The Pennsylvania State University

The Graduate School

EXPOSOMICS BASED METHODOLOGIES FOR THE IDENTIFICATION AND CORRELATION OF ORGANIC CONTAMINANTS CONTRIBUTING TO DISEASE IN SMALLMOUTH BASS POPULATIONS

A Dissertation in

Biochemistry, Microbiology, and Molecular Biology

by

Paige D. Teehan

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December 2021

The dissertation of Paige D. Teehan was reviewed and approved by the following:

Frank Dorman
Special Member
Dissertation Co-Advisor

Andrew Patterson
Professor of Molecular Toxicology, and Biochemistry and Molecular Biology
Dissertation Co-Advisor
Chair of Committee

Philip Bevilacqua
Distinguished Professor of Chemistry, and Biochemistry and Molecular Biology
Head of the Department of Chemistry

William Burgos
Professor of Civil and Environmental Engineering

Ross Hardison
T. Ming Chu Professor of Biochemistry and Molecular Biology
Associate Director of the Genome Sciences Institute

Wendy Hanna-Rose
Professor of Biochemistry and Molecular Biology
Head of the Department of Biochemistry and Molecular Biology

ABSTRACT

Since 2005, smallmouth bass (*Micropterus dolomieu*) throughout the Susquehanna River Basin have been observed exhibiting signs of disease; these observations include intersex adult males, high/chronic mortality in young-of-year fish, and generalized immunosuppression characterized as a single fish being co-infected with multiple pathogens. To date, no single factor or group of factors has been identified as a causative agent to explain the observed signs of disease, but environmental contaminants are suspected of being a contributing influence. Previous targeted studies have reported differing levels of contaminants, including polychlorinated biphenyls (PCBs), polybrominated diphenyl ethers (BDEs), legacy and current use pesticides, and pharmaceuticals and personal care products (PPCPs) in adult fish from sites throughout the river basin. Here, young-of-year fish are examined for both target and non-target compounds using an exposomics based approach to identify statistically significant features among populations experiencing disease traits.

Previous investigations have focused largely on the identification of target compounds within adult fish or water samples, but this targeted investigative approach provides limited information about the samples. The relationship between environmental conditions, contaminants, pathogens, and the health of living organisms throughout their lifetime is complex and requires a more in-depth analysis than identifying target compounds in adults. By examining young-of-year fish, it is possible to identify exposures that are occurring during developmental life stages that may contribute to immunosuppression and intersex conditions later in life. And by expanding the analysis from a strict target list to any extractable non-target compound, it is possible to identify contaminants not included on the target list, new contaminants of emerging concern, and even biomarkers that are associated with disease profiles.

Analyzing a complex matrix like whole young-of-year fish for both target and non-target compounds, required advancements in current sample preparation and data analysis techniques. Modifications to the quick, easy, cheap, effective, rugged, and safe (QuEChERS) extraction method have yielded a high-throughput sample preparation technique that is quicker, cheaper, and greener than current EPA methods for tissue analysis, while remaining suitable for both targeted and non-targeted analyses of fish tissue. This sample preparation has been combined with comprehensive two-dimensional gas chromatography coupled with time-of-flight mass spectrometry (GC×GC-TOFMS) as well as ultra-performance liquid chromatography coupled with high-resolution quadrupole time-of-flight mass spectrometry (UPLC-QTOF-MS) to analyze young-of-year smallmouth bass from fourteen affected sites throughout the Susquehanna River Basin. Sample data sets were then processed to identify target and non-target compounds. Target compounds were quantified using internal standard normalized external calibration curves, while significant non-target compounds were identified using multivariate and univariate data analyses. The resulting compounds and features were then examined for correlations among disease characteristics or locations.

A total of 146 young-of-year smallmouth bass were collected by electrofishing from 14 sampling sites throughout the Pennsylvania region of the Susquehanna River Basin, USA in 2015. The modified sample preparation methodology was used for the extraction and cleanup of all samples. Cleaned sample extracts were divided for analysis using both gas and liquid chromatography platforms. Targeted analyses using GC×GC-TOFMS and UPLC-QTOF-MS have revealed the presence of PCBs, BDEs, pesticides, plasticizers, polycyclic aromatic hydrocarbons, PPCPs, and perfluoroalkyl substances. Target compounds were analyzed for correlations with the disease characteristics observed at the time of young-of-year smallmouth

bass collection. Several slight to moderate correlations were observed, but few were statistically significant.

Non-targeted analyses were designed to elucidate differences between healthy fish and fish displaying outward signs of disease (lesions, parasitic infections, gill erosion, etc.) as well as site-specific features. The same data sets used for the targeted analyses, were mined for statistically significant non-targets. Chromatographic alignment was performed using commercially available software specific to either GC×GC or UPLC applications that produced aligned peak tables of normalized areas. For GC×GC-TOFMS data, Fisher ratio analysis was combined with principal component analyses or partial least squares discriminant analyses for feature reduction. This resulted in two final datasets, one for disease classification and one for collection sites, of tentatively identified, significant, non-target compounds. Comparisons between these non-target compounds revealed three chemical classes that were more abundant in fish displaying signs of disease and four compounds that were identified as significant in both datasets. For UPLC-QTOF-MS data, this optimized data reduction strategy was also applied, but with different results. No compounds or chemical classes were revealed to be significant to specific classifications of fish, either by disease state or collection site. Instead, the non-targeted analysis tentatively identified several pesticides, hormones, steroids, and pharmaceutical compounds not previously reported in smallmouth bass tissue.

This work has produced a high-throughput sample preparation, analysis, and data processing methodology that can be applied to large environmental sample sets. While targeted analyses continue to be the default method of environmental monitoring efforts, the ongoing discovery of contaminants of emerging concern illustrates the need for robust non-targeted analysis methods as well. Chapter 2 has been published in *Analytical Methods*. Chapter 3 has

been published in *Science of the Total Environment*. Data from Chapter 4 is currently in preparation for submission as a manuscript. I am first author on all these publications.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Meaning
ANOVA	analysis of variance
BDE	polybrominated diphenyl ether
CCV	continuing calibration verification
CEC	contaminant of emerging concern
DEET	N,N'-diethyl toluamide
DEHP	bis(2-ethylhexyl) phthalate
dSPE	dispersive solid phase extraction
EDC	endocrine disrupting compound
EPA	United States Environmental Protection Agency
ES+/ES-	electrospray ionization mode (+/-)
EtOAc	ethyl acetate
FAME	fatty acid methyl ester
FDA	United States Food and Drug Administration
GC	gas chromatography
GC×GC	comprehensive two-dimensional gas chromatography
HPLC	high-pressure/high-performance liquid chromatography
HR	high resolution
LC	liquid chromatography
LOQ	limit of quantitation
MDL	method detection limit
MeCN	acetonitrile
MS	mass spectrometry
ND	not detected (the analyte was not detected during chromatographic analysis)
NIST	National Institute of Standards and Technology
NR	not reported (the analyte was detected at a level below the LOQ or at a comparable level within the matrix blank sample)

OCP	organochlorine pesticide
PAH	polycyclic aromatic hydrocarbons
PCA	principal component analysis
PCB	polychlorinated biphenyl
PFAS	per- and polyfluoroalkyl substances
PFOS	perfluorooactanesulfonate
PLS-DA	partial least square discriminant analysis
POP	persistent organic pollutant
PPCP	pharmaceutical and personal care products
PSA	primary secondary amine
QC	quality control
qPCR	quantitative polymerase chain reaction
QTOF-MS	quadrupole time-of-flight mass spectrometry
QuEChERS	quick, easy, cheap, effective, rugged, and safe extraction methodology
SMB	smallmouth bass
sPLS-DA	sparse partial least squares discriminant analysis
TIC	total ion chromatogram
TOFMS	time-of-flight mass spectrometry
t_R	retention time
t_{R1}	first-dimension retention time
t_{R2}	second-dimension retention time
UPLC/UHPLC	ultra-performance liquid chromatography/ultra-high performance liquid chromatography
USGS	United States Geological Survey
VIP	variable importance in projection score
YOY	young of year

ACKNOWLEDGEMENTS

There are many people I would like to thank for supporting me throughout the years and helping me make it through my graduate career. First, thank you to Dr. Frank Dorman for taking me on as a student even though I had no idea what chromatography was or how it worked the first time I walked into your office. I have learned so much since then about both science and life from you. I am very grateful for all the opportunities you have provided me, from getting first-hand experience with brand-new technologies to meeting and growing an amazing network of wonderful scientists.

Thank you to the United States Geological Survey and Pennsylvania Sea Grant for making this research possible. I would like to thank Dr. Vicki Blazer and Dr. Megan Schall for all their help and guidance throughout my graduate career. I would like to thank my colleagues in the separation science community as well. Thank you to everyone I have had the pleasure of working with at LECO for always taking the time to answer my many questions and offering advice for problem solving when I needed it. And to Dr. Kari Organtini at Waters, thank you for making sure that despite a global pandemic I was still able to finish my research.

I would like to thank all the former Dorman lab members who have helped me become a better scientist and who have become great friends, especially Dr. Kyra Murrell and Dr. Conner Stultz. Kyra, I know I would not be where I am today without you and Miguel in my life. Thank you for always being there for me even after you moved far away. Conner you always encouraged me and pushed me to be better, do more, and to not stay in my comfort zone, even if that meant literally pushing me up a mountain. I'm so glad to have met you. And to my friends, Jimmy, Caitlyn, and Katie, thank you for all the fun times, workouts, walks, talks, and food that always seemed to happen right when I needed them most.

Dr. Zhicheng Dou, you are the reason I went to graduate school. Thank you for giving me the chance to learn how to be a scientist and the encouragement to pursue it.

To my family, I'm not sure words can express how much I appreciate everything you have done for me. Mom, you have always encouraged me to be the best I could in whatever I did. Travis and Miles, though we don't always see eye to eye on things, I know that no matter what you will always be there for me, through the good times and the bad. Bhavesh, I really appreciate everything you have done for me and the rest of the family since becoming a member. Thank you all for your endless support and for never letting me give up.

This dissertation is dedicated to the memory of Cubby. My sweet boy, you are dearly missed.

Chapter 1

Introduction

1.1 Anthropogenic Contamination in Aquatic Environments

Across the world, the harmful effects of anthropogenic contamination can be seen in aquatic environments. Anthropogenic contaminants are pollutants derived from human activities and are sourced from a variety of industries, such as manufacturing, agriculture, and healthcare. They can take the form of legacy persistent organic pollutants (POPs, e.g., aldrin, polychlorinated biphenyls [PCBs]), newly discovered chemicals of emerging concern (CECs, e.g., per- and polyfluoroalkyl substances [PFAS]), or physical debris (e.g., plastics).[1–4] The negative effects of many POPs have been known for decades, while negative effects associated with CECs are continually being uncovered through *in vitro*, *in vivo*, and *in silico* laboratory studies.[5–8] Environmental matrices are routinely monitored for known contaminants that may pose a threat to living beings, and flora and fauna may also be measured directly during exposure assessments.[9] Some organisms react strongly to changes within their environment and have been dubbed as indicator species. These species are monitored for changes in health/behavior to detect potential negative environmental changes, such as contamination events.[10]

Throughout recent years aquatic life has been increasingly impacted by anthropogenic contaminants and numerous studies have documented characteristics such as endocrine disruption,[11,12] immunosuppression,[13,14] and morphological changes[15,16] in a variety of fish species. Contaminant exposures have also been shown to alter fish behaviors. Male fathead minnows (*Pimephales promelas*) exposed to PCBs through the course of their life exhibited a reduction in paternal offspring care that may have been responsible for increased embryo

mortality.[17] Turquoise killifish (*Nothobranchius furzeri*) exposed to the antidepressant fluoxetine showed increased mating frequency and male sociability.[18] The timing of an exposure event during a fish's lifecycle is also important. Zebrafish (*Danio rerio*) exposed to estradiol during or before sex differentiation (embryo-larvae or juvenile fish) developed altered sex ratios with an increase in the number of female fish as well as the formation of a female-like retrogonadal cavity in presumptively male fish. Adult male zebrafish exposed to the same concentrations of estradiol exhibited modified secondary sexual characteristics and developed uro-genital papillae (a female-specific sexual characteristic). Development of uro-genital papillae was not seen in any male embryo-larvae or juvenile fish after sexual maturation.[19]

Dose-response studies are vital to understanding the impact of anthropogenic contaminants on exposed organisms, but unlike many laboratory investigations, environmental contamination events are complex involving multiple stressors, some of which may be unknown at the time of study. For example, male rainbow darters (*Etheostoma caeruleum*) from the Grand River in southern Ontario, Canada were experiencing intersex rates between 70 – 100% directly downstream of two wastewater treatment plants. In 2013, the wastewater treatment plants underwent upgrades and immediately the rates of intersex among male darters decreased to < 10% within the same areas. The cause of intersex in these populations pre-upgrade remains unknown, but post-upgrade measured levels of pharmaceuticals, ammonia, and estrogenic compounds decreased suggesting an improvement in the overall quality of the effluent waters.[20] It is unlikely that the initial cause of intersex in these populations will be uncovered as it is nearly impossible to replicate the necessary environmental conditions in a laboratory setting. For these rainbow darters, the upgrades to the wastewater treatment plants may have preserved the population, but for many fish where the cause of disease is unknown remediation is not possible. It is therefore crucial to continue developing better methodologies for the detection and monitoring of environmental contributors to disease.

1.2 Smallmouth Bass

While anthropogenic contamination contributes to negative health impairments of aquatic life on a global scale, the work presented here focuses on the challenges faced by smallmouth bass (*Micropterus dolomieu*, SMB) within the Chesapeake Bay Watershed, USA, with an emphasis on fish from the Susquehanna River Basin region.

A member of the Centrarchidae or sunfish family, SMB are closely related to other black bass species, the largemouth bass (*Micropterus salmoides*) and spotted bass (*Micropterus punctulatus*). The three can be distinguished from each other by their distinct size, coloration, and jaw positions. Largemouth bass are the largest and usually have a horizontal black stripe or broken splotches running from head to tail. The upper jaw of largemouth bass extends past the back edge of the eye and the lower jaw protrudes out. Spotted bass are smaller but similar in color to largemouth bass. They have additional spots above and below a dark horizontal stripe running from head to tail. The upper jaw of spotted bass does not extend past the eye and the upper and bottom jaws are almost equal. SMB coloration varies depending on environment and can range from dark brown or olive green to bronze on the dorsal area (giving rise to the nickname "bronzeback"). SMB have lighter sides with dark vertical bands as opposed to the horizontal bands seen in other black basses. They typically have white to pale yellow undersides. The upper jaw of SMB is in line with the eye and does not extend past the back edge. Juvenile SMB are similar in color to adults with more pronounced vertical bands and an orange caudal fin. There are no obvious sexual dimorphisms present in SMB, though females are generally larger than male fish of the same age.[21,22]

Depending on the location, SMB spawn during the spring and summer months when water temperatures reach 15 – 16 ℃. Male SMB form depressions in mud or firm sand for females to deposit eggs within. Males will guard nests until the eggs hatch and remain guarding

the fry until they disperse, approximately 3 – 4 weeks post spawning. SMB are freshwater predators. Larval SMB feed on zooplankton and aquatic insects as they grow larger. Adult SMB eat a diet composed of mostly insects, crayfish, and other smaller fishes. Their fusiform bodies allow for excellent acceleration and swift swim speeds in open water. It also makes SMB a prized sport fish for many anglers.[21,22]

SMB are not native to the Chesapeake Bay Watershed but have inhabited the area for over 100 years. Native to the Mississippi River Basin and Great Lakes, it was their popularity as sport fish that led to state and federal commissions stocking the Chesapeake Bay Watershed with SMB around the mid-19th century. According to a 2013 Chesapeake Bay Foundation report, "Angling for Healthier Rivers," SMB fishing accounts for $630 million annually across Pennsylvania, Maryland, Virginia, and West Virginia.[23] Although SMB are not native to the watershed, they are not considered an invasive species. Invasive species cause harmful levels of competition for resources in the area and ultimately lead to the death or displacement of native species. SMB are instead considered to be an exotic species that exist in harmony with the native populations. Additionally, SMB serve as indicator species for water quality and ecosystem health, as they are extremely sensitive to minute changes in the environment.[21,23] It is their designation as an indicator species and a series of observed health anomalies persisting over the last two decades that has made them the subject of this work.

1.2.1 Smallmouth Bass Disease in the Chesapeake Bay Watershed

The first indications of problems with the health of SMB within the Chesapeake Bay Watershed were documented in the Potomac River Basin in 2002[24] and the Susquehanna River Basin in 2005.[25,26] In 2002, a major mortality event affected multiple species of fish, including adult SMB, in the South Branch River of the Potomac River Basin. During this event, a wide

variety of dermal lesions were noted on the affected fish (Figure 1.1). Areas of the Potomac River Basin continued to suffer chronic low levels of mortality of SMB every spring, punctuated by further major mortality events in select years and areas. The rate of skin lesions varied with little consistency in appearance or severity between species or geographical locations. In SMB, skin lesions were observed in 20 – 100% of fish sampled during mortality events.[24]

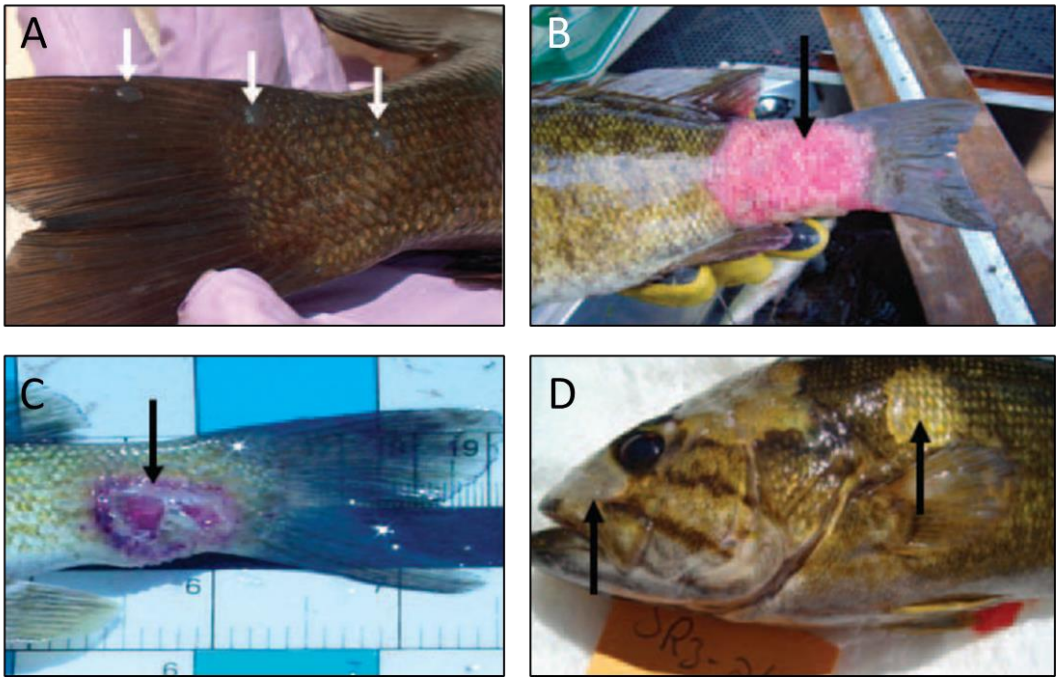

Figure 1.1 Examples of external lesions on SMB observed during mortality events: A) raised, pale lesions, B) large, reddened area, C) deep erosion extending into muscle, D) pale areas with patches of fungal hyphae. Adapted from Blazer et al.[24]

The Susquehanna River Basin experienced similar fish kills beginning in 2005, but mainly affecting young-of-year (YOY, age = 0) SMB instead of adults as was previously seen in the Potomac River Basin events. YOY SMB in the West Branch Susquehanna, Susquehanna, and Juniata rivers were observed with lesions and eroded fins while swimming weakly at the water's surface or deceased. Unlike in the Potomac River Basin events, other species of fish were largely

unaffected. These episodes have occurred annually from 2005 with varying severity, but similar occurrences have not been documented in adjacent Pennsylvania river basins.[25,26] In addition to the observed mortality and dermal lesions, adult male SMB have been exhibiting signs of endocrine disruption, characterized by the production of testicular oocytes and increased levels of plasma vitellogenin, both markers of intersex. Intersex male SMB have been documented in both the Potomac and Susquehanna River Basins.[27–30]

Investigations into the causes of mortality, lesions, and intersex observed in SMB have uncovered a variety of pathogens and endocrine disrupting compounds (EDCs), but to date, no unifying causative agent(s) has been identified that explains the similar observed signs of disease within either the Potomac or Susquehanna River basins. Microbial investigations of external lesions on YOY SMB resulted in the culturing of multiple bacteria, including *Flavobacterium columnare*, *Aeromonas veronii* bv. *sobria*, *A. hydrophila*, *A. popoffi*, *A. allosaccharophila*, *A. schubertii*, *A. caviae*, *Plesiomonas shigelloides*, *Citrobacter* species, and *Pseudomonas* species. Several of these bacteria were also cultured from internal tissue homogenates of the kidney, spleen, or swim bladder. Largemouth bass virus (*Ranavirus*, Iridoviridae) was also detected in YOY SMB from the Susquehanna River Basin using qPCR.[31] Further investigation into largemouth bass virus has shown that YOY SMB are not only susceptible to this virus, but that co-infection of the virus with a bacterial infection, such as *F. columnare* or *A. salmonicida*, increases the mortality rates of infected fish.[32] The parasite *Myxobolus inornatus* was documented in SMB for the first time in fish from the Susquehanna River.[33] Generally, fish in the wild do not incur infection by multiple pathogens at one time. The observed co-infections by multiple bacteria, parasites, and/or viruses in YOY SMB and their apparent inability to clear infections is suggestive of an overall immunosuppression condition of unknown origin.

While it is known that some anthropogenic contaminants can cause endocrine disruption and immunosuppression, investigations into chemical contaminants that may play a role in

smallmouth bass disease have not revealed a likely chemical or group of chemicals that explain the observed morbidity throughout the region. These investigations have focused on detecting specific sets of target compounds within water, sediment, and adult SMB tissue samples.[27,34,35] The use of targeted analyses is standard in environmental monitoring, but they limit the amount of information available to researchers. Therefore, even if the identified environmental contaminants can be correlated with the observed signs of disease, it may still only be a partial explanation of the situation as all other non-targeted factors are ignored. This issue is amplified in environmental samples over those from a controlled laboratory setting, as unknown or unaccounted for factors may be interacting in unique ways to produce the observed health characteristics. It is likely due to these complex interactions between unknown and known environmental stressors that a causative agent(s) has not been identified for the current disease characteristics in the Susquehanna River Basin. A broader investigative approach is needed to discern more about these complex interactions and make connections between different populations exhibiting the same signs of disease.

1.3 Exposomics

Exposomics is the study of environmental influences on organism health. The term exposome was first introduced by Christopher Wild in 2005 as a complement to genomics research. The exposome represents the sum of all exposures (external and internal) undergone by an individual from the point of conception until death and how those exposures relate to an individual's health.[36,37] The exposome is composed of three different domains: the general external environment, the specific external environment, and the internal environment. While defined as distinct domains, there is some overlap between the three. Starting at the level of the individual, the internal environment includes internal processes and reactions, such as

metabolism, microbiota, oxidative stress, and ageing. Specific external environmental factors include lifestyle choices (e.g., diet, alcohol use, exercise), occupation, and exposures to infectious pathogens and chemical contaminants. The general external environment includes broader socioeconomic factors, such as education, financial status, and climate.[38] Though not a comprehensive list, the aforementioned highlight the diversity of factors capable of influencing an individual's health.

Capturing the totality of exposures for a single individual is a daunting task, and as omics studies rely on numerous individuals the difficulty of this challenge only increases in application. Additionally, the exposome spans the entire range of the individual's life, therefore these factors are continuously changing from infancy through adulthood. There are two main approaches for exposomics studies, "bottom-up" and "top-down" (Figure 1.2). In bottom-up exposomics, researchers identify the important exogenous exposures by sampling their sources. This is an approach often used in environmental monitoring studies where samples (e.g., air, water, soil, food, etc.) are analyzed to identify hazardous contaminants.[25,39–41] In a top-down exposomics approach, the individuals are sampled (e.g., cells, tissues, body fluids) to identify important exposures, including endogenous exposures missed when using a bottom-up approach. However, using a top-down approach would not provide the insight about potential sources of exogenous exposures that a bottom-up approach would.[37]

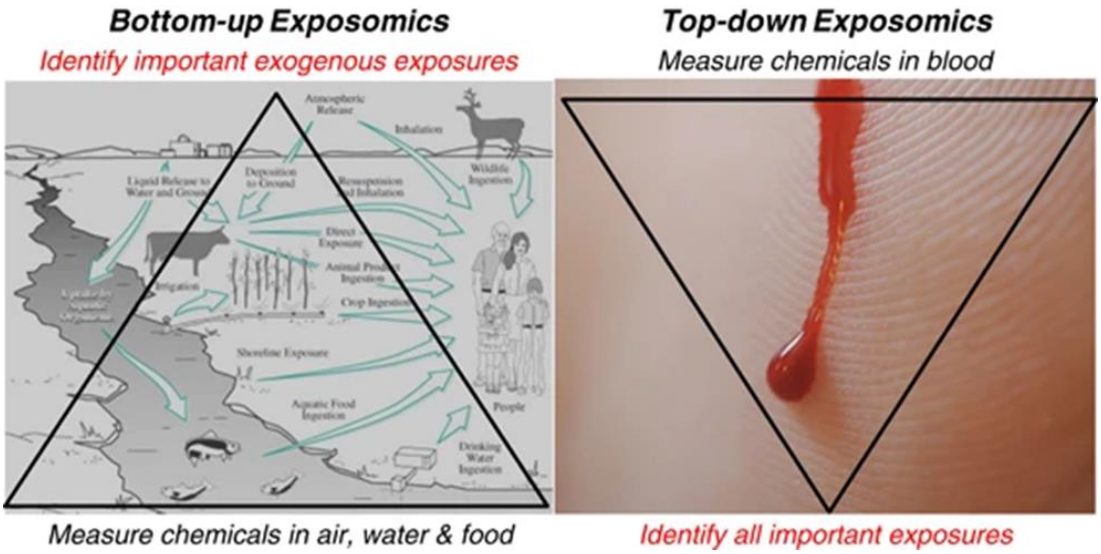

Figure 1.2 An overview of two approaches to exposomics. Adapted from Rappaport, S.[37]

1.3.1 Exposomics of Smallmouth Bass Disease

Exposomics has mainly been used in the context of humans, but this comprehensive approach to health can be applied to any species. Figure 1.3 illustrates how the human exposome can be mirrored for SMB, by reducing the specific external environment to the immediate habitat of SMB, broadening the general external environment to include surrounding land use, and maintaining similar approaches for the internal environment. Previous environmental monitoring efforts employing targeted analyses of water and sediment can be used to form the base of the bottom-up exposomics analysis. The direct analysis of fish tissues establishes a top-down approach.

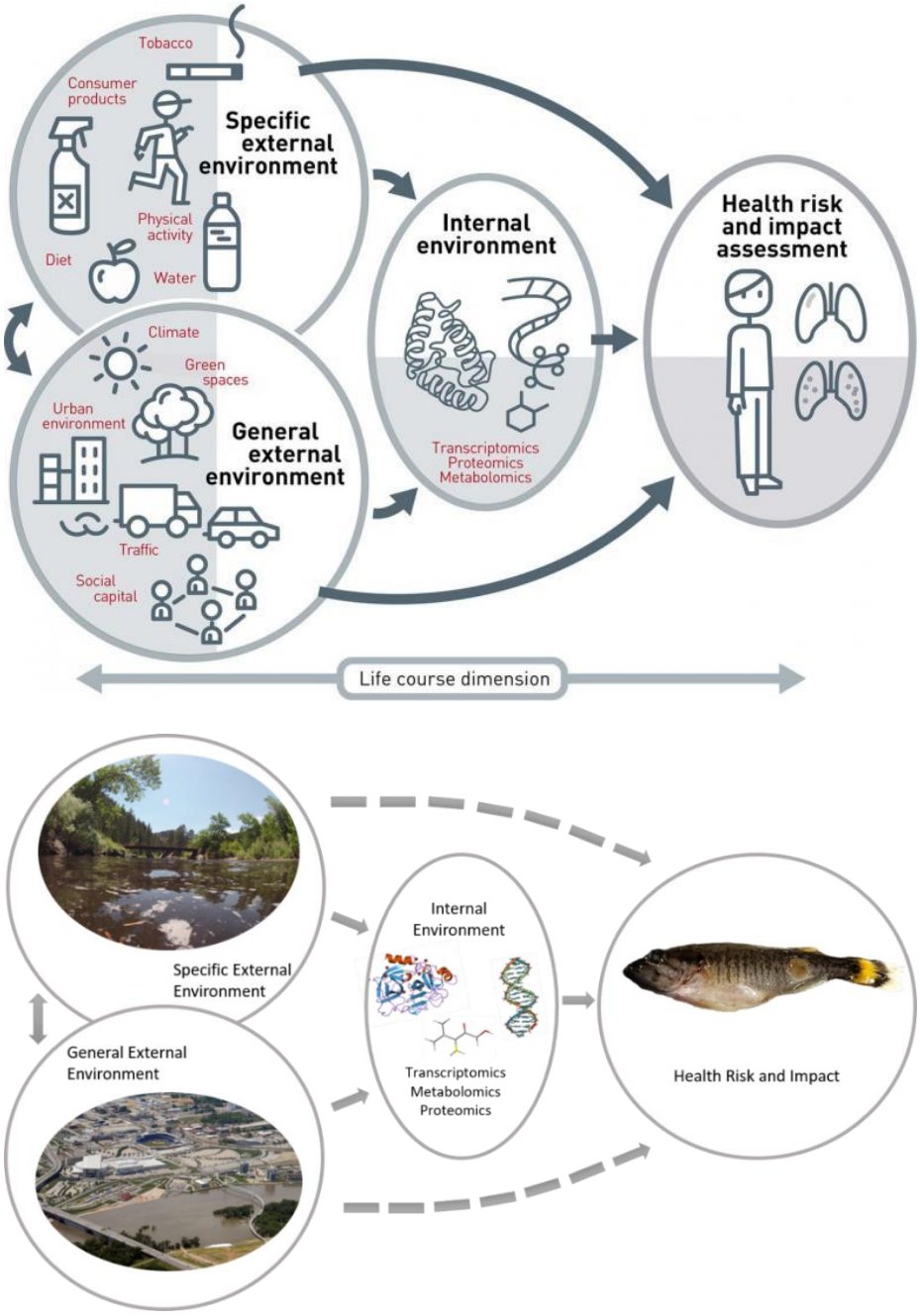

Figure 1.3 A comparison of the human exposome (top) with a proposed smallmouth bass model exposome (bottom). Adapted from Vrijheid, M.[42] *Photo credits: USACE HQ and USDAgov.*

A true exposomics analysis for most organisms would require multiple years of study following the organism from conception to death, a minimally invasive sample collection method, and a multidisciplinary sample analysis investigating every possible internal and external exposure (e.g., chemical contaminants, pathogens, microbiome, transcriptome, etc.). This study instead utilizes an exposomics based approach to help elucidate unknowns that may lead to a better understanding of the smallmouth bass health crisis currently happening within the Susquehanna River Basin. The YOY SMB were collected in summer 2015 from multiple sites exhibiting the disease characteristics described previously, but not all fish collected had visible health anomalies, such as dermal lesions, gill/fin erosion, or parasitic infections. Sexual maturation had yet to begin in these fish, so it is not possible to study the intersex condition observed in adult male SMB directly. However, previous studies have shown that exposure to endocrine disrupting compounds during development can lead to changes in sexual morphology in adult fish and changes in normal population sex ratios.[43,44] Similar studies have shown that immunosuppression can also result from exposing juveniles to environmental contaminants.[45,46] Therefore, studying YOY SMB from a single collection timepoint can still provide valuable insight into disease characteristics experienced by other members of the population.

The exposomics based approach relies on creating the broadest range sample analysis pipeline possible for a large sample set. The whole fish is analyzed as opposed to a single tissue type. Samples are not pooled but are prepared and analyzed as individuals to retain the most data possible for each specific YOY SMB. Targeted and non-targeted analyses are combined to identify as many significant features as possible from a single sample extraction (top-down exposomics). To employ an exposomics based approach, the development of better sample preparation and data analysis techniques were required. Current methodologies rely on expensive, time and labor intensive techniques, like Soxhlet, for sample preparation. Not only are these methods hard to apply for large scale studies, often they are highly optimized for the analysis of a

specific set of target compounds and prevent the discovery of non-target compounds. Chapter 2 details the development of the analytical pipeline, including optimization of the sample extraction/cleanup, two-dimensional gas chromatography analysis, and data processing methodology to quantify target contaminants and identify non-target contaminants specific to disease states. Chapter 3 expands on the developed methodology and applies it to the full sample set of 146 YOY SMB. Chapter 4 provides a complementary analysis using ultra-performance liquid chromatography to examine the same samples for compounds less amenable to gas chromatography. Chapter 5 reviews possible correlations between detected contaminants and disease characteristics as well as identifying contaminants that are significant among different collection sites.

1.4 Overview of Analytical Instrumentation

Analyzing biota for known or suspected target compounds is already a challenging feat, and the analysis is further complicated when expanded to include the discovery of unknowns. Biological matrices are complex, containing a broad range of chemical classes, and producing large background signals that overshadow compounds of interest to researchers. Targeted analyses of animal tissues often employ extensive sample extraction and cleanup procedures followed by analysis using either gas (GC) or liquid (LC) chromatography optimized specifically for the target compounds, thereby negating the discovery of non-targets.[47,48] For routine environmental monitoring situations, this can be considered an ideal approach, as the methods have been well validated and standardized by regulatory agencies, such as the United States Environmental Protection Agency. However, in instances where the organism of interest is experiencing a health crisis and routine monitoring efforts are unable to account for the phenomenon, it is necessary to look beyond the known or suspected contaminants and employ

more robust and sensitive analytical methodologies. GC is primarily used for the analysis of smaller, more volatile compounds that are not heat labile. LC can detect larger, more polar compounds, including proteins or other heat labile compounds that would be destroyed at the high temperatures utilized during GC analyses. By using both gas and liquid chromatography it is possible to identify a wider range of chemical components than by using a single analytical approach. An overview of the analytical techniques and instruments utilized for this purpose is presented here.

1.4.1 Comprehensive Two-Dimensional Gas Chromatography and Time-of-Flight Mass Spectrometry

Comprehensive two-dimensional gas chromatography (GC×GC) often coupled with time-of-flight mass spectrometry (TOFMS) is an increasingly effective method for the analysis of non-target compounds within complex sample matrices.[49] As a relatively new technique, GC×GC is not yet widely used in routine and commercial laboratories due to its increased cost and complexity over one-dimensional (1D) GC. Despite this, GC×GC provides significantly increased peak capacity (selectivity) and resolution as well as greater sensitivity (depending on the choice of modulator and detector).

A typical GC×GC system relies on the use of two complementary columns coupled together with a modulator. The sample is introduced to the primary column in the same manner as in 1D GC. After separation by the primary column the effluent enters the modulator, a device that uses thermal or valve controls to trap and release the effluent onto the secondary column. The secondary column is usually much shorter than the primary column and may be housed in a secondary oven held at a higher temperature than the primary oven. Second dimension separations are fast, in part to prevent losing the separation already achieved in the first

dimension. As the modulator traps small fractions of effluent leaving the primary column, it refocuses these discreet fractions and releases them onto the secondary column mimicking a second sample injection. After the second separation, effluent continues to the detector of choice as it would in traditional 1D GC analyses.

Unlike previous multi-dimensional techniques, GC×GC is a truly comprehensive technique, as the trap and release of effluent by the modulator is performed continuously throughout the entire analysis, resulting in the entire sample undergoing separation on two different columns. Ideally, two complementary stationary phases that utilize independent separation mechanisms are used to maximize chromatographic separation of analytes. The resulting output of a 2D separation is a series of second-dimension chromatograms that are placed alongside one another to create a single 2D chromatogram with the first dimension represented on the X-axis and the second dimension represented on the Y-axis. A color scale is normally used to indicate signal intensity (Figure 1.4). This visualization is accomplished through computer software, either laboratory written or as is increasingly common, by commercial programs. Data can also be viewed in a three-dimensional representation as well.

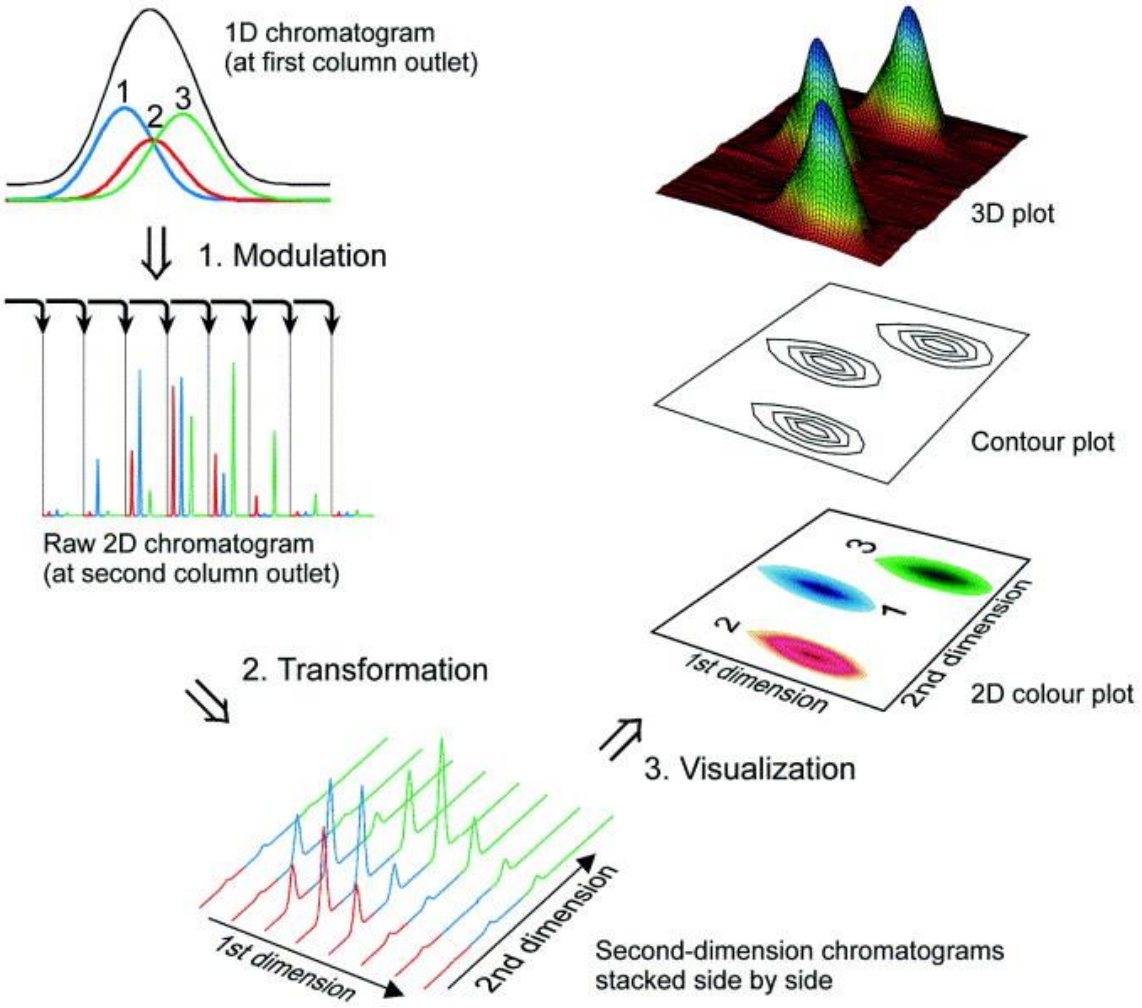

Figure 1.4 The creation and visualization of a two-dimensional chromatogram. Adapted from Dallüge et al.[50]

Comprehensive GC×GC-TOFMS was selected for this study for multiple reasons. The increased separation using a two-dimensional plane allows for analytes of interest to be separated from interfering background noise caused by the complex biological matrix. As opposed to other multidimensional techniques, with GC×GC-TOFMS the entire sample is analyzed in both dimensions, allowing a better chance for identifying statistically significant non-target compounds. The specific use of a TOFMS detector also aids in this. The TOFMS provides high sensitivity, simultaneous measurement of a wide range of mass-to-charge ratios, and the fast

acquisition speed required for GC×GC analyses. This technique allows for the detection of nearly every extractable compound from a sample under the optimal analysis conditions.

