mass spectral matching, specifically a) D-limonene, b) 6-[(*Z*)-1-butenyl]-1,4-cycloheptadiene, c) dianhydromannitol, d) 6, 10, 14-trimethyl-2-pentadecanone, e) *Z*-11-hexadecenoic acid, and f) n-hexadecanoic acid. For the monounsaturated fatty acid *Z*-11-hexadecenoic acid, it is not possible to distinguish the double bond position using EI fragmentation. Therefore, subsequent reference to it will be

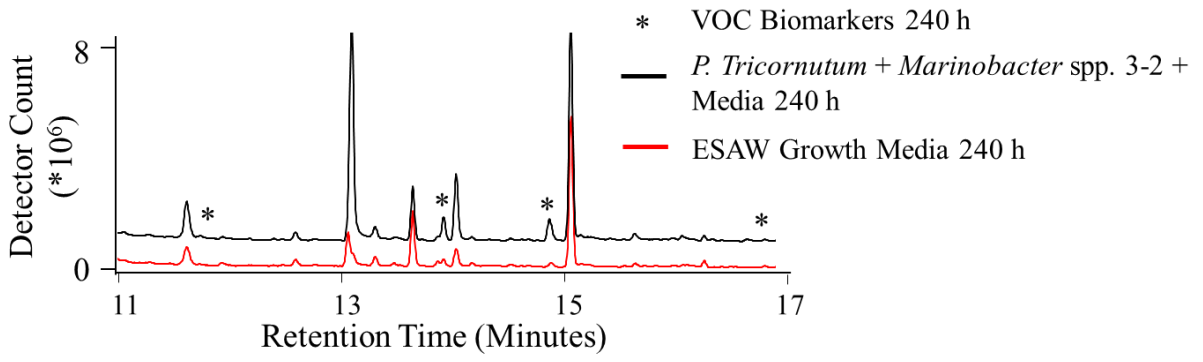


Figure 5.3 Example total ion current chromatograms for observed VOCs sampled from the **Co-cultures** and from **ESAW** at 240 hours post inoculation, with peaks indicated by * being potential VOC markers of algae-bacterial interactions.

"hexadecenoic acid (16:1) at RI 1963". An additional 3 biomarkers were assigned to classes based on evidence of a chemical functional group, and additional carotenoid compounds, β-cyclocitral and β-ionone, were identified as the substances annotated at *m/z* 137 at RI 1208 and *m/z* 177 at RI 1495 based on comparison to the work in Chapter 4.

Analysis of VOC data from **Algae**, **Bacteria**, and **Co-cultures** revealed groups of biomarkers that were indicative of each sample type as well as select compounds shared across sample types. Figure 5.4 details the overlap among **Co-cultures**, **Algae**, and **Bacteria**. VOC biomarkers found in **Algae** also observed in **Co-cultures** included a) 6-[(*Z*)-1-butenyl]-1,4-cycloheptadiene, b) 6, 10, 14-trimethyl-2-pentadecanone, c) β-ionone, d) an unidentified alkane at *m/z* 71, RI 1271, e) unidentified marker *m/z* 108 at RI 1021, and f) unidentified marker at *m/z* 73 at RI 2572. Markers found in **Bacteria** also observed in algal-bacterial **Co-cultures** included a) D-limonene, b) *m/z* 42 at RI 815, c) *m/z* 95 at RI 1151, and d) *m/z* 73 at RI 1912.

First, for this work we hypothesized that co-culturing *P. tricornutum* and *Marinobacter* spp. 3-2 would produce more chemical complexity in the headspace VOCs, manifesting in a larger total number of produced biomarkers and/or production of additional, different biomarkers, in the **Co-cultures** profile compared to monocultures. Twenty-one of the VOC markers from **Co-cultures** were novel signatures not observed in the monocultures, and thus suggested that the metabolism of **Algae** and **Bacteria** was altered because of their interaction, thus causing additional markers to be produced. The origins of several markers were traced back to VOC emissions of the monocultures, six from the **Algae** cultures and eleven from the **Bacteria** cultures, indicating the production of these markers was not altered.

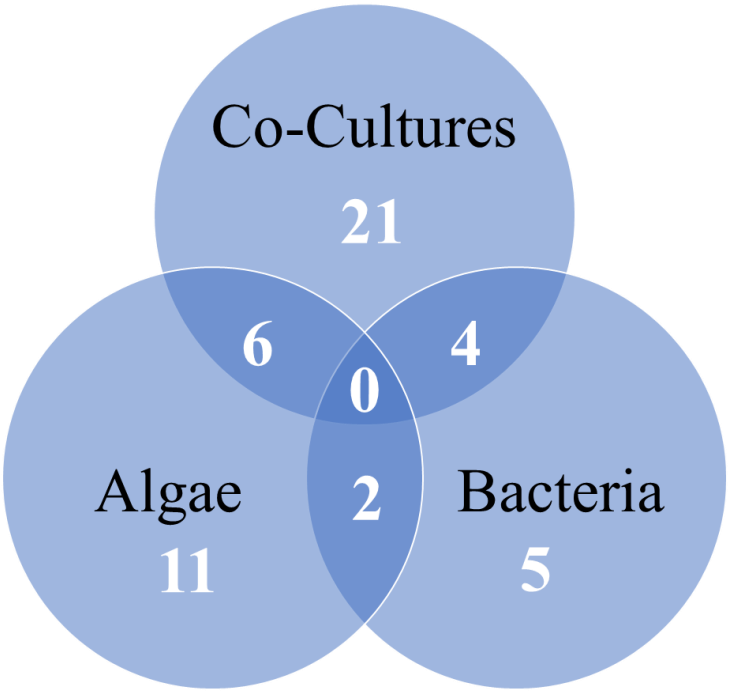


Figure 5.4 Venn diagram of the overlapping VOC biomarkers amongst **Co-cultures**, **Algae,** and **Bacteria** as inclusive of markers detected across all measured timepoints. Each group is derived from n=2 replicates.

Detected volatile biomarkers in **Co-cultures, Algae,** and **Bacteria** evolved over time within monocultures and co-cultures as observed when sampling at 72, 120, 168, and 240 hours post-inoculation. For example, ectocarpene was observed to be present and increase at all timepoints for **Algae** and **Co-cultures**. Previous studies have shown high levels of ectocarpene and other C11 unsaturated olefins result from metabolic cleavage of eicosanoid precursors in diatoms, of which *P. tricornutum* is a member [9,14,15]. Ectocarpene has been observed in *P. tricornutum* and in species of brown algae, where the biochemical function was related to pheromone production and sexual attraction [14]. These results are in agreement, as a higher cell count at each subsequent timepoint produced greater levels of ectocarpene.

Two biomarkers indicative of the later timepoints, where growth phase approaches stationary phase, were two 16-carbon fatty acids: hexadecenoic acid (16:1) at RI 1963 and n-hexadecanoic acid (16:0; also known as palmitic acid). Both biomarkers were observed only at the 240 hour sampling of **Co-cultures**. An observation or increase in the levels of saturated or monosaturated lipids for the stationary phase has been documented in previous studies with axenic *P. tricornutum* cultures [15-18], and this research further reports its presence when grown in the presence of additional microorganisms. Fatty acids such as this are components of triacylglycerols, which serve as a storage source for carbon and energy. When algae have depleted media nutrients during the stationary phase, internal triacylglycerols can be metabolized to produce energy for further growth, thus releasing fatty acids such as palmitic acid.

Interestingly, some biomarkers only observed at the 240 hour sampling time of **Algae** and **Co-cultures** were annotated as carotenoid degradation products. These signals were previously observed in the research presented in Chapter 4 specifically as early-warning VOC biomarkers for algal pond crashes caused by active predation of the algae *Microchloropsis salina* [13]. Conversely, β-ionone was present in **Algae** while β-cyclocitral and β-ionone were both observed in **Co-cultures.** The growth data presented in Figure 5.2 suggests *Marinobacter* exudates inhibit the growth of *P. tricornutum*, potentially indicating a toxic effect that could increase algal wounding and the release of carotenoid products. Additionally, different algae lineages have been shown to produce varying levels of endogenous carotenoids [19]. The presence of β-ionone only in the 240 hour timepoint of *P. tricornutum* cultures suggests that rising levels of algal stress or cell death during the start of stationary phase releases endogenous levels of carotenoids at a level detectable by SPME-GC-MS.

A potential infochemical, the methyl ketone 6,10,14-trimethyl-2-pentadecanone, common name hexahydrofarnesyl acetone, was also present and increase in abundance at all sampled timepoints for **Algae** and **Co-cultures**. This ketone is a fragrant compound observed in the green algae genera *Spirogyra* [20] and *Cladophora* [21], and brown alga *Padina pavonia* [22]. It has also been observed in higher plants, in particular an extract of the vining plant *Vitis setosa*, and demonstrated to contain antimicrobial properties [23].

Although exact identification was not possible, also alkane-like biomarkers were annotated in this work (based on characteristic fragmentations at *m/z* 43, 57, and 71). Different annotated alkanes were present in each sample type and represented at every sampled timepoint. Studies report *P. tricornutum* accumulates a range of hydrocarbons similar to those found petroleum-based fuels, such as octane, undecane, nonadecane, heneicosane, heptadecane, nonadecane, and eicosane, although the temperature at which the algae was grown played a role in whether shorter or longer chain hydrocarbons were produced [24]. The hydrocarbon content of *P. tricornutum* and other algae species has been extensively studied for determining which species would be optimal to use in biofuel cultivation.

The presence of a wide range of VOCs have been reported for *P. tricornutum* that were not observed in this work. In particular, isoprene [25], aldehydes [26], organohalogens comprised of chlorine and bromine [15], and sulfur-containing compounds [27] have been reported as originating from marine sources. If present in *P. tricornutum*, this method did not possess either the selectivity or sensitivity to observe these compounds. Multiple methodologies and detection schemes should be considered when performing untargeted biomarker profiling of algae and associated microorganisms.

5.3.3 Headspace VOCs for monocultures of one species exposed to exudates from the other

Biomarkers from cultures of one species supplemented with exudates of the respective other (**Alg+BacEx** and **Bac+AlgEx**) were profiled to check if any of the additional biomarkers from **Co-cultures** could be replicated when the presence of one species is simulated by adding exudates. If this is true, this would indicate soluble compounds produced by one species triggers production of the VOCs by the other species. First, the biomarker profiles for the exudates (**AlgEx** and **BacEx**) are described, followed by building up the analysis of one of the organisms exposed to exudates from the respective other organism (**Alg+BacEx**, **Bac+AlgEx**).

AlgEx: A total of 19 putative VOC biomarkers were detected in the **AlgEx** samples (Appendix Table A.5.4), covering the compound classes of alkanes and carboxylic acids. Two carboxylic acids, a) hexadecenoic acid (16:1) at RI 1963 and b) n-hexadecanoic acid were assigned identifications, while an additional three were assigned to classes based on a chemical functional group. Figure 5.5 illustrates the overlap between compounds detected in **Algae** and **AlgEx**, showing a 21% overlap of each profile based upon the following shared but unidentified biomarkers annotated as: a) *m/z* 71 at RI 1271, b) *m/z* 401 at RI 1424, c) *m/z* 135 at RI 1473, and d) *m/z* 73 at RI 2572.

BacEx: A total of 24 putative VOC biomarkers were detected in the **BacEx** Appendix Table A.5.5). While no compounds could be assigned a confident chemical identification, two markers were annotated as alkanes: *m/z* 57 at RI 1521 and *m/z* 71 at RI 1759. The volatile biomarker profile of **BacEx** contained the largest number of markers at 72 hours post inoculation, the first sampling timepoint, and numbers of markers decreased for each timepoint thereafter. There was one compound shared between **Bacteria** and **BacEx**, *m/z* 73 at RI 1912.

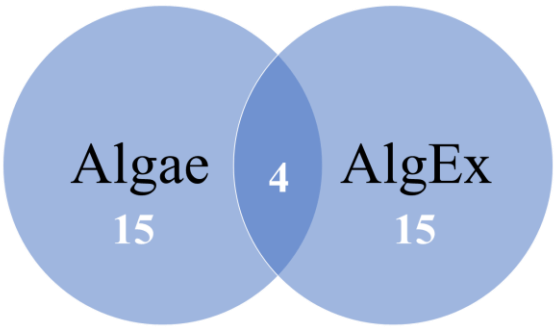


Figure 5.5 Venn diagram of the overlapping VOC biomarkers among **Algae** and **AlgEx** across all measured timepoints. Each group is derived from n=2 replicates. Greatest qualitative overlap occurred during sampling at the 72 hour timepoint.