1.4.2 Ultra-Performance Liquid Chromatography and High-Resolution Quadrupole Time-of-Flight Mass Spectrometry

Advancements in LC have led to more efficient chromatographic platforms with increased sensitivity and resolution as well as shorter analysis times and less solvent consumption. High-pressure (aka high-performance) LC, or HPLC systems are capable of handling pressures up to 6000 psi and particle sizes around 2.5 – 5 µm in columns. Ultra-performance LC or ultra-high-performance LC (UPLC/UHPLC) systems are capable of handling smaller particles, sub-2 µm, and pressures up to 15000 psi. As the goal of the exposomics based approach is to identify as many contributing factors as possible, it is important to utilize complementary analytical techniques. For this study, UPLC coupled with high-resolution quadrupole time-of-flight mass spectrometry (HR-QTOF-MS) was chosen to analyze samples for compounds that are not amenable to GC techniques.

Non-targeted environmental and metabolomic studies have increasingly relied upon UPLC methods in recent years for the discovery of CECs derived from pharmaceuticals, pesticides, and personal care products.[51,52] The increased sensitivity and resolution of UPLC make it ideal for the separation of complex biological matrices, and the faster sample analysis times allow for high-throughput study designs that garner large amounts of data. These are features that align well with the needs of this study. A high-resolution MS instrument was chosen to aid in the identification of unknown compounds as LC-MS lacks a universal standardized spectral database, like that available for GC-MS. Using a QTOF-MS combines the fragmentation efficiency of a

quadrupole with the fast analysis speed and resolution of a TOF. This allows for accurate quantification of target compounds as well as better identification of non-target compounds.

Though metabolomics is only a part of the total exposome of an organism, analytical and data processing methodologies from non-targeted metabolomics studies can be applied to this study. In metabolomics studies, many metabolites of interest have a medium to high polarity that is more compatible with reversed phase than normal phase LC.[53] In reversed phase LC, a hydrophobic stationary phase is used, usually in conjunction with a mobile phase gradient of decreasing polarity throughout the analysis. The mobile phase begins highly polar (often aqueous), and the polarity is decreased as the concentration of non-polar organic solvent increases. This means that hydrophilic molecules will elute first followed by molecules of increasing hydrophobicity. Here, reversed phase LC was utilized to increase the likelihood of identifying metabolites or biomarkers associated with the disease characteristics being investigated.

1.5 References

1. M. Compa, C. Alomar, C. Wilcox, E. van Sebille, L. Lebreton, B.D. Hardesty, S. Deudero, Risk assessment of plastic pollution on marine diversity in the Mediterranean Sea, Sci. Total Environ. 678 (2019) 188–196. doi:10.1016/j.scitotenv.2019.04.355.
2. S. Sauvé, M. Desrosiers, A review of what is an emerging contaminant, Chem. Cent. J. 8 (2014) 15. doi:10.1186/1752-153X-8-15.
3. J. Milic, M. Curcic, Z. Brnjas, H. Carapina, J. Randjelovic, K. Krinulovic, A. Jovovic, The socio-economic impact timeline in Serbia for persistent organic pollutants (POPs), Sci. Total Environ. 688 (2019) 486–493. doi:10.1016/j.scitotenv.2019.06.161.
4. D.P. Häder, A.T. Banaszak, V.E. Villafañe, M.A. Narvarte, R.A. González, E.W. Helbling, Anthropogenic pollution of aquatic ecosystems: Emerging problems with global implications, Sci. Total Environ. 713 (2020) 136586. doi:10.1016/j.scitotenv.2020.136586.
5. M. Dunier, A.K. Siwicki, Effects of pesticides and other organic pollutants in the aquatic

environment on immunity of fish: a review, Fish Shellfish Immunol. 3 (1993) 423–438. doi:10.1006/fsim.1993.1042.

6. F. Freeling, N.A. Alygizakis, P. von der Ohe, J. Slobodnik, P. Oswald, R. Aalizadeh, L. Cirka, N.S. Thomaidis, M. Scheurer, Occurrence and potential environmental risk of surfactants and their transformation products discharged by wastewater treatment plants, Sci. Total Environ. 681 (2019) 475–487. doi:10.1016/j.scitotenv.2019.04.445.
7. P. Xiong, X. Yan, Q. Zhu, G. Qu, J. Shi, C. Liao, G. Jiang, A Review of Environmental Occurrence, Fate, and Toxicity of Novel Brominated Flame Retardants, Environ. Sci. Technol. (2019) acs.est.9b03159. doi:10.1021/acs.est.9b03159.
8. M.E. Franco, R. Lavado, Applicability of in vitro methods in evaluating the biotransformation of polycyclic aromatic hydrocarbons (PAHs) in fish: Advances and challenges, Sci. Total Environ. 671 (2019) 685–695. doi:10.1016/j.scitotenv.2019.03.394.
9. R. Kocagöz, O. Onmus¸, I. Onat, B. C¸ag˘das¸, M. Siki, H. Orhan, Environmental and biological monitoring of persistent organic pollutants in waterbirds by non-invasive versus invasive sampling, Toxicol. Lett. 230 (2014) 208–217. doi:10.1016/j.toxlet.2014.01.044.
10. V. Carignan, M.-A. Villard, Selecting indicator species to monitor ecological integrity: a review., Environ. Monit. Assess. 78 (2002) 45–61. doi:10.1023/a:1016136723584.
11. Y. Li, R. Chen, J. He, H. Ma, F. Zhao, S. Tao, J. Liu, J. Hu, Triphenyl Phosphate at Environmental Levels Retarded Ovary Development and Reduced Egg Production in Japanese Medaka (Oryzias latipes), Environ. Sci. Technol. 53 (2019) 14709–14715. doi:10.1021/acs.est.9b05669.
12. M.Y. Gross-Sorokin, S.. Roast, Brighty G.C, Causes and Consequences of Feminisation of Male Fish in English Rivers, Bristol, 2004. www.environment-agency.gov.uk (accessed July 25, 2019).
13. L. Zhang, X. Zhao, S. Yan, J. Zha, X. Ma, The immune responses of the Chinese rare minnow (Gobiocypris rarus)exposed to environmentally relevant concentrations of cypermethrin and subsequently infected by the bacteria Pseudomonas fluorescens, Environ. Pollut. 250 (2019) 990–997. doi:10.1016/j.envpol.2019.03.126.
14. M. Dunier, Water pollution and immunosuppression of freshwater fish, Ital. J. Zool. 63 (1996) 303–309. doi:10.1080/11250009609356150.
15. A.G. Silva, C.B.R. Martinez, Morphological changes in the kidney of a fish living in an urban stream, Environ. Toxicol. Pharmacol. 23 (2007) 185–192. doi:10.1016/j.etap.2006.08.009.
16. W.-K. Yang, L.-F. Chiang, S.-W. Tan, P.-J. Chen, Environmentally relevant concentrations

of di(2-ethylhexyl)phthalate exposure alter larval growth and locomotion in medaka fish via multiple pathways, Sci. Total Environ. 640–641 (2018) 512–522. doi:10.1016/j.scitotenv.2018.05.312.

17. D.P. Coulter, K.E. Huff Hartz, M.S. Sepúlveda, A. Godfrey, J.E. Garvey, M.J. Lydy, Lifelong Exposure to Dioxin-Like PCBs Alters Paternal Offspring Care Behavior and Reduces Male Fish Reproductive Success, Environ. Sci. Technol. 53 (2019) 11507–11514. doi:10.1021/acs.est.9b03460.
18. E.S.J. Thoré, C. Philippe, L. Brendonck, T. Pinceel, Antidepressant exposure reduces body size, increases fecundity and alters social behavior in the short-lived killifish Nothobranchius furzeri, Environ. Pollut. 265 (2020) 115068. doi:10.1016/j.envpol.2020.115068.
19. F. Brion, C.R. Tyler, X. Palazzi, B. Laillet, J.M. Porcher, J. Garric, P. Flammarion, Impacts of 17β-estradiol, including environmentally relevant concentrations, on reproduction after exposure during embryo-larval-, juvenile- and adult-life stages in zebrafish (Danio rerio), Aquat. Toxicol. 68 (2004) 193–217. doi:10.1016/j.aquatox.2004.01.022.
20. K.A. Hicks, M.L.M. Fuzzen, E.K. McCann, M.J. Arlos, L.M. Bragg, S. Kleywegt, G.R. Tetreault, M.E. McMaster, M.R. Servos, Reduction of Intersex in a Wild Fish Population in Response to Major Municipal Wastewater Treatment Plant Upgrades, Environ. Sci. Technol. 51 (2017) 1811–1819. doi:10.1021/acs.est.6b05370.
21. T.G. Brown, B. Runciman, S. Pollard, A.D.A. Grant, M.J. Bradford, Biological Synopsis of Smallmouth Bass (Micropterus dolomieu), 2009.
22. L. Steiner, Pennsylvania Fishes, Pennsylvania Fish and Boat Commission, Harrisburg, PA, 2000.
23. T. Pelton, Angling for Healthier Rivers, 2013. https://www.cbf.org/document-library/cbf-reports/smallmouth-bass-report0358.pdf (accessed June 5, 2019).
24. V.S. Blazer, L.R. Iwanowicz, C.E. Starliper, D.D. Iwanowicz, P. Barbash, J.D. Hedrick, S.J. Reeser, J.E. Mullican, S.D. Zaugg, M.R. Burkhardt, J. Kelble, Mortality of Centrarchid Fishes in the Potomac Drainage: Survey Results and Overview of Potential Contributing Factors, J. Aquat. Anim. Health. 22 (2010) 190–218. doi:10.1577/H10-002.1.
25. J.J. Chaplin, J. Kent Crawford, R.A. Brightbill, Water-Quality Monitoring in Response to Young-of-the- Year Smallmouth Bass (Micropterus dolomieu) Mortality in the Susquehanna River and Major Tributaries, Pennsylvania: 2008, 2009. https://pubs.usgs.gov/of/2009/1216/ofr2009-1216.pdf.
26. G.D. Smith, Susquehanna River smallmouth bass disease investigation Annual Report to the

Susquehanna River Policy Committee 2010, 2010.

27. D.W. Kolpin, V.S. Blazer, J.L. Gray, M.J. Focazio, J.A. Young, D.A. Alvarez, L.R. Iwanowicz, W.T. Foreman, E.T. Furlong, G.K. Speiran, S.D. Zaugg, L.E. Hubbard, M.T. Meyer, M.W. Sandstrom, L.B. Barber, Chemical contaminants in water and sediment near fish nesting sites in the Potomac River basin: Determining potential exposures to smallmouth bass (Micropterus dolomieu), Sci. Total Environ. 443 (2013) 700–716. doi:10.1016/j.scitotenv.2012.09.063.
28. V.S. Blazer, L.R. Iwanowicz, D.D. Iwanowicz, D.R. Smith, J.A. Young, J.D. Hedrick, S.W. Foster, S.J. Reeser, Intersex (Testicular oocytes) in smallmouth bass from the potomac river and selected nearby drainages, J. Aquat. Anim. Health. 19 (2007) 242–253. doi:10.1577/H07-031.1.
29. V.S. Blazer, L.R. Iwanowicz, H. Henderson, P.M. Mazik, J.A. Jenkins, D.A. Alvarez, J.A. Young, Reproductive endocrine disruption in smallmouth bass (Micropterus dolomieu) in the Potomac River basin: Spatial and temporal comparisons of biological effects, Environ. Monit. Assess. 184 (2012) 4309–4334. doi:10.1007/s10661-011-2266-5.
30. V.S. Blazer, D.D. Iwanowicz, H.L. Walsh, A.J. Sperry, L.R. Iwanowicz, D.A. Alvarez, R.A. Brightbill, G. Smith, W.T. Foreman, R. Manning, Reproductive health indicators of fishes from Pennsylvania watersheds: association with chemicals of emerging concern, Environ. Monit. Assess. 186 (2014) 6471–6491. doi:10.1007/s10661-014-3868-5.
31. C. Starliper, V.S. Blazer, L.R. Iwanowicz, H.L. Walsh, Microbial isolates in diseased fishes, primarily smallmouth bass (Micropterus dolomieu), within the Chesapeake Bay drainage in 2009-2011, in: Proc. West VIrginia Acad. Sci. Submitt. Pap. Eight-Eighth Annu. Sess. 2013, 2013.
32. T. Boonthai, T.P. Loch, C.J. Yamashita, G.D. Smith, A.D. Winters, M. Kiupel, T.O. Brenden, M. Faisal, Laboratory investigation into the role of largemouth bass virus (Ranavirus, Iridoviridae) in smallmouth bass mortality events in Pennsylvania rivers, BMC Vet. Res. 14 (2018) 62. doi:10.1186/s12917-018-1371-x.
33. H.L. Walsh, V.S. Blazer, L.R. Iwanowicz, G. Smith, A Redescription of Myxobolus inornatus from Young-of-the-Year Smallmouth Bass (Micropterus dolomieu), J. Parasitol. 98 (2013) 1236–1242. doi:10.1645/ge-3081.1.
34. D. Shull, M. Pulket, Causal Analysis of the Smallmouth Bass decline in the Susquehanna and Juniata Rivers December 2015, 2015.
35. D.A. Alvarez, W.L. Cranor, S.D. Perkins, V.L. Schroeder, L.R. Iwanowicz, R.C. Clark, C.P.

Guy, A.E. Pinkney, V.S. Blazer, J.E. Mullican, Reproductive health of bass in the potomac, USA, drainage: part 2. Seasonal occurrence of persistent and emerging organic contaminants, Environ. Toxicol. Chem. 28 (2009) 1084–1095. http://www.scopus.com/inward/record.url?eid=2-s2.0-67249093449&partnerID=40.

36. C.P. Wild, Complementing the Genome with an “Exposome”: The Outstanding Challenge of Environmental Exposure Measurement in Molecular Epidemiology, Cancer Epidemiol. Biomarkers Prev. 14 (2005) 1847–1850. doi:10.1158/1055-9965.EPI-05-0456.
37. S.M. Rappaport, Implications of the exposome for exposure science, J. Expo. Sci. Environ. Epidemiol. 21 (2011) 5–9. doi:10.1038/jes.2010.50.
38. C.P. Wild, The exposome: from concept to utility, Int. J. Epidemiol. 41 (2012) 24–32. doi:10.1093/ije/dyr236.
39. C.C. Coffey, T.A. Pearce, Direct-reading methods for workplace air monitoring, J. Chem. Heal. Saf. 17 (2010) 10–21. doi:10.1016/j.jchas.2009.08.003.
40. C.N. Nunes, V. Egéa dos Anjos, S.P. Quináia, Are there pharmaceutical compounds in sediments or in water? Determination of the distribution coefficient of benzodiazepine drugs in aquatic environment, Environ. Pollut. 251 (2019) 522–529. doi:10.1016/j.envpol.2019.05.015.
41. P. Spirhanzlova, J.-B. Fini, B. Demeneix, S. Lardy-Fontan, S. Vaslin-Reimann, B. Lalere, N. Guma, A. Tindall, S. Krief, Composition and endocrine effects of water collected in the Kibale national park in Uganda, Environ. Pollut. 251 (2019) 460–468. doi:10.1016/j.envpol.2019.05.006.
42. M. Vrijheid, The exposome: a new paradigm to study the impact of environment on health, Thorax. 69 (2014) 876–878. doi:10.1136/thoraxjnl-2013-204949.
43. N.J. Niemuth, R.D. Klaper, Emerging wastewater contaminant metformin causes intersex and reduced fecundity in fish, Chemosphere. 135 (2015) 38–45. doi:10.1016/j.chemosphere.2015.03.060.
44. Luzio, S.M. Monteiro, S. Garcia-Santos, E. Rocha, A.A. Fontaínhas-Fernandes, A.M. Coimbra, Zebrafish sex differentiation and gonad development after exposure to 17α-ethinylestradiol, fadrozole and their binary mixture: A stereological study, Aquat. Toxicol. 166 (2015) 83–95. doi:10.1016/j.aquatox.2015.07.015.
45. R.H. Milston, M.S. Fitzpatrick, A.T. Vella, S. Clements, D. Gundersen, G. Feist, T.L. Crippen, J. Leong, C.B. Schreck, Short-term exposure of chinook salmon (Oncoryhnchus tshawytscha) to o,p′-DDE or DMSO during early life-history stages causes long-term

humoral immunosuppression, Environ. Health Perspect. 111 (2003) 1601–1607. doi:10.1289/ehp.6171.

46. C. Fang, Q. Huang, T. Ye, Y. Chen, L. Liu, M. Kang, Y. Lin, H. Shen, S. Dong, Embryonic exposure to PFOS induces immunosuppression in the fish larvae of marine medaka, Ecotoxicol. Environ. Saf. 92 (2013) 104–111. doi:10.1016/j.ecoenv.2013.03.005.
47. USEPA, Guidance for Assessing Chemical Contaminant Data for Use in Fish Advisories: Volume 1 - Fish Sampling and Analysis Third Edition, 2000. https://www.epa.gov/sites/production/files/2015-06/documents/volume1.pdf (accessed July 8, 2019).
48. USEPA, Guidance for assessing chemical contaminant data for use in fish advisories, volume 2: Risk assessment and fish consumption limits, 3rd edition, 2000.
49. B. Gruber, B.A. Weggler, R. Jaramillo, K.A. Murrell, P.K. Piotrowski, F.L. Dorman, Comprehensive two-dimensional gas chromatography in forensic science: A critical review of recent trends, TrAC - Trends Anal. Chem. 105 (2018) 292–301. doi:10.1016/j.trac.2018.05.017.
50. J. Dallüge, J. Beens, U.A.T. Brinkman, Comprehensive two-dimensional gas chromatography: a powerful and versatile analytical tool, J. Chromatogr. A. 1000 (2003) 69–108. doi:10.1016/S0021-9673(03)00242-5.
51. F. Gosetti, E. Mazzucco, M.C. Gennaro, E. Marengo, Contaminants in water: non-target UHPLC/MS analysis, Environ. Chem. Lett. 14 (2016) 51–65. doi:10.1007/s10311-015-0527-1.
52. S. Forcisi, F. Moritz, B. Kanawati, D. Tziotis, R. Lehmann, P. Schmitt-Kopplin, Liquid chromatography-mass spectrometry in metabolomics research: Mass analyzers in ultra high pressure liquid chromatography coupling, J. Chromatogr. A. 1292 (2013) 51–65. doi:10.1016/j.chroma.2013.04.017.
53. Z. Guo, S. Huang, J. Wang, Y.L. Feng, Recent advances in non-targeted screening analysis using liquid chromatography - high resolution mass spectrometry to explore new biomarkers for human exposure, Talanta. 219 (2020) 121339. doi:10.1016/j.talanta.2020.121339.

Chapter 2

Modified QuEChERS Extraction for the Analysis of Young-of-Year Smallmouth Bass Using GC×GC-TOFMS

2.1 Abstract

Signs of disease, such as external lesions, have been prevalent in smallmouth bass throughout the Susquehanna River Basin, USA. Previous targeted chemical studies in this system have identified known persistent organic pollutants, but a common explanatory link across multiple affected sites remains undetermined. A fast and robust extraction method that can be applied to young-of-year fish is needed to effectively screen for target and non-target compounds that may be impacting organism health. The quick, easy, cheap, effective, rugged, and safe (QuEChERS) extraction methodology was optimized to perform both targeted and non-targeted chemical analyses from a single extraction of whole young-of-year fish. Comprehensive two-dimensional gas chromatography coupled with time-of-flight mass spectrometry (GC×GC-TOFMS) was used for extract analysis. Sample extraction was performed using the solvent ethyl acetate, followed by a two-step cleanup in which samples were frozen for lipid removal and subjected to dispersive solid phase extraction using Florisil. A sample set of 21 young-of-year smallmouth bass collected from areas with disease, and exhibiting different types of external lesions, were evaluated for 233 target compounds. A total of 34 organic contaminants, including polychlorinated biphenyls, brominated diphenyl ethers, organochlorinated pesticides, and personal care products, were detected. Data from this sample set was then analyzed for non-targets. Using the Fisher ratio method and multivariate analysis, an additional 10 significant

features were identified specific to either fish with visible lesions or with no visible disease characteristics.

2.2 Introduction

Changes in organism health often indicate shifts in environmental suitability, which can be due to anthropogenic contamination such as the release of pesticides, manufacturing by-products, pharmaceuticals, and personal care products into the surroundings.[1–4] Smallmouth bass (*Micropterus dolomieu*, SMB) within the Chesapeake Bay Watershed, USA have exhibited signs of adverse health effects including sporadic fish kills, low chronic mortality, skin lesions, reproductive endocrine disruption, and population declines in certain areas.[5–11] Disease in both adults[12] and young-of-year[11,13] (YOY) is characterized by co-infections of bacterial, viral, and/or parasitic pathogens, suggesting immunosuppression. Studies have identified various environmental contaminants capable of inducing endocrine disruption or immune suppression, in water, sediment, and fish tissue. However, most of these have focused on the analysis of adult SMB biological endpoints[12,14–16] , with only one study reporting YOY SMB chemical contaminants.[11,12,14–16] This targeted analysis detected 32 contaminants in YOY from the Susquehanna River compared with ten from an out of basin reference site.[11] To date, no individual or group of chemicals has been definitively associated with the increased prevalence of disease. This is likely due to a combination of chemical contaminants and environmental stressors, but it is also possible that traditional targeted analytical approaches may miss important analytes not included on target lists. Additionally, previous studies have shown that both endocrine disruption and immunosuppression are conditions that can be induced in some species by exposing organisms to contaminants during key developmental stages, a factor that may be missed in the analysis of adult fish.[17–19]

Many organic contaminants persist within the environment, bioaccumulating within organisms and negatively impacting their health throughout their entire lifetime. Furthermore, top predators, including humans, are more severely impacted by biomagnification of these compounds through the food chain.[20,21] Monitoring efforts require accurate detection methods to identify new problems and facilitate remediation in affected areas. Traditional environmental monitoring approaches for chemical contamination utilize targeted analyses of air, water, or sediment collected through discrete or passive sampling methods or the direct analysis of biological tissues.[22–24] These analyses are optimized for the identification and quantification of known pollutants, which are preselected based on the suspected source of contamination, for example sites in proximity to wastewater treatment facilities or agricultural runoff.[15,25] However, focusing on known targets may inhibit the discovery of new contaminants of emerging concern as well as metabolites from contaminants within biological matrices. Non-targeted analyses using exhaustive extraction for a broad range of chemical classes would provide more information for unresolved environmental contamination events, but such extraction methods need to be capable of relatively high throughput to meet the needs of statistically relevant data reduction.

Quick, easy, cheap, effective, rugged, and safe (QuEChERS) is a multiresidue extraction technique originally developed for the analysis of incurred pesticides in produce.[26] In recent years, this method has been applied to analyze a variety of biological sample matrices for a wide range of contaminants, including legacy persistent organic pollutants (POPs), current use pesticides, and steroid hormones. [27–29] QuEChERS is often modified, adapting the method to specific challenges relative to the sample.[30–32] The versatility of QuEChERS displayed in previous studies demonstrates its potential as a suitable non-targeted extraction technique, and the nature of the technique lends itself to being able to parallel-process samples in a relatively efficient manner.

The high lipid content of whole fish extracts can be an analytically challenging issue for one dimensional gas chromatography (GC), resulting in large unresolved complex mixtures in the chromatogram if samples do not undergo extensive cleanup before analysis. However, extensive sample cleanup negates the benefit of a non-targeted analysis by potentially removing unknown compounds that may be important in the correlation of exposure to health. Background interference and analyte coelutions are common problems in the separations of complex matrices, but they can be significantly improved using comprehensive two-dimensional gas chromatography (GC×GC). Compared to traditional GC, the two-dimensional separation of GC×GC allows for much greater peak capacity detecting thousands of components per sample.[33,34] Due to the large amount of data generated with this technique, robust data reduction techniques are necessary to determine which components are significant to a sample set. Fisher ratios measure the class-to-class variation of samples in comparison to the within class variation and can be combined with multivariate statistics to identify components that contribute the most to between class variation.[35,36]

This study presents a modified QuEChERS protocol to analyze YOY SMB from an affected area of the Susquehanna River in Pennsylvania, USA using GC×GC-TOFMS. Method optimization focused on detecting a wide range of compounds while minimizing sample cleanup. YOY SMB were analyzed in both targeted and non-targeted applications to determine differences between fish from the same collection site that presented different visible signs of disease at the time of collection.

2.3 Experimental

2.3.1 Reagents and Chemicals

A comprehensive list of all standards used for both method development and the targeted analysis of YOY SMB tissues is presented in supporting information Tables 2.3 and 2.4. Method development was performed using a mixture of 64 target compounds consisting of polychlorinated biphenyls (PCBs), organochlorine pesticides (OCPs) and brominated diphenyl ethers (BDEs). All solvents were of analytical grade. Acetonitrile was obtained from Honeywell Burdick and Jackson (Muskegon, MI, USA). Ethyl acetate, HPLC water, and acetone were obtained from JT Baker (Phillipsburg, NJ, USA).

2.3.2 Sample Collection

Method development was performed using both farm-raised Atlantic salmon (*Salmo salar*) muscle tissue, sourced from a local market, and whole SMB raised in captivity at the Leetown Science Center, Kearneysville, WV. Captive fish were raised and euthanized according to the Leetown Science Centers Institutional Animal Care and Use Committee. The finalized method was applied to a sample set of 21 YOY SMB collected during July 2015 using electrofishing from two different streamflow sources (labeled as Juniata Flow and West Branch Flow) within the Susquehanna River at Harrisburg, PA. The streamflows were defined by Pennsylvania Department of Environmental Protection as distinct regions across the stream channel with divergent water quality and chemical profiles originating from upstream tributaries (e.g., Juniata River or West Branch Susquehanna River). The confluence of the Juniata River with the mainstem Susquehanna River is approximately 15 km upstream from the collection site. The confluence of the West Branch Susquehanna with the Susquehanna River mainstem is

approximately 75 km upstream from the collection site. Coordinates for the Juniata Flow and West Branch Flow collection sites are 40.334023, -76.91709 and 40.33538, -76.912562, respectively.

2.3.3 Sample Fortification and Extraction

Individual pieces of fish muscle (salmon) as well as individual whole SMB (from Leetown Science Center) were homogenized separately using a PRO Scientific PRO250 handheld homogenizer (Oxford, CT, USA) on ice. Whole YOY SMB were homogenized individually using a PRO Scientific Bio-Gen PRO200 homogenizer on ice. A volume of HPLC grade water equivalent to 25% of the weight of each sample was added to aid homogenization. Method development extractions used 1 g aliquots of whole SMB or fish muscle, as specified. For the targeted analysis application, extractions were performed using 1 g aliquots from individual fish weighing more than 1 g total or by using the entire homogenized sample of individual fish weighing less than 1 g.

A stock solution of the 64 target compounds listed above was prepared at a final concentration of 2 ng/µl in acetone. For the method development recovery studies, 1 g samples were spiked at a final concentration of 100 ng/g, vortexed, and held for 1 hour before extraction. Matrix blanks containing no spiked target compounds were analyzed under the same conditions.

QuEChERS extraction methodology was modified from European Standard EN 15662.[37] Spiked sample aliquots and matrix blanks were vortexed for 30 seconds with 1 ml of extraction solvent (ethyl acetate or acetonitrile). 0.75 g of buffered QuEChERS extraction salts (Restek Corp., Bellefonte, PA, USA) were added and vortexed for 30 seconds. Samples were centrifuged for 5 minutes at 3000 rpm. The organic layer was transferred to a clean vial for overnight freeze out of lipids at -20°C. Upon removal from freezer, samples were immediately centrifuged at 4°C

for 3 minutes at 5000 rpm causing precipitated lipids to aggregate at the bottom of the tubes. The supernatant was transferred to a vial containing either 0.05 g Florisil or a mixture of 0.15 g $MgSO_4$ + 0.025 mg primary secondary amine (PSA) + 0.025 mg C18 (Restek Corp.) for dispersive solid phase extraction (dSPE). Samples were vortexed for 1 minute, followed by centrifugation for 3 minutes at 8000 rpm. The supernatant was transferred to a clean vial and spiked with internal standards (acenaphthylene-d8, anthracene-d10, and pyrene-d10, Restek Corp.) at a final concentration of 10 pg/µl prior to analysis to monitor variation in instrumental performance.

The method was finalized using ethyl acetate as the extraction solvent and Florisil for dSPE cleanup. For YOY SMB samples used in application, a surrogate mix containing PCB congeners #18, 28, and 52, triphenyl phosphate, and tris(1,3-dichloroisopropyl) phosphate (Restek Corp.) was added before extraction to monitor extraction efficiency.

2.3.4 Method Validation

Recovery values were determined by 3 replicate extractions of each matrix using spiked and blank samples. Target analytes were normalized using internal standards added prior to instrumental analysis. Method detection limits (MDLs) were calculated as outlined in U.S. EPA Method 40 CFR 136 using 7 replicates of spiked and blank matrix.[38]

2.3.5 Instrumentation

Samples were analyzed using a Pegasus BT 4D GC×GC-TOFMS (LECO Corp., St. Joseph, MI, USA) equipped with a 7890A GC system (Agilent Technologies, DE, USA) and split/splitless injection port containing a Siltek deactivated straight liner with glass wool (Restek

Corp.). An Rtx-1 (60 m x 0.25 mm ID x 0.25 μm df) column was installed in the first dimension and an Rxi-17Sil MS (1 m x 0.18 mm ID x 0.18 μm df) column was installed in the second dimension (both columns obtained from Restek Corp.). Samples were introduced to the column by 1 μl pulsed splitless injection with a pulse pressure of 60 psi for a duration of 1 minute and splitless hold time of 1 minute. The inlet was held at 250°C for the entirety of the run. Helium carrier gas was maintained at a flow rate of 1.4 ml/min. The oven temperature program began at 60°C for 1 minute, increased 5°C/minute to 330°C and held for 10 minutes. Samples extracted in acetonitrile were run with an initial oven temperature of 100°C for 1 minute, increased 5°C/minute to 330°C, and held for 10 minutes. The initial oven temperature was increased to overcome incompatibility between the highly polar solvent and the non-polar stationary phase of the first-dimension column. This provided the best chromatography for acetonitrile extracts with the least deviation from the settings used for ethyl acetate extracts. The secondary oven offset was +5°C of the primary oven. The modulator temperature offset was +15°C of the primary oven. The modulation period was 4.00 s with a 1.00 s hot pulse and 1.00 s cold pulse. Data were collected in the mass range 45-750 amu at an acquisition rate of 200 spectra/second with an electron ionization energy of 70 eV.

2.3.6 Data Processing

Data collected for method optimization and the applied targeted analysis were processed with LECO software ChromaTOF BT v5.40.12.0 for basic peak finding, mass spectral deconvolution, peak area integration, and quantification using 7-point, internal standard-corrected, calibration curves for each target compound. Data acquired for the targeted analysis was exported as .netCDF files to ChromaTOF GC v4.74.2.0 for non-targeted data processing. All samples were reprocessed for basic peak finding, mass spectral deconvolution, and peak area

integration and input into the statistical compare function. Final parameters used for basic peak finding and statistical compare are listed in supplemental information. Statistical compare was used to align chromatograms and generate a common peak table. Artifact compounds (peaks resulting from solvent or column bleed) were manually removed. Fisher ratios were calculated for all analytes retained after alignment and all analytes with ratios above the F_{crit} value at $\alpha = 0.05$ were exported to RStudio (Boston, MA, USA) for further statistical analysis using R v3.5.3 (R Foundation for Statistical Computing, Vienna, Austria).

2.4 Results and Discussion

2.4.1 Method Optimization

For non-targeted analyses, sample extraction methods should be exhaustive, retaining a broad range of compounds, unlike targeted analyses which actively remove compounds not selected as targets. Targeted analyses may be more efficient for rapid sample screening, but limit researchers to only evaluating pre-selected compounds. Non-targeted analyses can result in the discovery of new contaminants of concern or when investigating biological samples, the discovery of new biomarkers or metabolites.[34] For this study, a non-targeted method was optimized to analyze fish tissues for both organic contaminants and metabolites. QuEChERS was chosen as the extraction method for its combined ability to detect a large range of chemical classes using a single extraction, while employing minimal sample cleanup and thereby preventing unnecessary analyte loss. Additionally, the low cost and simplicity of the method make it ideal for high-throughput extraction of large sample sets.

A stock solution composed of 64 BDEs, PCBs, and OCPs in a single mix was used for method optimization. These compound classes are routinely analyzed for fish intended for human

consumption and have been previously shown to impair the health of aquatic life.[39,40] QuEChERS traditionally employs acetonitrile (MeCN) as the extraction solvent, however, initial recovery studies using whole fish resulted in low or no recovery of target compounds. The high polarity of MeCN limits the co-extraction of lipids, providing cleaner final extracts, but it also deters extraction of analytes sequestered within these fatty tissues. Due to the fatty nature of whole fish homogenates, the extraction solvent was changed to ethyl acetate (EtOAc), a less polar solvent, which resulted in significantly higher extraction recoveries, shown in supporting information Table 2.5. The calibration for samples in EtOAc was linear for all targets over the range of 0.25 – 250 ng/g, having correlation coefficients (R) of 0.994 or greater, with the exception of 2,2',4,4',5,5'-hexabromobiphenyl. Relative standard deviation of the calibration response factors ranged from 2 – 25% and was determined to be satisfactory for this study. EtOAc resulted in high levels of lipid co-extraction. As seen in Figure 1a, large fatty acid peaks were present on the total ion chromatogram (TIC) despite freezing the samples overnight and removing coagulated lipids by centrifugation as well as performing dSPE using $MgSO_4$ + PSA + C18. Substituting Florisil as the dSPE sorbent removed considerably more lipid compounds, as determined by the visual reduction in large fatty acid peaks in Figure 1b, without impairing target analyte recovery. Therefore, Florisil was chosen as the dSPE sorbent.

The finalized method maintains the speed and ease of extraction associated with QuEChERS. Studies examining fish tissue using QuEChERS often add cleanup, through syringe filters or columns, and concentration steps, prolonging the total extraction time per sample and increasing costs.[31,41,42] The presented method avoids these complications while extracting a broad range of analytes from a single extraction. Here, the original QuEChERS protocol was scaled down to accommodate a 1 g sample and extraction was performed using the solvent ethyl acetate. Excess lipids in the sample were removed through an overnight freezing step and the use of Florisil as the dSPE sorbent. Use of a single extraction for sample analysis was emphasized

during optimization as the small size of individual YOY SMB limits most samples to a single extraction. While pooling samples would allow multiple extractions on a single batch, information about the individuals would be lost, hindering the non-targeted analysis by preventing the association of significant compounds with individuals exhibiting different states of health.

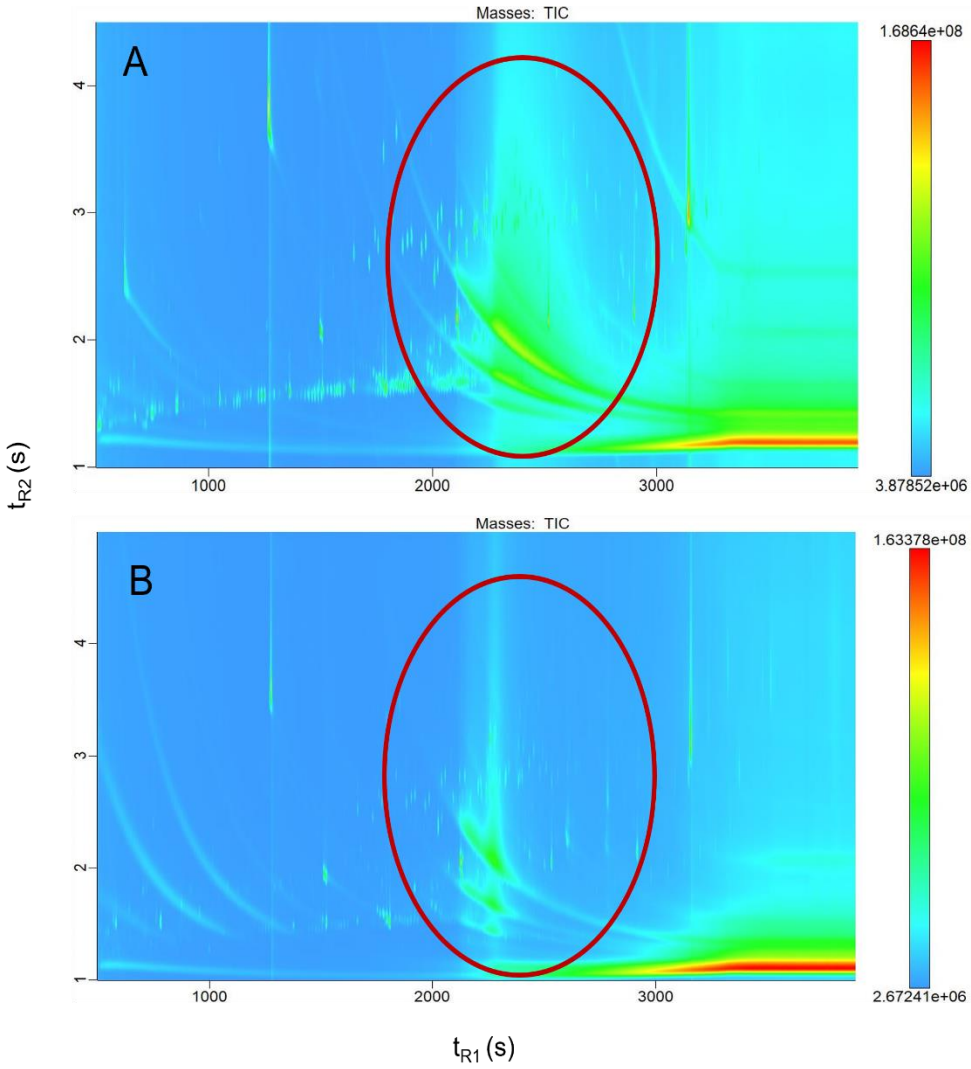

Figure 2.1 Comparison of total ion chromatograms (TIC) using PSA + C18 vs. Florisil for dSPE. Circled regions show the elution of fatty acid compounds. A) Fish muscle extract subjected to dSPE with PSA + C18. B) Fish muscle extract subjected to dSPE with Florisil. X-axis is first dimension retention time in seconds. Y-axis is second dimension retention time in seconds. Z-axis is peak intensity.

2.4.2 Targeted Analysis of YOY SMB

The finalized method was applied to a sample set of 21 YOY SMB from the Susquehanna River in Harrisburg, PA, USA at a site where multiple streamflows result in distinct water quality regions across the stream channel. Ten fish were collected from the Juniata Flow and 11 from the West Branch Flow. SMB spawning in this region occurs during May – June, therefore fish examined in this study were estimated to be between 1 – 2 months old at the time of collection.

A total of 34 of 233 targets were detected in at least one sample (Table 2.1). Out of the 34 chemicals detected, 28 were detected at least once in the Juniata Flow fish and 31 were detected at least once in the West Branch Flow fish. Only a single compound, indole, was detected in every sample. Indole was also present at the highest concentration of any detected compound averaging 389.5 ng/g in Juniata Flow fish and 507 ng/g in West Branch Flow fish. Three samples had indole values above 1 μg/g. Indole is a widely used compound and can be found in fragrances, food additives, pesticides, herbicides, and is naturally produced by select bacteria during protein decomposition, specifically of the amino acid tryptophan.[43,44] The ubiquitous nature of indole in the environment could explain the concentrations detected within these samples despite its rapid metabolization rate.

Most detected target compounds (23 of 34) were PCBs. PCBs are legacy persistent organic pollutants (POPs) known to negatively impact organism health through endocrine disruption and immune suppression.[45] Here, individual PCB congener concentration averages are reported (Table 2.1), however, PCB concentrations in fish are commonly reported as the sum of all PCBs. Juniata Flow fish averaged a total of 3.3 ng/g for all PCBs, while West Branch Flow fish averaged 4.4 ng/g. These averages are far below the “no consumption” advisory level of the Pennsylvania Department of Environmental Protection, which is set at > 1.9 ppm (1,900 ng/g).

Fish tissues containing less than 50 ng/g total PCBs are considered safe for unrestricted consumption.[46] A single dioxin-like PCB, congener 118, was detected above the MDL. Dioxin-like PCBs are a group of 12 PCBs that possess the same mechanism of action as dioxins. These compounds have a high affinity for aryl hydrocarbon receptors and are toxic to humans and other organisms.[47]

The role of PCBs and BDEs in SMB disease is unclear. A previous study compared YOY SMB collected from a reference experimental pond at the Leetown Science Center, Kearneysville, WV and individuals collected from affected areas of the Susquehanna River in 2011.[11] Several compounds detected in this study, including PCB 170, PCB 180, PCB 187, BDE-100, BDE-99, and p,p'-DDE, were detected in fish from the reference pond that showed no signs of disease. PCB congeners 118, 149, 174, 177, 183, and 206 were also detected in the affected populations of both studies in similar concentration ranges (see Table 2.1), along with several other PCBs and BDEs that were not found in common. Though several compounds were found in common between these two studies, it is difficult to determine if these were the only similarities as different target lists would reveal different findings.

Table 2.1 Summary of target analytes identified within young-of-year smallmouth bass.