Alg+BacEx: A total of 49 VOC biomarkers were detected in **Alg+BacEx** cultures (Table A.5.6), a larger number than that produced in the **Algae** or **BacEx** alone. Chemically identified markers included the cyclic olefin ectocarpene, the methyl ketone 2-pentadecanone, 6,10,14-trimethyl-, and the carboxylic acid n-hexadecanoic acid, all of which had been previously observed in co-cultures. 2,5-di-tert-butyl-1,4-benzoquinone was identified at the 72 hour sampling timepoint, identified by mass spectral fragments *m/z* 220, 205, 192, 177, 163 and a RI of 1475. An additional nine biomarkers could be annotated as alkanes (*m/z* fragments at 55 and 71) or carotenoids (*m/z* fragments 177 and 137). While 18 and 24 biomarkers were previously observed in monocultures of **Algae** and **BacEx**, respectively, unique production of 31 VOC biomarkers were observed in **Alg+BacEx**. Overlaps in biomarkers present in the sample types combined to create **Alg+BacEx** are presented in a Venn diagram in Figure 5.6a.

Bac+AlgEx: In **Bac+AlgEx** (Table A.5.7), a total of 29 biomarkers remained after data filtering, a larger number of markers than the sum of numbers of markers found for **Bacteria** and **AlgEx**. None of these 29 biomarkers could be given confident chemical identifications, although

several markers displayed fragmentation patterns consistent with alkanes. Unique production of 22 biomarkers were observed in **Bac+AlgEx** while 7 markers were previously observed in either monocultures of **Bacteria** and **AlgEx**, (Figure 5.6b). Finally, a comparison of the **Co-cultures** against that of **Alg+BacEx** and **Bac+AlgEx** (Figure 5.6c) revealed that a majority of the shared substances were produced by *P. tricornutum* and/or the presence of exudates of *Marinobacter*.

Our second hypothesis that exuded volatile metabolites from one organism can be sensed by and affect the growth of the other organism, demonstrated here for 1) *P. tricornutum* exposed to *Marinobacter* spp. 3-2 exudates and for 2) *Marinobacter* spp. 3-2 exposed to *P.* tricornutum exudates, was supported by results from this study (Table 5.2). A greater complexity in VOC biomarker profiles was observed in **Alg+BacEx** compared to monocultures of **Algae**. However, the biomarker count in **Bac+AlgEx** was comparable to the biomarker count for monocultures of **Bacteria**. **Alg+BacEx** also contained more of the markers observed in **Co-cultures**. The VOC marker profiles are hypothesized to be more similar because the algae cell (average size 26-28 μm length x 2.5 μm width x 2.8 μm height [28]) is much larger in volume compared to the bacterial cell (average size 2-3 μm length x 0.3-0.5 μm diameter); thus, the emission profile of **Co-cultures and Alg+BacEx** are similar because the live algae biomarkers dominate those of *Marinobacter*.

There was a low overlap of biomarker profiles between the following sample types: a) **Algae** vs. **AlgEx** and b) **Bacteria** vs. **BacEx**. Four biomarkers were confidently shared between **Algae** vs **AlgEx** samples while one shared between **Bacteria** and **BacEx** samples, with highest similarity during the 72 hour timepoint. Initially, the biomarker profiles from exudates were hypothesized to have a large overlap with the biomarker profiles of live organisms. As described in the methods, stock cultures were filtered after 144 hours (early stationary phase) and the

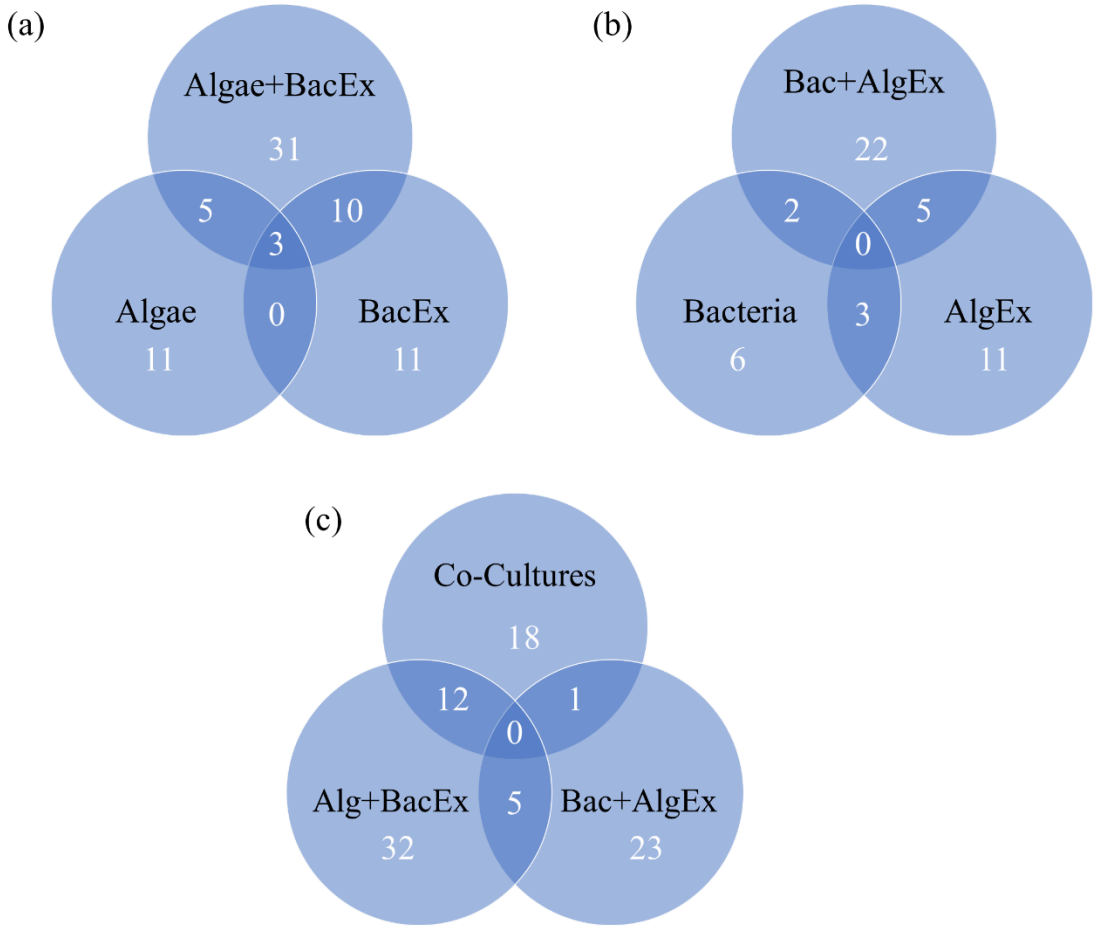


Figure 5.6 Venn diagram of the overlapping VOC biomarkers for organisms exposed to exudates from the respective other species: (a) *P. tricornutum* exposed to *Marinobacter* exudates, (b) *Marinobacter* exposed to *P. tricornutum* exudates. Panel (c) evaluates the overlap of **Co-cultures**, with **Alg+BacEx** and **Bac+AlgEx**. All circles are summation of VOC biomarkers across all measured timepoints. The number of replicates for each sample type can be referenced in Table 5.1.

exudates were stored. With further consideration on the methodology, the vacuum filtration process may have expedited depletion of total VOC levels in the headspace of the filtrate while separating the biomass. Therefore, less abundant and more volatile VOCs may have been inadvertently removed. It is unknown whether the method of filtration and storage introduced VOC contaminants, but should be addressed in future experiments by performing the same

procedures with a media blank control. Subsequent to the filtration process, storage of exudates for an extended period of time may have allowed additional chemistry to take place, e.g.,. secreted non-volatile enzymes in the exudates could have broken down expected biomarkers within the media, even with storage at 20°C. Finally, once the experiments were started, the exudates were not sampled for VOC emissions until 72 hours post-inoculation. Therefore, it is possible that biomarkers of interest were depleted in the media (e.g., evaporation, increased O_2 or light exposure) prior to this time. The absence of living microorganisms from the exudates did not allow VOC markers to be replenished. Experimental evidence for additional chemistry taking place over the SPME sampling periods was seen in the biomarker profiles for **AlgEx** (Table A.5.4), where the largest number of compounds were observed at the 240 hour sampling timepoint. Regardless, the exuded materials from algae and bacteria are comprised of both volatile and non-volatile compounds, including proteins and lipids. Therefore, even though some volatile biomarkers were depleted or degraded prior to sampling, compounds necessary to affect microorganism metabolism in **Alg+BacEx** and **Bac+AlgEx** were hypothesized to be present. Therefore, a concurrent measurement of non-volatile compounds via Liquid Chromatography (LC) – MS would be an appropriate further step to more deeply profile exuded materials. Future experiments should prepare exudate samples from the corresponding microorganism as close as possible to the start of the experiment, preferably the same day.

While the VOC profiles for different growth states of *P. tricornutum* and related alga have been previously reported, to my knowledge, no study has concurrently studied volatile emission of *P. tricornutum* interacting with a bacterial partner. Scarce research has been conducted on the VOC profiles of the genus *Marinobacter*, with a recent study by Lawson et al. (2020) [29] profiling VOCs from *M. adhaerens* HP15 interacting with the exudates from the

algae *Symbiodiniaceae* found to be the only published materials. Using thermal desorption, their work identified VOC biomarkers from *M. adhaerens* to include alcohols, aromatic hydrocarbons, esters, ethers, halogenated hydrocarbons, ketones, and sulfur-containing compounds. While no VOC biomarkers were common between their study and this work, their work observed a greater complexity of bacterial biomarkers in the presence of algal exudates. Similarity to the work presented here underscores the importance of expanding research into phycosphere-related interactions of algae-bacteria and the effects of exchanged metabolites in ecological systems and/or commercial systems (e.g., biofuels industry).

5.4 CONCLUSION

This work presented in this chapter aimed to detect and identify VOC biomarkers related to the micro-scale interactions of a model system of *P. tricornutum* and *Marinobacter* spp. 3-2 and associated cultures. The presence of *Marinobacter* spp. 3-2, either the primary bacterium or its metabolites in form of exudates from bacterial cultures, caused unexpected, modest inhibition in the growth rates of *P. tricornutum*. Substantial differences in VOC biomarker profiles were observed in 1) co-cultures of both organisms, 2) *P. tricornutum* exposed to *Marinobacter* spp. 3-2 exudates, and 3) *Marinobacter* spp. 3-2 exposed to *P. tricornutum* exudates, all relative to the VOC biomarker profiles of corresponding monocultures. Increasing the knowledge base of algae-bacterial interactions at the phycosphere and alterations in microorganism physiology will enable better prediction and/or manipulation of these interactions for commercial purposes as well as a deeper understanding of the basic science of microorganism signaling.

Further experiments are necessary to determine if individual biomarkers or compound classes are consistently indicative of monocultures of algae, bacteria, or algae-bacterial

interactions. Additionally, a changing complexity was observed in the qualitative profiles of each sample type over time and warrants further investigation. This information can in turn provide researchers further details on the transfer of biomarkers between organisms in the phycosphere. Increasing the number of and combinations of algal and bacterial species, and a larger number of experimental replicates, will strengthen and solidify the results presented here. Additionally, a large portion of detected biomarkers could not be chemically identified. Therefore, utilization of more sensitive and/or selective sampling methodologies and volatile detection methods will undoubtedly expand upon the VOC biomarker profiles described in this work. Additionally, future work in this area could explore complementary techniques to SPME-GC-MS, such as the stable isotopic labeling of metabolic products to trace the origins of emitted VOCs. Finally, the approach and technologies developed in this field would also be applicable and transferable to other microbial communities under study, such as cyanobacteria, rhizosphere communities and biofilms.