Analyte	Juniata Flow		West Branch Flow	
	# of Detections	*Mean (ng/g)*	*# of Detections*	*Mean (ng/g)*
2,6-dimethylnaphthalene	3	0.25	5	0.34
p,p'-DDE	5	NR	7	0.74
9,10-anthraquinone	0	ND	1	6.62
BDE-100	0	ND	2	0.59
BDE-99	5	1.30	4	1.92
Celestolide	1	0.12	1	0.14
Cotinine	0	ND	2	3.88
Galaxolide	5	0.67	7	0.87
Indole	10	389.53	11	507.01
Isoquinoline	0	ND	1	5.83
Menthol	7	2.60	3	2.97
PCB 002	0	ND	1	0.06
PCB 008	7	NR	7	0.06
PCB 118	3	0.15	4	0.13
PCB 138	7	0.63	8	1.11
PCB 141	1	NR	3	0.23
PCB 147	1	0.13	4	0.34
PCB 149	5	0.18	4	0.35
PCB 153	6	0.86	9	1.02
PCB 170	2	0.33	0	ND
PCB 174	1	0.19	1	0.63
PCB 175	1	0.78	2	1.05
PCB 177	1	0.19	2	0.23
PCB 180	5	0.69	10	1.06
PCB 183	0	ND	1	0.22
PCB 187	6	0.36	8	0.37
PCB 190	2	0.31	2	0.65
PCB 196	1	0.22	0	ND
PCB 199	4	0.39	2	0.46
PCB 202	1	0.10	0	ND
PCB 203	5	0.32	3	0.40
PCB 206	6	1.05	2	1.41
PCB 208	1	1.04	1	0.41
PCB 209	3	0.90	2	0.93

PCB = polychlorinated biphenyl, BDE = brominated diphenyl ether, ND = Not Detected, NR = Not Reported (calculated value was below method detection limit)

2.4.3 Non-targeted Analysis of YOY SMB

After completion of the targeted analysis, chromatographic data for YOY SMB was exported to ChromaTOF GC to utilize this software's statistical compare function in a non-targeted analysis. Statistical compare allows multiple chromatograms to be aligned based on peak retention times and mass spectral data. Accurate alignment of sample chromatograms is an essential first step in data reduction. Comprehensive GC×GC may identify thousands of analytes within a single sample, however, compounds found in only one or very few samples are unlikely to be statistically significant and should be removed from further investigation. To accurately identify peaks representing the same compound throughout multiple samples, individual chromatograms are aligned and normalized against internal standards to account for the inherent instrumental variation between each run. The following criteria were implemented for the retention of a compound within the aligned peak list: a) the compound was found within at least 60% of samples from the same class or b) the compound was found within at least 12 samples (~60%) regardless of classification. Sample classification was determined based on the condition of the fish at the time of collection. A definitive diagnosis on cause of lesions (other than leeches) observed on individual SMB in this study could not be made because the whole fish was used for chemical analyses. However, visual observations were used to characterize external disease characteristics including lesions. Microscopic analyses of similar lesions have indicated the pale and eroded areas are most often associated with bacterial or viral pathogens,[11,48] while the raised lesions are associated with a myxozoan parasite, *Myxobolus inornatus*.[10,49] Fish with no observed disease characteristics were classified as Normal and fish presenting any of the listed lesions were classified as Lesioned (Table 2.2).

After alignment, 1374 analytes were retained in the aligned peak list. To reduce the data set for further statistical analysis, Fisher ratios were calculated for each analyte. Fisher ratios are a

measure of between class variation compared to within class variation. These ratios can be used to identify compounds that contribute the most variation between sample classes.[35,50] Analytes were considered significant if their calculated Fisher ratio value was above the critical value for $\alpha = 0.05$ ($F_{crit} = 4.38$). This reduced the data set to 91 analytes, which were searched against the NIST MS Database (2014) for identification. 32 of the 91 analytes were presumptively identified with a minimum match similarity score of 700 (see supporting information Table 2.6).

The reduced data set was visualized using principal component analysis (Figure 2). A tight cluster was seen for Lesioned fish. Eighteen analytes were identified within the 95% ellipse of the Lesioned group cluster; eight of these analytes only occurred within Lesioned class samples. Within the 95% ellipse of Normal fish, excluding the region covered by the ellipse for Lesioned fish, two analytes were found to only occur within Normal class samples. The final data set was reduced to the ten total unique analytes identified in both Lesioned and Normal class fish (see supporting information Table 2.6). Only one of the ten unique analytes was presumptively identified using the above match criteria. It was assigned the tentative identification of myristic acid glycidyl ester and was found in Lesioned fish. Identification of the unknown analytes was not a priority for this study, as the focus was on developing a high-throughput method and data reduction techniques for targeted and non-targeted analyses. Here, only a limited sample set from a single collection was analyzed to demonstrate the validity of this method. To address the widespread occurrence of the SMB disease, both temporally and spatially, a significantly larger data set that includes samples collected from multiple dates and locations is required. This analysis is in progress and will employ high resolution mass spectrometry for identification of unknown compounds determined to be significant using the methodology developed and presented herein. GC×GC coupled with high resolution TOFMS (HR-TOFMS) is a powerful tool that can be used to identify molecular formulas for unknown compounds. However, by reducing the number of unknowns to a small set of statistically significant compounds, time intensive

investigation of irrelevant unknowns may be avoided, and future efforts can be focused on meaningful compounds of interest.

Table 2.2 Collection information for YOY SMB. All fish were collected on July 14, 2015.

Collection Streamflow	Fish #	Wet Weight (g)	Classification	Lesion Type
West Branch	001	1.00	Lesioned	Pale Area
West Branch	002	2.30	Normal	N/A
West Branch	003	1.00	Lesioned	Leech
West Branch	004	1.70	Lesioned	Erosion
West Branch	005	1.10	Lesioned	Pale Area
West Branch	006	2.20	Lesioned	Raised
West Branch	007	2.00	Lesioned	Erosion
West Branch	008	1.20	Normal	N/A
West Branch	009	0.60	Normal	N/A
West Branch	010	0.90	Normal	N/A
West Branch	011	2.40	Lesioned	Raised
Juniata	001	2.00	Normal	N/A
Juniata	002	4.50	Lesioned	Leech
Juniata	003	2.30	Normal	N/A
Juniata	004	2.80	Lesioned	Frayed Fins
Juniata	005	2.80	Lesioned	Raised
Juniata	006	2.00	Lesioned	Erosion
Juniata	007	2.60	Lesioned	Leech
Juniata	008	3.30	Normal	N/A
Juniata	009	1.20	Normal	N/A
Juniata	010	3.60	Lesioned	Frayed Fins

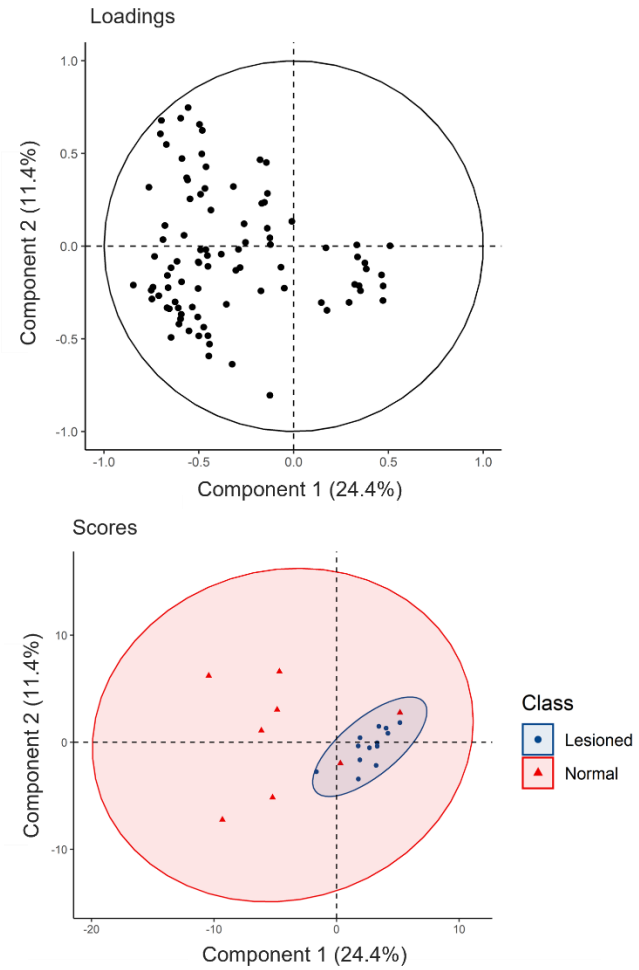

Figure 2.2 Principal component analysis loadings (top) and scores (bottom) of YOY SMB classified as either Lesioned or Normal. Ellipses represent 95% confidence interval. Black circles on loadings plots represent individual analytes retained after data reduction by Fisher ratio analysis. Blue circles of scores plot represent individual samples classified as Lesioned. Red triangles represent individual samples classified as Normal.

2.5 Conclusion

The modifications to the QuEChERS method presented in this study provide a high-throughput extraction technique that can be coupled with GC×GC-TOFMS for the targeted and non-targeted analysis of YOY fish. The speed, cost, and efficiency of this method are well suited for rapid screening of affected populations, and the same data can be directly evaluated for non-targets with no additional sample preparation or instrumental analysis. The statistical compare

function of ChromaTOF GC successfully aligned and normalized all 21 chromatograms, and the Fisher ratio method effectively reduced the data set for multivariate analysis demonstrating an efficient data analysis technique for non-target screening.

For the YOY SMB in this study, a targeted analysis detected 34 organic contaminants, including legacy PCBs, BDEs, and OCPs, in fish with an estimated age of 1-2 months. While a non-targeted analysis identified an additional 10 significant compounds occurring in either Normal or Lesioned fish. Due to the limited sample set evaluated, it is not possible to directly correlate identified compounds with SMB disease throughout the Chesapeake Bay Watershed. Additional analyses of multiple sites are needed to apply these findings to other populations. Identified contaminants also likely exhibit synergistic effects with environmental conditions and other contaminants that are as yet unknown but contribute to the expression of SMB disease signs.

2.6 Acknowledgements

Funding for this work was provided by Pennsylvania Sea Grant and the U.S. Geological Survey's Environmental Health (Contaminants Biology) Program. Industrial support for this project was provided by LECO Corporation and Restek Corporation. The authors would like to specially thank Kelly Smalling for assistance with acquisition of standard materials and Timothy Wertz, Pennsylvania Department of Environmental Protection and Geoff Smith, Pennsylvania Fish and Boat Commission, for assistance with fish collections. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

2.7 References

1. D.P. Häder, A.T. Banaszak, V.E. Villafañe, M.A. Narvarte, R.A. González, E.W. Helbling, Anthropogenic pollution of aquatic ecosystems: Emerging problems with global implications, Sci. Total Environ. 713 (2020) 136586. doi:10.1016/j.scitotenv.2020.136586.
2. O.R. Ibor, A.O. Adeogun, O.A. Fagbohun, A. Arukwe, Gonado-histopathological changes, intersex and endocrine disruptor responses in relation to contaminant burden in Tilapia species from Ogun River, Nigeria, Chemosphere. 164 (2016) 248–262. doi:10.1016/j.chemosphere.2016.08.087.
3. K.A. Hicks, M.L.M. Fuzzen, E.K. McCann, M.J. Arlos, L.M. Bragg, S. Kleywegt, G.R. Tetreault, M.E. McMaster, M.R. Servos, Reduction of Intersex in a Wild Fish Population in Response to Major Municipal Wastewater Treatment Plant Upgrades, Environ. Sci. Technol. 51 (2017) 1811–1819. doi:10.1021/acs.est.6b05370.
4. A.R. Fonseca, L.F. Sanches Fernandes, A. Fontainhas-Fernandes, S.M. Monteiro, F.A.L. Pacheco, From catchment to fish: Impact of anthropogenic pressures on gill histopathology, Sci. Total Environ. 550 (2016) 972–986. doi:10.1016/j.scitotenv.2016.01.199.
5. V.S. Blazer, L.R. Iwanowicz, D.D. Iwanowicz, D.R. Smith, J.A. Young, J.D. Hedrick, S.W. Foster, S.J. Reeser, Intersex (Testicular oocytes) in smallmouth bass from the potomac river and selected nearby drainages, J. Aquat. Anim. Health. 19 (2007) 242–253. doi:10.1577/H07-031.1.
6. J.J. Chaplin, J. Kent Crawford, R.A. Brightbill, Water-Quality Monitoring in Response to Young-of-the- Year Smallmouth Bass (Micropterus dolomieu) Mortality in the Susquehanna River and Major Tributaries, Pennsylvania: 2008, 2009. https://pubs.usgs.gov/of/2009/1216/ofr2009-1216.pdf.
7. L.R. Iwanowicz, V.S. Blazer, C.P. Guy, A.E. Pinkney, J.E. Mullican, D.A. Alvarez, Reproductive health of bass in the Potomac, USA, drainage: Part 1. Exploring the effects of proximity to wastewater treatment plant discharge, Environ. Toxicol. Chem. 28 (2009) 1072–1083. http://dx.doi.org/10.1897%2F08-433.1.
8. V.S. Blazer, L.R. Iwanowicz, H. Henderson, P.M. Mazik, J.A. Jenkins, D.A. Alvarez, J.A. Young, Reproductive endocrine disruption in smallmouth bass (Micropterus dolomieu) in the Potomac River basin: Spatial and temporal comparisons of biological effects, Environ. Monit. Assess. 184 (2012) 4309–4334. doi:10.1007/s10661-011-2266-5.
9. V.S. Blazer, D.D. Iwanowicz, H.L. Walsh, A.J. Sperry, L.R. Iwanowicz, D.A. Alvarez, R.A.

Brightbill, G. Smith, W.T. Foreman, R. Manning, Reproductive health indicators of fishes from Pennsylvania watersheds: association with chemicals of emerging concern, Environ. Monit. Assess. 186 (2014) 6471–6491. doi:10.1007/s10661-014-3868-5.

10. M.K. Schall, V.S. Blazer, H.L. Walsh, G.D. Smith, T. Wertz, T. Wagner, Spatial and temporal variability of myxozoan parasite, Myxobolus inornatus, prevalence in young of the year smallmouth bass in the Susquehanna River Basin, Pennsylvania, J. Fish Dis. 41 (2018) 1689–1700. doi:10.1111/jfd.12878.
11. H.L. Walsh, V.S. Blazer, G.D. Smith, M. Lookenbill, D.A. Alvarez, K.L. Smalling, Risk Factors Associated with Mortality of Age-0 Smallmouth Bass in the Susquehanna River Basin, Pennsylvania, J. Aquat. Anim. Health. 30 (2018) 65–80. doi:10.1002/aah.10009.
12. V.S. Blazer, L.R. Iwanowicz, C.E. Starliper, D.D. Iwanowicz, P. Barbash, J.D. Hedrick, S.J. Reeser, J.E. Mullican, S.D. Zaugg, M.R. Burkhardt, J. Kelble, Mortality of centrarchid fishes in the potomac drainage: Survey results and overview of potential contributing factors, J. Aquat. Anim. Health. 22 (2010) 190–218. doi:10.1577/H10-002.1.
13. G.D. Smith, V.S. Blazer, H.L. Walsh, L.R. Iwanowicz, C. Starliper, A.J. Sperry, The effects of disease-related mortality of young-of-year smallmouth bass on the population characteristics in the Susquehanna River basin, Pennsylvania and potential implications to conservation of black bass diversity., in: Am. Fish. Soc. Symp., 2015: pp. 82:319-332.
14. D.A. Alvarez, W.L. Cranor, S.D. Perkins, V.L. Schroeder, L.R. Iwanowicz, R.C. Clark, C.P. Guy, A.E. Pinkney, V.S. Blazer, J.E. Mullican, Reproductive health of bass in the potomac, USA, drainage: part 2. Seasonal occurrence of persistent and emerging organic contaminants, Environ. Toxicol. Chem. 28 (2009) 1084–1095. http://www.scopus.com/inward/record.url?eid=2-s2.0-67249093449&partnerID=40.
15. D.W. Kolpin, V.S. Blazer, J.L. Gray, M.J. Focazio, J.A. Young, D.A. Alvarez, L.R. Iwanowicz, W.T. Foreman, E.T. Furlong, G.K. Speiran, S.D. Zaugg, L.E. Hubbard, M.T. Meyer, M.W. Sandstrom, L.B. Barber, Chemical contaminants in water and sediment near fish nesting sites in the Potomac River basin: Determining potential exposures to smallmouth bass (Micropterus dolomieu), Sci. Total Environ. 443 (2013) 700–716. doi:10.1016/j.scitotenv.2012.09.063.
16. A.G. Reif, J.K. Crawford, C.A. Loper, A. Proctor, R. Manning, R. Titler, Occurrence of Pharmaceuticals, Hormones, and Organic Wastewater Compounds in Pennsylvania Waters, 2006 – 09, 2012.
17. Y. Lor, A. Revak, J. Weigand, E. Hicks, D.R. Howard, T.C. King-Heiden, Juvenile exposure

to vinclozolin shifts sex ratios and impairs reproductive capacity of zebrafish, Reprod. Toxicol. 58 (2015) 111–118. doi:10.1016/j.reprotox.2015.09.003.

18. M. Bayley, M. Junge, E. Baatrup, Exposure of juvenile guppies to three antiandrogens causes demasculinization and a reduced sperm count in adult males, Aquat. Toxicol. 56 (2002) 227–239. doi:10.1016/S0166-445X(01)00210-7.
19. R.H. Milston, M.S. Fitzpatrick, A.T. Vella, S. Clements, D. Gundersen, G. Feist, T.L. Crippen, J. Leong, C.B. Schreck, Short-term exposure of chinook salmon (Oncoryhnchus tshawytscha) to o,p′-DDE or DMSO during early life-history stages causes long-term humoral immunosuppression, Environ. Health Perspect. 111 (2003) 1601–1607. doi:10.1289/ehp.6171.
20. L. Yao, Y.Z. Lv, L.J. Zhang, W.R. Liu, J.L. Zhao, Y.Y. Yang, Y.W. Jia, Y.S. Liu, L.Y. He, G.G. Ying, Bioaccumulation and risks of 24 personal care products in plasma of wild fish from the Yangtze River, China, Sci. Total Environ. 665 (2019) 810–819. doi:10.1016/j.scitotenv.2019.02.176.
21. G. Boisvert, C. Sonne, F.F. Rigét, R. Dietz, R.J. Letcher, Bioaccumulation and biomagnification of perfluoroalkyl acids and precursors in East Greenland polar bears and their ringed seal prey, Environ. Pollut. 252 (2019) 1335–1343. doi:10.1016/j.envpol.2019.06.035.
22. USEPA, Method 3500C: Organic Extraction and Sample Preparation, 2007.
23. W. Sanchez, J.M. Porcher, Fish biomarkers for environmental monitoring within the Water Framework Directive of the European Union, TrAC - Trends Anal. Chem. 28 (2009) 150–158. doi:10.1016/j.trac.2008.10.012.
24. K. Booij, C.D. Robinson, R.M. Burgess, P. Mayer, C.A. Roberts, L. Ahrens, I.J. Allan, J. Brant, L. Jones, U.R. Kraus, M.M. Larsen, P. Lepom, J. Petersen, D. Pröfrock, P. Roose, S. Schäfer, F. Smedes, C. Tixier, K. Vorkamp, P. Whitehouse, Passive Sampling in Regulatory Chemical Monitoring of Nonpolar Organic Compounds in the Aquatic Environment, Environ. Sci. Technol. 50 (2016) 3–17. doi:10.1021/acs.est.5b04050.
25. S. Furihata, A. Kasai, K. Hidaka, M. Ikegami, H. Ohnishi, K. Goka, Ecological risks of insecticide contamination in water and sediment around off-farm irrigated rice paddy fields, Environ. Pollut. 251 (2019) 628–638. doi:10.1016/j.envpol.2019.05.009.
26. M. Anastassiades, S.J. Lehotay, D. Stajbaher, F.J. Schenck, Fast and Easy Multiresidue Method Employing Acetonitrile, J. AOAC Int. 86 (2003) 412–431. doi:10.2478/s11687-011-0011-9.

27. T. Rejczak, T. Tuzimski, A review of recent developments and trends in the QuEChERS sample preparation approach, Open Chem. 13 (2015) 980–1010. doi:10.1515/chem-2015-0109.
28. Á. Santana-Mayor, B. Socas-Rodríguez, A. V. Herrera-Herrera, M.Á. Rodríguez-Delgado, Current trends in QuEChERS method. A versatile procedure for food, environmental and biological analysis, TrAC - Trends Anal. Chem. 116 (2019) 214–235. doi:10.1016/j.trac.2019.04.018.
29. J. Camilleri, E. Vulliet, Determination of steroid hormones in sediments based on quick, easy, cheap, effective, rugged, and safe (modified-QuEChERS) extraction followed by liquid chromatography-tandem mass spectrometry (LC-MS/MS), Anal. Methods. 7 (2015) 9577–9586. doi:10.1039/c5ay01972a.
30. T.V. Madureira, S. Velhote, C. Santos, C. Cruzeiro, E. Rocha, M.J. Rocha, A step forward using QuEChERS (Quick, Easy, Cheap, Effective, Rugged, and Safe) based extraction and gas chromatography-tandem mass spectrometry-levels of priority polycyclic aromatic hydrocarbons in wild and commercial mussels, Environ. Sci. Pollut. Res. 21 (2014) 6089–6098. doi:10.1007/s11356-014-2510-y.
31. M. Yang, Y. Ma, W. Gui, Y. Ren, G. Zhu, Y. Liu, Determination of 26 endocrine disrupting chemicals in fish and water using modified QuEChERS combined with solid-phase extraction and UHPLC-MS/MS, Anal. Methods. 7 (2015) 8418–8431. doi:10.1039/c5ay01082a.
32. J. Manuelmolina-Ruiz, E. Cieslik, I. Cieslik, I. Walkowska, Determination of pesticide residues in fish tissues by modified QuEChERS method and dual-d-SPE clean-up coupled to gas chromatography–mass spectrometry, Environ. Sci. Pollut. Res. 22 (2014) 369–378. doi:10.1007/s11356-014-3361-2.
33. P.H. Stefanuto, K.A. Perrault, R.M. Lloyd, B. Stuart, T. Rai, S.L. Forbes, J.F. Focant, Exploring new dimensions in cadaveric decomposition odour analysis, Anal. Methods. 7 (2015) 2287–2294. doi:10.1039/c5ay00371g.
34. S. Fernando, A. Renaguli, M.S. Milligan, J.J. Pagano, P.K. Hopke, T.M. Holsen, B.S. Crimmins, Comprehensive Analysis of the Great Lakes Top Predator Fish for Novel Halogenated Organic Contaminants by GC×GC-HR-ToF Mass Spectrometry, Environ. Sci. Technol. 52 (2018) 2909–2917. doi:10.1021/acs.est.7b05999.
35. P.H. Stefanuto, K.A. Perrault, L.M. Dubois, B. L'Homme, C. Allen, C. Loughnane, N. Ochiai, J.F. Focant, Advanced method optimization for volatile aroma profiling of beer using two-dimensional gas chromatography time-of-flight mass spectrometry, J. Chromatogr. A.

1507 (2017) 45–52. doi:10.1016/j.chroma.2017.05.064.

36. K.M. Pierce, J.C. Hoggard, J.L. Hope, P.M. Rainey, A.N. Hoofnagle, R.M. Jack, B.W. Wright, R.E. Synovec, Fisher ratio method applied to third-order separation data to identify significant chemical components of metabolite extracts, Anal. Chem. 78 (2006) 5068–5075. doi:10.1021/ac0602625.
37. Foods of plant origin - Multimethod for the determination of pesticide residues using GC- and LC- based analysis following acetonitrile extraction/partitioning and clean-up by dispersive SPE - Modular QuEChERS-method, European Committee for Standardization (CEN), Belgium, 2008.
38. U.S. EPA, Definition and Procedure for the Determination of the Method Detection Limit, Revision 2: EPA 821-R - 16-0 06, 2016. www.epa.gov (accessed July 31, 2019).
39. M.L. Trego, E. Hoh, A. Whitehead, N.M. Kellar, M. Lauf, D.O. Datuin, R.L. Lewison, Contaminant Exposure Linked to Cellular and Endocrine Biomarkers in Southern California Bottlenose Dolphins, Environ. Sci. Technol. 53 (2019) 3810–3822. doi:10.1021/acs.est.8b06487.
40. P. Vasseur, C. Cossu-Leguille, Linking molecular interactions to consequent effects of persistent organic pollutants (POPs) upon populations, Chemosphere. 62 (2006) 1033–1042. doi:10.1016/j.chemosphere.2005.05.043.
41. S. Mandal, R. Poi, S. Bhattacharyya, I. Ansary, S.D. Roy, D.K. Hazra, R. Karmakar, Multiclass Multipesticide Residue Analysis in Fish Matrix by a Modified QuEChERS Method Using Gas Chromatography with Mass Spectrometric Determination, J. AOAC Int. 103 (2020) 62–67. doi:10.5740/jaoacint.19-0205.
42. K. Kalachova, J. Pulkrabova, L. Drabova, T. Cajka, V. Kocourek, J. Hajslova, Simplified and rapid determination of polychlorinated biphenyls, polybrominated diphenyl ethers, and polycyclic aromatic hydrocarbons in fish and shrimps integrated into a single method, Anal. Chim. Acta. 707 (2011) 84–91. doi:10.1016/j.aca.2011.09.016.
43. U. EPA Office of Pesticide Programs, BIOPESTICIDES REGISTRATION ACTION DOCUMENT Indole PC Code 025000, 2009.
44. J. Kim, W. Park, Indole: a signaling molecule or a mere metabolic byproduct that alters bacterial physiology at a high concentration?, J. Microbiol. 53 (2015) 421–428. doi:10.1007/s12275-015-5273-3.
45. U.S.D. of H. and H. Services, Toxicological Profile for Polychlorinated Biphenyls (PCBs), 2002. doi:10.1201/9781420061888_ch129.

46. Pennsylvania Department of Environmental Protection, Office of water programs bureau of clean water assessment methodology for rivers and streams 2018, 2018.
47. M. Van Den Berg, L. Birnbaum, A.T.C. Bosveld, B. Brunström, P. Cook, M. Feeley, J.P. Giesy, A. Hanberg, R. Hasegawa, S.W. Kennedy, T. Kubiak, J.C. Larsen, F.X.R. Van Leeuwen, A.K.D. Liem, C. Nolt, R.E. Peterson, L. Poellinger, S. Safe, D. Schrenk, D. Tillitt, M. Tysklind, M. Younes, F. Wærn, T. Zacharewski, Toxic equivalency factors (TEFs) for PCBs, PCDDs, PCDFs for humans and wildlife, Environ. Health Perspect. 106 (1998) 775–792. doi:10.1289/ehp.98106775.
48. C. Starliper, V.S. Blazer, L.R. Iwanowicz, H.L. Walsh, Microbial isolates in diseased fishes, primarily smallmouth bass (Micropterus dolomieu), within the Chesapeake Bay drainage in 2009-2011, in: Proc. West VIrginia Acad. Sci. Submitt. Pap. Eight-Eighth Annu. Sess. 2013, 2013.
49. H.L. Walsh, V.S. Blazer, L.R. Iwanowicz, G. Smith, A Redescription of Myxobolus inornatus from Young-of-the-Year Smallmouth Bass (Micropterus dolomieu), J. Parasitol. 98 (2013) 1236–1242. doi:10.1645/ge-3081.1.
50. J. Heim, Utilization of Statistical Compare Software and Fisher Ratios Prior to Multivariate Analysis for Complex GCxGC-TOFMS Data in Order to Define Statistical Variation Between the Small Molecule Metabolite Profiles of Different Fish Species, 2012. http://www.leco.com/support/application-support-analytical/app-notes

2.8 Supporting Information

2.8.1 Targeted Analysis Data Processing

Data was processed using LECO ChromaTOF BT software v5.40.12.0. The following parameters were used following optimization of the data processing method. Initial peak finding was accomplished using a signal-to-noise ratio of 20 and a minimum match of 500 required to combine slices. Analytes were quantitated using extracted ion chromatograms with a mass tolerance of 500 ppm.

2.8.2 Non-Targeted Analysis Data Processing

Manufacturer guidelines were followed to import ChromaTOF BT data into ChromaTOF GC software v4.74.2.0 as .netCDF files. Parameters for basic peak finding and statistical compare were determined based on these guidelines. Initial peak finding was accomplished using a baseline offset of 1.0, no smoothing of data points, minimum signal-to-noise ratio of 200, and minimum match of 500 required to combine slices. Compound identification using the NIST MS Database (2014) required a minimum similarity match of 700. During the alignment process of statistical compare a second peak finding method is employed. Second peak finding was accomplished using a signal-to-noise ratio of 50. Alignment required a minimum match similarity of 600 for compounds. For aligned compounds to be retained within the final aligned peak table, compounds had to occur in at least 60% of samples within a class or within at least 60% of all samples.

2.8.3 Tables

Table 2.3 Standards used for method development and targeted analysis application. Standards used for method development are listed in red.

Compounds	Cas #	Manufacturer
1-methylnaphthalene	90-12-0	Chem Service
2,2',4,4',5,5'-Hexabromobiphenyl	59080-40-9	Restek
2,4'-Methoxychlor	30667-99-3	Restek
2,6-dimethylnaphthalene	581-42-0	Chem Service
2-methylnaphthalene	91-57-6	Chem Service
4,4'-Dichlorobenzophenone	90-98-2	Restek
4,4'-Methoxychlor olefin	2132-70-9	Restek
9,10-anthraquinone	84-65-1	Chem Service
Acetophenone	98-86-2	Restek
Aldrin	309-00-2	Restek
anthracene	120-12-7	Restek
BDE-100	189084-64-8	Restek
BDE-153	68631-49-2	Restek
BDE-47	5436-43-1	Restek
BDE-99	60348-60-9	Restek
benfluralin	1861-40-1	Chem Service
benzophenone	119-61-9	Chem Service
caffeine	58-08-2	Restek
Celestolide	13171-00-1	TCI America
Chlorbenside	103-17-3	Restek
Chlorfenson (Ovex)	80-33-1	Restek
Chloroneb	2675-77-6	Restek
chlorpyrifos	2921-88-2	Restek
cis-Chlordane	5103-71-9	Restek
cis-Nonachlor	5103-73-1	Restek
cotinine	486-56-6	Restek
Dacthal	1861-32-1	Restek
DEET	134-62-3	Chem Service
DEHP Bis(2-ethylhexyl)phthalate	117-81-7	Restek
diazinon	333-41-5	Chem Service
dichlorvos	62-73-7	Chem Service
Dieldrin	60-57-1	Restek
diethyl phthalate	84-66-2	Chem Service
dl-camphor	76-22-2	Chem Service
Endosulfan ether	3369-52-6	Restek
Endosulfan I	959-98-8	Restek

Compounds	Cas #	Manufacturer
Endosulfan II	33213-65-9	Restek
Endosulfan sulfate	1031-07-8	Restek
Endrin	72-20-8	Restek
Endrin aldehyde	7421-93-4	Restek
Endrin ketone	53494-70-5	Restek
Ethylan (Perthane)	72-56-0	Restek
Fenson	80-38-6	Restek
fluoranthene	206-44-0	Restek
Galaxolide	1222-05-5	TCI America
Heptachlor	76-44-8	Restek
Heptachlor epoxide	1024-57-3	Restek
Hexachlorobenzene	118-74-1	Restek
Indole	120-72-9	Chem Service
Isodrin	465-73-6	Restek
Isophorone	78-59-1	Chem Service
isoquinoline	119-65-3	Chem Service
menthol	1490-04-6	Chem Service
metalaxyl	57837-19-1	Chem Service
Methoxychlor	72-43-5	Restek
methyl salicylate	119-36-8	Chem Service
methyl triclosan	4640_01_1	Dr. Ehrenstorfer GmbH
Mirex	2385-85-5	Restek
naphthalene	91-20-3	Restek
o,p'-DDD	53-19-0	Restek
o,p'-DDE	3424-82-6	Restek
o,p'-DDT	789-02-6	Restek
oxychlordane	27304-13-8	Chem Service
p,p'-DDD	72-54-8	Restek
p,p'-DDE	72-55-9	Restek
p,p'-DDT	50-29-3	Restek
Pentachloroanisole	1825-21-4	Restek
Pentachlorobenzene	608-93-5	Restek
Pentachlorothioanisole	1825-19-0	Restek
Phantolid	15323-35-0	Key Organics
Phenanthrene	85-01-8	Restek
pyrene	129-00-0	Restek
Tetradifon	116-29-0	Restek
Tonalid	21145-77-7	Key Organics
trans-Chlordane	5103-74-2	Restek
trans-Nonachlor	39765-80-5	Restek

Compounds	Cas #	Manufacturer
Traseolide	68140-48-7	Chemcruz
tributoxyethyl phosphate	78-51-3	Chem Service
tributyl phosphate	126-73-8	Chem Service
triethyl citrate	77-93-0	Chem Service
Trifluralin	1582-09-8	Restek
triphenyl phosphate	115-86-6	Chem Service
tris(1,3-dichloroisopropyl)phosphate	13674-87-8	Restek
tris(2-chloroethyl)phosphate	115-96-8	Chem Service
α-HCH	319-84-6	Restek
β-HCH	319-85-7	Restek
γ-HCH	58-89-9	Restek
δ-HCH	319-86-8	Restek

Table 2.4 Polychlorinated biphenyl congeners used for method development and targeted analysis application. Congeners used in method development are listed in red.

PCB Congener	Cas #	Manufacturer	PCB Congener	Cas #	Manufacturer
1	2051-60-7	AccuStandard	46	41464-47-5	AccuStandard
2	2051-61-8	AccuStandard	47	2437-79-8	AccuStandard
3	2051-62-9	AccuStandard	48	70362-47-9	AccuStandard
4	13029-08-8	AccuStandard	49	41464-40-8	AccuStandard
5	16605-91-7	AccuStandard	51	68194-04-7	AccuStandard
6	34883-39-1	AccuStandard	52	41464-39-5	AccuStandard
7	33284-50-3	AccuStandard	53	41464-41-9	AccuStandard
8	25569-80-6	AccuStandard	54	15968-05-5	AccuStandard
9	34883-43-7	AccuStandard	56	41464-46-4	AccuStandard
10	33146-45-1	AccuStandard	59	74472-33-6	AccuStandard
12	2974-92-7	AccuStandard	60	33025-41-1	AccuStandard
13	2974-90-5	AccuStandard	63	74472-34-7	AccuStandard
14	34883-41-5	AccuStandard	64	52663-58-8	AccuStandard
15	2050-68-2	AccuStandard	66	73557-53-8	AccuStandard
16	38444-73-4	AccuStandard	67	32690-93-0	AccuStandard
17	37680-66-3	AccuStandard	69	60233-24-1	AccuStandard
18	37680-65-2	AccuStandard	70	32598-11-1	AccuStandard
19	38444-78-9	AccuStandard	71	32598-10-0	AccuStandard
20	38444-84-7	AccuStandard	73	74338-23-1	AccuStandard
22	55712-37-3	AccuStandard	74	41464-43-1	AccuStandard
24	55702-45-9	AccuStandard	75	32598-12-2	AccuStandard
25	7012-37-5	AccuStandard	77	32598-13-3	AccuStandard
26	38444-81-4	AccuStandard	81	70362-50-4	AccuStandard
27	38444-76-7	AccuStandard	82	38380-01-7	AccuStandard
28	38444-85-8	AccuStandard	83	60145-20-2	AccuStandard
29	15862-07-4	AccuStandard	84	52663-60-2	AccuStandard
31	16606-02-3	AccuStandard	85	65510-45-4	AccuStandard
32	38444-77-8	AccuStandard	87	38380-02-8	AccuStandard
33	38444-86-9	AccuStandard	90	68194-07-0	AccuStandard
34	37680-68-5	AccuStandard	91	68194-05-8	AccuStandard
35	37680-69-6	AccuStandard	92	52663-61-3	AccuStandard
37	38444-90-5	AccuStandard	93	73575-56-1	AccuStandard
40	38444-93-8	AccuStandard	95	38379-99-6	AccuStandard
41	52663-59-9	AccuStandard	97	41464-51-1	AccuStandard
42	36559-22-5	AccuStandard	99	38380-03-9	AccuStandard
44	35693-99-3	AccuStandard	100	39485-83-1	AccuStandard
45	70362-45-7	AccuStandard	101	37680-73-2	AccuStandard

PCB Congener	Cas #	Manufacturer	PCB Congener	Cas #	Manufacturer
103	60145-21-3	AccuStandard	164	74472-45-0	AccuStandard
104	56558-16-8	AccuStandard	165	74472-46-1	AccuStandard
105	32598-14-4	AccuStandard	167	52663-72-6	AccuStandard
107	70424-68-9	AccuStandard	170	35065-30-6	AccuStandard
110	52663-62-4	AccuStandard	171	52663-71-5	AccuStandard
114	74472-37-0	AccuStandard	172	52663-74-8	AccuStandard
115	74472-38-1	AccuStandard	173	35065-28-2	AccuStandard
117	68194-11-6	AccuStandard	174	52663-68-0	AccuStandard
118	31508-00-6	AccuStandard	175	40186-70-7	AccuStandard
119	56558-17-9	AccuStandard	176	52663-65-7	AccuStandard
122	76842-07-4	AccuStandard	177	38411-25-5	AccuStandard
123	65510-44-3	AccuStandard	178	52663-67-9	AccuStandard
124	70424-70-3	AccuStandard	179	52663-70-4	AccuStandard
128	38380-07-3	AccuStandard	180	68194-16-1	AccuStandard
129	55215-18-4	AccuStandard	183	52663-69-1	AccuStandard
130	52663-66-8	AccuStandard	185	52712-05-7	AccuStandard
131	61798-70-7	AccuStandard	187	35065-29-3	AccuStandard
132	38380-05-1	AccuStandard	189	39635-31-9	AccuStandard
134	52704-70-8	AccuStandard	190	41411-64-7	AccuStandard
135	52744-13-5	AccuStandard	191	74472-50-7	AccuStandard
136	38411-22-2	AccuStandard	193	69782-91-8	AccuStandard
137	35694-06-5	AccuStandard	194	52663-75-9	AccuStandard
138	68194-13-8	AccuStandard	195	52663-76-0	AccuStandard
141	52712-04-6	AccuStandard	196	42740-50-1	AccuStandard
144	68194-14-9	AccuStandard	197	33091-17-7	AccuStandard
146	51908-16-8	AccuStandard	199	52663-78-2	AccuStandard
147	35065-27-1	AccuStandard	200	52663-73-7	AccuStandard
149	38380-04-0	AccuStandard	201	40186-71-8	AccuStandard
151	52663-63-5	AccuStandard	202	2136-99-4	AccuStandard
153	52663-64-6	AccuStandard	203	35694-08-7	AccuStandard
154	60145-22-4	AccuStandard	205	74472-53-0	AccuStandard
156	38380-08-4	AccuStandard	206	40186-72-9	AccuStandard
157	69782-90-7	AccuStandard	207	52663-79-3	AccuStandard
158	74472-42-7	AccuStandard	208	52663-77-1	AccuStandard
163	74472-44-9	AccuStandard	209	2051-24-3	AccuStandard

Table 2.5 Summary of recovery values using acetonitrile as the extraction solvent vs. ethyl acetate. Method detection limit (MDL) calculated using final optimization of method.