APPENDIX

Appendix Tables

Table A.5.1 Annotations of VOC biomarkers measured from *P. tricornutum* samples (**Algae**, n=2) detected at several timepoints spanning 240 hours of sample growth

Compound #	Base Peak *m/z*	Tentative Compound Class	NIST 14 ID	NIST % Match	Experimental Retention Index	Day 3	Day 5	Day 7	Day 10
1	44	Amine	2-Propen-1-amine, N-ethyl-N-methyl-	72	773		X		
2	108				1021				X
3	57	**Alkane			1085	X[a]	X[a]		X
4	57	**Alkane			1134			X	
5	91	Cyclic Olefin	6-[(*Z*)-1-Butenyl]-1,4-cycloheptadiene	75	1137	X[a]	X	X[a]	X
6	339				1199	X[a]	X		
7	71	**Alkane			1231		X[a]	X	
8	125				1239			X	
9	71	**Alkane			1271			X	
10	415				1285		X	X[a]	
11	57	**Alkane			1316	X[a]	X[a]	X	
12	401				1423	X[a]	X		
13	135				1473	X			
14	177	**Carotenoid	(β-ionone)		1495				X

Table A.5.1 (cont'd)

15	58	Methyl Ketone	2-Pentadecanone, 6,10,14-trimethyl-	80	1863	X[a]	X	X	X
16	149				1893	X			
17	149				2495	X[a]			X
18	73				2572	X[a]	X	X	X[a]
19	73				2626		X		

X[a]: Detection in 1 of 2 replicates

**: Compound class assigned using characteristic fragmentation pattern

Table A.5.2 Annotations of VOC biomarkers measured from *Marinobacter* spp. 3-2 samples (**Bacteria**, n=2) detected at several timepoints spanning 240 hours of sample growth

Compound #	**Base Peak *m/z***	**Tentative Compound Class**	**NIST 14 ID**	**NIST % Match**	**Experimental Retention Index**	**Day 3**	**Day 5**	**Day 7**	**Day 10**
1	42				815				X
2	68	Terpene	D-Limonene	74	1015				X
3	119				1065	X			
4	57	**Alkane			1085	X[a]			X
5	117				1133	X			
6	95				1151			X[a]	X
7	119				1159	X	X[a]		
8	401				1423	X[a]	X		
9	135				1473	X			
10	57	**Alkane			1816				X
11	73				1912		X	X	X

X[a]: Detection in 1 of 2 replicates

**: Compound class assigned using characteristic fragmentation pattern

Table A.5.3 Annotations of VOC biomarkers measured from *P. tricornutum* and *Marinobacter* spp. 3-2 samples (**Co-cultures**, n=3) detected at several timepoints spanning 240 hours of sample growth

Compound #	Base Peak *m/z*	Tentative Compound Class	NIST 14 ID	NIST % Match	Experimental Retention Index	Day 3	Day 5	Day 7	Day 10
1	42				815				X[b]
2	133				909				X[b]
3	68	Aliphatic	D-Limonene	74	1015	X[b]			
4	108				1021				X[b]
5	99				1028				X[b]
6	91	Cyclic Olefin	6-[(*Z*)-1-Butenyl]-1,4-cycloheptadiene	75	1137	X	X	X	X
7	95				1151			X[a]	X[b]
8	119				1158				X[b]
9	133				1163	X[b]			
10	86	Diol	Dianhydromannitol	73	1195				X[b]
11	137	**Carotenoid	(β-cyclocitral)		1208			X[a]	X
12	71	**Alkane			1271			X[b]	
13	57	**Alkane			1372	X[b]			X[b]
14	55				1473				X[b]
15	177	**Carotenoid	(β-ionone)		1495				X[b]
16	135				1644	X			

Table A.5.3 (cont'd)

17	73				1770				X^b
18	58	Methyl Ketone	2-Pentadecanone, 6,10,14-trimethyl-	80	1863	X^b	X^a	X	X
19	100				1900				X^b
20	73				1912	X^a			X^b
21	71	**Alkane			1962	X^b			
22	55	Carboxylic acid	Hexadecenoic acid (16:1)	80	1963				X^b
23	73	Carboxylic acid	n-Hexadecanoic acid	80	1982				X^b
24	73				2454		X^b		
25	73				2572	X		X^a	
26	73				2572			X^a	X^b
27	208				2678	X^a	X^b		
28	97				2814				X^b
29	248				3026	X^a			X^b
30	97				3057		X^b		
31	248				3304	X^b			

X^a: Detection in 1 of 3 replicates

X^b: Detection in 2 of 3 replicates

**: Compound class assigned using characteristic fragmentation pattern

Table A.5.4 Annotations of VOC biomarkers measured from *P. tricornutum* exudates (**AlgEx**, n=2) detected at several timepoints spanning 240 hours of sample growth

Compound #	Base Peak *m/z*	Tentative Compound Class	NIST 14 ID	NIST % Match	Experimental Retention Index	Day 3	Day 5	Day 7	Day 10
1	79				744			X	
2	41				814				X
3	45				958				X
4	95				1151		X	X[a]	X
5	285				1220				X
6	68				1244	X			
7	71	**Alkane			1271		X[a]	X	
8	57	**Alkane			1290				X
9	92				1340	X	X[a]		
10	154				1372				X
11	401				1424				X
12	135				1473	X			
13	57	**Alkane			1692	X[a]	X[a]		X
14	57	**Alkane			1811	X[a]	X[a]		X
15	73				1912		X[a]	X	X
16	55	Carboxylic Acid	Hexadecenoic acid (16:1)	80	1963				X
17	73	Carboxylic Acid	n-Hexadecanoic acid	80	1982				X

Table A.5.4 (cont'd)

18	73		2572	X^a	X	X^a
19	248		2901			X^a

X^a: Detection in 1 of 3 replicates

X^b: Detection in 2 of 3 replicates

**: Compound class assigned using characteristic fragmentation pattern

Table A.5.5 Annotations of VOC biomarkers measured from *Marinobacter* spp. 3-2 exudates (**BacEx**, n=2) detected at several timepoints spanning 240 hours of sample growth

Compound #	Base Peak *m/z*	Tentative Compound Class	NIST 14 ID	NIST % Match	Experimental Retention Index	Day 3	Day 5	Day 7	Day 10
1	103				976			X	
2	121				981	X	X	X	
3	109				998	X	X	X	X
4	123				1078	X	X	X[a]	
5	267				1101		X	X[a]	
6	127				1126	X			
7	68				1244	X	X	X	X
8	401				1424			X	
9	55				1479	X			
10	57	**Alkane			1521				X
11	216				1617	X			
12	253				1627	X[a]	X	X[a]	X
13	405				1690	X			
14	71	**Alkane			1759	X			
15	194				1889	X	X[a]		
16	73				1912	X[a]			X
17	58				1924	X			

Table A.5.5 (cont'd)

18	71		1945		X		
19	73		2142	X			
20	73		2572	X	X	X[a]	X[a]
21	73		2572				
22	208		2697				X
23	248		3095				X
24	97		3153			X	

X[a]: Detection in 1 of 2 replicates

**: Compound class assigned using characteristic fragmentation pattern

Table A.5.6 Annotations of VOC biomarkers measured from *P. tricornutum* + *Marinobacter* spp. 3-2 exudates (**Alg+BacEx**, n=3) detected at several timepoints spanning 240 hours of sample growth

Compound #	Base Peak *m/z*	Tentative Compound Class	NIST14 ID	NIST % Match	Experimental Retention Index	Day 3	Day 5	Day 7	Day 10
1	45				768		X[b]		
2	69				773			X[b]	
3	42				815	X[a]			X[b]
4	71	**Alkane			816		X[b]		
5	42				945		X[b]		
6	105				963		X[a]	X[a]	X[b]
7	121				981	X	X[a]	X[b]	X[a]
8	57				988	X[b]			X[a]
9	109				998	X	X	X[b]	X[b]
10	108				1021	X[b]			X
11	71				1029			X[b]	
12	123				1078	X	X[b]	X[a]	X[a]
13	57	**Alkane			1085			X[b]	
14	57	**Alkane			1086	X[a]			X[b]
15	91	Cyclic Olefin	6-[(*Z*)-1-Butenyl]-1,4-cycloheptadiene	75	1137	X[b]	X[b]	X[b]	X[b]
16	95				1151			X[b]	X[b]

Table A.5.6 (cont'd)

17	105				1158	X[b]			
18	119				1159			X[a]	X[b]
19	57	**Alkane			1185		X[b]		
20	137	**Carotenoid	(β-cyclocitral)		1208				X[b]
21	68				1244	X	X	X	X
22	122				1253	X[b]	X[b]	X	X
23	71	**Alkane			1271		X[b]	X[b]	
24	57	**Alkane			1290	X[a]			X[b]
25	133				1328	X[b]			X[a]
26	109				1418	X[b]			
27	401				1423	X[a]	X[b]	X[a]	
28	135				1473	X[b]			
29	55				1473		X[a]		X[b]
30	205	Quinone	2,5-di-tert-Butyl-1,4-benzoquinone	86	1475	X[b]			
31	91				1500	X[b]			
32	216				1617	X			
33	405				1690	X[b]			X[b]
34	73				1771				X[b]
35	57	**Alkane			1816				X[b]

Table A.5.6 (cont'd)

36	421				1837				X^b
37	58	Methyl Ketone	2-Pentadecanone, 6,10,14-trimethyl-	80	1863	X^a	X^a	X^a	X^b
38	73				1912			X^a	X^b
39	58				1924	X^b		X^a	X^a
40	71				1945		X^b		
41	71	**Alkane			1948		X^a	X^b	
42	73	Carboxylic Acid	n-Hexadecanoic acid	80	1982	X^a			X^b
43	73				2142	X			
44	135				2472	X^b			
45	73				2572	X^b	X		X^a
46	73				2572	X^a		X	X^b
47	97				2859			X^a	X^b
48	97				3106				X^b
49	97				3159	X^b			X^a

X^a: Detection in 1 of 3 replicates

X^b: Detection in 2 of 3 replicates

**: Compound class assigned using characteristic fragmentation pattern

Table A.5.7 Annotations of VOC biomarkers measured from *Marinobacter* spp. 3-2 + *P. tricornutum* exudates (**Bac+AlgEx**, n=3) detected at several timepoints spanning 240 hours of sample growth

Compound #	Base Peak *m/z*	Tentative Compound Class	NIST 14 ID	NIST % Match	Experimental Retention Index	Day 3	Day 5	Day 7	Day 10
1	42				815	X[b]			X[a]
2	91				947				X[b]
3	103				976			X[b]	
4	121				981	X[b]	X	X[a]	X[a]
5	267				998		X	X[b]	
6	59				1003	X[b]	X[a]		
7	57	**Alkane			1021				X[b]
8	99				1028				X[b]
9	57	**Alkane			1085		X[b]	X[a]	
10	61				1118			X[b]	
11	95				1131			X[b]	
12	68				1244	X[b]	X	X[b]	
13	327				1318				X[b]
14	92				1340	X[b]			
15	57	**Alkane			1348	X[b]			
16	154				1372				X
17	170				1400				X[b]

Table A.5.7 (cont'd)

18	401		1424				X^b
19	343		1452			X^b	X^a
20	91		1476		X^b		
21	405		1634		X^b		
22	236		1676	X^b			
23	57	**Alkane	1692	X^b			X^a
24	71	**Alkane	1759	X^b			
25	58		1924	X^b			X^a
26	73		2142	X	X^a	X^a	
27	97		2610		X^b		
29	208		2689				X^b

X^a: Detection in 1 of 3 replicates

X^b: Detection in 2 of 3 replicates

**: Compound class assigned using characteristic fragmentation pattern

REFERENCES

REFERENCES

1 Amin, S. A., Parker, M. S. & Armbrust, E. V. Interactions between diatoms and bacteria. *Microbiol Mol Biol Rev* **76**, 667-684, doi:10.1128/MMBR.00007-12 (2012).

2 Amin, S. A. *et al.* Interaction and signalling between a cosmopolitan phytoplankton and associated bacteria. *Nature* **522**, 98-101, doi:10.1038/nature14488 (2015).