Analyte	Acetonitrile		Ethyl Acetate		MDL (ng/g)
	Mean % Recovery	*Std. Error*	*Mean % Recovery*	*Std. Error*	
BDE-47	10.67	0.83	16.76	1.98	0.04
BDE-99	6.11	0.74	7.37	1.37	0.22
BDE-100	4.24	1.76	7.33	1.79	0.31
BDE-153	0.00	N/A	41.42	2.12	2.17
2,2',4,4',5,5'-Hexabromobiphenyl	0.00	N/A	48.57	2.83	0.97
PCB 001	15.56	1.32	59.10	4.85	0.18
PCB 002	18.96	1.52	68.37	4.45	0.20
PCB 003	19.47	1.52	74.52	5.75	0.54
PCB 004	12.04	0.89	57.23	4.10	0.37
PCB 006	16.78	0.77	76.80	3.10	0.33
PCB 008	16.33	0.65	78.11	3.48	0.32
PCB 009	15.31	0.72	71.69	1.76	0.29
PCB 016	11.82	1.17	40.69	4.95	0.16
PCB 018	11.08	1.07	39.58	6.26	0.15
PCB 019	15.31	1.12	31.82	3.19	0.28
PCB 022	13.96	0.74	54.79	8.45	0.11
PCB 025	13.64	0.87	50.70	7.81	0.13
PCB 028	13.22	0.87	54.73	8.48	0.11
PCB 044	13.03	0.88	18.97	4.34	0.10
PCB 052	12.88	0.71	22.15	4.53	0.10
PCB 056	7.78	3.61	56.68	1.98	0.23
PCB 066	13.23	0.73	55.16	0.91	0.25
PCB 067	12.55	1.17	34.71	5.43	0.16
PCB 071	13.31	1.02	31.27	6.64	0.10
PCB 074	11.77	1.35	39.54	8.24	0.16
PCB 082	8.19	0.23	30.93	0.76	0.15
PCB 087	4.97	1.94	18.86	0.59	0.20
PCB 099	7.73	0.38	30.63	1.41	0.29
PCB 110	0.00	N/A	55.68	4.39	0.66
PCB 138	9.27	0.11	21.39	11.04	0.46
PCB 146	7.27	0.03	29.59	3.89	0.11
PCB 147	7.36	0.16	24.54	5.57	0.08
PCB 153	7.68	0.26	20.67	9.35	0.19
PCB 173	7.44	1.95	11.27	11.27	0.22
PCB 174	9.29	0.28	28.28	9.42	0.40

Analyte	Acetonitrile		Ethyl Acetate		MDL (ng/g)
	Mean % Recovery	*Std. Error*	*Mean % Recovery*	*Std. Error*	
PCB 177	8.62	0.17	12.00	11.41	0.25
PCB 179	8.11	0.21	5.25	1.98	0.31
PCB 180	5.17	2.72	35.37	1.29	0.06
PCB 194	3.88	2.03	78.10	1.35	0.08
PCB 195	5.85	1.97	71.56	8.35	0.26
PCB 199	8.04	0.50	60.75	1.11	0.11
PCB 203	6.24	0.29	66.44	2.72	0.11
PCB 206	4.67	0.48	87.21	1.39	0.10
α-HCH	15.81	0.84	60.72	9.06	0.64
β-HCH	23.95	1.37	74.31	10.07	0.40
γ-HCH	17.01	1.22	60.28	9.05	0.33
δ-HCH	14.46	1.43	70.55	9.26	0.18
p,p'-DDD	13.73	0.06	60.56	6.84	0.50
p,p'-DDE	7.67	0.41	76.73	7.19	0.52
p,p'-DDT	15.77	0.05	47.24	4.34	0.16
Aldrin	17.19	3.37	62.29	5.26	0.29
cis-Chlordane	2.46	2.46	65.52	3.94	0.45
Dieldrin	0.00	N/A	61.49	7.80	5.26
Endosulfan I	9.81	0.45	11.42	2.39	2.32
Endosulfan II	0.59	0.15	12.96	1.78	0.35
Endosulfan sulfate	7.36	4.00	38.69	2.79	0.57
Endrin	11.12	3.11	73.92	14.47	3.70
Endrin aldehyde	0.00	N/A	34.81	4.51	2.34
Endrin ketone	11.47	0.44	21.51	11.07	6.05
Heptachlor	14.43	0.63	65.50	6.35	0.32
Heptachlor epoxide	19.65	5.10	70.57	7.57	0.22
Methoxychlor	19.56	0.94	28.00	4.61	0.20
trans-Chlordane	1.55	0.15	11.99	3.23	0.91

Table 2.6 List of all 91 features included in multivariate analysis of YOY SMB samples. Named compounds were tentatively identified based on mass spectral match with NIST MS Database (2014) only. Compounds identified as significant features are listed in red.

Name	Mass	Avg t_{R1} (s)	StDev t_{R1}	Avg t_{R2} (s)	StDev t_{R2}	Match	Reverse	Probability
1,3-Oxathiolane, 2-acetyl-2-methyl-	103	2321.54	2.03	2.27	0.01	720	826	4393
11-(3,4-Dimethyl-5-pentyl-2-furyl)-dodecanoic acid, methyl ester	364	2607.60	1.23	2.37	0.01	736	869	9794
11-(3,4-Dimethyl-5-propyl-2-furyl)-undecanoic acid, methyl ester	307	2448.63	1.50	2.31	0.01	816	882	9720
13-Docosenamide, (Z)-	59	2852.40	1.23	2.80	0.01	855	866	9396
2,3,3,4,7-Pentamethyl-2,3-dihydro-benzofuran	175	1359.81	0.87	2.04	0.01	730	741	1636
2,4,4-Trimethyl-2-butenolide	111	1147.81	0.87	2.64	0.01	718	843	5002
2-Hexyldecanol	280	2204.00	1.46	1.72	0.10	718	770	2478
2-Piperidinone, 1-methyl-	113	939.06	2.25	2.54	0.01	859	872	8963
2-Pyrrolidinone, 1-methyl-	99	753.40	2.98	2.38	0.01	875	897	9342
2-Tetradecanone	58	1631.81	0.87	1.86	0.01	788	806	5305
3-Buten-2-one, 4-(6,6-dimethyl-1-cyclohexen-1-yl)-	59	1668.60	1.96	2.10	0.01	708	787	1896
4-(2,6,6-Trimethylcyclohexa-1,3-dienyl)but-3-en-2-one	175	1471.58	1.26	2.15	0.01	830	842	6690
4,7,10,13,16,19-Docosahexaenoic acid, methyl ester, (all-Z)-	79	2581.00	1.78	2.46	0.01	924	939	6582
7-Acetyl-6-ethyl-1,1,4,4-tetramethyltetralin	243	2119.76	0.97	2.19	0.01	727	771	8746
8-Hydroxyquinoline, 2-methylpropionate	145	2780.00	0.00	2.42	0.01	712	806	2872
9,12-Octadecadienoic acid (Z,Z)-, methyl ester	294	2231.58	1.84	2.10	0.12	892	900	3636
9-Octadecenamide, (Z)-	262	2478.12	3.20	2.53	0.32	723	766	2988
9-Octadecenoic acid (Z)-, 2,3-dihydroxypropyl ester	264	2759.27	3.93	2.84	0.04	724	795	8050
à-Hydroxyisobutyric acid, acetate	59	520.19	0.87	1.47	0.01	710	752	5046
Cholesta-3,5-diene	260	3182.67	3.60	2.44	0.03	772	824	2899
Cholesterol	276	3162.35	4.49	1.74	0.18	719	729	6634
Ethylparaben	121	1568.94	2.25	2.22	0.01	733	836	5738
Heptadecane, 2-methyl-	85	1715.71	4.83	1.71	0.08	780	847	2118

Name	Mass	Avg t_{R1} (s)	StDev t_{R1}	Avg t_{R2} (s)	StDev t_{R2}	Match	Reverse	Probability
Indole	90	1160.38	1.20	2.72	0.01	936	944	6169
Methyl 3-hydroxyoctadecanoate	103	2873.78	2.05	2.61	0.01	736	817	4832
Myristic acid glycidyl ester	297	2623.50	1.41	2.34	0.01	760	835	8324
n-Hexadecanoic acid	60	2625.45	3.24	1.73	0.15	792	826	6135
Octacosane	351	2884.57	2.14	2.02	0.02	760	821	1415
Octadecanoic acid	330	2613.85	2.08	1.72	0.14	757	762	2383
Palmitinic acid	270	3084.42	3.75	1.53	0.04	836	842	6947
Pentadecanoic acid, methyl ester	74	1928.00	0.00	1.89	0.01	910	924	7007
Vitamin E	430	3145.26	1.91	3.00	0.02	863	906	4371
Analyte 29	71	535.81	0.87	1.62	0.01	N/A	N/A	N/A
Analyte 157	367	783.11	4.01	2.13	0.02	N/A	N/A	N/A
Analyte 171	367	814.89	3.31	2.03	0.01	N/A	N/A	N/A
Analyte 219	441	897.41	3.14	2.68	0.02	N/A	N/A	N/A
Analyte 256	441	964.44	2.33	2.38	0.02	N/A	N/A	N/A
Analyte 266	441	989.20	1.93	2.30	0.02	N/A	N/A	N/A
Analyte 273	367	1000.33	1.15	1.68	0.02	N/A	N/A	N/A
Analyte 282	367	1019.71	1.07	1.65	0.01	N/A	N/A	N/A
Analyte 348	516	1116.71	2.11	2.72	0.02	N/A	N/A	N/A
Analyte 386	113	1175.58	1.26	2.71	0.01	N/A	N/A	N/A
Analyte 445	441	1246.13	2.56	1.73	0.01	N/A	N/A	N/A
Analyte 448	472	1251.81	0.87	1.47	0.01	N/A	N/A	N/A
Analyte 466	441	1276.24	2.22	1.69	0.01	N/A	N/A	N/A
Analyte 535	355	1371.56	1.29	1.75	0.01	N/A	N/A	N/A
Analyte 580	125	1443.80	0.89	3.03	0.01	N/A	N/A	N/A
Analyte 722	491	1616.00	0.00	1.44	0.01	N/A	N/A	N/A
Analyte 810	429	1754.53	4.26	1.77	0.12	N/A	N/A	N/A
Analyte 850	149	1824.60	1.96	2.27	0.01	N/A	N/A	N/A
Analyte 1210	321	2100.00	0.00	3.71	0.01	N/A	N/A	N/A

Name	Mass	Avg t_{R1} (s)	StDev t_{R1}	Avg t_{R2} (s)	StDev t_{R2}	Match	Reverse	Probability
Analyte 1878	322	2340.71	1.57	2.28	0.01	N/A	N/A	N/A
Analyte 1938	176	2358.33	2.06	3.37	0.02	N/A	N/A	N/A
Analyte 1949	165	2363.50	2.48	3.10	0.01	N/A	N/A	N/A
Analyte 2110	336	2416.67	1.53	2.30	0.01	N/A	N/A	N/A
Analyte 2186	273	2444.00	2.67	1.62	0.01	N/A	N/A	N/A
Analyte 2252	266	2464.92	1.75	3.16	0.02	N/A	N/A	N/A
Analyte 2272	326	2472.00	1.57	2.07	0.01	N/A	N/A	N/A
Analyte 2307	527	2480.80	3.68	1.35	0.00	N/A	N/A	N/A
Analyte 2378	131	2504.00	3.58	2.08	0.05	N/A	N/A	N/A
Analyte 2383	215	2504.00	1.30	2.67	0.03	N/A	N/A	N/A
Analyte 2384	296	2504.29	1.07	2.71	0.01	N/A	N/A	N/A
Analyte 2666	393	2581.45	2.02	2.33	0.01	N/A	N/A	N/A
Analyte 2679	393	2587.00	1.85	2.12	0.01	N/A	N/A	N/A
Analyte 2683	103	2587.20	1.79	2.44	0.10	N/A	N/A	N/A
Analyte 2858	248	2629.00	3.10	2.48	0.03	N/A	N/A	N/A
Analyte 2951	191	2652.22	0.94	1.98	0.01	N/A	N/A	N/A
Analyte 3047	377	2671.11	3.33	0.07	0.02	N/A	N/A	N/A
Analyte 3065	300	2675.67	1.15	3.56	0.01	N/A	N/A	N/A
Analyte 3137	377	2692.00	2.83	3.92	0.03	N/A	N/A	N/A
Analyte 3240	464	2721.00	3.15	1.35	0.01	N/A	N/A	N/A
Analyte 3361	476	2752.00	0.00	1.77	0.00	N/A	N/A	N/A
Analyte 3391	412	2762.00	2.19	2.35	0.02	N/A	N/A	N/A
Analyte 3393	215	2763.06	1.75	2.93	0.02	N/A	N/A	N/A
Analyte 3474	286	2780.29	2.92	1.75	0.01	N/A	N/A	N/A
Analyte 3495	103	2788.33	1.15	2.56	0.01	N/A	N/A	N/A
Analyte 3557	534	2808.00	0.00	1.76	0.01	N/A	N/A	N/A
Analyte 3562	131	2812.73	3.00	3.02	0.06	N/A	N/A	N/A
Analyte 3836	112	2894.91	1.87	1.83	0.03	N/A	N/A	N/A

Name	Mass	Avg t_{R1} (s)	StDev t_{R1}	Avg t_{R2} (s)	StDev t_{R2}	Match	Reverse	Probability
Analyte 3871	331	2912.40	2.27	1.97	0.02	N/A	N/A	N/A
Analyte 3872	414	2909.25	2.82	2.10	0.16	N/A	N/A	N/A
Analyte 3977	398	2950.00	4.62	1.85	0.08	N/A	N/A	N/A
Analyte 3997	99	2959.11	1.76	1.76	0.00	N/A	N/A	N/A
Analyte 4031	476	2971.67	2.67	1.97	0.17	N/A	N/A	N/A
Analyte 4476	451	3155.00	2.83	2.18	0.01	N/A	N/A	N/A
Analyte 4556	429	3192.00	3.58	2.60	0.23	N/A	N/A	N/A
Analyte 4611	165	3224.33	2.06	3.75	0.05	N/A	N/A	N/A
Analyte 4732	579	3295.69	3.04	2.07	0.21	N/A	N/A	N/A
Analyte 4925	466	3428.62	3.95	2.06	0.05	N/A	N/A	N/A
Analyte 5317	430	3705.14	3.02	2.29	0.21	N/A	N/A	N/A
Analyte 5470	579	3816.57	4.43	1.82	0.18	N/A	N/A	N/A

Chapter 3

Targeted and Non-targeted Analysis of Young-of-Year Smallmouth Bass Using Comprehensive Two-Dimensional Gas Chromatography Coupled with Time-of-Flight Mass Spectrometry

3.1 Abstract

Smallmouth bass in the Susquehanna River Basin, Chesapeake Bay Watershed, USA, have been exhibiting clinical signs of disease and reproductive endocrine disruption (e.g., intersex, male plasma vitellogenin) for over fifteen years. Previous histological and targeted chemical analyses have identified infectious agents and pollutants in fish tissues including organic contaminants, mercury, and perfluorinated compounds, but a common causative link for the observed signs of disease across this widespread area has not been determined. This study examines 146 young-of-year smallmouth bass collected from 14 sampling sites in the Susquehanna River Basin, Pennsylvania, USA with varying levels of disease prevalence. Whole fish were extracted by a recently developed modification to the quick, easy, cheap, effective, rugged, and safe extraction method and analyzed by comprehensive two-dimensional gas chromatography coupled with time-of-flight mass spectrometry. A targeted analysis was conducted to identify the presence and quantity of 127 known contaminants, including polychlorinated biphenyls, brominated diphenyl ethers, organochlorinated pesticides, and pharmaceutical and personal care products. A non-targeted analysis was conducted on the same data set to identify analytes of interest not included on routine target compound lists. Chromatographic alignment through Statistical Compare (ChromaTOF GC) was followed by Fisher ratio and principal component analysis to reduce the data set from thousands of peaks per sample to a final data set of 65 analytes of interest. Comparisons of these 65 compounds between

Normal (no observed health anomalies) and Lesioned (observed health anomaly at time of collection) fish revealed increased levels of three chemical families in Lesioned fish including esters, ketones, and nitrogen containing compounds.

3.2 Introduction

The effects of anthropogenic contamination, through the release of manufacturing byproducts, pharmaceuticals, and pesticides, are evidenced by their adverse effects on the health of indicator species worldwide.[1,2] Persistent organic pollutants (POPs) are contaminants that do not readily breakdown but persist within the environment and may bioaccumulate in the fatty tissues of animals. Numerous POPs have been recognized as carcinogens, endocrine disruptors, and immunosuppressants capable of impacting the health of multiple species.[3,4] However, not all contaminants are environmentally persistent. Additionally, many harmful substances have not yet been fully recognized for their effects, but rather fall into the category of contaminants of emerging concern (CECs). CECs can generally be defined as new compounds that were previously unknown or known compounds with newly recognized environmental concerns.[5–7] Regardless of a compound's designation as a POP or CEC, the detection of anthropogenic contaminants remains a top priority in environmental monitoring studies, especially those where there are threats to organism health. In the Susquehanna and Potomac River basins of the Chesapeake Bay Watershed, USA, smallmouth bass (*Micropterus dolomieu*, SMB) have exhibited signs of adverse health effects such as dermal lesions, reproductive endocrine disruption, and population declines due to low chronic mortality and sporadic fish kills.[8–14] Adult male SMB have been documented expressing the egg precursor protein vitellogenin and having testicular oocytes, both markers of intersex due to endocrine disruption. Increased rates of chronic mortality and co-infections of bacterial, viral, and/or parasitic pathogens within both adult and

young-of-year (YOY) fish may be indicative of immunosuppression.[14–16] SMB, an indicator species, are extremely sensitive to environmental changes, and previous analyses of water, sediment, and fish tissues have identified the presence of varying levels of environmental contaminants capable of inducing endocrine disruption and immunosuppression.[14,17–19] The majority of previous studies have focused on adult SMB and the identification of estrogenic compounds in surface waters through targeted analyses. However, no one causative agent or contaminant group has yet been linked with the disease characteristics identified. This is likely due to the complex relationship between chemical contaminants, environmental stressors, pathogens, and organism health that requires a broader analytical approach to uncover connections.

Targeted analyses are an ideal approach when contaminants are known or suspected to be in an area. In developed communities, it is possible to predict which chemical contaminants will be present at a site based on the land use around the area (e.g., farmland, manufacturing complexes, or wastewater treatment plants).[17,20] However, target compounds are not always detected and those that are, may not correlate with the observed health impairments, such as intersex and immunosuppression in SMB. This could be due to the limited nature of targeted analyses, where only a select number of compounds are examined in each study. Additionally, even when correlations are drawn from targeted studies, it is likely that this is only a partial explanation as the non-target analytes are ignored in these studies. This precludes the identification of known compounds that were not included on the target list as well as the discovery of CECs. By expanding on the targeted analysis and looking at non-target compounds, it may be possible to detect previously unknown contaminants or metabolites of contaminants that explain the same observed negative health impacts across such a widespread area in the Chesapeake Bay Watershed. Furthermore, the use of an advanced analytical technique like comprehensive two-dimensional gas chromatography (GC × GC) will provide more information

about both target and non-target compounds. As opposed to one-dimensional gas chromatography (1DGC), GC × GC provides increased separation and greater peak capacity resulting in the detection of a significantly greater number of unique analytes per sample.[21,22] This increased separation power can be harnessed to directly examine complex matrices, like fish tissues, for contaminants that would otherwise be obscured due to coelutions and background matrix interference. When combined with a broad range sample preparation technique, like the quick, easy, cheap, effective, rugged, and safe (QuEChERS) extraction,[23–25] researchers can obtain an immense wealth of knowledge about what an individual fish has been exposed to. A high-throughput analytical methodology also allows for the analysis of larger sample sets, adding to the likelihood of successfully identifying correlations between populations.

This study presents a continuation of the work presented by Teehan et al. (2020) on understanding the role of environmental contaminants in SMB disease. An analytical pipeline was developed for the completion of targeted and non-targeted analyses of YOY SMB and applied to a sample set of 21 individuals from two collection sites from the Susquehanna River Basin.[26] YOY SMB were chosen for analysis due to observed mortality events in affected areas and their susceptibility to lesions characteristic of multiple pathogens and parasites, suggesting immunosuppression. Additionally, exposure of juvenile fish to contaminants during key developmental life stages may contribute to lifelong health impairments in adults, even when removed from the source of exposure.[27–29] Therefore, identifying chemical contaminants that YOY SMB are exposed to may provide valuable insight into the observed fish health impairments documented in adult SMB. This study utilizes the developed methodology for sample preparation and data analysis to examine 146 YOY SMB collected from 14 sites throughout the Susquehanna River Basin.

3.3 Materials and Methods

3.3.1 Reagents and Chemicals

A total of 127 compounds were selected for the targeted analysis of YOY SMB (SI Table 3.3). Surrogate compounds [polychlorinated biphenyl (PCB) congeners CB-18, CB-28, and CB-52, triphenyl phosphate, and tris(1,3-dichloroisopropyl)phosphate] as well as internal standard compounds [acenaphthylene-d8, anthracene-d10, and pyrene-d10] were obtained from Restek Corp. (Bellefonte, PA, USA). Analytical grade ethyl acetate and HPLC water were obtained from JT Baker (Phillipsburg, NJ, USA). Buffered QuEChERS extraction salts and Florisil were obtained from Restek Corp.

3.3.2 Sample Collection

Young-of-year SMB were collected during June and July of 2015 from 14 sampling locations throughout the Susquehanna River Basin in Pennsylvania, USA (Figure 3.1 and Table 3.1). Between 9 – 12 fish were collected from each site by electrofishing. Collection site details and observed health anomalies at the time of collection for each fish are provided in SI Table 3.4. For matrix blanks, SMB raised in captivity at the Eastern Ecological Science Center, Kearneysville, WV were used. Captive fish were raised and euthanized in accordance with the Science Center's Institutional Animal Care and Use Committee guidance. Research ponds used for rearing captive fish are fed by well water. These fish represent an out-of-basin control site with a low level of known contaminants detected during previous analyses but displaying none of the observed signs of disease being investigated.[14]

Table 3.1 List of sampling sites and their associated abbreviations and coordinates.

Sampling Location	Abbreviation	Longitude	Latitude
Conodoquinet Creek	CC	-77.040419	40.259451
Juniata River - Howe Township Park	JR	-77.096804	40.49193
Kettle Creek Leidy	KC	-77.92192	41.40506
Loyalsock Creek	LC	-76.93471	41.24238
North Branch Susquehanna Danville	NBSD	-76.598441	40.943603
North Branch Susquehanna Falls	NBSR	-75.86773	41.46166
Pine Creek Ramsey	PCHB	-77.40161	41.34870
Pine Creek Hamilton Bottom	PC	-77.31870	41.27751
Susquehanna Harrisburg Juniata	JF	-76.91709	40.334023
Susquehanna River Harrisburg North Branch	SRNB	-76.906361	40.33630
Susquehanna Harrisburg West Branch	WB	-76.912562	40.33538
Wyalusing Creek	WC	-76.23094	41.69748
Bald Eagle Creek Unionville	UBEC	-77.86952	40.90636
West Branch Susquehanna Lewisburg	WBSL	-76.877843	40.966325

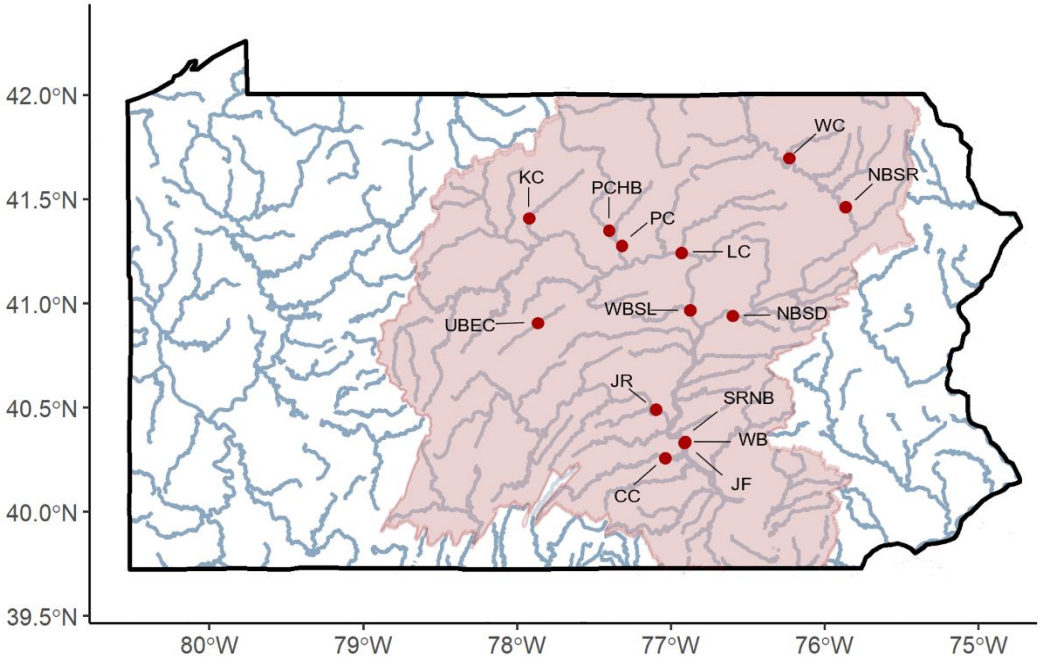

Figure 3.1 Illustration of young-of-year smallmouth bass collection sites within Pennsylvania, USA. The shaded region indicates the boundary of the Susquehanna River Basin. Red circles indicate individual collection sites.

3.3.3 Sample Extraction and Cleanup

All YOY SMB were extracted using a previously described modified QuEChERS protocol.[26] In brief, whole YOY SMB were individually homogenized, and 1 g aliquots of homogenized tissue (or the entire homogenate of samples weighing less than 1 g) were spiked with surrogate compounds at an intended final concentration of 25 ng g^{-1}. Samples were extracted using ethyl acetate and buffered QuEChERS salts, followed by overnight freezing at -20 ℃ to remove excess lipids. Extracts were then cleaned by dispersive solid phase extraction using Florisil. Internal standards were spiked into extracts immediately preceding instrumental analysis at a final concentration of 10 pg μl^{-1} to monitor for variation in instrumental performance. To monitor for contaminants introduced during sample storage or preparation, samples and matrix blanks were stored together at -20 ℃ until batched by location and extracted. A homogenized matrix blank spiked with surrogate compounds was extracted with every sample batch.

3.3.4 Data Acquisition

Samples were analyzed using a Pegasus BT 4D GC × GC-TOFMS (LECO Corp., St. Joseph, MI, USA) equipped with a 7890A GC system (Agilent Technologies, DE, USA). The split/splitless injection port contained a Siltek deactivated straight liner with glass wool (Restek Corp.) A 1 µl pulsed splitless injection, with pulse pressure of 60 psi for 1 minute and splitless hold time of 1 minute, was used to introduce samples onto a DB-5ms (60 m × 0.25 mm ID × 0.25 µm df) column (Agilent Technologies) in the first dimension coupled to an Rxi-17Sil MS (1 m × 0.18 mm ID × 0.18 µm df) column (Restek Corp.) in the second dimension. The inlet was maintained at 250 ℃ for the entirety of the run, and the helium carrier gas flow was maintained at 1.4 ml min^{-1}. The oven temperature program was as follows: 60 ℃ for 1 minute, increasing at 5

℃ per minute to 330 ℃, final hold at 330 ℃ for 10 minutes. The secondary oven offset was +5 ℃ relative to the primary oven, and the modulator temperature offset was +15 ℃ relative to the primary oven for the entirety of the run. A 4-second modulation period with hot pulse of 1 second and cold pulse of 1 second was used. Data were collected over the mass range of 45 – 750 amu with an acquisition rate of 200 spectra/second using an electron ionization energy of 70 eV.

Samples were randomly batched into groups of 16 for analysis. Daily inlet maintenance was performed between every batch of samples and consisted of cleaning the inlet with methanol, replacing the liner, and removing 22 cm from the front of the first-dimension column. Continuing calibration verification standards (CCVs) were analyzed daily to monitor and account for any changes in analytical performance throughout the series.

3.3.5 Data Analysis

Data analysis was split into two sections. First, data were processed for the targeted analysis using ChromaTOF BT v5.40.12.0 for basic peak finding, mass spectral deconvolution, peak area integration, and quantification using a 7-point, internal standard-corrected calibration curve. Initial peak finding was completed using a signal-to-noise ratio of 20:1 with a minimum match requirement of 500 to re-combine slices. Target analytes were quantified using their extracted ion chromatograms with a mass tolerance of 500 ppm.

Second, the same data was processed to identify non-targets. Chromatographic alignment is not a standard feature of the ChromaTOF BT software but is a necessary step in non-targeted. ChromaTOF GC v4.74.2.0 contains the Statistical Compare function that can be utilized for chromatographic alignment and data reduction. Before exporting samples to ChromaTOF GC, a manual correction was implemented to correct an overall -28 second shift in the first-dimension retention time across analyses that occurred. Sample chromatograms were divided into thirds and

the first-dimension retention time was adjusted to ensure the alignment of internal standards and surrogate compounds [CB-18, CB-28, CB-52, triphenyl phosphate, tris(1,3-dichloroisopropyl)phosphate, acenaphthylene-d8, anthracene-d10, and pyrene-d10] present within each section across all samples. Samples were then exported as .netCDF files and imported into ChromaTOF GC. All samples were reprocessed using ChromaTOF GC for basic peak finding, mass spectral deconvolution, and peak integration. Initial peak finding was completed using a higher signal-to-noise ratio of 200:1 as the data was originally collected using ChromaTOF BT and has less noise than data normally collected with ChromaTOF GC. The minimum match requirement of 500 to combine slices was retained.

A second peak finding step is employed during alignment which utilized a signal-to-noise ratio of 100:1 to detect peaks not found during initial peak finding. By using a lower signal-to-noise ratio on the second pass, peaks that were below the initial signal-to-noise ratio in some samples but found in other samples can be added to the aligned peak table. Aligned analytes required a minimum match similarity of 600 and were only retained if present within at least 66% of samples within a class or within at least 96 (66%) samples total. Artifact compounds resulting from column bleed or solvents were manually removed, and analytes were searched against the NIST MS Database (2014 and 2017) for tentative identification, with a minimum similarity match of 700 for compound identification. To further reduce the data set, Fisher ratios were calculated for every compound and those with values above the F_{crit} value for $\alpha = 0.05$ were exported as .csv files for further statistical analysis using R v3.5.3 (R Foundation for Statistical Computing, Vienna, Austria) in the platform, RStudio (Boston, MA, USA). A list of all R packages used in this study is provided in the SI.

3.4 Results and Discussion

Data analyses of all 146 YOY SMB were divided into two distinct approaches. First, a targeted analysis to determine and quantify known contaminants was conducted. Second, a non-targeted analysis to identify any significant unknown contaminants or metabolites within the samples was performed on the same data set. For the non-targeted analysis, fish were assigned to one of two sample classes, "Normal" or "Lesioned," based on external fish health observations.

3.4.1 Targeted Analysis

For the targeted analysis, a target list of 127 compounds of interest (SI Table 3.3) was generated, containing compounds from the following chemical classes: PCBs, polybrominated diphenyl ethers (PBDEs), organochlorine pesticides (OCPs), polycyclic aromatic hydrocarbons (PAHs), and pharmaceutical and personal care products (PPCPs). These are compound classes that are routinely monitored in the environment and are known to pose health risks to organisms. Reportable compounds detected at 12 collection sites are summarized in Table 3.2. Results from the targeted analysis of two of the collection sites, Susquehanna Harrisburg Juniata (JF) and Susquehanna Harrisburg West Branch (WB), have previously been reported,[26] and the repeated GC×GC analysis using a DB-5ms first-dimension column did not present any significantly different findings.

Table 3.2 Summary of analytes found during targeted analysis of young-of-year smallmouth bass from 12 collection sites.

	Bald Eagle Creek Unionville (UBEC)		Conodoquinet Creek (CC)		Juniata River Howe Township Park (JR)		Kettle Creek Leidy (KC)		Loyalsock Creek (LC)		North Branch Susquehanna Danville (NBSD)	
Analyte	# Detects	Mean (ng g^{-1})	# Detects	Mean (ng g^{-1})	# Detects	Mean (ng g^{-1})	# Detects	Mean (ng g^{-1})	# Detects	Mean (ng g^{-1})	# Detects	Mean (ng g^{-1})
2-Methylnaphthalene	12	NR	10	NR	11	NR	10	NR	10	7.12	10	NR
9,10-Anthraquinone	1	0.45	0	ND	2	0.62	1	0.30	0	ND	3	0.48
Anthracene	10	0.11	7	NR	11	0.16	1	NR	5	NR	8	NR
BDE-100	0	ND	8	2.66	5	2.32	0	ND	2	1.98	8	2.15
BDE-47	2	1.34	9	5.02	8	2.95	0	ND	4	3.08	9	2.26
BDE-99	0	ND	0	ND	0	ND	0	ND	0	ND	0	ND
DEHP	6	NR	5	114.97	7	106.89	3	82.70	5	NR	5	NR
DEET	12	78.62	2	1.26	1	0.42	1	0.21	0	ND	0	ND
Indole	12	NR	10	NR	11	24.40	10	NR	10	NR	10	NR
Menthol	1	0.75	0	ND	0	ND	1	1.78	0	ND	1	1.05
Methyl triclosan	0	ND	0	ND	0	ND	0	ND	0	ND	1	0.21
Oxychlordane	0	ND	0	ND	1	0.83	0	ND	0	ND	0	ND
Tonalid	0	ND	2	0.47	1	0.34	0	ND	1	0.78	0	ND
trans-Nonachlor	0	ND	0	ND	1	2.18	0	ND	0	ND	0	ND
Tri(butoxyethyl) phosphate	0	ND	0	ND	3	9.56	0	ND	0	ND	2	2.67
Total PCBs	-	0.07	-	7.42	-	13.17	-	0.04	-	0.57	-	3.39

	North Branch Susquehanna Falls (NBSR)		Pine Creek Ramsey (PC)		Pine Creek Hamilton Bottom (PCHB)		Susquehanna Harrisburg North Branch (SRNB)		West Branch Susquehanna Lewisburg (WBSL)		Wyalusing Creek (WC)	
Analyte	# Detects	Mean (ng g^{-1})	# Detects	Mean (ng g^{-1})	# Detects	Mean (ng g^{-1})	# Detects	Mean (ng g^{-1})	# Detects	Mean (ng g^{-1})	# Detects	Mean (ng g^{-1})
2-Methylnaphthalene	11	NR	9	NR	12	NR	10	NR	10	NR	10	NR
9,10-Anthraquinone	2	0.34	0	ND	2	0.61	0	ND	1	0.42	0	ND
Anthracene	9	0.12	4	NR	5	NR	8	NR	10	0.14	10	0.12
BDE-100	6	2.32	0	ND	0	ND	8	3.21	2	1.70	0	ND
BDE-47	7	4.79	0	ND	0	ND	9	5.07	8	2.58	1	1.30
BDE-99	0	ND	0	ND	0	ND	1	1.17	0	ND	0	ND
DEHP	5	NR	4	51.53	2	51.61	4	49.30	7	71.48	3	36.61
DEET	0	ND	0	ND	1	0.18	0	ND	0	ND	10	48.98
Indole	11	NR	9	NR	12	25.45	10	22.50	10	NR	10	27.76
Menthol	1	1.38	1	2.70	1	0.93	1	1.42	0	ND	0	ND
Methyl triclosan	0	ND	0	ND	0	ND	1	0.49	0	ND	0	ND
Oxychlordane	0	ND	0	ND	0	ND	0	ND	0	ND	0	ND
Tonalid	0	ND	0	ND	0	ND	2	0.30	0	ND	0	ND
trans-Nonachlor	0	ND	0	ND	0	ND	0	ND	0	ND	0	ND
Tri(butoxyethyl) phosphate	1	1.15	0	ND	2	2.38	0	ND	1	4.95	1	14.13
Total PCBs	-	3.54	-	ND	-	5.56	-	7.38	-	0.82	-	1.09

BDE = brominated diphenyl ether, PCB = polychlorinated biphenyl, ND = not detected, NR = not reported (calculated value was below method quantification limit)

Target compounds within samples were identified by retention time and mass spectral matching with analytical standards. A 7-pt calibration curve with internal standard normalization was used to quantify each detected compound. Values reported in Table 2 were greater than 10 x any matching signal detected within the matrix blanks. Compounds detected within samples and not within the matrix blanks, but at levels lower than the method quantification limit or those not exceeding 10 x the signal in the matrix blanks are not reported (NR). The total PCB level for a location represents the average sum of all PCBs detected within samples at that collection site. PCB congeners without an exact analytical standard match were semi-quantified using the congener with the closest retention time and an equivalent number of chlorine substitutions.

Of the 127 target compounds, 39 were PCB congeners and 88 were a combination of BDEs, OCPs, PAHs, and PPCPs. The persistence and abundance of PCBs in aquatic organisms, as well as their known ability to interrupt normal endocrine and immune function,[30] have made them a routine part of environmental analyses. PCBs are commonly reported as sum totals instead of individual congener concentrations. As such, this study reports only the average total per collection site and did not confirm the identify of each congener. PCB average totals across sites ranged from 0.00 to 13.17 ng g^{-1}. PCB advisory levels are usually given in terms for human consumption. The Pennsylvania Department of Environmental Protection advises that fish containing < 50 ng g^{-1} are safe for unrestricted consumption, while those containing > 1.9 ppm (1900 ng g^{-1}) should not be consumed.[31] In terms of organism health, few studies have examined the effects of environmentally relevant concentrations of multiple congeners in juvenile fish. One study of zebrafish exposed to environmentally realistic mixtures of PCBs and PBDEs found that exposed fish had delays in reaching reproductive maturity while also growing to larger sizes. Additionally, the progeny of these exposed fish had decreased likelihoods of survival under starvation conditions.[32] PCBs bioaccumulate in fatty tissues and biomagnify through the trophic

chain, so though the YOY SMB examined currently have “safe” levels, this value will likely be higher in adult SMB from the same areas.

Of the remaining 88 target compounds, 15 were detected in at least one sample at a reportable concentration. Two compounds, 2-methylnaphthalene and indole, were detected within every sample, but mostly at levels equal to or below those detected in the matrix blanks (recorded as NR in Table 2). Only one site, Loyalsock Creek (LC), had a reportable concentration of 2-methylnapthalene (7.12 ng g^{-1}), while four sites had reportable concentrations of indole ($mean_{reportable}$ = 25.03 ng g^{-1}, n = 4). The detection of indole in every sample was not surprising, but the low concentrations detected were unexpected given the higher concentrations previously reported at the JF and WB locations. Previous analysis found indole levels in excess of 380 and 500 ng g^{-1} at these two sites, respectively.[26] Both indole and 2-methylnaphthalene are widely used compounds in industrial settings. Indole is found in fragrances, pesticides, food additives, and is a by-product of bacterial protein decomposition.[33,34] 2-Methylnaphthalene naturally occurs within crude oils and coal and is produced by pyrolysis and combustion reactions as well as being used in the manufacturing of vitamin K and insecticides.[35] Due to the ubiquity of indole naturally occurring in the environment, little research has been conducted into the effects of pure indole on organism health. However, indole forms the base of many compounds with biological activity, and indole derivatives have been explored as antimicrobials, antivirals, and anticancer agents.[36–38] Due to its occurrence in crude oil, more information is available for 2-methylnaphthalene, but studies on the singular effects of this compound are still limited. One study, comparing the toxicity of 2-methylnaphthalene in arctic and temperate marine organisms, reported a 96-hr LC_{50} of 3.06 mg l^{-1} for European seabass,[39] showing that it can produce toxic effects in some species of fish.

The two compounds detected at the highest concentrations were N,N-diethyl-*m*-toluamide (DEET) and bis(2-ethylhexyl) phthalate (DEHP). DEET is an active ingredient of

insect repellants used to protect against biting insects and the diseases carried by them. It is the most commonly used insect repellant worldwide and was registered for use by the general public in 1957 by the U.S. Environmental Protection Agency.[40] As DEET is considered to be neither persistent nor bioaccumulative, with an estimated half-life of 5 – 15 days in surface water,[41] it is possible that sampling occurred recently after a contamination event at two of the collection sites where there were reportable levels of DEET in all fish sampled. Lower levels were reported in 1 – 2 YOY SMB each at four other sampling sites. The variation in DEET levels across sampling sites highlights one of the hardest challenges with environmental analyses, relating the detected compounds to potential exposure sources. Some level of exposure occurred, as DEET was detected within the fish tissues, but whether the exposure was due to a major contamination, minor contamination, or is caused by continuous influx through wastewater treatment plant effluent or other waste runoff sources is unclear. Continuous water monitoring would help to identify the source but would not be able to show uptake of the chemical in exposed fish.

DEHP is a plasticizer commonly used in the manufacturing of everyday plastic products, such as food packaging, toys, and medical equipment. In this study, DEHP was detected at the highest concentration of any target compound, with average concentrations ranging from 36.61 – 114.97 ng g^{-1} across seven sampling sites. Nine individual fish had DEHP concentrations above 100 ng g^{-1}, with 400.4 ng g^{-1} as the highest concentration. Long term exposure to DEHP during embryonic and juvenile developmental stages at environmentally relevant concentrations has been linked with endocrine disruption and changes in growth and development in some species of fish.[29,42,43] One study in Atlantic salmon (*Salmo salar*) found that when fed a diet contaminated with DEHP for four weeks immediately following yolk sac resorption followed by a clean diet for four months there was a small increase in intersex fish. However, a significant difference between control and exposed fish was only observed at the highest concentration tested, 1500 mg kg^{-1}

DEHP in food resulting in an average of 2.55 mg kg^{-1} wet weight DEHP (2550 ng g^{-1}) detected in fish.[44]

The detection of DEHP provides another example of the difficulty in discerning the source of contamination and how that changes the potential biological consequences. Exposure to YOY SMB throughout their developmental life stages could be associated with the observed reproductive endocrine disruption (intersex) in adult male fish, but further investigation into the effects of DEHP specifically on SMB is needed to draw any conclusions. Many of the adverse effects documented in the literature from DEHP may not be manifested as observable external characteristics, such as lesions; therefore, the role of DEHP in SMB disease cannot be confirmed at this time.