3 Seyedsayamdost, M. R., Wang, R., Kolter, R. & Clardy, J. Hybrid biosynthesis of roseobacticides from algal and bacterial precursor molecules. *Journal of the American Chemical Society* **136**, 15150-15153, doi:10.1021/ja508782y (2014).

4 Seymour, J. R., Amin, S. A., Raina, J. B. & Stocker, R. Zooming in on the phycosphere: the ecological interface for phytoplankton-bacteria relationships. *Nature microbiology* **2**, 17065, doi:10.1038/nmicrobiol.2017.65 (2017).

5 Le Chevanton, M. *et al.* Screening and selection of growth-promoting bacteria for Dunaliella cultures. *Algal Research* **2**, 212-222, doi:10.1016/j.algal.2013.05.003 (2013).

6 Kim, B.-H., Ramanan, R., Cho, D.-H., Oh, H.-M. & Kim, H.-S. Role of Rhizobium, a plant growth promoting bacterium, in enhancing algal biomass through mutualistic interaction. *Biomass and Bioenergy* **69**, 95-105, doi:10.1016/j.biombioe.2014.07.015 (2014).

7 Wang, H., Hill, R. T., Zheng, T., Hu, X. & Wang, B. Effects of bacterial communities on biofuel-producing microalgae: stimulation, inhibition and harvesting. *Critical Reviews in Biotechnology* **36**, 341-352, doi:10.3109/07388551.2014.961402 (2016).

8 Mayali, X. & Azam, F. Algicidal bacteria in the sea and their impact on algal blooms1. *Journal of Eukaryotic Microbiology* **51**, 139-144, doi:10.1111/j.1550-7408.2004.tb00538.x (2004).

9 Paul, C., Mausz, M. A. & Pohnert, G. A co-culturing/metabolomics approach to investigate chemically mediated interactions of planktonic organisms reveals influence of bacteria on diatom metabolism. *Metabolomics* **9**, 349-359, doi:10.1007/s11306-012-0453-1 (2013).

10 Zerrifi, S. E. A., El Khalloufi, F., Oudra, B. & Vasconcelos, V. Seaweed Bioactive Compounds against Pathogens and Microalgae: Potential Uses on Pharmacology and Harmful Algae Bloom Control. *Mar Drugs* **16**, 55, doi:10.3390/md16020055 (2018).

11 Yang, Y. H. *et al.* Fatty acid and lipid class composition of the microalga Phaeodactylum tricornutum. *Journal of oleo science* **66**, 363-368, doi:10.5650/jos.ess16205 (2017).

12 Gauthier, M. J. *et al.* Marinobacter hydrocarbonoclasticus gen. nov., sp. nov., a new, extremely halotolerant, hydrocarbon-degrading marine bacterium. *International Journal*

of Systematic and Evolutionary Microbiology **42**, 568-576, doi:10.1099/00207713-42-4-568 (1992).

13 Reese, K. L. *et al.* Chemical profiling of volatile organic compounds in the headspace of algal cultures as early biomarkers of algal pond crashes. *Scientific Reports* **9**, 13866, doi:10.1038/s41598-019-50125-z (2019).

14 Pohnert, G. & Boland, W. The oxylipin chemistry of attraction and defense in brown algae and diatoms. *Natural product reports* **19**, 108-122, doi:10.1039/a806888g (2002).

15 Prestegard, S. K. *et al.* Specific metabolites in a Phaeodactylum tricornutum strain isolated from western norwegian fjord water. *Mar Drugs* **14**, 9 (2016).

16 Lupette, J. *et al.* The architecture of lipid droplets in the diatom Phaeodactylum tricornutum. *Algal Research* **38**, 101415, doi:10.1016/j.algal.2019.101415 (2019).

17 Breuer, G., Lamers, P. P., Martens, D. E., Draaisma, R. B. & Wijffels, R. H. The impact of nitrogen starvation on the dynamics of triacylglycerol accumulation in nine microalgae strains. *Bioresource technology* **124**, 217-226, doi:10.1016/j.biortech.2012.08.003 (2012).

18 Liang, Y., Beardall, J. & Heraud, P. Effects of nitrogen source and UV radiation on the growth, chlorophyll fluorescence and fatty acid composition of Phaeodactylum tricornutum and Chaetoceros muelleri (Bacillariophyceae). *Journal of photochemistry and photobiology. B, Biology* **82**, 161-172, doi:10.1016/j.jphotobiol.2005.11.002 (2006).

19 Cezare-Gomes, E. A. *et al.* Potential of microalgae carotenoids for industrial application. *Applied Biochemistry and Biotechnology* **188**, 602-634, doi:10.1007/s12010-018-02945-4 (2019).

20 Abdel-Aal, E. I., Haroon, A. M. & Mofeed, J. Successive solvent extraction and GC–MS analysis for the evaluation of the phytochemical constituents of the filamentous green alga Spirogyra longata. *The Egyptian Journal of Aquatic Research* **41**, 233-246, doi:10.1016/j.ejar.2015.06.001 (2015).

21 Kamenarska, Z. *et al.* Chemical composition and biological activity of the brackish-water green alga Cladophora rivularis (L.) Hoek. **47**, 215, doi:10.1515/BOT.2004.022 (2004).

22 Kamenarska, Z. *et al.* Chemical composition of the brown alga Padina pavonia (L.) Gaill. from the Adriatic sea. *Botanica Marina* **45**, 339-345, doi:10.1515/BOT.2002.034 (2002).

23 Gobalakrishnan, R. Antimicrobial potential and bioactive constituents from aerial parts of Vitis setosa Wall. *Journal of Medicinal Plants Research* **8**, 454-460, doi:10.5897/JMPR2013.5356 (2014).

24 Dodson, V. J. & Leblond, J. D. Now you see it, now you don't: differences in hydrocarbon production in the diatom Phaeodactylum tricornutum due to growth

temperature. *Journal of Applied Phycology* **27**, 1463-1472, doi:10.1007/s10811-014-0464-6 (2015).

25 Srikanta Dani, K. G. *et al.* Relationship between isoprene emission and photosynthesis in diatoms, and its implications for global marine isoprene estimates. *Marine Chemistry* **189**, 17-24, doi:10.1016/j.marchem.2016.12.005 (2017).