It should be noted that the QuEChERS extraction employed in this study did use multiple plastic products capable of leaching plasticizers into the sample extracts as lab contaminants. This is not believed to have happened with DEHP. If it were introduced by part of the sample preparation process, it would be expected to be found in all or most samples at some minimum concentration value. There was no sampling location where DEHP was found in all samples, and it was only detected in a single matrix blank sample at 3.28 ng g^{-1}. Conversely, it is believed that at least some form of contamination occurred with diethyl phthalate, another plasticizer commonly used in everyday plastic products. Diethyl phthalate was found in 98.5% of samples at an average concentration 1.4 ng g^{-1} for the matrix blanks and 2.9 ng g^{-1} for the samples, with a total range of 0.1 – 27.6 ng g^{-1}. This indicates that although some amount was likely introduced during the sample preparation, there could still be small exposures to fish at collection sites through other means.

3.4.2 Non-targeted Analysis

Comprehensive GC × GC analysis is a data rich analytical technique. It is not uncommon for an individual sample, especially one with a complex matrix like whole fish homogenate, to have 10,000 – 20,000 peaks per analysis. However, it is unlikely that all these peaks will correspond to compounds that warrant further investigation. Therefore, it is necessary to employ data reduction strategies to identify any statistically significant compounds of interest within the samples. For this study, the Statistical Compare function of the ChromaTOF GC software was the first step in data reduction. Statistical Compare uses pairwise comparisons between every chromatogram for accurate alignment and generates a peak list of aligned analytes that meet the user specified criteria. Samples are first divided into user defined groupings, or sample classes, to facilitate the identification of notable differences between different sample types. Here, samples were assigned to either the Normal or Lesioned sample class, and analytes had to be present in at least 66% of samples within a class or within 96 samples (66%) of any class. As the goal is to identify compounds that are associated with the sample class, it is unlikely that a compound found in only a few or one sample would be a main contributor to the differences between classes. The sample classes were determined based on gross observation of the health of the fish at the time of collection. Normal fish were observed to have no obvious signs of external infection or deformity; however microscopic changes could have been present. Lesioned fish exhibited some form of external health anomaly. A definitive diagnosis for these samples was not possible as the entire sample was needed for chemical analysis, but visual observations were used to characterize the types of lesions based on previous histological analyses of similar patterns (SI Table 3.4). Raised lesions are often associated with the parasite *Myxobolus inornatus*,[13,45] while pale and eroded areas are typically found to be caused by bacterial or viral pathogens.[14,15]

Accurate alignment of all chromatograms is essential to identifying truly unique compounds versus those present within multiple samples. This data set of 146 YOY SMB extracts presented two unique challenges during chromatographic alignment. First, during data acquisition there was an overall -28 second drift in the first-dimension retention time (t_{R1}) occurring in 4 second increments between sample batches. Second, the pair wise comparison process of such data dense chromatograms as those produced by this complex matrix required an enormous amount of computing power to complete data processing. The first challenge was overcome using daily CCVs and the inclusion of both internal standards and surrogate compounds. A daily CCV was run prior to beginning each sample batch to monitor for any changes in data acquisition following daily inlet maintenance. On day five, a -8 second shift in the t_{R1} across the chromatogram was discovered. A total of five more -4 second shifts occurred during the remaining analysis. This likely occurred because of the shortening of the first-dimension column as a result of trimming for inlet maintenance. While the flow calculator should, in theory, allow for linear velocity adjustment to maintain retention times, the data in this study reflected a likely limitation of the pressure control for the needed level of accuracy to completely correct for the change in column lengths. After the first shift was detected, a second CCV was added at the end of each batch of samples to monitor for drifting throughout the sample batch; none was observed. While the t_{R1} drift was easy to overcome during the targeted analysis, the Statistical Compare function used for non-targeted analysis could not account for such a large difference in retention times during alignment. Therefore, each chromatogram's t_{R1} was manually corrected by aligning the internal standards and surrogate compounds spiked within. The second challenge was solved concurrently with this correction, as the chromatogram was sectioned into three parts, each containing surrogate and internal standard compounds. These smaller chromatogram sections did not overwhelm the computer during the pair wise comparisons and alignment of all 146 chromatograms was successful for each section. Correcting the t_{R1} also

allowed for narrow retention shift windows to be employed in the final data processing method. Large data sets are prone to false positives due to misalignment. Narrowing the retention shift window helps to reduce false positive matches during alignment. Additionally, this analysis utilized a second peak finding step since ChromaTOF GC required higher initial signal-to-noise ratios for processing ChromaTOF BT data. While this step decreases the likelihood of missing peaks during alignment, it presents another opportunity to introduce false positives during alignment. The second peak finding step used a lower signal-to-noise ratio but required a higher match threshold to retain the compound and to counteract potential false positives. For each aligned peak list generated, samples were normalized to either an internal standard or surrogate compound. Background artifact peaks, such as solvent and column bleed, were manually removed, and Fisher ratios were calculated for all remaining peaks. The aligned peak lists and normalized areas were exported to Excel and recombined for further statistical analysis. The authors would like to emphasize the importance of having appropriate controls (i.e., internal standards, surrogates, and CCVs) as it is unlikely this non-targeted analysis would have been possible without them.

After alignment and removal of artifact compounds, a total of 512 features were retained in the recombined peak list. An unsupervised principal component analysis (PCA) was applied to this data set to look for signs of batch effect or similar interference with the samples. There was no apparent clustering (SI Figure 3.4). This combined with the results of the targeted analysis indicates there were no exogenous factors such as sample storage or preparation contributing to differences between the Normal and Lesioned classes. The targeted analysis did reveal varying amounts of target compounds among the different collection sites. However, the focus of this study is to differentiate between Normal and Lesioned class fish across all sites. Therefore, the non-targeted analysis did not focus on site-specific non-target compounds. To further reduce the data set and identify features that are significant in differentiating between the Normal and

Lesioned classes, only compounds whose Fisher ratio value were above the critical value for α = 0.05 (F_{crit} = 3.89) were retained for further analysis. The Fisher ratio is a measure of variance, specifically the between class variation compared to the within class variation. A higher Fisher ratio value indicates a stronger contribution to the differences between classes as opposed to within a class. Only 95 features had calculated Fisher ratio values above the F_{crit} value. A PCA of these remaining features showed clustering of Lesioned YOY SMB with slight overlap of Normal YOY SMB (Figure 3.2). A final data reduction step was employed based on the variable contribution to each principal component. Features were plotted based on their contribution to each principal component and those that contributed more than the expected value for uniform contribution by all variables were retained in the final data set. This reduced the final data set to 65 features. An additional PCA of these, as expected, showed further separation between the Lesioned and Normal classes (SI Figure 3.5).

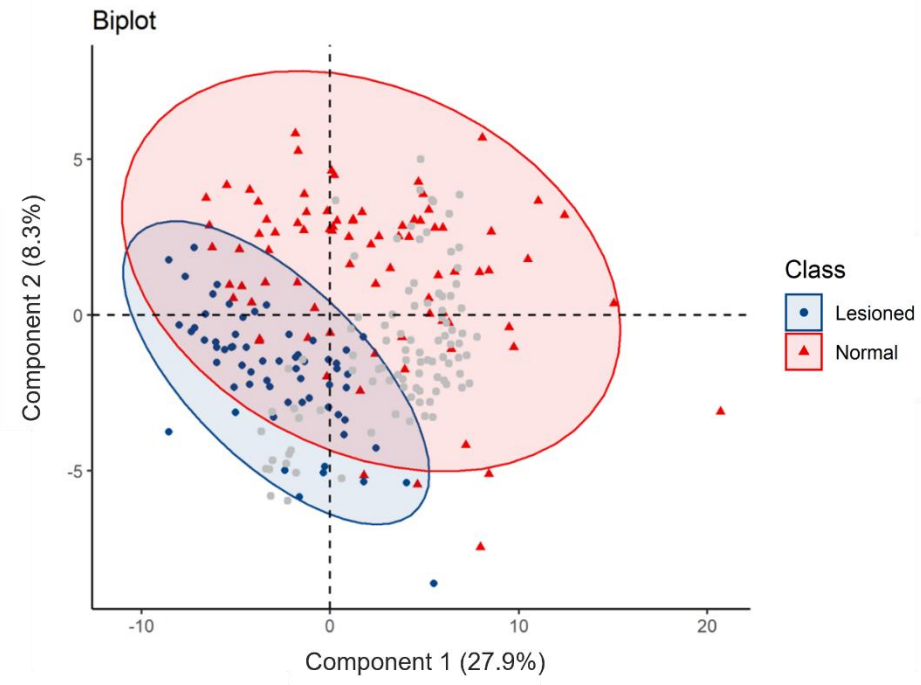

Figure 3.2 Principal component analysis biplot of reduced data set for 146 YOY SMB. PCA was performed on normalized areas of 95 analytes whose Fisher ratios were above the Fcrit value for α = 0.05. Ellipses are 95%

confidence intervals around group centered mean. Grey circles represent variable loadings, the analytes retained after Fisher ratio analysis. Red triangles (Normal class) and blue circles (Lesioned class) represent individual samples.

The 65 features identified as the most significant were tentatively identified (SI Table 3.5) by mass spectral comparison with the NIST MS Database (2014 and 2017). One analyte, 2-oxo-1-methyl-3-isopropylpyrazine, was found to occur only in Normal class samples. However, no analytes were found to be specific to Lesioned class samples. This was unexpected as a previous non-targeted analysis of 21 YOY SMB (from JF and WB) using the same data processing methodology found eight unique features that occurred only within Lesioned class samples.[26] It was expected that different features would be identified as significant when the sample set was increased. By using a data reduction methodology focused on identifying significantly different features between sample classes it was hoped that class specific compounds would be revealed for each sample class. To further investigate the differences between the two sample classes, tentatively identified analytes were grouped into chemical classes based on their mass spectra. The chemical groups included alcohols, aldehydes, carbonates, esters, hydrocarbons, ketones, nitrogen containing, sulfur containing, or miscellaneous other. The percent composition of both Normal and Lesioned fish was calculated based on the normalized areas used during the statistical analysis and compared (Figure 3.3). As further confirmation of the identity of analytes is not possible for the current study, comparing chemical groups can provide more information into the differentiating factors between sample classes. Three chemical groups, esters, ketones, and nitrogen containing compounds, were more abundant in Lesioned fish. All three of these groups have been associated with oxidative stress in organisms. Aromatic nitrogen compounds have the capacity to cause oxidative stress, while ketones and esters, specifically fatty acid methyl esters (FAMEs), have been identified as possible biomarkers of oxidative stress and diseases in a variety of species.[46,47] Additional examination of these

tentatively identified compounds is needed to recognize their full potential as disease biomarkers, including confirmation of their identity by external analytical standards.

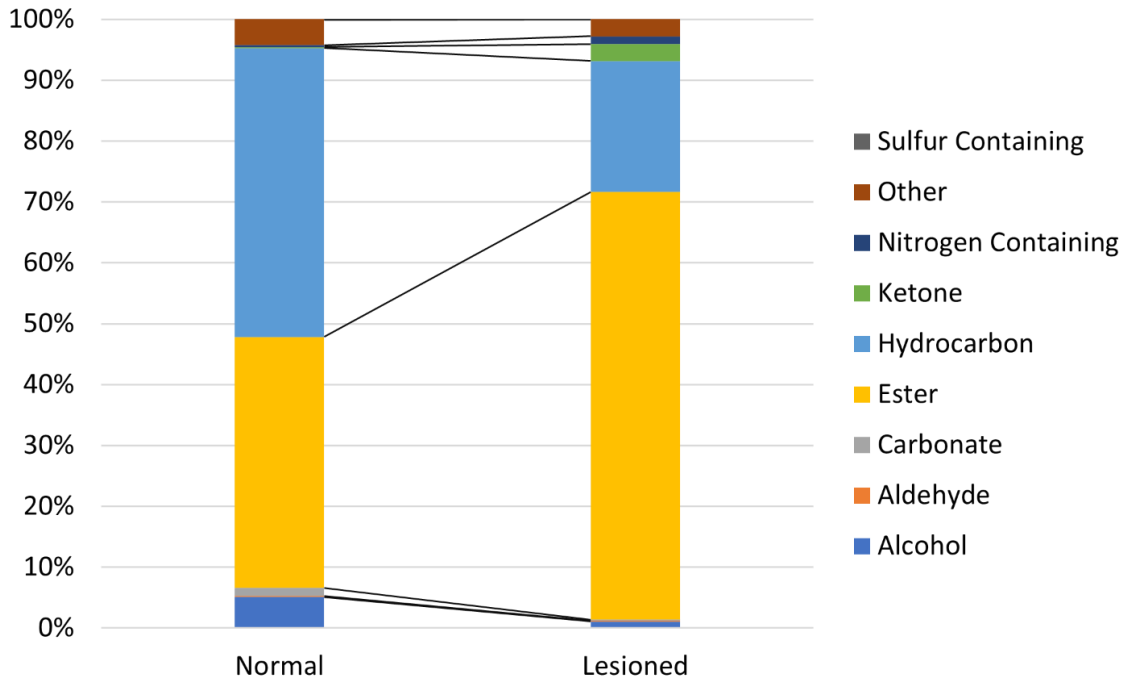

Figure 3.3 Comparison of percent composition of final 65 analytes identified in non-targeted analysis by chemical class.

3.5 Conclusions

Targeted analysis of the 146 YOY SMB revealed the presence or absence of 127 selected target analytes within samples. These target compounds are known contaminants in aquatic environments and are routinely investigated. This analysis detected two target compounds, indole and 2-methylnaphthalene, in every sample at low to trace level concentrations. DEET and DEHP were the two most abundant target compounds identified, with DEHP levels being the highest for any single target compound detected in this study. Although no clear correlations were revealed between Lesioned YOY SMB and the reported targets, compounds detected in the YOY may

impact their future development and health but be undetectable during adulthood. Determining past exposure to non-persistent compounds remains a challenge to environmental studies worldwide. The non-targeted analysis was used to help overcome this by examining all detected analytes within samples and determining the most significant features present. It is not currently feasible to fully identify the thousands of analytes detected through GC × GC-TOFMS analysis, nor would it benefit the researcher to spend time identifying insignificant background compounds. This problem becomes compounded when combined with large sample sets of complex matrices. Additionally, the complex mixtures of contaminants at the various sites complicates our understanding of what exposures the fish encounter. Therefore, using data reduction techniques that focus on identifying differences between sample classes provides the best insight into which features are significant, however it may need to be conducted with larger sample sizes at individual sites or with a focus on discovering non-target compounds specific to individual sites. We now have a more thorough understanding of contaminant compounds found in YOY SMB tissues in the Susquehanna River Basin, many of which would not have been identified using traditional laboratory chemical schedules. However, future studies are needed to investigate compounds detected in relation to fish disease characteristics observed and fish biology. For example, field studies quantifying YOY SMB movement may help identify possible sources of exposure relative to collection sites. Multiple routes of exposures should also be considered including from water, soil, or maternal egg source. Laboratory studies will be important to investigate biological and disease response to compounds and compound mixtures identified which currently cannot be done by through the detection of compounds in tissues alone. Additionally, consideration of spatial attributes and site-specific variables, including land use in combination with fish health and chemical analytes present, may be an important future research need to investigate the site-specific differences observed and quantify underlying relationships. The large sample set analyzed in this study did reveal three chemical classes of interest that

warrant future investigation into their part in SMB disease and/or the response of fish to exposures of chemicals, pathogens, or a combination of the two. The 65 tentatively identified analytes provide a starting point for that investigation.

3.6 Acknowledgements

The authors would like to thank Kevin Siek of LECO Corporation for help with data analysis. We also thank Geoff Smith and others in Pennsylvania Fish & Boat Commission and Tim Wertz and others in Pennsylvania Department of Environmental Protection for fish collections. Funding for this work was provided by Pennsylvania Sea Grant and by the U.S. Geological Survey's Environmental Health and Species Management Research (Fisheries) programs within the Ecosystems Mission Area. Any mention of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. Industrial support for this project was provided by LECO Corporation and Restek Corporation.

3.7 References

1. D.P. Häder, A.T. Banaszak, V.E. Villafañe, M.A. Narvarte, R.A. González, E.W. Helbling, Anthropogenic pollution of aquatic ecosystems: Emerging problems with global implications, Sci. Total Environ. 713 (2020) 136586. doi:10.1016/j.scitotenv.2020.136586.
2. I.Y. López-Pacheco, A. Silva-Núñez, C. Salinas-Salazar, A. Arévalo-Gallegos, L.A. Lizarazo-Holguin, D. Barceló, H.M.N. Iqbal, R. Parra-Saldívar, Anthropogenic contaminants of high concern: Existence in water resources and their adverse effects, Sci. Total Environ. 690 (2019) 1068–1088. doi:10.1016/j.scitotenv.2019.07.052.
3. M. Van Den Berg, L. Birnbaum, A.T.C. Bosveld, B. Brunström, P. Cook, M. Feeley, J.P. Giesy, A. Hanberg, R. Hasegawa, S.W. Kennedy, T. Kubiak, J.C. Larsen, F.X.R. Van Leeuwen, A.K.D. Liem, C. Nolt, R.E. Peterson, L. Poellinger, S. Safe, D. Schrenk, D. Tillitt, M. Tysklind, M. Younes, F. Wærn, T. Zacharewski, Toxic equivalency factors (TEFs) for

PCBs, PCDDs, PCDFs for humans and wildlife, Environ. Health Perspect. 106 (1998) 775–792. doi:10.1289/ehp.98106775.

4. P. Vasseur, C. Cossu-Leguille, Linking molecular interactions to consequent effects of persistent organic pollutants (POPs) upon populations, Chemosphere. 62 (2006) 1033–1042. doi:10.1016/j.chemosphere.2005.05.043.
5. S. Sauvé, M. Desrosiers, A review of what is an emerging contaminant, Chem. Cent. J. 8 (2014) 15. doi:10.1186/1752-153X-8-15.
6. Y. Aminot, S.J. Sayfritz, K. V. Thomas, L. Godinho, E. Botteon, F. Ferrari, V. Boti, T. Albanis, M. Köck-Schulmeyer, S. Diaz-Cruz, M. Farré, D. Barceló, A. Marques, J.W. Readman, Environmental risks associated with contaminants of legacy and emerging concern at European aquaculture areas, Environ. Pollut. 252 (2019) 1301–1310. doi:10.1016/j.envpol.2019.05.133.
7. M. Salimi, A. Esrafili, M. Gholami, A. Jonidi Jafari, R. Rezaei Kalantary, M. Farzadkia, M. Kermani, H.R. Sobhi, Contaminants of emerging concern: a review of new approach in AOP technologies, Environ. Monit. Assess. 189 (2017) 414. doi:10.1007/s10661-017-6097-x.
8. V.S. Blazer, L.R. Iwanowicz, D.D. Iwanowicz, D.R. Smith, J.A. Young, J.D. Hedrick, S.W. Foster, S.J. Reeser, Intersex (Testicular oocytes) in smallmouth bass from the potomac river and selected nearby drainages, J. Aquat. Anim. Health. 19 (2007) 242–253. doi:10.1577/H07-031.1.
9. J.J. Chaplin, J. Kent Crawford, R.A. Brightbill, Water-Quality Monitoring in Response to Young-of-the- Year Smallmouth Bass (Micropterus dolomieu) Mortality in the Susquehanna River and Major Tributaries, Pennsylvania: 2008, 2009. https://pubs.usgs.gov/of/2009/1216/ofr2009-1216.pdf.
10. L.R. Iwanowicz, V.S. Blazer, C.P. Guy, A.E. Pinkney, J.E. Mullican, D.A. Alvarez, Reproductive health of bass in the Potomac, USA, drainage: Part 1. Exploring the effects of proximity to wastewater treatment plant discharge, Environ. Toxicol. Chem. 28 (2009) 1072–1083. http://dx.doi.org/10.1897%2F08-433.1.
11. V.S. Blazer, L.R. Iwanowicz, H. Henderson, P.M. Mazik, J.A. Jenkins, D.A. Alvarez, J.A. Young, Reproductive endocrine disruption in smallmouth bass (Micropterus dolomieu) in the Potomac River basin: Spatial and temporal comparisons of biological effects, Environ. Monit. Assess. 184 (2012) 4309–4334. doi:10.1007/s10661-011-2266-5.
12. V.S. Blazer, D.D. Iwanowicz, H.L. Walsh, A.J. Sperry, L.R. Iwanowicz, D.A. Alvarez, R.A. Brightbill, G. Smith, W.T. Foreman, R. Manning, Reproductive health indicators of fishes

from Pennsylvania watersheds: association with chemicals of emerging concern, Environ. Monit. Assess. 186 (2014) 6471–6491. doi:10.1007/s10661-014-3868-5.

13. M.K. Schall, V.S. Blazer, H.L. Walsh, G.D. Smith, T. Wertz, T. Wagner, Spatial and temporal variability of myxozoan parasite, Myxobolus inornatus, prevalence in young of the year smallmouth bass in the Susquehanna River Basin, Pennsylvania, J. Fish Dis. 41 (2018) 1689–1700. doi:10.1111/jfd.12878.
14. H.L. Walsh, V.S. Blazer, G.D. Smith, M. Lookenbill, D.A. Alvarez, K.L. Smalling, Risk Factors Associated with Mortality of Age-0 Smallmouth Bass in the Susquehanna River Basin, Pennsylvania, J. Aquat. Anim. Health. 30 (2018) 65–80. doi:10.1002/aah.10009.
15. C. Starliper, V.S. Blazer, L.R. Iwanowicz, H.L. Walsh, Microbial isolates in diseased fishes, primarily smallmouth bass (Micropterus dolomieu), within the Chesapeake Bay drainage in 2009-2011, in: Proc. West VIrginia Acad. Sci. Submitt. Pap. Eight-Eighth Annu. Sess. 2013, 2013.
16. G.D. Smith, V.S. Blazer, H.L. Walsh, L.R. Iwanowicz, C. Starliper, A.J. Sperry, The effects of disease-related mortality of young-of-year smallmouth bass on the population characteristics in the Susquehanna River basin, Pennsylvania and potential implications to conservation of black bass diversity., in: Am. Fish. Soc. Symp., 2015: pp. 82:319-332.
17. D.W. Kolpin, V.S. Blazer, J.L. Gray, M.J. Focazio, J.A. Young, D.A. Alvarez, L.R. Iwanowicz, W.T. Foreman, E.T. Furlong, G.K. Speiran, S.D. Zaugg, L.E. Hubbard, M.T. Meyer, M.W. Sandstrom, L.B. Barber, Chemical contaminants in water and sediment near fish nesting sites in the Potomac River basin: Determining potential exposures to smallmouth bass (Micropterus dolomieu), Sci. Total Environ. 443 (2013) 700–716. doi:10.1016/j.scitotenv.2012.09.063.
18. C.M. McClure, K.L. Smalling, V.S. Blazer, A.J. Sperry, M.K. Schall, D.W. Kolpin, P.J. Phillips, M.L. Hladik, T. Wagner, Spatiotemporal variation in occurrence and co-occurrence of pesticides, hormones, and other organic contaminants in rivers in the Chesapeake Bay Watershed, United States, Sci. Total Environ. 728 (2020) 138765. doi:10.1016/j.scitotenv.2020.138765.
19. D.A. Alvarez, W.L. Cranor, S.D. Perkins, V.L. Schroeder, L.R. Iwanowicz, R.C. Clark, C.P. Guy, A.E. Pinkney, V.S. Blazer, J.E. Mullican, Reproductive health of bass in the potomac, USA, drainage: part 2. Seasonal occurrence of persistent and emerging organic contaminants, Environ. Toxicol. Chem. 28 (2009) 1084–1095. http://www.scopus.com/inward/record.url?eid=2-s2.0-67249093449&partnerID=40.

20. S. Furihata, A. Kasai, K. Hidaka, M. Ikegami, H. Ohnishi, K. Goka, Ecological risks of insecticide contamination in water and sediment around off-farm irrigated rice paddy fields, Environ. Pollut. 251 (2019) 628–638. doi:10.1016/j.envpol.2019.05.009.
21. P.H. Stefanuto, K.A. Perrault, R.M. Lloyd, B. Stuart, T. Rai, S.L. Forbes, J.F. Focant, Exploring new dimensions in cadaveric decomposition odour analysis, Anal. Methods. 7 (2015) 2287–2294. doi:10.1039/c5ay00371g.
22. S. Fernando, A. Renaguli, M.S. Milligan, J.J. Pagano, P.K. Hopke, T.M. Holsen, B.S. Crimmins, Comprehensive Analysis of the Great Lakes Top Predator Fish for Novel Halogenated Organic Contaminants by GC×GC-HR-ToF Mass Spectrometry, Environ. Sci. Technol. 52 (2018) 2909–2917. doi:10.1021/acs.est.7b05999.
23. M. Anastassiades, S.J. Lehotay, D. Stajbaher, F.J. Schenck, Fast and Easy Multiresidue Method Employing Acetonitrile, J. AOAC Int. 86 (2003) 412–431. doi:10.2478/s11687-011-0011-9.
24. T. Rejczak, T. Tuzimski, A review of recent developments and trends in the QuEChERS sample preparation approach, Open Chem. 13 (2015) 980–1010. doi:10.1515/chem-2015-0109.
25. Á. Santana-Mayor, B. Socas-Rodríguez, A. V. Herrera-Herrera, M.Á. Rodríguez-Delgado, Current trends in QuEChERS method. A versatile procedure for food, environmental and biological analysis, TrAC - Trends Anal. Chem. 116 (2019) 214–235. doi:10.1016/j.trac.2019.04.018.
26. P. Teehan, M.K. Schall, V.S. Blazer, B. Gruber, F.L. Dorman, Modified QuEChERS extraction for the analysis of young-of-year smallmouth bass using GC × GC-TOFMS, Anal. Methods. 12 (2020) 3697–3704. doi:10.1039/D0AY00880J.
27. Y. Lor, A. Revak, J. Weigand, E. Hicks, D.R. Howard, T.C. King-Heiden, Juvenile exposure to vinclozolin shifts sex ratios and impairs reproductive capacity of zebrafish, Reprod. Toxicol. 58 (2015) 111–118. doi:10.1016/j.reprotox.2015.09.003.
28. M. Bayley, M. Junge, E. Baatrup, Exposure of juvenile guppies to three antiandrogens causes demasculinization and a reduced sperm count in adult males, Aquat. Toxicol. 56 (2002) 227–239. doi:10.1016/S0166-445X(01)00210-7.
29. V.R.T. Zanotelli, S.C.F. Neuhauss, M.U. Ehrengruber, Long-term exposure to bis(2-ethylhexyl)phthalate (DEHP) inhibits growth of guppy fish (Poecilia reticulata), J. Appl. Toxicol. 30 (2010) 29–33. doi:10.1002/jat.1468.
30. U.S.D. of H. and H. Services, Toxicological Profile for Polychlorinated Biphenyls (PCBs),

2002. doi:10.1201/9781420061888_ch129.

31. Pennsylvania Department of Environmental Protection, Office of water programs bureau of clean water assessment methodology for rivers and streams 2018, 2018.
32. K. Horri, S. Alfonso, X. Cousin, C. Munschy, V. Loizeau, S. Aroua, M.L. Bégout, B. Ernande, Fish life-history traits are affected after chronic dietary exposure to an environmentally realistic marine mixture of PCBs and PBDEs, Sci. Total Environ. 610–611 (2018) 531–545. doi:10.1016/j.scitotenv.2017.08.083.
33. J. Kim, W. Park, Indole: a signaling molecule or a mere metabolic byproduct that alters bacterial physiology at a high concentration?, J. Microbiol. 53 (2015) 421–428. doi:10.1007/s12275-015-5273-3.
34. U. EPA Office of Pesticide Programs, BIOPESTICIDES REGISTRATION ACTION DOCUMENT Indole PC Code 025000, 2009.
35. U. EPA ORD NCEA Integrated Risk Information System, Draft toxicological review of 2-Methylnaphthalene: In support of summary information on the integrated risk information system (IRIS), Fed. Regist. 76 (2011) 57033–57034. www.epa.gov/iris (accessed April 5, 2021).
36. Kumari, R.K. Singh, Medicinal chemistry of indole derivatives: Current to future therapeutic prospectives, Bioorg. Chem. 89 (2019) 103021. doi:10.1016/j.bioorg.2019.103021.
37. Srivastava, S.N. Pandeya, "INDOLE" A VERSATILE NUCLEUS IN PHARMACEUTICAL FIELD, 2010. www.ijcpr.com (accessed May 16, 2021).
38. M.-Z. Zhang, Q. Chen, G.-F. Yang, A review on recent developments of indole-containing antiviral agents, Eur. J. Med. Chem. 89 (2015) 421–441. doi:10.1016/j.ejmech.2014.10.065.
39. G.H. Olsen, M.G.D. Smit, J. Carroll, I. Jæger, T. Smith, L. Camus, Arctic versus temperate comparison of risk assessment metrics for 2-methyl-naphthalene, Mar. Environ. Res. 72 (2011) 179–187. doi:10.1016/j.marenvres.2011.08.003.
40. S.D. Costanzo, A.J. Watkinson, E.J. Murby, D.W. Kolpin, M.W. Sandstrom, Is there a risk associated with the insect repellent DEET (N,N-diethyl-m-toluamide) commonly found in aquatic environments?, Sci. Total Environ. 384 (2007) 214–220. doi:10.1016/j.scitotenv.2007.05.036.
41. J. Weeks, P. Guiney, A. Nikiforov, Assessment of the environmental fate and ecotoxicity of N,N-diethyl-m-toluamide (DEET), Integr. Environ. Assess. Manag. 8 (2012) 120–134. doi:10.1002/ieam.1246.
42. W.-K. Yang, L.-F. Chiang, S.-W. Tan, P.-J. Chen, Environmentally relevant concentrations

of di(2-ethylhexyl)phthalate exposure alter larval growth and locomotion in medaka fish via multiple pathways, Sci. Total Environ. 640–641 (2018) 512–522. doi:10.1016/j.scitotenv.2018.05.312.

43. T. Ye, M. Kang, Q. Huang, C. Fang, Y. Chen, H. Shen, S. Dong, Exposure to DEHP and MEHP from hatching to adulthood causes reproductive dysfunction and endocrine disruption in marine medaka (Oryzias melastigma), Aquat. Toxicol. 146 (2014) 115–126. doi:10.1016/j.aquatox.2013.10.025.
44. Norman, H. Börjeson, F. David, B. Tienpont, L. Norrgren, Studies of Uptake, Elimination, and Late Effects in Atlantic Salmon (Salmo salar) Dietary Exposed to Di-2-Ethylhexyl Phthalate (DEHP) During Early Life, Arch. Environ. Contam. Toxicol. 52 (2007) 235–242. doi:10.1007/s00244-005-5089-y.
45. H.L. Walsh, V.S. Blazer, L.R. Iwanowicz, G. Smith, A Redescription of Myxobolus inornatus from Young-of-the-Year Smallmouth Bass (Micropterus dolomieu), J. Parasitol. 98 (2013) 1236–1242. doi:10.1645/ge-3081.1.
46. Valavanidis, T. Vlahogianni, M. Dassenakis, M. Scoullos, Molecular biomarkers of oxidative stress in aquatic organisms in relation to toxic environmental pollutants, Ecotoxicol. Environ. Saf. 64 (2006) 178–189. doi:10.1016/j.ecoenv.2005.03.013.
47. R. Pesesse, P.-H. Stefanuto, F. Schleich, R. Louis, J.-F. Focant, Multimodal chemometric approach for the analysis of human exhaled breath in lung cancer patients by TD-GC × GC-OFMS, J. Chromatogr. B. 1114–1115 (2019) 146–153. doi:10.1016/j.jchromb.2019.01.029.

3.8 Supporting Information

3.8.1 R Packages Used for Analysis and Figure Creation

- Figure 3.1
  - Package: “ggplot2”
    - H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.
  - Package: “sf”
    - Pebesma, E., 2018. Simple Features for R: Standardized Support for Spatial Vector Data. The R Journal 10 (1), 439-446, https://doi.org/10.32614/RJ-2018-009
  - Package: “usmap”
    - Paolo Di Lorenzo (2021). usmap: US Maps Including Alaska and Hawaii. R package version 0.5.2. https://CRAN.R-project.org/package=usmap
  - Package: “maps”
    - Original S code by Richard A. Becker, Allan R. Wilks. R version by Ray Brownrigg. Enhancements by Thomas P Minka and Alex Deckmyn. (2018). maps: Draw Geographical Maps. R package version 3.3.0. https://CRAN.R-project.org/package=maps
  - Package: “mapdata”
    - Original S code by Richard A. Becker and Allan R. Wilks. R version by Ray Brownrigg. (2018). mapdata: Extra Map Databases. R package version 2.3.0. https://CRAN.R-project.org/package=mapdata

- Figure 3.2 and SI Figures 3.4-3.5
  - Package: “ade4”
    - Dray S, Dufour A (2007). “The ade4 Package: Implementing the Duality Diagram for Ecologists.” _Journal of Statistical Software_, *22*(4), 1-20. doi: 10.18637/jss.v022.i04 (URL: http://doi.org/10.18637/jss.v022.i04).
    - Bougeard S, Dray S (2018). “Supervised Multiblock Analysis in R with the ade4 Package.” _Journal of Statistical Software_, *86*(1), 1-17. doi: 10.18637/jss.v086.i01 (URL: http://doi.org/10.18637/jss.v086.i01).
    - Chessel D, Dufour A, Thioulouse J (2004). “The ade4 Package - I: One-Table Methods.” _R News_, *4*(1), 5-10. <URL:https://cran.r-project.org/doc/Rnews/>.

- Dray S, Dufour A, Chessel D (2007). “The ade4 Package - II: Two-Table and K-Table Methods.” _R News_, *7*(2), 47-52. <URL: https://cran.r-project.org/doc/Rnews/>.
- Thioulouse J, Dray S, Dufour A, Siberchicot A, Jombart T, Pavoine S (2018). _Multivariate Analysis of Ecological Data with ade4_. Springer. doi: 10.1007/978-1-4939-8850-1 (URL: http://doi.org/10.1007/978-1-4939-8850-1).

- Package: “factoextra”
  - Alboukadel Kassambara and Fabian Mundt (2020). factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.7. https://CRAN.R-project.org/package=factoextra

3.8.2 Tables

Table 3.3 Standards used for targeted analysis of young-of-year smallmouth bass tissue.

Compound	Cas #	Manufacturer
1-methylnapthalene	90-12-0	Chem Service
2,2',3,3',4,4',5,5',6-Nonachlorobiphenyl	40186-72-9	AccuStandard
2,2',3,3',4,4',5,5'-Octachlorobiphenyl	35694-08-7	AccuStandard
2,2',3,3',4,4',5,6-Octachlorobiphenyl	52663-78-2	AccuStandard
2,2',3,3',4,5,5',6'-Octachlorobiphenyl	52663-75-9	AccuStandard
2,2',3,3',4,5,6'-Heptachlorobiphenyl	38411-25-5	AccuStandard
2,2',3,3',4,5,6-Heptachlorobiphenyl	68194-16-1	AccuStandard
2,2',3,3',4',5,6-Heptachlorobiphenyl	52663-70-4	AccuStandard
2,2',3,3',4-Pentachlorobiphenyl	52663-62-4	AccuStandard
2,2',3,3',5,6,6'-Heptachlorobiphenyl	52663-64-6	AccuStandard
2,2',3,4,4',5,5',6-Octachlorobiphenyl	52663-76-0	AccuStandard
2,2',3,4,4',5,5'-Heptachlorobiphenyl	35065-29-3	AccuStandard
2,2',3,4,4',5'-Hexachlorobiphenyl	35065-28-2	AccuStandard
2,2',3,4',5,5',6-Heptachlorobiphenyl	52663-68-0	AccuStandard
2,2',3,4',5,5'-Hexachlorobiphenyl	51908-16-8	AccuStandard
2,2',3,4',5,6-Hexachlorobiphenyl	68194-13-8	AccuStandard
2,2',3,4,5'-Pentachlorobiphenyl	38380-02-8	AccuStandard
2,2',3,5'-Tetrachlorobiphenyl	41464-39-5	AccuStandard
2,2',3-Trichlorobiphenyl	38444-78-9	AccuStandard
2,2',4,4',5,5'-Hexabromobiphenyl	59080-40-9	Restek
2,2',4,4',5,5'-Hexachlorobiphenyl	35065-27-1	AccuStandard
2,2',4,4',5-Pentachlorobiphenyl	38380-01-7	AccuStandard
2,2',5,5'-Tetrachlorobiphenyl	35693-99-3	AccuStandard
2,2',5-Trichlorobiphenyl	37680-65-2	AccuStandard
2,2',6-Trichlorobiphenyl	38444-73-4	AccuStandard
2,2'-Dichlorobiphenyl	13029-08-8	AccuStandard
2,3,3',4',6-Pentachlorobiphenyl	38380-03-9	AccuStandard
2,3,3',4'-Tetrachlorobiphenyl	41464-43-1	AccuStandard
2,3',4,4'-Tetrachlorobiphenyl	32598-10-0	AccuStandard
2,3',4,5-Tetrachlorobiphenyl	73575-53-8	AccuStandard
2,3',4',6-Tetrachlorobiphenyl	41464-46-4	AccuStandard
2,3,4'-Trichlorobiphenyl	38444-85-8	AccuStandard
2,3',4-Trichlorobiphenyl	55712-37-3	AccuStandard
2,3'-Dichlorobiphenyl	25569-80-6	AccuStandard
2,4,4',5-Tetrachlorobiphenyl	32690-93-0	AccuStandard
2,4,4'-Trichlorobiphenyl	7012-37-5	AccuStandard
2,4'-DDD	53-19-0	Restek

Compound	Cas #	Manufacturer
2,4'-DDE	3424-82-6	Restek
2,4'-DDT	789-02-6	Restek
2,4'-Dichlorobiphenyl	34883-43-7	AccuStandard
2,4'-Methoxychlor	30667-99-3	Restek
2,5-Dichlorobiphenyl	34883-39-1	AccuStandard
2,6-dimethylnaphthalene	581-42-0	Chem Service
2-Chlorobiphenyl	2051-60-7	AccuStandard
2-methylnaphthalene	91-57-6	Chem Service
3-Chlorobiphenyl	2051-61-8	AccuStandard
4,4'-DDD	72-54-8	Restek
4,4'-DDE	72-55-9	Restek
4,4'-DDT	50-29-3	Restek
4,4'-Dichlorobenzophenone	90-98-2	Restek
4,4'-Methoxychlor olefin	2132-70-9	Restek
4-Chlorobiphenyl	2051-62-9	AccuStandard
9,10-anthraquinone	84-65-1	Chem Service
Acetophenone	98-86-2	Restek
Aldrin	309-00-2	Restek
anthracene	120-12-7	Restek
BDE-100	189084-64-8	Restek
BDE-153	68631-49-2	Restek
BDE-47	5436-43-1	Restek
BDE-99	60348-60-9	Restek
benfluralin	1861-40-1	Chem Service
benzophenone	119-61-9	Chem Service
caffeine	58-08-2	Restek
celestolide	13171-00-1	TCI America
Chlorbenside	103-17-3	Restek
Chlorfenson	80-33-1	Restek
Chloroneb	2675-77-6	Restek
chlorpyrifos	2921-88-2	Restek
cis-Chlordane	5103-71-9	Restek
cis-Nonachlor	5103-73-1	Restek
cotinine	486-56-6	Restek
Dacthal	1861-32-1	Restek
Deet	134-62-3	Chem Service
DEHP Bis(2-ethylhexyl)phthalate	117-81-7	Restek
diazinon	333-41-5	Chem Service
dichlorvos	62-73-7	Chem Service
Dieldrin	60-57-1	Restek
diethyl phthalate	84-66-2	Chem Service
dl-camphor	76-22-2	Chem Service

Compound	Cas #	Manufacturer
Endosulfan ether	3369-52-6	Restek
Endosulfan I	959-98-8	Restek
Endosulfan II	33213-65-9	Restek
Endosulfan sulfate	1031-07-8	Restek
Endrin	72-20-8	Restek
Endrin aldehyde	7421-93-4	Restek
Endrin ketone	53494-70-5	Restek
Ethylan	72-56-0	Restek
Fenson	80-38-6	Restek
fluoranthene	206-44-0	Restek
Galaxolide	1222-05-5	TCI America
Heptachlor	76-44-8	Restek
Heptachlor epoxide	1024-57-3	Restek
Hexachlorobenzene	118-74-1	Restek
Indole	120-72-9	Chem Service
Isodrin	465-73-6	Restek
Isophorone	78-59-1	Chem Service
isoquinoline	119-65-3	Chem Service
menthol	1490-04-6	Chem Service
metalaxyl	57837-19-1	Chem Service
Methoxychlor	72-43-5	Restek
methyl salicylate	119-36-8	Chem Service
methyl triclosan	4640_01_1	GmbH
Mirex	2385-85-5	Restek
naphthalene	91-20-3	Restek
oxychlordane	27304-13-8	Chem Service
Pentachloroanisole	1825-21-4	Restek
Pentachlorobenzene	608-93-5	Restek
Pentachlorothioanisole	1825-19-0	Restek
Phantolid	15323-35-0	Key Organics
phenanthrene	85-01-8	Restek
pyrene	129-00-0	Restek
Tetradifon	116-29-0	Restek
tonalid	21145-77-7	Key Organics
trans-Chlordane	5103-74-2	Restek
trans-Nonachlor	39765-80-5	Restek
Traseolide	68140-48-7	Chemcruz
tributoxyethyl phosphate	78-51-3	Chem Service
tributyl phosphate	126-73-8	Chem Service
triethyl citrate	77-93-0	Chem Service
Trifluralin	1582-09-8	Restek
Triphenylphosphate	115-86-6	Chem Service

Compound	Cas #	Manufacturer
Tris(1,3-dichloroisopropyl)phosphate	13674-87-8	Restek
tris(2-chloroethyl)phosphate	115-96-8	Chem Service
α-BHC	319-84-6	Restek
β-BHC	319-85-7	Restek
γ-BHC	58-89-9	Restek
δ-BHC	319-86-8	Restek

Table 3.4 Collection site, health observation, and classification of young-of-year smallmouth bass.

Collection Site	Longitude	Latitude	Collection Date	Fish	Weight (g)	Health Observation at Collection							Classification
						Normal	Myxobolus Infection	Leech Attached	Dermal Lesion	Frayed Fins	Pale Areas	Other	
Conodoquinet Creek (CC)	-77.0404	40.2595	22-Jul-15	CC 001	1.6	x							Normal
				CC 002	1.7	x							Normal
				CC 003	1.8	x							Normal
				CC 004	1.8	x							Normal
				CC 005	1.7	x							Normal
				CC 006	1.5	x							Normal
				CC 007	1.5	x							Normal
				CC 008	1.7							x	Lesioned
				CC 009	1.3	x							Normal
				CC 010	1.6	x							Normal
Susquehanna Harrisburg Juniata (JF)	-76.9171	40.334	14-Jul-15	JF-01	2.0	x							Normal
				JF-02	4.5			x					Lesioned
				JF-03	2.3	x							Normal
				JF-04	2.8					x			Lesioned
				JF-05	2.8		x						Lesioned
				JF-06	2.0				x				Lesioned
				JF-07	2.6			x					Lesioned
				JF-08	3.3	x							Normal
				JF-09	1.2	x							Normal
				JF-10	3.6					x			Lesioned

Collection Site	Longitude	Latitude	Collection Date	Fish	Weight (g)	Health Observation at Collection							Classification
						Normal	Myxobolus Infection	Leech Attached	Dermal Lesion	Frayed Fins	Pale Areas	Other	
Juniata River - Howe Township Park (JR)	-77.0968	40.49193	8-Jul-15	JR 001	1.1						x		Lesioned
				JR 002	1.8		x						Lesioned
				JR 003	1.6						x		Lesioned
				JR 004	1.2	x							Normal
				JR 005	0.9	x							Normal
				JR 006	1.3						x	x	Lesioned
				JR 007	2.7	x							Normal
				JR 008	1.1					x			Lesioned
				JR 009	1.4				x				Lesioned
				JR 010	1.8		x				x		Lesioned
				JR 011	0.9	x							Normal
Kettle Creek Leidy (KC)	-77.9219	41.4051	28-Jul-15	KC 001	5.6			x					Lesioned
				KC 002	5.8	x							Normal
				KC 003	5.5			x					Lesioned
				KC 004	4.9		x	x					Lesioned
				KC 005	4.5			x					Lesioned
				KC 006	4	x							Normal
				KC 007	4.7		x	x					Lesioned
				KC 008	4			x					Lesioned
				KC 009	3.5	x							Normal
				KC 010	4.6			x					Lesioned

Collection Site	Longitude	Latitude	Collection Date	Fish	Weight (g)	Health Observation at Collection							Classification
						Normal	Myxobolus Infection	Leech Attached	Dermal Lesion	Frayed Fins	Pale Areas	Other	
Loyalsock Creek (LC)	-76.9347	41.2424	10-Jul-15	LC 001	1.8		x						Lesioned
				LC 002	1.2	x							Normal
				LC 003	2.2		x						Lesioned
				LC 004	1.4	x							Normal
				LC 005	1		x						Lesioned
				LC 006	0.9	x							Normal
				LC 007	0.9	x							Normal
				LC 008	0.8	x							Normal
				LC 009	0.6	x							Normal
				LC 010	0.5	x							Normal
North Branch Susquehanna Danville (NBSD)	-76.5984	40.9436	24-Jul-15	NBSD 001	4.9					x			Lesioned
				NBSD 002	1.9	x							Normal
				NBSD 003	5.6		x						Lesioned
				NBSD 004	3.8	x							Normal
				NBSD 005	2.3		x						Lesioned
				NBSD 006	2.3		x						Lesioned
				NBSD 007	2.5		x						Lesioned
				NBSD 008	2.9	x							Normal
				NBSD 009	2	x							Normal
				NBSD 010	1.3	x							Normal

Collection Site	Longitude	Latitude	Collection Date	Fish	Weight (g)	Health Observation at Collection: Normal	Myxobolus Infection	Leech Attached	Dermal Lesion	Frayed Fins	Pale Areas	Other	Classification
North Branch Susquehanna Falls (NBSR)	-75.8677	41.46166	15-Jul-15	NBSR 001	2.6	x							Normal
				NBSR 002	2.5	x							Normal
				NBSR 003	0.6	x							Normal
				NBSR 004	1.5	x							Normal
				NBSR 005	1.4	x							Normal
				NBSR 006	0.8	x							Normal
				NBSR 007	0.8	x							Normal
				NBSR 008	1.9	x							Normal
				NBSR 009	3.8	x							Normal
				NBSR 010	3.5	x							Normal
				NBSR 011	2.7						x		Lesioned
Pine Creek Ramsey - Early (PC)	-77.3187	41.2775	24-Jun-15	PC 001	0.7	x							Normal
				PC 002	0.8	x							Normal
				PC 003	0.7	x							Normal
				PC 004	0.6	x							Normal
				PC 005	0.6	x							Normal
				PC 006	0.5	x							Normal
				PC 007	0.5	x							Normal
				PC 008	0.5	x							Normal
				PC 009	0.5	x							Normal

Collection Site	Longitude	Latitude	Collection Date	Fish	Weight (g)	Health Observation at Collection: Normal	Myxobolus Infection	Leech Attached	Dermal Lesion	Frayed Fins	Pale Areas	Other	Classification
Pine Creek Hamilton Bottom (PCHB)	-77.40161	41.34870	21-Jul-15	PCHB 001	1.2	x							Normal
				PCHB 002	0.7	x							Normal
				PCHB 003	1.4		x						Lesioned
				PCHB 004	0.7		x						Lesioned
				PCHB 005	0.7		x						Lesioned
				PCHB 006	0.8	x							Normal
				PCHB 007	2.0	x							Normal
				PCHB 008	1.9		x						Lesioned
				PCHB 009	3.3	x							Normal
				PCHB 010	1.5	x							Normal
				PCHB 011	0.9		x						Lesioned
				PCHB 012	1.0		x				x		Lesioned
Susquehanna River Harrisburg North Branch (SRNB)	-76.9064	40.3363	14-Jul-15	SRNB 001	5		x						Lesioned
				SRNB 002	4.2	x							Normal
				SRNB 003	3.5	x							Normal
				SRNB 004	1.5							x	Lesioned
				SRNB 005	2.6				x				Lesioned
				SRNB 006	2.1	x							Normal
				SRNB 007	2		x						Lesioned
				SRNB 008	4.6					x			Lesioned
				SRNB 009	4.7			x					Lesioned
				SRNB 010	2.5					x			Lesioned

Collection Site	Longitude	Latitude	Collection Date	Fish	Weight (g)	Health Observation at Collection: Normal	Myxobolus Infection	Leech Attached	Dermal Lesion	Frayed Fins	Pale Areas	Other	Classification
Bald Eagle Creek Unionville (UBEC)	-77.86952	40.90636	29-Jul-15	UBEC 001	2.6	x							Normal
				UBEC 002	2.1	x							Normal
				UBEC 003	0.6	x							Normal
				UBEC 004	1.6		x						Lesioned
				UBEC 005	0.8			x					Lesioned
				UBEC 006	1.7		x						Lesioned
				UBEC 007	1.8		x						Lesioned
				UBEC 008	2.3	x							Normal
				UBEC 009	1.6		x						Lesioned
				UBEC 010	1.5	x							Normal
				UBEC 011	1.2	x							Normal
				UBEC 012	1.4	x							Normal
Susquehanna Harrisburg West Branch (WB)	-76.9126	40.33538	14-Jul-15	WB-01	1						x		Lesioned
				WB-02	2.3	x							Normal
				WB-03	1			x					Lesioned
				WB-04	1.7				x				Lesioned
				WB-05	1.1						x		Lesioned
				WB-06	2.2		x						Lesioned
				WB-07	2				x				Lesioned
				WB-08	1.2	x							Normal
				WB-09	0.6	x							Normal
				WB-10	0.9	x							Normal
				WB-11	2.4		x						Lesioned

Collection Site	Longitude	Latitude	Collection Date	Fish	Weight (g)	Health Observation at Collection							Classification
						Normal	Myxobolus Infection	Leech Attached	Dermal Lesion	Frayed Fins	Pale Areas	Other	
West Branch Susquehanna Lewisburg (WBSL)	-76.8778	40.9663	13-Jul-15	WBSL 001	1.5		x						Lesioned
				WBSL 002	2.5		x						Lesioned
				WBSL 003	2.7		x						Lesioned
				WBSL 004	1.9	x							Normal
				WBSL 005	3.5	x							Normal
				WBSL 006	1.3		x						Lesioned
				WBSL 007	1.7		x						Lesioned
				WBSL 008	1.4	x							Normal
				WBSL 009	4.1					x			Lesioned
				WBSL 010	1.6		x						Lesioned
Wyalusing Creek (WC)	-76.2309	41.6975	15-Jul-15	WC 001	2.6	x							Normal
				WC 002	1.6	x							Normal
				WC 003	1.7	x							Normal
				WC 004	1.5	x							Normal
				WC 005	1.1	x							Normal
				WC 006	1.5	x							Normal
				WC 007	2.9			x					Lesioned
				WC 008	3.5	x							Normal
				WC 009	2.4	x							Normal
				WC 010	1.7	x							Normal

Table 3.5 Tentative identification of final reduced data set for non-targeted analysis. Analytes were searched against the NIST MS Database (2014 – black, 2017 – red). First and second retention time are averaged value across all runs. Spectral match values are from randomly selected sample containing analyte of interest.

Analyte	Class	Fisher Ratio	Mass	t_{R1} (s)	t_{R2} (s)	Similarity	Reverse	Probability
Desmosterol	alcohol	19.17	271	3163.64	0.58	826	831	6775
2,4,7,9-Tetramethyl-5-decyn-4,7-diol	alcohol	5.35	109	1352.25	1.13	839	841	6912
2H-Pyranmethanol, tetrahydro-2,5-dimethyl-	alcohol	7.37	113	638.87	0.93	733	741	3273
3-Methoxy-3-methylbutanol	alcohol	6.08	73	753.09	0.89	720	748	2400
4-Methyl-5-decanol	alcohol	4.46	83	1579.70	1.65	800	842	1991
Benzaldehyde, 2-methyl-	aldehyde	4.14	91	832.71	1.30	937	947	4474
2-Octenal, (E)-	aldehyde	4.45	70	779.01	1.07	956	963	7535
Isobutyl propane-1,3-diyl dicarbonate	carbonate	5.18	103	2812.26	2.14	764	784	6277
Phosphonofluoridic acid, (1-methylethyl)-, cyclohexyl ester	ester	8.25	127	1360.84	1.30	706	806	29.9
Benzoic acid, 3-amino-, ethyl ester	ester	4.19	120	1530.09	1.76	917	959	9031
Fumaric acid, 2-chlorophenyl ethyl ester	ester	9.50	127	1304.33	1.51	807	828	2659
Methyl tetradecanoate	ester	15.55	74	1803.29	1.17	955	958	8149
Pentanedioic acid, dimethyl ester	ester	4.53	59	900.88	1.26	813	887	8394
Octadecanoic acid, 3-hydroxy-, methyl ester	ester	23.79	103	2605.33	1.78	762	809	49.7
Docosapentaenoic acid, methyl ester	ester	38.78	79	2627.40	1.91	888	888	5172
Dodecanoic acid, methyl ester	ester	7.56	74	1526.39	1.12	935	941	6652
Methyl 3-hydroxyhexadecanoate	ester	6.28	103	2625.56	1.81	768	781	5382
4,7,10,13,16,19-Docosahexaenoic acid, methyl ester, (all-Z)-	ester	16.15	79	2612.88	1.91	909	910	7461
9-Octadecenoic acid (Z)-, methyl ester	ester	18.13	55	1862.05	1.36	849	861	4170
Perhydrophenalene, (3aα, 6aα, 9aα, 9bβ)-	hydrocarbon	8.96	81	1308.15	1.17	736	863	52.3
1H-Indene, 2,3-dihydro-1,6-dimethyl-	hydrocarbon	4.13	131	1028.88	1.52	825	828	35.9
Benzene, 1-ethyl-3,5-dimethyl-	hydrocarbon	7.62	119	830.14	1.05	904	912	1692

Analyte	Class	Fisher Ratio	Mass	t_{R1} (s)	t_{R2} (s)	Similarity	Reverse	Probability
Benzene, 1-methyl-3-(1-methylethyl)-	hydrocarbon	7.99	119	730.27	0.97	904	915	3190
Benzene, 2-ethyl-1,4-dimethyl-	hydrocarbon	5.45	119	864.86	1.11	892	912	1315
Biphenyl	hydrocarbon	5.44	154	1331.19	1.47	877	905	7243
1-Ethyldecahydronaphthalene , (Z,E)	hydrocarbon	9.18	137	1249.58	1.06	776	846	886
Cycloeicosane	hydrocarbon	29.62	55	2455.77	1.45	880	886	641
Decane	hydrocarbon	10.83	57	679.18	0.79	944	952	6360
Decane, 2,6,8-trimethyl-	hydrocarbon	15.23	57	1016.48	0.88	796	860	2068
Decane, 2,9-dimethyl-	hydrocarbon	6.25	57	893.84	0.84	923	929	5740
Dodecane	hydrocarbon	5.23	57	769.88	0.81	827	850	1599
Heptadecane	hydrocarbon	8.24	57	1752.00	1.00	854	868	1005
Heptadecane, 2,6,10,15-tetramethyl-	hydrocarbon	3.99	57	1475.56	0.95	848	863	1065
Hexadecane, 2-methyl-	hydrocarbon	7.82	57	1727.54	1.00	929	937	4223
Hexadecane	hydrocarbon	5.17	57	1140.63	0.89	856	873	3271
Limonene	hydrocarbon	5.00	68	738.41	0.92	932	941	3531
Naphthalene, 1-(2-propenyl)-	hydrocarbon	5.60	153	1355.73	1.42	723	763	6206
Naphthalene, 2-decyldecahydro-	hydrocarbon	4.95	137	1725.04	1.14	715	718	4122
Naphthalene, decahydro-1,5-dimethyl-	hydrocarbon	5.82	95	1008.00	0.98	799	801	1434
Octane, 2,4,6-trimethyl-	hydrocarbon	23.57	57	636.77	0.77	833	844	937
Pentacosane	hydrocarbon	68.67	57	2916.00	1.76	919	931	2052
2-Methylhexacosane	hydrocarbon	5.81	57	2915.48	1.76	928	935	1438
6-Methylundecane	hydrocarbon	6.02	57	1417.47	0.95	736	873	6154
Pentadecane	hydrocarbon	4.47	57	1493.78	0.98	894	908	4103
Hexadecahydropyrene	hydrocarbon	4.82	218	1779.87	1.35	788	806	5185
Undecane	hydrocarbon	9.02	57	845.31	0.83	855	872	1287
Isogermacrene D	hydrocarbon	5.69	161	1488.20	1.20	835	885	1849
1,19-Eicosadiene	hydrocarbon	8.39	55	2983.84	2.37	912	929	1470
1-Heptadecene	hydrocarbon	4.55	55	1764.18	1.46	926	933	1324

Analyte	Class	Fisher Ratio	Mass	t_{R1} (s)	t_{R2} (s)	Similarity	Reverse	Probability
2,2'-Dimethylbiphenyl	hydrocarbon	4.05	167	1496.59	1.43	793	868	2705
7-Octylidenebicyclo[4.1.0]heptane	hydrocarbon	4.68	135	1319.85	1.15	724	793	995
Ethanone, 1-[4-(1-hydroxy-1-methylethyl)phenyl]-	ketone	9.56	163	1491.88	1.53	828	841	67.5
6,10,14-Trimethyl-2-pentadecanone	ketone	5.86	58	1952.98	1.20	918	922	7920
3-Methylene-1-oxa-spiro[4.5]decan-2-one	ketone	4.13	123	1397.15	1.40	734	754	4347
Semioxamazide	nitrogen containing	4.63	103	2703.05	1.92	831	846	97.2
Benzenamine, 4,4'-ethylidenebis[N,N-diethyl-3-methyl-	nitrogen containing	5.30	337	2723.62	2.09	717	783	7461
Caprolactam	nitrogen containing	8.68	113	1118.49	1.72	909	927	9728
2-Oxo-1-methyl-3-isopropylpyrazine	nitrogen containing	8.98	137	1219.96	1.44	721	747	1206
3,5-Di-tert-butyl-2-hydroxybenzonitrile	nitrogen containing	8.10	216	1635.28	1.32	829	871	7972
9-Octadecenamide, (Z)-	nitrogen containing	5.49	59	2532.30	1.90	776	776	8391
2-tert-Butoxytetrahydrofuran	other	4.87	71	629.28	0.96	747	774	36.7
Ethylene glycol - Adipate - Diethylene glycol	other	5.51	173	1701.10	1.85	702	707	63.9
Costunolide	other	137.80	81	1876.34	1.37	753	808	2550
Tripropylene glycol monomethyl ether	other	4.80	59	1162.41	1.19	904	919	9657
2-(Methylmercapto)benzothiazole	sulfur containing	5.15	181	1668.20	1.82	792	799	7186

3.8.3 Figures

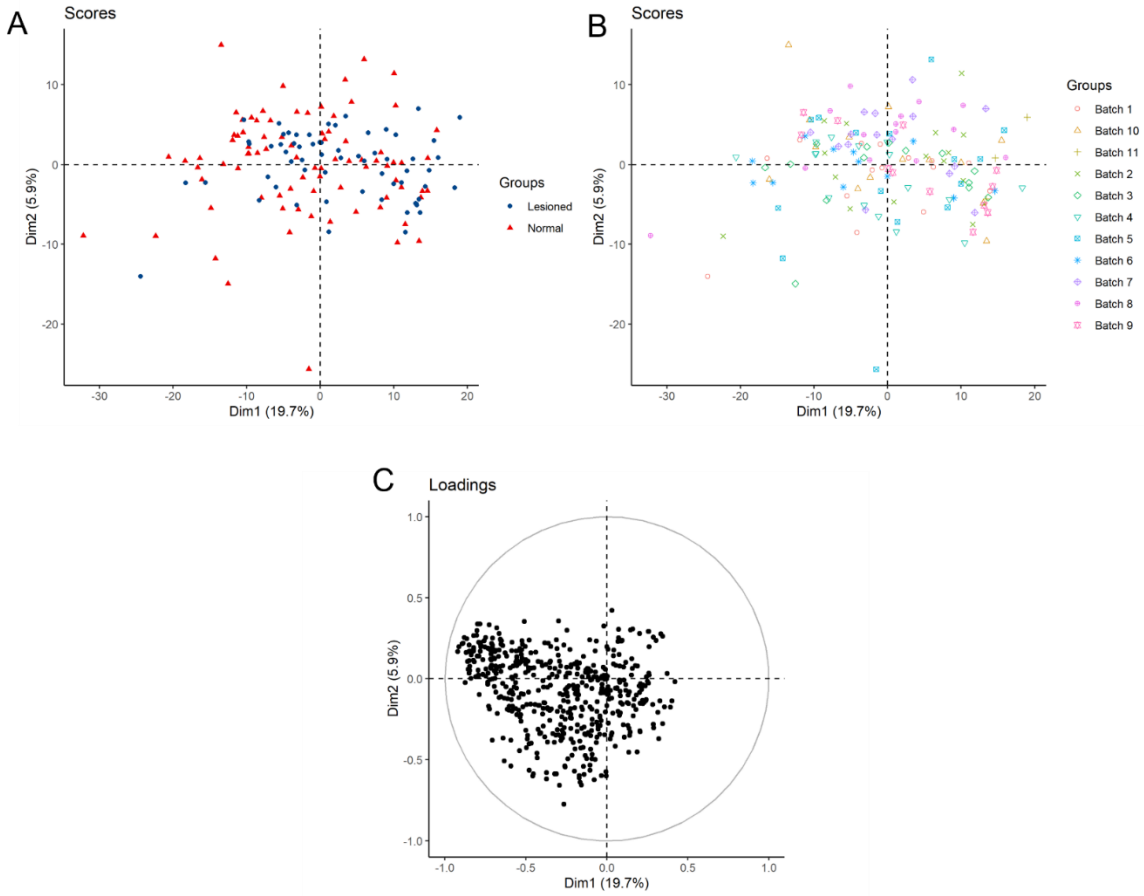

Figure 3.4 Unsupervised PCA scores (A and B) and loadings (C) of aligned peaks from 146 YOY SMB. PCA was performed on 512 analytes with no data filtering by Fisher ratio analysis. The scores plots (A and B) are the same individual samples assigned to different sample groups. A) Scores plot of individual samples grouped by disease state as either Lesioned sample class (blue) or Normal sample class (red). B) Scores plot of individual samples grouped by sample preparation batch 1-11. C) Loadings plot corresponding to the shown scores plot of all 512 analytes. Dim1 is principal component 1. Dim2 is principal component 2. Unsupervised PCA did not indicate any clustering among samples when defined by sample classes used in non-targeted analysis or by sample batch.

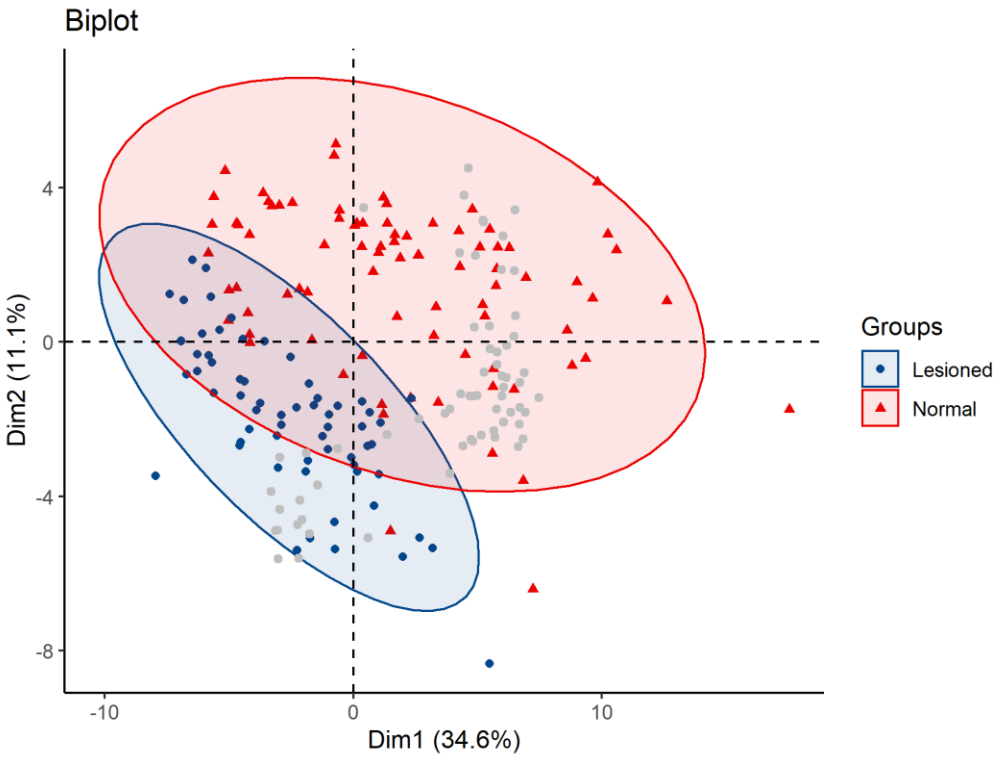

Figure 3.5 PCA biplot of final reduced data set for 146 YOY SMB containing 65 analytes. Ellipses are 95% confidence intervals around group centered mean Dim1 is principal component 1. Dim2 is principal component 2. Grey circles represent variable loadings of the 65 final analytes. Red triangle (Normal class) and blue circle (Lesioned class) represent individual sample scores.

Chapter 4

Targeted and Non-Targeted Analysis of Young-of-Year Smallmouth Bass by Ultra-Performance Liquid Chromatography Coupled with Quadrupole Time-of-Flight Mass Spectrometry

4.1 Introduction

The negative impacts of anthropogenic contamination can be observed worldwide in aquatic life. These contaminants are released into the environment through various sources, including farming, manufacturing, wastewater treatment plants, or run off from other industrialized areas.[1,2] Negative health characteristics have been associated with contaminants like legacy organochlorine pesticides and current use pesticides in multiple species of fish for decades. Acute exposures to pesticides can be lethal or lead to morphological and behavioral changes.[3] Long-term exposures, especially during developmental life stages, can also cause long term health impairments such as endocrine disruption or immunosuppression.[4,5] Newly recognized contaminants of emerging concern (CECs) also pose health risks. Per- and polyfluoroalkyl substances (PFAS) have become an increasing concern in recent years due to negative health associations in humans.[6] These compounds are also being detected in fish tissues more frequently as they transition into a routinely monitored chemical class in the environment.[7,8] In addition to chemical analyses, a common way to monitor for the negative effects of anthropogenic contaminants is to monitor the health of indicator species who are sensitive to changes in environmental conditions.[9]

Smallmouth bass (*Micropterus dolomieu*, SMB) are a popular sport fish in North America that also serve as excellent indicators for water quality and ecosystem health due to their extreme sensitivity to environmental changes. Throughout the Susquehanna River Basin of the Chesapeake Bay Watershed, USA, SMB have been exhibiting signs of disease for over fifteen years.[10–13] Reproductive endocrine disruption, limited to adult male SMB, is characterized by indicators of intersex such as testicular oocytes and the expression of the egg precursor protein vitellogenin.[10,14] Immunosuppression, which can impact

any SMB including young-of-year (YOY), is characterized by co-infection with multiple bacterial, viral, fungal, and/or parasitic pathogens.[11,15–17] These infections can cause a variety of dermal and internal lesions.[13,16,18] Previous analyses have explored chemical contaminants in water, sediment, and fish tissues and have identified varying levels of pollutants capable of inducing endocrine disruption and immunosuppression, but to date no unifying contaminant or group of contaminants capable of inducing these conditions have been identified across multiple regions exhibiting the same signs of disease. This is likely due to the complex interactions between environmental stressors, pathogens, and chemical contaminants in relation to SMB health. Although, as most previous studies have relied upon targeted analyses and focused on the study of adult SMB, it is possible that the unifying factor has been missed due to limitations in these study designs.

Targeted analyses are a standardized practice in environmental monitoring situations and use well-validated methodologies to identify known contaminants. Target lists are curated based on specific features of the surrounding areas (e.g., manufacturing complexes, farmland) and the suspected contaminants associated with these areas.[19,20] However, the suspected target compounds are not always detected during analysis. Furthermore, the target compounds detected may not explain the observed health characteristics or may only represent a partial explanation. This is due to targeted analyses actively excluding non-target compounds of interest from identification, usually through extensive sample clean up or by using highly specific analysis conditions. This provides researchers with a limited view of what is happening in the environment and prevents the discovery of CECs or other known anthropogenic contaminants that were not included on the target list.[21] By expanding investigations to include the discovery of non-target compounds, it is possible to identify unknown pollutants, CECs, and potential biomarkers that correlate with the observed signs of disease.[22,23]

Simultaneous analysis of samples for both target and non-targeted compounds can be aided by using advanced analytical techniques, such as ultra-performance liquid chromatography coupled with high resolution quadrupole time-of-flight mass spectrometry (UPLC-QTOF-MS). The increased sensitivity and resolution of UPLC-QTOF-MS, over other LC-MS systems, is well suited to separations

in complex biological matrices. To retain the most information about non-target compounds, minimal sample cleanup procedures are used, thus, such advanced chromatographic platforms are necessary to identify compounds. The combination quadrupole/time-of-flight MS allows for the detection of both molecular ions and fragment ions and provides more information for the identification of unknowns, while still capturing a broad range of m/z ratios.

This study presents the simultaneous targeted and non-targeted analysis of 146 YOY SMB collected from 14 sites throughout the Susquehanna River Basin in Pennsylvania, USA. YOY SMB were chosen for this study due to their high/chronic mortality in the river basin as well as their susceptibility to infections by multiple pathogens suggestive of immunosuppression. Additionally, studies in other fish species have linked exposure to contaminants during juvenile life stages with long term health impairments.[24–26] Therefore, identifying contaminant exposures in YOY SMB may provide information about the observed signs of disease in adult SMB for future studies. UPLC-QTOF-MS is combined with a modified quick, easy, cheap, effective, rugged, and safe (QuEChERS) methodology, previously optimized for the analysis of YOY SMB,[27] to provide a quantitative analysis of 33 target pesticides and 5 target PFAS as well as a qualitative investigation of non-target compounds.

4.2 Materials and Methods

4.2.1 Reagents and Chemicals

External and internal calibration standards used for the targeted analysis are presented in Table 4.1. Target pesticide standards were obtained from Restek Corp. (Bellefonte, PA, USA). Target PFAS standards and labeled internal standard PFAS were obtained from Wellington Laboratories (Guelph, ON, CA). Labeled internal standard pesticides were obtained from Sigma-Aldrich (St. Louis, MO, USA). All solvents used in this analysis were of analytical grade. Ethyl acetate and HPLC water used for extraction were obtained from JT Baker (Phillipsburg, NJ, USA). Acetonitrile used during extraction was obtained

from Honeywell Burdick and Jackson (Muskegon, MI, USA). Buffered QuEChERS extraction salts were obtained from Restek Corp. Surrogate compounds [polychlorinated biphenyl congeners: CB-18, CB-28, CB-52, triphenyl phosphate, and tris(1,3-dichloroisopropyl)phosphate] were obtained from Restek Corp. LCMS QC Reference Standard mix was obtained from Waters Corp. (Milford, MA, USA) and included the following compounds: sulfaguanidine, acetaminophen, caffeine, val-tyr-val, leucine-enkelphalin, sulfadimethoxine, verapamil, reserpine, and terfenadine. For sample analysis, formic acid was from Fisher Scientific (Waltham, MA, USA). Water and acetonitrile were both LC-MS grade and from Honeywell (Charlotte, NC, USA).

Table 4.1 Summary of target and labeled internal standard compounds used during targeted analysis.

Target Compound	Labeled Internal Standard	Ionization	t_R (min)	Exact Mass	Molecular Ion	Fragment Ion 1	Fragment Ion 2
Acetamiprid	Acetamiprid-d3	+	2.68	222.0672245	223.075	223.0745	187.0978
Acetochlor	Acetochlor-(2-ethyl-6-methylphenyl-d11)	+	7.1	269.118256	270.1261	148.1126	133.0891
Alachlor	Alachlor-d13	+	7.1	269.118256	270.1261	147.1048	238.0999
Atrazine	Atrazine-d5	+	4.76	215.093773	216.1016	174.0546	96.0562
Boscalid	Dimethomorph-(dimethyoxy-d6)	+	6.54	342.03267	343.0405	307.0638	139.9903
Chlorantraniliprole	Dimethomorph-(dimethyoxy-d6)	+	5.63	480.970793	481.9786	283.9226	450.9364
Clofentezine	Dimethomorph-(dimethyoxy-d6)	+	8.04	302.012602	303.0204	138.0111	120.0436
Cyanazine (Bladex)	Atrazine-d5	+	3.8	240.089022	241.0968	214.0859	104.0015
Desethylatrazine	Atrazine-d5	+	2.45	187.062473	188.0703	188.0698	146.0228
Desisopropylatrazine	Atrazine-d5	+	1.72	173.046823	174.0546	174.0541	146.0228
Dimethenamid	Thiacloprid-(thiazolidin ring-d4)	+	6.22	275.074678	276.0825	244.0563	168.0847
Dimethomorph	Dimethomorph-(dimethyoxy-d6)	+	5.7/5.95	387.123735	388.1316	165.0107	301.0631
Etoxazole	Propiconazole-(phenyl-d3)	+	9.6	359.169685	360.1775	141.0152	304.1144
Fenhexamid	Fenhexamid-(methyl-d3)	+	6.58	301.063635	302.0715	97.1017	55.0548
Fipronil	Fipronil-(pyrazole-13C3, cyano-13C)	+/-	7.5	435.938707	436.9465 (+) / 434.9309 (-)	367.9513 (+) / 329.959 (-)	249.9813 (+) / 248.9813 (-)
Flonicamid	Imidacloprid-d4	+	1.89	229.046296	230.0541	174.0167	-
Fludioxonil	Fipronil-(pyrazole-13C3, cyano-13C)	-	5.96	248.03974	247.0319	180.0324	126.0344
Hexythiazox	Thiamethoxam-d3	+	9.4	352.101227	353.109	228.025	-
Imazalil	Imazalil-(allyl-d5)	+	3.96	296.048319	297.0561	158.9768	69.0453
Imidacloprid	Imidacloprid-d4	+	2.35	255.052302	256.0601	209.0594	175.0978
Metolachlor	Metolachlor-(2-ethyl-6-methylphenyl-d11)	+	7.18	283.133907	284.1417	176.1439	252.1155
Metribuzin	Metribuzin-(S-methyl-d3)	+	3.9	214.088832	215.0967	84.0813	215.0961
Myclobutanil	Myclobutanil-(phenyl-d4)	+	6.35	288.114174	289.112	70.0405	125.0158
Paclobutrazol	Paclobutrazol-(phenyl-d4)	+	5.8	293.12949	294.1373	125.0158	70.0405
Pendimethalin	Acetochlor-(2-ethyl-6-methylphenyl-d11)	+	9.36	281.137556	282.1454	212.0671	282.1448
Prometon	Acetamiprid-d3	+	3.17	225.15896	226.1668	142.0729	184.1198

Target Compound	Labeled Internal Standard	Ionization	t_R (min)	Exact Mass	Molecular Ion	Fragment Ion 1	Fragment Ion 2
Propachlor	Atrazine-d5	+	5.3	211.076392	212.0842	170.0373	94.0657
Propazine	Atrazine-d5	+	5.7	229.109424	230.1172	146.0233	230.1167
Pyridaben	Thiamethoxam-d3	+	10.16	364.137611	365.1454	147.1174	309.0828
Simazine	Atrazine-d5	+	3.75	201.078123	202.086	124.0875	132.0328
Tebuconazole	Tebuconazole-(trimethyl-13C3)	+	6.75	307.14514	308.1549	70.0405	125.0158
Thiacloprid	Thiacloprid-(thiazolidin ring-d4)	+	3.12	252.023645	253.0315	126.0111	253.0309
Thiamethoxam	Thiamethoxam-d3	+	1.85	291.019288	292.0271	211.0654	181.0548
Perfluorodecanoic acid (PFDA)	Perfluoro-n-[1,2-13C2]decanoic acid (MPFDA)	-	7.16	513.967316	512.9595	468.9697	218.9856
Perfluorononanoic acid (PFNA)	Perfluoro-n-[1,2,3,4,5-13C5]nonanoic acid (MPFNA)	-	6.4	463.970509	462.9627	418.9729	218.9856
Perfluorooctanesulfonate (PFOS)	Sodium perfluoro-1-[1,2,3,4-13C4]octanesulfonate (MPFOS)	-	7.2	499.937494	498.9297	79.9568	98.9552
Perfluorooctanoic acid (PFOA)	Perfluoro-n-[1,2-13C2]octanoic acid (M2PFOA)	-	5.6	413.973703	412.9659	368.976	168.9888
Perfluoroundecanoic acid (PFUnA)	Perfluoro-n-[1,2-13C2]undecanoic acid (MPFUdA)	-	7.95	563.964121	562.9563	518.9665	168.9888

4.2.2 Sample Collection and Preparation

Sample collection and preparation were previously described in section 3.3 Methods and Materials. In brief, YOY SMB were collected from 14 sites in Pennsylvania, USA located within the Susquehanna River Basin (Figure 3.1, Tables 3.1 and 3.4), several weeks post-spawning (June – July 2015). A total of 9 – 12 fish were collected per site by electrofishing. Matrix blank samples were YOY SMB raised in captivity at the Eastern Ecological Science Center, Kearneysville, WV. Captive fish were raised and euthanized according to the Science Center's Institutional Animal Care and Use Committee guidelines.

Samples were previously extracted using the described modification to the QuEChERS extraction protocol in Chapter 2.[27] Prior to analysis by GC×GC-TOFMS each sample was divided for analysis on multiple chromatographic platforms. For samples undergoing analysis by UPLC-QTOF-MS, the sample extracts underwent a solvent exchange from ethyl acetate to acetonitrile. The volume of each sample extract was measured, then samples were blown to dryness by gentle nitrogen blow-down. Samples were reconstituted using an equal volume of acetonitrile. Extracts were transferred to clean vials to prevent interference from any precipitated lipids that did not go back into solution during reconstitution with acetonitrile. This provided a cleaner extract for the one-dimensional LC analysis. Extracts were spiked with internal standards at 10 ng ml^{-1} final concentration before instrumental analysis to monitor for instrumental variation. Aliquots from each sample were combined to create pooled samples for quality control purposes.

4.2.3 Data Acquisition

Samples were analyzed using a Xevo G2-XS QTof (Waters Corp.) equipped with an Acquity UPLC I-Class Fixed Loop sample manager (Waters Corp.) and an Acquity HSS C18 column (2.1 mm x

150 mm, 1.8 µm; Waters Corp.). Mobile phase A consisted of 0.1% formic acid in water. Mobile phase B consisted of 0.1% formic acid in acetonitrile. The total analysis time was 22 minutes, and the following gradient program was used: initial – 12 minutes 15% B; 12 – 15 minutes 100 % B; 15.1 – 22 minutes 15% B. A flow rate of 0.5 ml min^{-1} was maintained for the entirety of the run, and the column was held at 50 ℃. Sample extracts were maintained at 4 ℃. A 1 µl injection by partial loop needle overfill was used to introduce sample extracts to the system.

Sample analysis was conducted using an electrospray ionization source in both positive (ES+) and negative (ES-) ion modes. The acquisition mode was set to MS^e using two acquisition functions, low energy where the collision energy was off and high energy where the collision energy ramped from 10 – 50 eV. The collected mass range was 50 – 1200 amu with a scan time of 0.2 seconds. The source temperature was maintained at 120 ℃. The capillary voltage was 1.0 kV and the cone voltage was 20 V. The desolvation gas temperature was 500 ℃ with flow rate of 1000 l hr^{-1}. The cone gas flow was 50 l hr^{-1}. The lock mass compound was leucine enkelphalin (ES+ *m/z* 556.2771, ES- *m/z* 554.2615) and a reference scan was obtained every 30 seconds. Data was acquired using MassLynx v4.2 SCN 1018 (Waters Corp.). Samples were randomly analyzed in duplicate for both ion modes. Pooled samples without LCMS QC Reference mix were included in the randomized sample queue. Pooled samples spiked with LCMS QC Reference mix at a 1:20 dilution from stock (compound concentrations 0.01 – 0.125 µg ml^{-1}) were analyzed after every 30 samples.

4.2.4 Data Analysis

The targeted analysis was completed using TargetLynx XS within the MassLynx software. Internal standard normalization was applied (per Table 4.1) to a 7-point external standard calibration curve and all detected target compounds within samples. External calibration concentrations ranged from 0.01 to 100 ng ml^{-1}. Linearity was considered good for target pesticides, with 29 out of 32 having correlation coefficients > 0.998. Three pesticides, fenhexamid, etoxazole, and fipronil, were > 0.982. For

PFAS targets, only PFOS was linear over the given concentration range. The remaining four PFAS targets were linear over the highest three concentrations (1 – 100 ng ml^{-1}). Their concentrations were approximated using this reduced calibration curve and reported as either above or below 1 ng g^{-1} wet weight (final calculated concentration in homogenate). Target compounds were identified using retention time, molecular ion, and fragment ion matching. Targets were quantified on their molecular ion trace using extracted ion chromatograms.