26 Wolfram, S. *et al.* A Metabolic probe-enabled strategy reveals uptake and protein targets of polyunsaturated aldehydes in the diatom Phaeodactylum tricornutum. *PLoS One* **10**, e0140927, doi:10.1371/journal.pone.0140927 (2015).

27 Spielmeyer, A. & Pohnert, G. Influence of temperature and elevated carbon dioxide on the production of dimethylsulfoniopropionate and glycine betaine by marine phytoplankton. *Marine environmental research* **73**, 62-69, doi:10.1016/j.marenvres.2011.11.002 (2012).

28 Renzi, M., Roselli, L., Giovani, A., Focardi, S. E. & Basset, A. Early warning tools for ecotoxicity assessment based on Phaeodactylum tricornutum. *Ecotoxicology (London, England)* **23**, 1055-1072, doi:10.1007/s10646-014-1249-z (2014).

29 Lawson, C. A., Seymour, J. R., Possell, M., Suggett, D. J. & Raina, J.-B. The volatilomes of Symbiodiniaceae-associated bacteria Are influenced by chemicals derived from their algal partner. *Frontiers in Marine Science* **7**, doi:10.3389/fmars.2020.00106 (2020).

Chapter 6: CONCLUSIONS AND BROADER IMPACTS

From biosecurity to biodefense, SPME-GC-MS has proven useful for profiling volatile biomarkers emitted from actively growing microbial cultures. The presence and physiological growth states for monocultures of *Bacillus anthracis* and *Francisella tularensis* in optimal growth media were reflected in changes in the corresponding headspace composition. VOCs diagnostic of the health states of biofuel-relevant algae were detected in the contexts of stress/predator grazing and for interactions with neighboring bacteria, both with the end goal of improving the yields of biomass from algae cultures. However, there are additional aspects to consider before the results of this work could be implemented.

While the data acquisition and data processing employed throughout this dissertation has proven capable of VOC marker collection, separation, detection, and annotation, a number of biomarkers could not be identified compared to known entries in the NIST14 mass spectral database. An inability to identify these biomarkers may be a result of poor spectral quality for either the experimental data or the reference library, though it is also likely that many VOCs are novel compounds yet to be identified or added to spectral databases. Standardization of sample collection and data processing practices within the respective fields would facilitate inter-laboratory comparisons, thus increasing success rates in de-replicating whether biomarkers observed in this research have been previously observed. In a similar vein, exploring the use of a high-resolution mass spectrometer could improve confidence in annotations of unknown compounds' identities. Determination of exact mass for molecular and/or fragment ions would allow estimation of chemical formula, even if fragmentation patterns do not match up with empirical databases or spectra predicted for unknowns *in silico*. Moreover, co-eluting

compounds with the same nominal mass but different chemical formulas (isobars) could be distinguished.

Generalization of VOCs as diagnostic biomarker profiles for a given taxa requires investigation of a wider range of model microbes. In the research of biological warfare agents, defined surrogate organisms are used in place of a target pathogen due to safety requirements as well as strict regulations and laws. As seen in Chapter 3, VOC biomarker profiles were more similar for fully virulent Risk Group 3 (RG3) *Bacillus anthracis* Ames and its surrogate Risk Group 2 (RG2) *Bacillus anthracis* Sterne compared to RG3 *Francisella tularensis* SCHU S4 and its surrogate RG2 *Francisella tularensis* novicida. Future VOC profiling should focus on the predictive strength of profiles generated for RG2 surrogates, in particular those RG2 species that closely reproduce the phenotypes of RG3 virulent pathogens. The use of surrogates has implications in the protocols and personal protective equipment (PPE) used to safely handle different microorganisms. If suitable surrogates are not available, the protocols used in this work for the safe transfer of materials from BSL-3 to BSL-2 laboratories, which preserved key metabolites while avoiding destructive decontamination procedures, would be useful to BSL-3 practitioners. Additional studies into the effects of different cellular environments on pathogen VOC profiles, in particular for growth conditions mimicking human-derived materials (e.g., cell lines), should be investigated. For algal research, better understanding of the relationship between volatile emissions, growth rates, and lipid production in different biofuel-relevant algae species would allow commercial producers to tailor biomass output. Different types of stress on algal ponds should be investigated to determine how various triggers activate intrinsic metabolic pathways. Finally, research into changes in algal metabolism for different algal genotypes, both with and without the presence of bacteria, should be conducted, as some combinations may

produce protective measures (e.g., combat the actions of a predator, etc.). Analysis from a larger number of algal taxa should explore whether diagnostic markers are specific to a given species or more generalizable to a genus. Experimental results can be validated using analytical standards to confirm the identities of biomarkers detected via mass spectrometry for inclusion in targeted methodologies.

Future work should explore metabolic profiling using conditions more analogous to real-world applications. Experiments in laboratory settings are the most easily controlled, where conditions can be well-defined and any extraneous influences can be minimized. Translating that knowledge to microorganisms in their natural habitat or in large-scale, commercial production, where conditions are ill-defined and can be externally contaminated, can provide researchers with selective and specific tools to target unhealthy ponds. Experiments are necessary to determine if individual VOC markers or their compound classes are consistently indicative of the pathogens, algae, or bacteria studied across this work. Subsequently, knowledge gained and biomarkers annotated from untargeted metabolomics efforts may guide development of targeted metabolite monitoring for changes in diagnostic chemical signatures and detecting volatile signals of metabolic changes in real-time. Miniaturized GC-MS systems for field deployable detector systems provide one such technology currently under development. Applications utilizing such systems would require initial validation of accuracy and reproducibility of qualitative and quantitative analyses, for both analytical standards and for complex mixtures.

The results presented in this dissertation present further research avenues and have implications for the wider research communities with interests in microbial detection and volatilomics. Although the methodology based on SPME-GC-quadrupole MS in scanning mode does not represent the pinnacle of sensitivity, a number of relatively robust and reproducible

putative volatile biomarkers could be detected in each study. Non-invasive monitoring of VOC biomarkers from the headspace of liquid cultures allows direct sampling with minimal perturbation of the culture. While the methods here applied to the headspace of liquid cultures, this method could also apply for the headspace of other microbial communities, such as biofilms, cyanobacteria, and rhizosphere communities.