For the non-targeted analysis, chromatographic alignment, automatic peak picking, deconvolution, and compound identification were completed using Progenesis QI v. 3.0 (Waters Corp.). Alignment was performed separately for each ionization mode using a reference chromatogram, chosen by the software out of a set of pooled sample analyses, and a minimum match score of 90% was required. The following adducts were included during ES+ analysis: M+H, M+H-H_2O, M+Na, M+K, M+2H, and 2M+H; for ES- analysis: M-H, M-H_2O-H, M+Na-2H, and M+k-2H. Analysis of variance (ANOVA) and fold change calculations were used to reduce the initial aggregate compound lists generated after peak picking and deconvolution. Features retained after reduction were exported to Microsoft Excel (Excel v.2107) for further statistical analysis using MetaboAnalyst v.5.0.[28]

4.3 Results and Discussion

4.3.1 Targeted Analysis

A targeted analysis was conducted to determine the presence and quantity of known contaminants that YOY SMB may have been exposed to. A list of 33 pesticides and 5 PFAS was generated based on previous analyses of SMB tissues.[17,20,29] Pesticides are routinely monitored in non-target organisms in the environment as many are known to negatively impact organism health. PFAS are being increasingly investigated in the environment due to growing concerns around human exposures to these compounds.

In total, eight target pesticides were detected at reportable concentrations in the YOY SMB sampled (Table 4.2). Duplicate sample runs were averaged, and the average of the detected target analytes had to be above the limit of quantification and exceed any background concentration detected within matrix blanks by 10 x to be reported. This is to ensure that reported analytes are not the result of contamination during sample handling and processing. Pesticides were detected at low levels, with few compounds exceeding 1 ng g^{-1} wet weight. Fludioxonil was detected at the highest concentration (avg. 2.43 ng g^{-1} at WB) but was only detected at a reportable concentration at one sampling site, occurring in three samples. Though the average concentration of pesticides per site was low, the frequency of detected pesticides was high. For the reported targets, five compounds were detected in over 60% of samples. Simazine, metolachlor, and propazine were detected in 100%, 98%, and 97% of samples, respectively. Hexythiazox and deisopropylatrazine were found in 72% and 64% of samples, respectively.

The five most frequently detected target compounds are all currently approved for use by the U.S. Environmental Protection Agency, as their risk to human health is considered minimal. However, several of these compounds are suspected or known of causing negative health effects in non-target species like freshwater fish as well as other terrestrial and aquatic organisms.[30–34] Deisopropylatrazine is a degradation product of the herbicide atrazine. Together, simazine, propazine, and atrazine are part of a chemical class of herbicides, known as triazines, primarily used in agriculture. Triazine toxicity has been studied in a variety of fish species, although not specifically in SMB, and toxic effects are highly dependent on the species of fish. For example, studies of atrazine have found it to be "practically non-toxic" in crucian carp (*Carassius carassius*) with a 96-hr LC_{50} value > 100 mg l^{-1}, while being "moderately toxic" in brook trout (*Salvelinus fontinalis*) and bluegill (*Lepomis macrochirus*) with 96-hr LC_{50} values of 6.3 mg l^{-1} and 8 mg l^{-1}, respectively. Additionally, triazines can induce behavioral changes in some species, that result in negative effects on the population level.[35]

Table 4.2 Average concentration of detected target pesticides at each sampling site.

Site		**Boscalid**	**Hexythiazox**	**Deisopropylatrazine**	**Metolachlor**	**Propachlor**	**Propazine**	**Simazine**	**Fludioxonil**
CC	*Mean (ng g^{-1})*	NR	0.32	0.17	0.28	NR	0.11	0.36	ND
	Frequency	3	8	9	10	7	8	10	0
JF	*Mean (ng g^{-1})*	NR	0.97	0.82	0.18	NR	0.04	0.48	ND
	Frequency	1	8	7	10	5	9	10	0
JR	*Mean (ng g^{-1})*	0.08	0.11	0.12	0.30	NR	0.05	0.37	ND
	Frequency	1	6	7	11	9	11	11	0
KC	*Mean (ng g^{-1})*	NR	0.06	NR	0.22	NR	0.03	0.17	ND
	Frequency	2	5	7	10	6	10	10	0
LC	*Mean (ng g^{-1})*	0.06	0.44	NR	0.61	NR	0.05	0.51	ND
	Frequency	1	8	6	10	3	10	10	0
NBSD	*Mean (ng g^{-1})*	NR	0.86	NR	0.11	NR	0.02	0.19	ND
	Frequency	1	9	5	7	2	10	10	0
NBSR	*Mean (ng g^{-1})*	NR	1.28	0.14	0.30	NR	0.04	0.40	NR
	Frequency	2	9	5	11	9	11	11	1
PC	*Mean (ng g^{-1})*	ND	0.12	0.12	0.32	NR	0.04	0.38	ND
	Frequency	0	5	5	9	2	9	9	0
PCHB	*Mean (ng g^{-1})*	NR	0.19	0.11	0.17	NR	0.05	0.47	NR
	Frequency	3	8	6	12	1	11	12	1
SRNB	*Mean (ng g^{-1})*	NR	0.11	NR	0.33	NR	0.05	0.53	ND
	Frequency	1	5	6	10	5	10	10	0
UBEC	*Mean (ng g^{-1})*	0.31	0.33	0.19	NR	NR	0.09	0.89	NR
	Frequency	1	10	7	12	5	12	12	1
WB	*Mean (ng g^{-1})*	NR	0.52	0.46	0.21	NR	NR	0.45	2.43
	Frequency	1	9	10	11	7	11	11	3
WBSL	*Mean (ng g^{-1})*	NR	0.40	0.15	0.52	NR	0.05	0.46	ND
	Frequency	2	5	6	10	7	10	10	0
WC	*Mean (ng g^{-1})*	NR	0.79	0.15	0.51	0.21	0.10	1.29	ND
	Frequency	2	10	8	10	3	10	10	0

PFAS were also detected in a large proportion of the samples analyzed in this study (Table 4.3), with PFOS being detected in 99% of samples and PFDA in 79%. Unfortunately, the generalized design of this study for a non-targeted analysis, meant that the conditions were not optimal for the detection and quantification of PFAS, especially the perfluorinated acid compounds. As such, though PFDA was detected in 79% of samples, it could not be accurately quantified. Estimated concentrations for PFDA were below 1 ng g^{-1} wet weight in most samples. Estimated concentrations were also calculated for PFOA, PFNA, and PFUdA, but these compounds were not frequently detected (or for PFOA, detected at all) in samples. It should be emphasized that under the analytical conditions employed for this study, the absence of these compounds within samples cannot be confirmed at this time. Previously, PFDA, PFUdA, and perfluorododecanoic acid (PFDoA) were detected in the serum of adult SMB from the Susquehanna River Basin[29] and further investigation into the presence of PFAS in YOY SMB is still needed.

PFOS, unlike the perfluorinated acid compounds, could be accurately quantified using an external calibration curve under these conditions. PFAS are persistent and bioaccumulate in wildlife, both terrestrial and aquatic. A study in wild fish of South Carolina, USA found that PFOS levels were the highest of any PFAS found in fish muscle and whole fish and exceeded the protective guidelines in 83% of the whole fish sampled.[36] PFOS has also been associated with immunosuppression in marine medaka larvae (*Oryzias melastigma*) exposed during early embryonic stages.[37] At this time, it is not possible to draw conclusions between PFOS exposure and immunosuppression in SMB and more investigation into the detection, quantification, and associated health risks of PFAS is needed. However, given the increasing frequency of detection of these compounds within the Susquehanna River Basin and documented associations of negative health effects in other species of fish, it is possible that future connections between PFAS and the observed disease characteristics may be made.

Table 4.3 Summary of PFAS detected at each sampling site.

	Perfluorooctanesulfonate (PFOS)		**Perfluoroundecanoic acid (PFUdA)**		**Perfluorodecanoic acid (PFDA)**		**Perfluorononanoic acid (PFNA)**	
Site	*Mean (ng g^{-1})*	*Frequency*	*Detection Level*	*Frequency*	*Detection Level*	*Frequency*	*Detection Level*	*Frequency*
CC	21.73	10	< 1 ng g^{-1}	1	< 1 ng g^{-1}	10	ND	0
JF	3.52	10	ND	0	< 1 ng g^{-1}	10	ND	0
JR	4.58	11	ND	0	< 1 ng g^{-1}	11	ND	0
KC	NR	10	ND	0	< 1 ng g^{-1}	10	ND	0
LC	NR	10	ND	0	< 1 ng g^{-1}	10	ND	0
NBSD	6.32	10	ND	0	< 1 ng g^{-1}	6	ND	0
NBSR	1.26	10	> 1 ng g^{-1}	1	< 1 ng g^{-1}	11	ND	0
PC	NR	9	ND	0	< 1 ng g^{-1}	9	ND	0
PCHB	NR	12	ND	0	< 1 ng g^{-1}	5	< 1 ng g^{-1}	1
SRNB	6.31	10	ND	0	< 1 ng g^{-1}	6	ND	0
UBEC	1.48	12	ND	0	< 1 ng g^{-1}	3	ND	0
WB	3.62	11	> 1 ng g^{-1}	1	> 1 ng g^{-1}	11	> 1 ng g^{-1}	1
WBSL	2.46	10	ND	0	< 1 ng g^{-1}	7	ND	0
WC	NR	10	ND	0	< 1 ng g^{-1}	7	ND	0

4.3.2 Non-Targeted Analysis

After identifying known contaminants, samples were examined for unknowns, or non-target compounds, that may be contributing factors to the observed signs of disease. The crucial part of any non-targeted analysis using chromatographic separations, is the accurate alignment of all sample chromatograms. Before determining if features are specific to a sample or sample class, it needs to be clear whether each feature being compared is the same between different samples. This allows researchers to identify features who differ significantly between sample classes. Data reduction is also a critical element of non-targeted analyses, as large sample datasets can pose considerable challenges to identifying significant features.

For this analysis, the commercially available software Progenesis QI was used for chromatographic alignment and automatic peak finding. Alignment is based on matching the retention times of sample ion intensity maps to a reference map. The software compares every sample, or a user defined subgroup, for similarity and chooses the one most similar to all other analyses. Each sample's ion intensity map is then directly compared to the chosen reference and the retention times are adjusted to align the features. Pooled samples without QC were included during analysis and used for the selection of an alignment reference. The pooled samples included extract from every sample, thus providing the most similarity for accurate alignment across all samples. After alignment, each sample was manually reviewed for accuracy. Any sample failing to meet the 90% alignment match score was manually corrected.

Following alignment, automatic peak finding and deconvolution were performed on a single analysis of each sample (duplicate runs were excluded). Using the entire sample set helps ensure that all the features are identified for the dataset, but it also greatly increases the likelihood of false positives. Optimization of the peak finding method found that using a reduced data set decreased the number of double-identified, false positive features. Progenesis QI offers a few data reduction strategies to help identify significant features. The user can create experimental designs that classify samples into similar

groups. The focus of this analysis was to identify significant features between YOY SMB with disease characteristics and those without. The same Normal and Lesioned sample classes employed during GC×GC data analysis were used here. Based on gross observations of their health at the time of collection, YOY SMB with no obvious sign of infection or external abnormalities were classified as Normal, while all other YOY SMB were classified as Lesioned. As previously discussed, definitive diagnosis of the observed maladies was not possible, due to the whole fish being used for chemical analysis. However, the lesions or other conditions were able to be classified based on previous histological analyses of SMB and visual observations (Table 3.4).[12,16,18,38] Data was filtered using analysis of variance (ANOVA, $p < 0.01$) and fold change (> 2) to reduce the dataset and identify significant features between the sample classes. This reduced the ES+ dataset from 10571 to 2432 features and the ES- dataset from 8105 to 1685 features. The normalized abundances of these features were exported as .csv files and combined into a single dataset, then uploaded to MetaboAnalyst for further statistical analysis.

MetaboAnalyst is an open access, web-based program that specializes in metabolomics data analysis. Using the Statistical Analysis function, data were filtered by interquantile range (IQR), then scaled using log transformation and mean centering. Both the ES+ and ES- data were visualized using principal component analysis (PCA). There was no apparent clustering between Normal or Lesioned fish for either ionization mode (Figure 4.1). In previous analyses using GC×GC data, PCA was used as a feature reduction tool after Fisher ratio analysis. The Progenesis QI software uses ANOVA and fold change for data reduction instead of Fisher ratios. Fisher ratios were calculated for both datasets and only features whose ratio were above the F_{crit} value for $\alpha=0.05$ were retained, reducing the number of features to 348 for the ES+ data and 234 for the ES- data. These reduced datasets were also visualized by PCA, but there was no improvement in clustering or separation of the Normal and Lesioned sample classes (Figure 4.1). The Statistical Analysis function of MetaboAnalyst provides a suite of statistical tests for data analysis. In addition to PCA, data were visualized using volcano plots, partial least squares discriminant analysis (PLS-DA), and sparse PLS-DA (sPLS-DA). Volcano plots are scatterplots that plot

the statistical significance versus the fold change between groups. PLS-DA is a multivariate regression model similar to PCA, but it focuses on covariance as opposed to variance between inputs. PLS-DA still uses linear combinations to build models but includes the classification information associated with each feature. No significant features were highlighted by the volcano plot and both PLS-DAs methods resembled the PCAs in lack of clustering and separation (Figure 4.2)

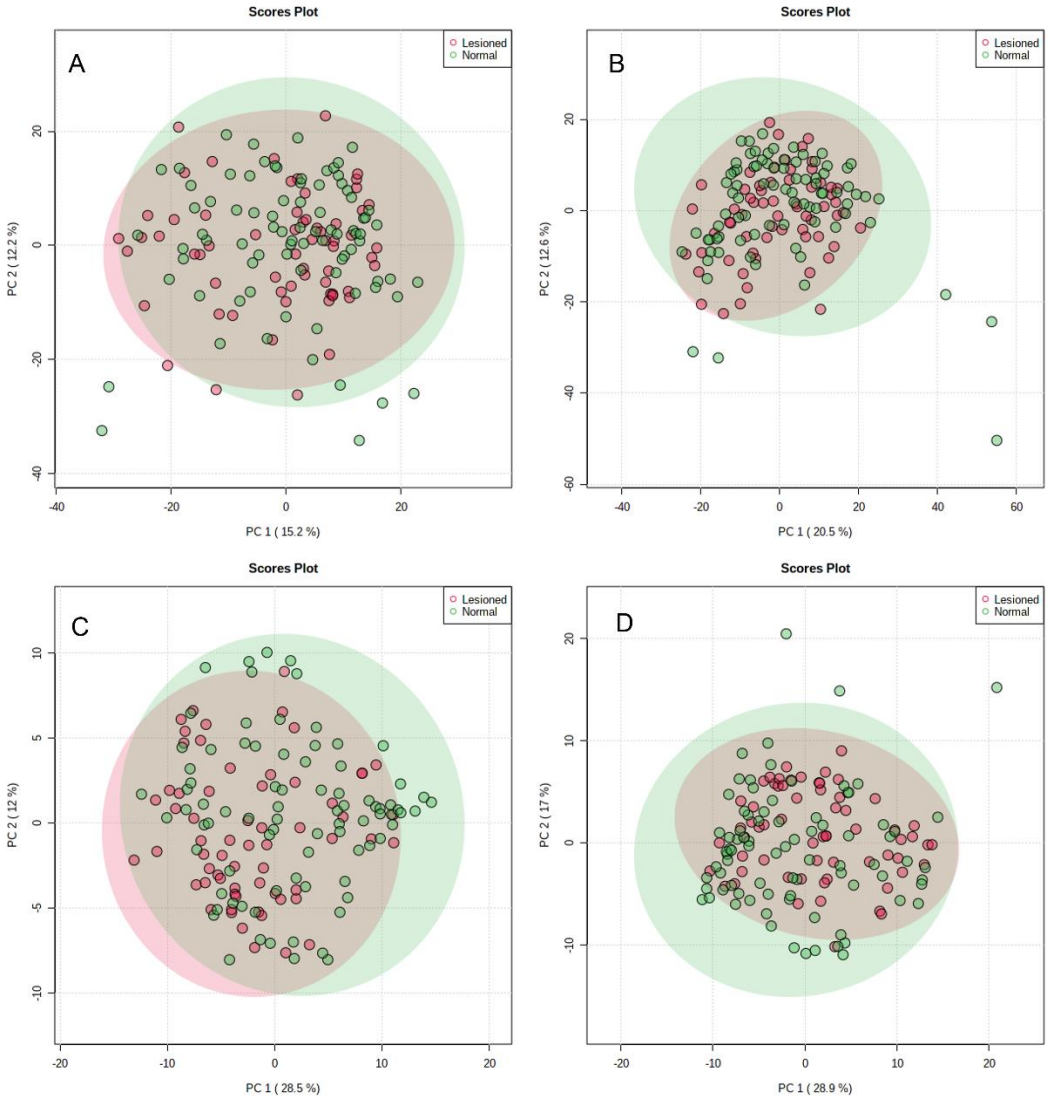

Figure 4.1 PCA scores plots of 146 YOY SMB. The top shows plots obtained by A)ES+ data and B)ES- data before Fisher ratio feature reduction. The bottom shows a repeated analysis of the reduced C)ES+ and D)ES- datasets. Ellipses are 95% confidence intervals. Lesioned class samples are represented in red. Normal class samples are represented in green.

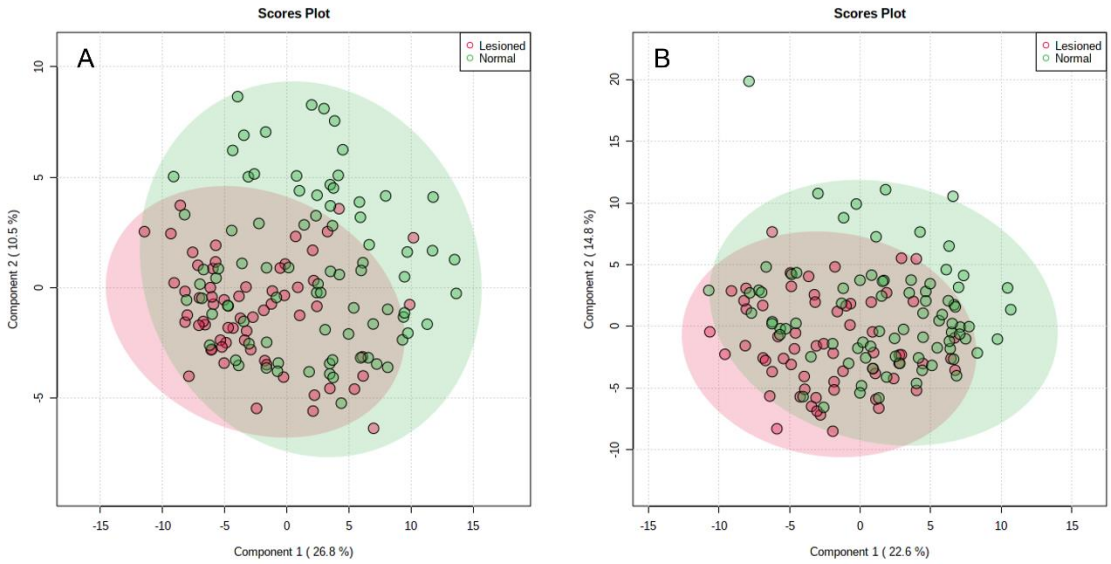

Figure 4.2 Scores plots for PLS-DA of 146 YOY SMB. PLS-DA was performed on the Fisher ratio reduced data set. A)ES+ and B)ES-. Ellipses are 95% confidence intervals. Lesioned class samples are represented in red. Normal class samples are represented in green.

As no features were determined to significantly contribute to differences between the sample classes Normal and Lesioned, samples were reclassified to try to identify if any features were significant to collection sites instead. A new experimental design was used with each sample classified by collection site. The same peak finding parameters were used. Data were filtered by ANOVA ($p < 0.01$) and fold change (> 2). After filtering, a total of 2931 features were exported for ES- and 6617 features for ES+ data. After export, Fisher ratio analysis was performed to further reduce the datasets to 2546 for ES- and 5604 for ES+ data. These datasets were uploaded to MetaboAnalyst and following filtering by IQR and scaling by log transformation and mean centering, data were visualized by ANOVA, PCA, PLS-DA, and sPLS-DA. Given that both data filtering within Progenesis QI and the Fisher ratio analysis had a minimum impact on feature reduction for these datasets, the lack of any clear clustering or detection of significant differences among these collection site-based datasets was not unexpected. Figure 4.3 shows the results of the PLS-DA of both the ES+ (A) and ES- (B) datasets. The PCA and sPLS-DA showed a similar lack of clustering or separation among sample classes.

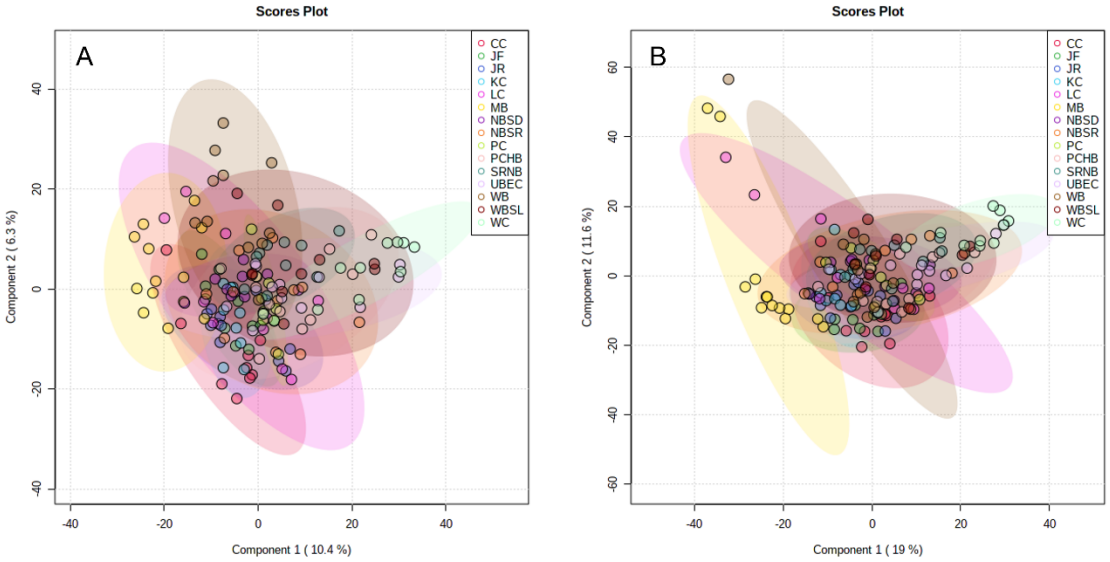

Figure 4.3 Scores plots of PLS-DA conducted on YOY SMB classified by collection site. A)ES+ dataset. B) ES- dataset. Ellipses are 95% confidence intervals.

Though it is still possible to calculate which features are responsible for the most variation among samples in the PCA and PLS-DAs presented here, due to the lack of separation among sample classes, it was determined that these features are not contributing to enough difference between the sample classes to warrant further investigation. Instead, the entire non-target datasets were searched for compound identifications to identify features of interest through suspect screening methods. Compounds were searched against the following ChemSpider databases: EPA Toxcast, FDA, Pesticide Common Names, enviPath, and Exposome Explorer. The precursor tolerance was set to 5 ppm and theoretical fragmentation was performed for each database search with a fragment tolerance of 5 ppm. Tentative identifications were accepted if a single match above a minimum match score of 40 could be determined. Steroids, pharmaceuticals, and pesticides were chosen as the focus of the non-targeted suspect screening (Table 4.4).

Table 4.4 Summary of tentatively identified non-target compounds detected in YOY SMB.

Ionization	ChemSpider ID	Tentative ID	Main Application	Formula	Score	Isotope Similarity	Mass	t_R (min)
Positive	CSID1064812	(+)-Nootkatone	Insect Repellant	C15H22O	40.3	96.20816	241.1571	8.338183
Positive	CSID111359	Icaridin	Insect Repellant	C12H23NO3	41	81.25555	268.1312	11.3478
Negative	CSID389877	(−)-Strychnine	Pesticide	C21H22N2O2	42.6	91.69328	355.1416	9.83865
Negative	CSID36552	(E)-Buthiobate	Pesticide	C21H28N2S2	42.4	88.79628	353.1509	3.588533
Negative	CSID13079489	Ametoctradin	Pesticide	C15H25N5	41	82.35748	549.4157	13.58523
Negative	CSID36543	Buminafos	Pesticide	C18H38NO3P	40.8	96.6927	368.2338	11.87897
Negative	CSID16737125	Cisanilide	Pesticide	C13H18N2O	41.3	87.28004	435.2744	11.82007
Negative	CSID171124	Dinotefuran	Pesticide	C7H14N4O3	40.2	94.63823	469.2549	11.0801
Negative	CSID6537	Diquat	Pesticide	C12H12N2+2	41.7	88.89301	367.189	9.756767
Negative	CSID2338984	Flucarbazone	Pesticide	C12H11F3N4O6S	40.1	86.34435	395.0281	2.646417
Negative	CSID391334	Flumetover	Pesticide	C19H20F3NO3	40	97.48406	733.2719	9.39865
Negative	CSID16738632	Formparanate	Pesticide	C12H17N3O2	42.5	92.40665	469.2549	11.0801
Negative	CSID21442049	Indaziflam	Pesticide	C16H20FN5	40.9	94.80347	601.3317	10.0252
Negative	CSID23078568	Isopyrazam	Pesticide	C20H23F2N3O	40.8	89.36964	380.1555	10.5912
Positive	CSID13684	Metolcarb	Pesticide	C9H11NO2	40.5	87.34688	188.0687	12.08335
Negative	CSID82795	Pencycuron	Pesticide	C19H21ClN2O	40.4	86.26521	349.1094	9.104
Negative	CSID8261320	Simeconazole	Pesticide	C14H20FN3OSi	41.2	93.69895	585.2632	10.53088
Positive	CSID10469196	Tralkoxydim	Pesticide	C20H27NO3	40.7	91.34461	241.1571	8.338183
Positive	CSID30142	Tridemorph	Pesticide	C19H39NO	42	98.86718	298.3094	11.8832
Positive	CSID54785	(-)-Arbutamine	Pharmaceutical	C18H23NO4	40.1	99.16207	300.1583	8.249333
Positive	CSID4447641	(−)-Nabilone	Pharmaceutical	C24H36O3	41.4	98.33629	355.2617	5.624167
Negative	CSID49179	(+)-Simvastatin	Pharmaceutical	C25H38O5	43.2	98.59494	399.2527	10.74073
Negative	CSID2038	Altretamine	Pharmaceutical	C9H18N6	43.1	94.79099	247.1073	2.325483
Positive	CSID3553	Batilol	Pharmaceutical	C21H44O3	41.1	97.19771	367.3174	11.9404
Negative	CSID2297864	Buprenorphine hydrochloride	Pharmaceutical	C29H41ClNO4	42.4	98.68844	247.1073	2.325483
Negative	CSID2866	Deferiprone	Pharmaceutical	C7H9NO2	40.5	96.54776	160.0381	2.339783
Negative	CSID8008478	Doramectin	Pharmaceutical	C50H74O14	41.8	84.47545	935.4606	9.104
Positive	CSID3129	Epirizole	Pharmaceutical	C11H14N4O2	40.6	98.5233	298.1275	12.3757

Ionization	ChemSpider ID	Tentative ID	Main Application	Formula	Score	Isotope Similarity	Mass	t_R (min)
Positive	CSID4445565	Epoprostenol sodium	Pharmaceutical	C20H31NaO5	43.4	96.48581	188.1115	10.12622
Positive	CSID3231	Fexofenadine	Pharmaceutical	C32H39NO4	42.4	98.94146	519.3221	12.95398
Negative	CSID6485	Methylprednisolone	Pharmaceutical	C22H30O5	46.4	95.89921	355.1924	5.844633
Positive	CSID96956	Perindopril	Pharmaceutical	C19H32N2O5	40.4	97.7325	369.2389	8.575383
Positive	CSID10439927	Salinazid	Pharmaceutical	C13H11N3O2	40.2	96.02634	298.3094	11.8832
Negative	CSID4968	Salmeterol	Pharmaceutical	C25H37NO4	40.6	97.03891	396.2531	9.90125
Positive	CSID5335	Trepibutone	Pharmaceutical	C16H22O6	40.8	95.95992	156.0782	7.324017
Positive	CSID5463	Vigabatrin	Pharmaceutical	C6H11NO2	40	98.34543	152.0688	1.875967
Negative	CSID72414	Diethyl dipropylmalonate	Pharmaceutical Intermediate	C13H24O4	40.6	98.2724	265.1411	8.140333
Positive	CSID4444059	(-)-Prostaglandin E2	Hormone	C20H32O5	43.3	97.4365	369.2389	8.575383
Positive	CSID4573597	Ecdysterone	Hormone	C27H44O7	44.7	68.16291	463.3049	11.3764
Positive	CSID5254715	Androstenone	Pheromone	C19H28O	40.2	97.33546	255.2097	10.1548
Positive	CSID62175	Oxabolone cipionate	Steroid	C26H38O4	41	96.13564	478.293	11.59083
Negative	CSID5667	Oxandrolone	Steroid	C19H30O3	43.8	99.44428	287.2001	9.615217
Positive	CSID14308	Pregnenolone 16alpha-carbonitrile	Steroid	C22H31NO2	41	96.52041	324.2312	9.093
Positive	CSID217360	Primobolan	Steroid	C27H42O3	43.3	97.59443	456.3461	12.44717
Positive	CSID83823	16-Dehydropregnenolone acetate	Steroid Intermediate	C23H32O3	40.3	95.17994	379.224	9.825083
Positive	CSID23383	17beta-Trenbolone	Steroid Metabolite	C18H22O2	40.3	97.49712	288.1945	11.71952

Pesticides were already a major focus of the targeted analysis for this study. Through suspect screening, several non-target pesticides were identified. Additionally, most of these non-target pesticides were not chlorinated, as was the focus of the targeted analysis. Several steroids, hormones, and pharmaceuticals were also identified. These are exogenous compounds that can enter the environment through runoff and wastewater treatment plants, and though previous studies have focused on identifying similar pharmaceuticals and hormones,[20,39,40] no studies have previously included the detected features on target lists. Therefore, to the author's knowledge, this is the first time any of these compounds have been reported in SMB tissues. This study illustrates the benefit of using non-targeted analyses to identify all possible contaminants instead of focusing only on suspected or known contaminants. Confirmation of these tentative identifications and monitoring of confirmed contaminants should be a priority in future studies, as this data represents a currently unexplored area of research into factors associated with disease in SMB.

4.4 Conclusions

Targeted analysis by UPLC-QTOF-MS allowed for the accurate detection and quantification of eight chlorinated pesticides and one PFAS compound. This was the first time any PFAS have been reported in YOY SMB. A previous study of adult SMB detected several PFAS in serum using HPLC-MS/MS, and PFAS have also been associated with immunosuppression in other species of fish. Therefore, more investigation into the presence of PFAS within YOY SMB should be conducted using an optimized methodology for their detection. It is possible this CEC could be an influencing factor on the disease characteristics observed in SMB. The datasets were also analyzed for non-target compounds. Though no statistically significant features were identified as contributing to a specific classification of

samples, several never before reported pesticides, steroids, hormones, and pharmaceuticals were tentatively identified. While these non-target features require confirmation of their identities, by matching with external analytical standards, they still represent a new direction for research efforts focusing on identifying environmental contaminants associated with disease in SMB.

4.5 Acknowledgements

Thank you to Geoff Smith and others in Pennsylvania Fish & Boat Commission as well as Tim Wertz and others in Pennsylvania Department of Environmental Protection for fish collections. Funding for this work was provided by Pennsylvania Sea Grant and by the U.S. Geological Survey's Environmental Health and Species Management Research (Fisheries) programs within the Ecosystems Mission Area. Any mention of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. Industrial support for this project was provided by Waters Corporation, Restek Corporation, and Wellington Laboratories.

4.6 References

1. D.P. Häder, A.T. Banaszak, V.E. Villafañe, M.A. Narvarte, R.A. González, E.W. Helbling, Anthropogenic pollution of aquatic ecosystems: Emerging problems with global implications, Sci. Total Environ. 713 (2020) 136586. doi:10.1016/j.scitotenv.2020.136586.
2. I.Y. López-Pacheco, A. Silva-Núñez, C. Salinas-Salazar, A. Arévalo-Gallegos, L.A. Lizarazo-Holguin, D. Barceló, H.M.N. Iqbal, R. Parra-Saldívar, Anthropogenic contaminants of high concern: Existence in water resources and their adverse effects, Sci. Total Environ. 690 (2019) 1068–1088. doi:10.1016/j.scitotenv.2019.07.052.
3. S. Ullah, M.J. Zorriehzahra, Ecotoxicology: A Review of Pesticides Induced Toxicity in Fish, Adv. Anim. Vet. Sci. 3 (2015) 40–57. doi:10.14737/journal.aavs/2015/3.1.40.57.

4. M. Dunier, A.K. Siwicki, Effects of pesticides and other organic pollutants in the aquatic environment on immunity of fish: a review, Fish Shellfish Immunol. 3 (1993) 423–438. doi:10.1006/fsim.1993.1042.
5. B. Lal, M.K. Sarang, P. Kumar, Malathion exposure induces the endocrine disruption and growth retardation in the catfish, Clarias batrachus (Linn.), Gen. Comp. Endocrinol. 181 (2013) 139–145. doi:10.1016/j.ygcen.2012.11.004.
6. K.E. Pelch, A. Reade, T.A.M. Wolffe, C.F. Kwiatkowski, PFAS health effects database: Protocol for a systematic evidence map, Environ. Int. 130 (2019) 104851. doi:10.1016/j.envint.2019.05.045.
7. S. Valsecchi, M. Babut, M. Mazzoni, S. Pascariello, C. Ferrario, B. De Felice, R. Bettinetti, B. Veyrand, P. Marchand, S. Polesello, Perfluoroalkyl Substances (PFAS) in Fish from European Lakes: Current Contamination Status, Sources, and Perspectives for Monitoring, Environ. Toxicol. Chem. (2020) etc.4815. doi:10.1002/etc.4815.
8. E. Zafeiraki, W.A. Gebbink, S.P.J. van Leeuwen, E. Dassenakis, P. Megalofonou, Occurrence and tissue distribution of perfluoroalkyl substances (PFASs) in sharks and rays from the eastern Mediterranean Sea, Environ. Pollut. 252 (2019) 379–387. doi:10.1016/j.envpol.2019.05.120.
9. V. Carignan, M.-A. Villard, Selecting indicator species to monitor ecological integrity: a review., Environ. Monit. Assess. 78 (2002) 45–61. doi:10.1023/a:1016136723584.
10. V.S. Blazer, D.D. Iwanowicz, H.L. Walsh, A.J. Sperry, L.R. Iwanowicz, D.A. Alvarez, R.A. Brightbill, G. Smith, W.T. Foreman, R. Manning, Reproductive health indicators of fishes from Pennsylvania watersheds: association with chemicals of emerging concern, Environ. Monit. Assess. 186 (2014) 6471–6491. doi:10.1007/s10661-014-3868-5.
11. J.J. Chaplin, J. Kent Crawford, R.A. Brightbill, Water-Quality Monitoring in Response to Young-of-the- Year Smallmouth Bass (Micropterus dolomieu) Mortality in the Susquehanna River and Major Tributaries, Pennsylvania: 2008, 2009. https://pubs.usgs.gov/of/2009/1216/ofr2009-1216.pdf.
12. H.L. Walsh, V.S. Blazer, G.D. Smith, M. Lookenbill, D.A. Alvarez, K.L. Smalling, Risk Factors Associated with Mortality of Age-0 Smallmouth Bass in the Susquehanna River Basin, Pennsylvania, J. Aquat. Anim. Health. 30 (2018) 65–80. doi:10.1002/aah.10009.
13. G.D. Smith, V.S. Blazer, H.L. Walsh, L.R. Iwanowicz, C. Starliper, A.J. Sperry, The effects of disease-related mortality of young-of-year smallmouth bass on the population characteristics in the Susquehanna River basin, Pennsylvania and potential implications to

conservation of black bass diversity., in: Am. Fish. Soc. Symp., 2015: pp. 82:319-332.

14. L.R. Iwanowicz, V.S. Blazer, A.E. Pinkney, C.P. Guy, A.M. Major, K. Munney, S. Mierzykowski, S. Lingenfelser, A. Secord, K. Patnode, T.J. Kubiak, C. Stern, C.M. Hahn, D.D. Iwanowicz, H.L. Walsh, A. Sperry, Evidence of estrogenic endocrine disruption in smallmouth and largemouth bass inhabiting Northeast U.S. national wildlife refuge waters: A reconnaissance study, Ecotoxicol. Environ. Saf. 124 (2016) 50–59. doi:10.1016/j.ecoenv.2015.09.035.
15. T. Boonthai, T.P. Loch, C.J. Yamashita, G.D. Smith, A.D. Winters, M. Kiupel, T.O. Brenden, M. Faisal, Laboratory investigation into the role of largemouth bass virus (Ranavirus, Iridoviridae) in smallmouth bass mortality events in Pennsylvania rivers, BMC Vet. Res. 14 (2018) 62. doi:10.1186/s12917-018-1371-x.
16. H.L. Walsh, V.S. Blazer, L.R. Iwanowicz, G. Smith, A Redescription of Myxobolus inornatus from Young-of-the-Year Smallmouth Bass (Micropterus dolomieu), J. Parasitol. 98 (2013) 1236–1242. doi:10.1645/ge-3081.1.
17. C.M. McClure, K.L. Smalling, V.S. Blazer, A.J. Sperry, M.K. Schall, D.W. Kolpin, P.J. Phillips, M.L. Hladik, T. Wagner, Spatiotemporal variation in occurrence and co-occurrence of pesticides, hormones, and other organic contaminants in rivers in the Chesapeake Bay Watershed, United States, Sci. Total Environ. 728 (2020) 138765. doi:10.1016/j.scitotenv.2020.138765.
18. C. Starliper, V.S. Blazer, L.R. Iwanowicz, H.L. Walsh, Microbial isolates in diseased fishes, primarily smallmouth bass (Micropterus dolomieu), within the Chesapeake Bay drainage in 2009-2011, in: Proc. West VIrginia Acad. Sci. Submitt. Pap. Eight-Eighth Annu. Sess. 2013, 2013.
19. S. Furihata, A. Kasai, K. Hidaka, M. Ikegami, H. Ohnishi, K. Goka, Ecological risks of insecticide contamination in water and sediment around off-farm irrigated rice paddy fields, Environ. Pollut. 251 (2019) 628–638. doi:10.1016/j.envpol.2019.05.009.
20. D.W. Kolpin, V.S. Blazer, J.L. Gray, M.J. Focazio, J.A. Young, D.A. Alvarez, L.R. Iwanowicz, W.T. Foreman, E.T. Furlong, G.K. Speiran, S.D. Zaugg, L.E. Hubbard, M.T. Meyer, M.W. Sandstrom, L.B. Barber, Chemical contaminants in water and sediment near fish nesting sites in the Potomac River basin: Determining potential exposures to smallmouth bass (Micropterus dolomieu), Sci. Total Environ. 443 (2013) 700–716. doi:10.1016/j.scitotenv.2012.09.063.
21. S. Sauvé, M. Desrosiers, A review of what is an emerging contaminant, Chem. Cent. J. 8

(2014) 15. doi:10.1186/1752-153X-8-15.

22. Ccanccapa-Cartagena, Y. Pico, X. Ortiz, E.J. Reiner, Suspect, non-target and target screening of emerging pollutants using data independent acquisition: Assessment of a Mediterranean River basin, Sci. Total Environ. 687 (2019) 355–368. doi:10.1016/j.scitotenv.2019.06.057.
23. S. Lee, K. Kim, J. Jeon, H.B. Moon, Optimization of suspect and non-target analytical methods using GC/TOF for prioritization of emerging contaminants in the Arctic environment, Ecotoxicol. Environ. Saf. 181 (2019) 11–17. doi:10.1016/j.ecoenv.2019.05.070.
24. Y. Lor, A. Revak, J. Weigand, E. Hicks, D.R. Howard, T.C. King-Heiden, Juvenile exposure to vinclozolin shifts sex ratios and impairs reproductive capacity of zebrafish, Reprod. Toxicol. 58 (2015) 111–118. doi:10.1016/j.reprotox.2015.09.003.
25. M. Bayley, M. Junge, E. Baatrup, Exposure of juvenile guppies to three antiandrogens causes demasculinization and a reduced sperm count in adult males, Aquat. Toxicol. 56 (2002) 227–239. doi:10.1016/S0166-445X(01)00210-7.
26. V.R.T. Zanotelli, S.C.F. Neuhauss, M.U. Ehrengruber, Long-term exposure to bis(2-ethylhexyl)phthalate (DEHP) inhibits growth of guppy fish (Poecilia reticulata), J. Appl. Toxicol. 30 (2010) 29–33. doi:10.1002/jat.1468.
27. P. Teehan, M.K. Schall, V.S. Blazer, B. Gruber, F.L. Dorman, Modified QuEChERS extraction for the analysis of young-of-year smallmouth bass using GC × GC-TOFMS, Anal. Methods. 12 (2020) 3697–3704. doi:10.1039/D0AY00880J.
28. Z. Pang, J. Chong, G. Zhou, D.A. de Lima Morais, L. Chang, M. Barrette, C. Gauthier, P.-É. Jacques, S. Li, J. Xia, MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights, Nucleic Acids Res. 49 (2021) W388–W396. doi:10.1093/NAR/GKAB382.
29. V.S. Blazer, S.E. Gordon, H.L. Walsh, C.R. Smith, Perfluoroalkyl Substances in Plasma of Smallmouth Bass from the Chesapeake Bay Watershed, Int. J. Environ. Res. Public Health. 18 (2021) 5881. doi:10.3390/ijerph18115881.
30. U. EPA, US EPA - Pesticides - Fact Sheet for Metolachlor, 1995.
31. U. EPA, Simazine Interim Registration Review Decision - Case Number 0070, 2020.
32. U. EPA, Propazine Interim Registration Review Decision Case 0230, 2020. www.regulations.gov (accessed August 16, 2021).
33. U. EPA, Atrazine: Interim Registration Review Decision Case: 0062, 2020. www.regulations.gov (accessed August 16, 2021).
34. U. EPA, Hexythiazox Summary Document: Registration Review, 2007. www.regulations.gov

(accessed August 16, 2021).

35. M. Hostovsky, J. Blahova, L. Plhalova, V. Kopriva, Z. Svobodova, Effects of the exposure offish to triazine herbicides, Neuroendocrinol. Lett. 35 (2014) 3–25. www.nel.edu (accessed August 16, 2021).
36. P.A. Fair, B. Wolf, N.D. White, S.A. Arnott, K. Kannan, R. Karthikraj, J.E. Vena, Perfluoroalkyl substances (PFASs) in edible fish species from Charleston Harbor and tributaries, South Carolina, United States: Exposure and risk assessment, Environ. Res. 171 (2019) 266–277. doi:10.1016/J.ENVRES.2019.01.021.
37. Fang, Q. Huang, T. Ye, Y. Chen, L. Liu, M. Kang, Y. Lin, H. Shen, S. Dong, Embryonic exposure to PFOS induces immunosuppression in the fish larvae of marine medaka, Ecotoxicol. Environ. Saf. 92 (2013) 104–111. doi:10.1016/j.ecoenv.2013.03.005.
38. M.K. Schall, V.S. Blazer, H.L. Walsh, G.D. Smith, T. Wertz, T. Wagner, Spatial and temporal variability of myxozoan parasite, Myxobolus inornatus, prevalence in young of the year smallmouth bass in the Susquehanna River Basin, Pennsylvania, J. Fish Dis. 41 (2018) 1689–1700. doi:10.1111/jfd.12878.
39. V.S. Blazer, L.R. Iwanowicz, H. Henderson, P.M. Mazik, J.A. Jenkins, D.A. Alvarez, J.A. Young, Reproductive endocrine disruption in smallmouth bass (Micropterus dolomieu) in the Potomac River basin: Spatial and temporal comparisons of biological effects, Environ. Monit. Assess. 184 (2012) 4309–4334. doi:10.1007/s10661-011-2266-5.
40. L.R. Iwanowicz, A.E. Pinkney, C.P. Guy, A.M. Major, K. Munney, V.S. Blazer, D.A. Alvarez, H.L. Walsh, A. Sperry, R. Braham, L.R. Sanders, D.R. Smith, Temporal evaluation of estrogenic endocrine disruption markers in smallmouth bass (Micropterus dolomieu) reveals seasonal variability in intersex, Sci. Total Environ. 646 (2019) 245–256. doi:10.1016/j.scitotenv.2018.07.167.

Chapter 5

Correlations Among Contaminants and Smallmouth Bass

5.1 Introduction

The goals when investigating signs of disease in an environmental setting is to remediate those factors responsible for the negative environmental effects and to prevent future recurrences. The first step is identifying the environmental factors or contaminants responsible for the observed health anomalies. In many cases, chemical contaminants are to blame when there are dramatic shifts in ecosystem health.[1–3] For remediation efforts to be successful, these contaminants need to be accurately identified within the environment, correlated with the observed health effects, and confirmed as the causative agent of disease. Otherwise, costly remediation efforts may be undertaken with no beneficial result.

For over 15 years, researchers have been working to identify factors associated with disease characteristics (immunosuppression, mortality, and intersex) in smallmouth bass (SMB) within the Susquehanna River Basin of the Chesapeake Bay Watershed.[4–9] An organism's health has a complex relationship with its surrounding environment. In the Susquehanna River Basin, chemical contaminants are suspected of contributing to the observed negative health effects in SMB. However, to date, no causative agent(s) has been definitively associated with the disease characteristics observed in SMB populations, though multiple studies have examined SMB tissues and environmental matrices (e.g., water and sediment) for target compounds,[4,10,11]

Most previous studies have focused on adult SMB or SMB nesting sites and have not explored potential exposures to young-of-year (YOY) SMB. Exposures to chemical contaminants during key developmental life stages have been shown to induce lifelong health impairments in

many species of fish.[12–14] Yet, only a single study has analyzed YOY SMB tissues for chemical contaminants.[6] Despite this study collecting samples for multiple years from over 20 sites within the Pennsylvania portion of the Susquehanna River Basin, only 16 individual YOY SMB from two collection sites were analyzed for chemical contaminants. The small sample set analyzed only provided a limited view of the contaminant exposures encountered by YOY SMB and gave little understanding of the health impacts these exposures may cause.

The work presented here has evaluated 146 YOY SMB, collected from 14 different sites, for target and non-target compounds of interest using both GC×GC-TOFMS and UPLC-QTOF-MS. Through these evaluations many environmental contaminants have been identified directly within fish tissues, with several contaminants being reported for the first time. This data has been further analyzed to determine if any correlations between disease characteristics and the detected target compounds could be made. Additionally, the need for an investigation into non-target compounds significant to collection sites within the GC×GC data was previously discussed in Chapter 3. This investigation has been performed and the results were compared with the non-targeted analysis from Chapter 3 that focused on identifying features that were significant between the Lesioned and Normal sample classes.

5.2 Materials and Methods

Chemical analysis data presented in this chapter was previously collected using the described materials and methods of Chapters 2 – 4. Statistical correlations were calculated in Microsoft Excel (Excel v.2107) or R v3.5.3 (R Foundation for Statistical Computing, Vienna, Austria) in the RStudio platform (Boston, MA, USA). No additional R packages were required.

A non-targeted analysis focused on site-specific features was performed on the GC×GC-TOFMS data presented in Chapter 3. The same data processing parameters established in section

3.3.5 for chromatographic alignment by Statistical Compare (ChromaTOF GC v4.74.2.0, LECO Corp., St. Joseph, MI, USA) were used for this analysis; only the user defined sample classes were modified. Feature reduction was performed by Fisher ratio analysis. Significant features were further analyzed using the Statistical Analysis feature of MetaboAnalyst v5.0.[7]

5.3 Results and Discussion

5.3.1 Associations with Disease Status

As one of the main priorities of this study was to identify environmental contaminants associated with disease characteristics in SMB, the sample classifications were designed on this basis. At the time of collection, gross observations were recorded for each YOY SMB. Samples were then classified based on these gross observations of their health as either Normal or Lesioned. Normal class fish had no observable external disease characteristics, while Lesioned class fish were observed to have visible lesions, gill/fin erosion, an attached parasite, or other abnormalities. For the chemical analyses, the whole fish was homogenized preventing the observed health anomalies from being definitively diagnosed. However, as previous histological data is available, it was possible to distinguish infections by the parasite *Myxobolus inornatus* from other dermal lesions typically caused by either bacterial or viral pathogens.[4,8]

To date, previous targeted analyses have been unable to definitively associate any identified target chemicals with the observed signs of disease. Therefore, the target compounds detected by GC×GC-TOFMS and UPLC-QTOF-MS in this study were examined for correlations with disease characteristics. Point-biserial correlation was used to determine correlations between the disease state of fish (Normal or Lesioned) and the detected target compounds. This correlation test allows for comparison of binary variables with continuous variables without needing further

data manipulation. Disease state was compared with both the frequency of target compound detection and the total body burden (ng g^{-1}) of all target compounds. Correlation coefficients (r_{pb}) of -0.03 and -0.01 were calculated for frequency and body burden, respectively, indicating no correlation between these variables. For point-biserial correlations, the correlation coefficient results in a value between -1 and 1, where 0 is considered to have no correlation, -1 is a perfect negative correlation, and 1 is a perfect positive correlation. On this scale, values < 0.3 can be considered slight correlations, 0.3 – 0.7 moderate correlations, and > 0.7 strong correlations.

Due to the lack of correlation between observed disease characteristics and the totality of detected target compounds, possible correlations between disease status and the body burden of specific chemical classes were investigated. Targets were grouped by chemical family or common use as follows: polychlorinated biphenyls (PCBs), polybrominated diphenyl ethers (BDEs), polycyclic aromatic hydrocarbons (PAHs), perfluoroalkyl substances (PFAS), pesticides, plasticizers, and pharmaceutical and personal care products (PPCPs). Most of the chemical classes showed no correlation with disease state (Table 5.1). Slight negative correlations were observed in the PAH, PFAS, and Pesticide groups (r_{pb} = -0.16, -0.11, and -0.10, respectively). However, none of these correlations were determined to be statistically significant. Although these slight correlations were not significant, their negative association was the opposite of any expected relationship between contaminants and disease state. A negative correlation indicates that increasing levels of these contaminants would occur with decreasing levels of disease characteristics. Even if chemical contaminants are not directly responsible for the observed signs of disease, it is unlikely that increasing the amount of environmental contamination would improve the health of SMB.

Table 5.1 Summary of point-biserial correlations between disease state and chemical classes.

Chemical Class	*PCB*	*BDE*	*PAH*	*PFAS*	*PPCP*	*Pesticide*	*Plasticizer*
Lesioned Group Mean (ng g^{-1})	3.10	2.50	3.07	5.20	62.99	5.64	22.40
Normal Group Mean (ng g^{-1})	2.56	1.94	1.73	2.75	93.87	15.60	23.75
r_{pb}	0.01	0.05	-0.16	-0.11	-0.02	-0.10	0.02
p-value	0.925	0.582	0.061	0.206	0.834	0.246	0.769

Throughout this study, all fish with external abnormalities were classified as Lesioned, regardless of the specified type of abnormality. To better understand the relationship between contaminants and different disease characteristics, the Lesioned group samples were subdivided into two groups. Group 1, Dermal Lesions, consisted of fish observed to have any of the following: *Mxyobolus* infection, pale areas, or dermal lesions. Group 2, Other, consisted of all remaining non-lesion-based health abnormalities. These were compared with the previously discussed chemical classes (Table 5.2). Overall, slight negative correlations were observed between lesion type and several chemical classes (PCBs, BDEs, PAHs, and plasticizers). However, none of these correlations were statistically significant. PPCPs showed a slight positive correlation with lesion type but was also not significant. Two chemical groups were found to have statistically significant correlations with the subdivided Lesioned groups. There was a slight negative correlation with PFAS (r_{pb} = -0.25) and a moderately positive correlation with pesticides (r_{pb} = 0.40). This would indicate that increasing levels of PFAS are more likely to occur with fish classified as exhibiting dermal lesions, while increasing levels of pesticides are more likely to occur with fish exhibiting other health abnormalities (Figure 5.1).

Table 5.2 Summary of point-biserial correlations between lesion type and chemical classes.

Chemical Class	*PCB*	*BDE*	*PAH*	*PFAS*	*PPCP*	*Pesticide*	*Plasticizer*
Dermal Lesion Mean (ng g^{-1})	4.09	3.34	3.03	7.25	87.53	1.54	26.70
Other Mean (ng g^{-1})	1.72	2.23	1.73	3.12	141.17	21.00	17.06
r_{pb}	-0.15	-0.15	-0.19	-0.25	0.10	0.40	-0.09
p-value	0.246	0.231	0.131	0.049	0.429	0.001	0.502

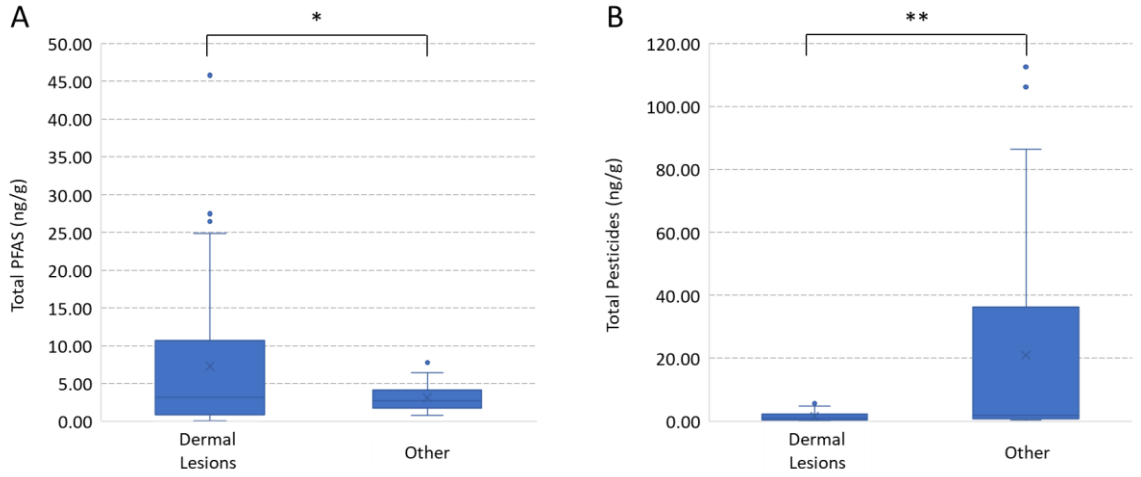

*Figure 5.1 Box and whisker plots representing point-biserial correlations between Lesioned class groups and selected chemical classes. A) Correlation of PFAS with Lesioned class groups shows a slight negative correlation (*p < 0.05). Fish with higher levels of PFAS are more likely to exhibit dermal lesions. B) Correlation of Pesticides with Lesioned class groups shows a moderate positive correlation (**p < 0.005). Fish with higher levels of pesticides are more likely to exhibit other negative health anomalies.*

The relationship between PFAS, pesticides, and lesion type was further investigated. The Lesioned sample class was subdivided by the specific health observations recorded in Table 3.4. The groups were defined as myxo, dermal, leech, or other. Myxo included all samples with a *Myxobolus* infection, the most dominant observed disease feature. Leech was any fish with a leech attached at the time of collection. Dermal was a combination of the dermal lesions and pale areas groups, as these are both dermal abnormalities typically caused by bacterial or viral infection. Other consisted of fish with frayed fins or other abnormalities, not specified, at the time of collection. These classes were analyzed by ANOVA (analysis of variance) against both the target PFAS and pesticide body burdens. ANOVA can be used to determine if the means of individual groups significantly differ from one another. Both ANOVA tests resulted in no significant difference between group means and therefore, no additional information was gained about these relationships.

As previously detailed, many of these compounds have been associated with negative health effects, such as immunosuppression, endocrine disruption, morphological changes, and altered behavioral patterns in different species of fish.[9–12] However, little research has been done on the effects in SMB specifically. Therefore, even if correlations between specific compounds and disease characteristics could be established, it would still be impossible to determine a cause-and-effect relationship at this time. This study could not confirm any strong correlations between the target chemicals and the observed signs of disease in YOY SMB, but weaker correlations were found between PFAS and pesticides with broadly defined lesion or non-lesion type health impairments. However, additional investigation into correlations between the specific chemical classes and lesion types could not be confirmed.

The relationship between fish size (g) and other factors was also examined. The size of YOY SMB is largely dependent on age. Fish in this study were collected between June 24, 2015 to July 29, 2015, but exact spawn and hatch dates for each sampling site are unknown, so age can only be estimated. The YOY SMB in this study are estimated to be between 1 – 2 months. A statistically significant, slight positive correlation was observed between fish size and disease state ($r_{pb} = 0.24$, $p < 0.005$), indicating that larger fish were slightly more likely to be classified as part of the Lesioned sample class (Figure 5.2). Pearson's *r* correlation was used to examine the relationship between fish size and total body burden of target chemicals, but no correlations were found (Figure 5.3).

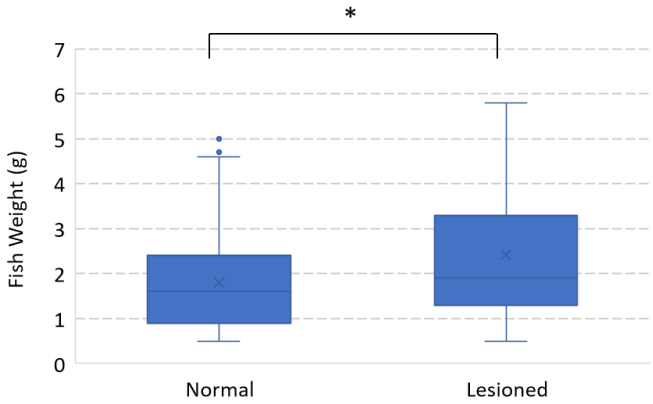

*Figure 5.2 Box and whisker plot representing point-biserial correlation of fish size by weight (g) versus disease status classification (Normal or Lesioned). A slight positive correlation was observed (*p<0.005) indicating that larger fish were more likely to be classified as Lesioned based on gross observations at the time of collection.*

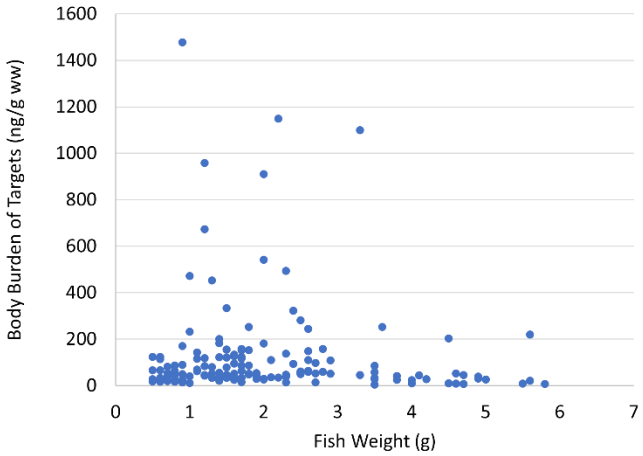

Figure 5.3 Scatterplot of fish size by weight (g) vs total body burden of all target compounds (ng g-1 wet weight). No correlation was found between fish size and total body burden of target compounds.

While no correlations were detected between the size of the YOY SMB and the total amount of target compounds detected, the slight correlation between size and disease state is an interesting observation. As the fish in this study were not allowed to mature further and no additional samples of older YOY SMB were taken in 2015 for comparison, it is unknown whether the smaller fish would have exhibited more signs of disease as they matured. YOY SMB are normally sampled in June and July of each study year. If samples were collected from the same

populations at multiple time points and analyzed for the same factors, it may be possible to elucidate more about the relationship between contaminants, size, and disease characteristics. Future studies should consider collecting fewer samples per site initially, and instead resample each site multiple times throughout the year. This would provide more information about changes within the population's disease status, and specifically, the changes in observable disease characteristics present on YOY SMB while maturing.

The investigation into correlations between the disease state of the YOY SMB and the detected target chemicals did not reveal any strong connections. However, this was not wholly unexpected as numerous targeted analyses with adult SMB and their surroundings have been conducted with the same lack of result. There were some limitations to the correlation studies presented here that should be considered. As discussed previously, only a single targeted chemical analysis of YOY SMB can be found in the literature predating this study, and the chemical analysis conducted was limited to only a fraction of the total samples collected.[3] Therefore, correlation analyses were limited to data collected in this study, and subsequently, all the chemical data presented comes from a single collection time point. Additional data from multiple collection time points, either within the same year or the same period across consecutive years would add a wealth of information to these correlations and allow for the building of more robust statistical models to examine trends between the target chemicals and disease characteristics. As this work has presented an improved sample preparation methodology for high-throughput analysis, future investigations into chemical contaminants should focus on incorporating multiple sampling time points along with sampling locations into the study design.

5.3.2 Reexamination of Non-Targeted GC×GC Data

Though the data collected by UPLC-QTOF-MS was previously analyzed for significant features specific to collection sites, this comparison was not originally performed for the GC×GC-TOFMS data. The GC×GC data were reprocessed using a new experimental setup designed to highlight differences between collection sites. All the alignment parameters optimized for previous non-targeted analyses were kept, including the manual corrections for first-dimension retention time shifts, but the sample classes were redefined. For retainment within the alignment peak table, features still needed to present in either 66% of their new sample class or within 66% of all samples. This helps to eliminate false positives, as it is unlikely that a feature detected in only a few or a single sample will be significantly contributing to differences between sample classes. Background artifacts, such as column bleed, were manually removed. Fisher ratios were calculated for each retained feature, and the normalized peak area of those above the F_{crit} value for $\alpha = 0.05$ were exported as .csv files for further analysis. This reduced the dataset from 2078 analytes retained after alignment to 529 analytes.

The previous non-targeted GC×GC analysis used unsupervised principal component analysis (PCA) to check for possible batch effects and supervised PCAs for analyte reduction. A supervised PCA was performed on the Fisher ratio reduced dataset of 529 analytes, using the Statistical Analysis feature of MetaboAnalyst, a web-based software platform specializing in metabolomics data analysis (Figure 5.4). Clustering was observed among several of the sample classes, but no class was completely separated from another. It should be noted that increasing the number of sample classes led to a large increase in the overall number of analytes retained. The previous non-targeted GC×GC analysis using disease state as the defining feature resulted in 512 analytes being retained after alignment and only 95 analytes remaining after reduction by Fisher ratio analysis. This increase is due to the ability of more analytes to meet the 66% threshold per

sample class, as each class has a small total number of samples. Therefore, analytes need only be present within only 6 – 8 samples per class for retainment, as opposed to either 24 or 55 samples in the disease state design. This large increase in retained analytes prompted the use of additional data reduction techniques.

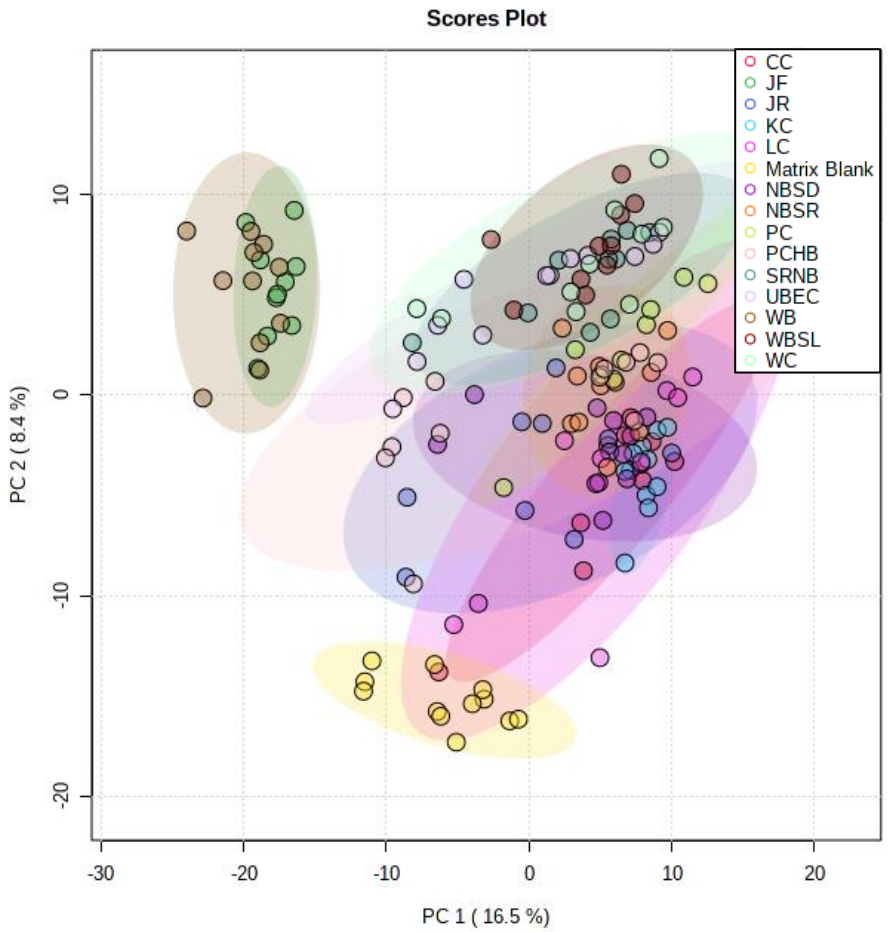

Figure 5.4 PCA scores plot of GC×GC-TOFMS non-targeted data grouped by collection site after analyte reduction by Fisher ratio analysis. Ellipses are 95% confidence intervals for each sample group.

Using MetaboAnalyst, a partial least squares discriminant analysis (PLS-DA) was performed (Figure 5.5). Similar to PCA, PLS-DA is a multivariate dimensionality-reduction tool that uses linear combinations. Where PCA focuses on maximizing the variance between samples,

PLS-DA focuses on maximizing covariance between samples and sample classifications.[13] Analysis by PLS-DA showed better clustering and separation for nearly every sample class. Variable importance in projection (VIP) scores were generated for each analyte of the PLS-DA model and those analytes with VIP scores > 2 were retained as the final reduced non-target dataset, resulting in a total of 38 analytes.

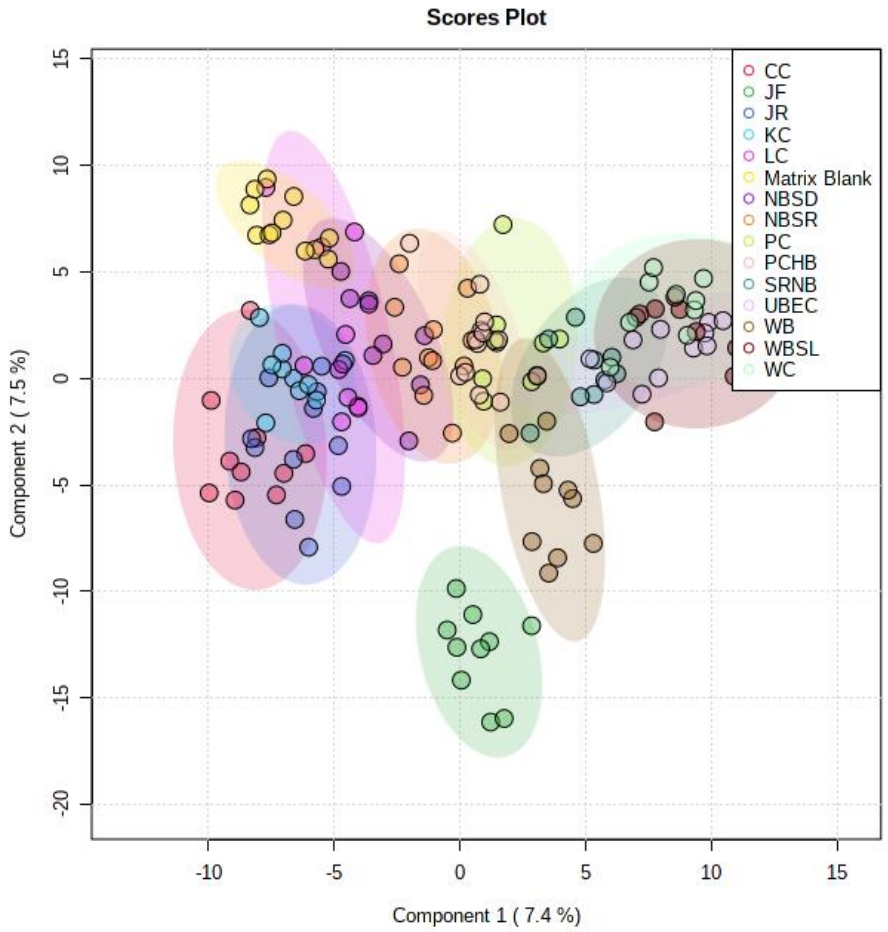

Figure 5.5 Scores plot of PLS-DA of GC×GC-TOFMS non-targeted data grouped by collection site after analyte reduction by Fisher ratio analysis. Ellipses are 95% confidence intervals for each sample group.

Analysis by PLS-DA has become commonly used in chemometrics and metabolomics studies due to wider availability in omics software platforms. However, this technique suffers from the possibility of creating overfitted classification models. In addition to the internal

standard normalization performed on the raw signal intensity data during data processing in ChromaTOF GC, the data set was reduced using Fisher ratio analysis and normalized by mean centering in MetaboAnalyst. These normalization measures were intended to help prevent interference due to background noise and reduce the likelihood of false positives among clusters. A further investigation into the clustering of data by collection sites was conducted using k-means clustering, an unsupervised machine learning technique.

K-means clustering uses an iterative algorithm to determine a clustering pattern that produces the lowest within-cluster variance across all groups. In k-means clustering, the desired number of clusters is determined by the user. For this dataset, a total of 15 clusters was desired (one per collection site and matrix blanks) to match the number of groups used in the PCA and PLS-DA. Figure 5.6 and Table 5.3 show the results of the site-specific non-targeted GC×GC-TOFMS data analyzed by k-means clustering in MetaboAnalyst. In k-means clustering each sample is assigned to a single cluster. Not all sample groups formed unique clusters, but samples were generally clustered by collection site. While the k-means clustering pattern was different from the clustering pattern produced by PLS-DA it does support the PLS-DA model. Each analysis uses different statistical methodologies to examine the pattern among samples, with k-means clustering being driven by reducing the variance within-clusters and PLS-DA being driven by the covariance of samples. Here, both analyses produced clustering patterns that were largely determined by sample collection site, indicating that the site-specific non-targeted analysis of the data did identify the significant analytes present within the samples.

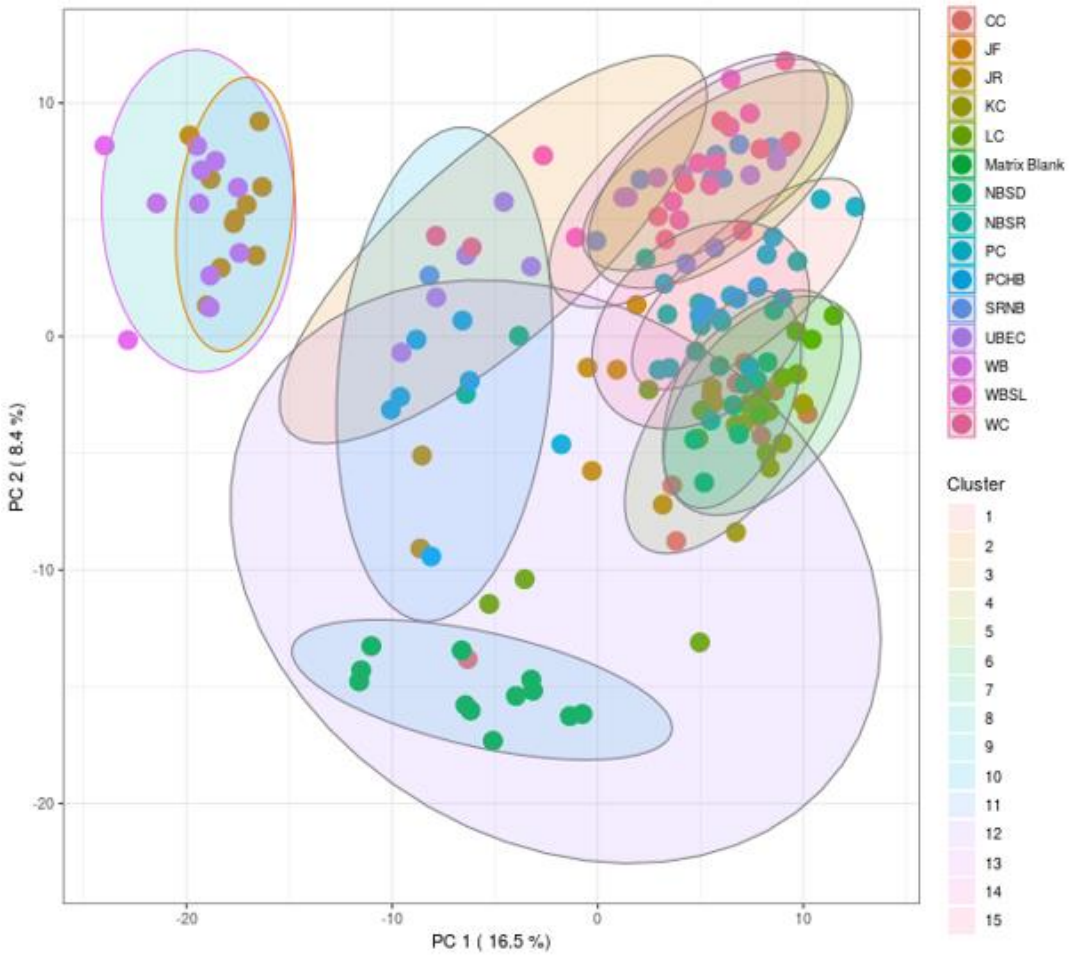

Figure 5.6 K-means clustering plot of non-targeted GC×GC-TOFMS data. Each sample point is colored based on the collection site it was taken from. The ellipses represent the cluster that each sample belongs to. Overlap of the clusters in the two-dimensional representation does not indicate that a sample belongs to more than one cluster.

Table 5.3 Results of k-means clustering analysis. Each sample is assigned to a single cluster.

Cluster #	Sample						
1	PC 001	PC 002	PC 003	PC 004	PC 005	PC 007	PC 008
	PC 009	PCHB 001	PCHB 002	PCHB 004	PCHB 005	PCHB 006	
2	UBEC 007	UBEC 008	UBEC 009	UBEC 010	UBEC 011	UBEC 012	WBSL 001
3	UBEC 001	UBEC 003	UBEC 005	WBSL 002	WBSL 003	WBSL 004	WBSL 005
	WBSL 006	WBSL 007	WBSL 008	WBSL 010			
4	SRNB 002	SRNB 003	SRNB 004	SRNB 005	SRNB 006	SRNB 007	SRNB 008
	SRNB 009	SRNB 010	WC 002	WC 007			
5	CC 002	CC 003	CC 004	CC 005	CC 006	CC 007	CC 008
	CC 009	CC 010	JR 001	JR 002	JR 003		
6	JR 010	JR 011	KC 001	KC 002	KC 003	KC 004	KC 005
	KC 006	KC 007	KC 008	KC 009	KC 010	LC 001	LC 002
	LC 003	LC 004	LC 005	LC 006	LC 007		
7	NBSD 001	NBSD 002	NBSD 003	NBSD 004	NBSD 005	NBSD 006	NBSD 007
	NBSD 008	NBSR 001					
8	WB-01	WB-02	WB-03	WB-04	WB-05	WB-06	WB-07
	WB-08	WB-09	WB-10	WB-11			
9	CC 001	MB CC	MB JR	MB KC	MB LC	MB NBSD	MB NBSR
	MB PC	MB PCHB	MB SRNB	MB UBEC	MB WBSL	MB WC	
10	JR 007	JR 008	JR 009	NBSD 009	NBSD 010	PCHB 007	PCHB 008
	PCHB 009	PCHB 010	PCHB 011	PCHB 012	SRNB 001	WC 001	WC 003
11	JF-01	JF-02	JF-03	JF-04	JF-05	JF-06	JF-07
	JF-08	JF-09	JF-10				
12	LC 008	LC 009	LC 010	PC 006			
13	UBEC 002	UBEC 004	UBEC 006	WBSL 009	WC 004	WC 005	WC 006
	WC 008	WC 009	WC 010				
14	JR 004	JR 005	JR 006				
15	NBSR 002	NBSR 003	NBSR 004	NBSR 005	NBSR 006	NBSR 007	NBSR 008
	NBSR 009	NBSR 010	NBSR 011	PCHB 003			

Tentative identifications of this final dataset were obtained by mass spectral comparison with the NIST MS Database (2014 and 2017). A minimum match similarity score of 700 was required for name assignment, and 35 of the 38 compounds were tentatively identified (Table 5.4). Four compounds from this non-targeted analysis had matching tentative identifications, retention times, and masses with compounds from the previous GC×GC non-targeted analysis.

These compounds include pentacosane; benzenamine,4,4’-ethylidenebis[N,N-diethyl-3-methyl-; 9-octadecenamide,(Z)-; and isobutyl propane-1,3-diyl dicarbonate. As these four compounds have been identified as statistically significant when samples are classified by both disease state and collection site, they warrant further investigation into their role in SMB health. Confirmation of these tentative identifications should be a priority in future studies.

In addition to having several compounds in common, the chemical families detected in each analysis were also similar. Alcohols, aldehydes, esters, hydrocarbons, ketones, and nitrogen-containing compounds were detected in both analyses, with esters and hydrocarbons accounting for the largest percent abundance of each sample class. As esters, ketones, and nitrogen-containing compounds were found to be more abundant in Lesioned class fish, these groups were examined for possible correlations between their total abundance at each collection site and the total number of Lesioned class fish from each site. Spearman’s rank correlation, ρ, was used for the analysis. This examines the monotonic relationship between the two variables, as opposed to a linear relationship such as when using Pearson’s correlation coefficient. Moderate positive correlations were found between two chemical groups, esters ($\rho = 0.50$) and nitrogen-containing compounds ($\rho = 0.43$), and the number of Lesioned class fish per site, but these correlations were not statistically significant ($p = 0.066$ and $p = 0.121$, respectively). Ketones showed a slight negative correlation ($\rho = -0.28$) with the Lesioned class fish but was also not statistically significant ($p = 0.335$).

Table 5.4 Tentative identifications of final reduced dataset for non-targeted analysis with samples classified by collection site. First- and second-dimension retention times (t_R) are averaged value across all runs. Spectral match values are from randomly selected sample containing analyte of interest.

Tentative ID	Chemical Family	Fisher Ratio	t_{R1} (sec)	t_{R2} (sec)	Mass	Similarity	Reverse	Probability
4,4,6-Trimethyl-cyclohex-2-en-1-ol	alcohol	3.3051	1252.8	1.43783	84	737	760	2313
trans-2-Nonenal	aldehyde	7.03	951.61	1.10695	55	849	865	1862
9,12,15-Octadecatrienal	aldehyde	28.36	1584.32	1.24592	79	794	813	1027
Isobutyl propane-1,3-diyl dicarbonate	carbonate ester	7.2418	2394.93	1.56807	103	723	746	2329
4,7,10,13,16,19-Docosahexaenoic acid, methyl ester, (all-Z)-	ester	2.0654	2612.8	1.87517	79	945	946	4646
Methyl 8-(2-furyl)octanoate	ester	21.734	1710.81	1.15399	81	721	756	7970
Orthoformic acid, tri-sec-butyl ester	ester	25.757	2702.67	1.88761	55	703	743	2595
Pentacosane	hydrocarbon	2.9157	2917.32	1.76035	57	924	930	2501
Hentriacontane	hydrocarbon	3.9199	3280.19	3.44727	57	833	886	1460
Pentadecane, 2,6,10-trimethyl-	hydrocarbon	3.9907	1699.96	0.998443	57	886	905	1813
1,7-Dimethyl-4-(1-methylethyl)cyclodecane	hydrocarbon	6.1331	1280.62	0.981117	69	729	782	592
Octadecane	hydrocarbon	9.2902	1903.5	1.06031	57	913	926	3018
Hexadecane	hydrocarbon	12.026	1728.47	0.988349	57	867	879	1216
Octane, 6-ethyl-2-methyl-	hydrocarbon	12.381	883.3	0.834228	71	845	867	991
Benzene, octyl-	hydrocarbon	13.877	1454.3	1.15476	91	726	834	5782
7-Pentadecyne	hydrocarbon	14.086	1496.43	1.08277	67	722	745	899
8-Hexadecyne	hydrocarbon	17.978	1532	1.08939	67	742	748	750
Octacosane	hydrocarbon	18.492	1976.33	1.07557	71	888	897	1115
Decane, 2,4,6-trimethyl-	hydrocarbon	19.549	924	0.836577	57	877	887	2633
Benzene, (1-butylheptyl)-	hydrocarbon	21.049	1707.66	1.19827	91	826	862	8710
Eicosane	hydrocarbon	22.231	2154	1.33001	57	856	862	2580
4-Undecene, 4-methyl-	hydrocarbon	24.296	1039.8	0.924004	55	746	799	480
2-Hexene, 3,5-dimethyl-	hydrocarbon	24.888	1096	0.945444	70	730	819	3112
cis-1,1,3,5-Tetramethylcyclohexane	hydrocarbon	27.893	1224.48	0.988333	125	739	758	5199
2,6,10-Trimethyltridecane	hydrocarbon	35.588	1135.43	0.954478	57	857	872	800

Tentative ID	Chemical Family	Fisher Ratio	t_{R1} (sec)	t_{R2} (sec)	Mass	Similarity	Reverse	Probability
Benzene, 1,2,4-trimethyl-	hydrocarbon	36.79	682.128	0.975817	105	710	767	2366
Ethanone, 1-(4-methylphenyl)-	ketone	3.5961	1005.68	1.34247	119	923	949	3300
1-Propanone, 1-phenyl-	ketone	15.337	971.628	1.29469	105	862	865	2386
2(1H)-Naphthalenone, 3,4,4a,5,6,7,8,8aà-octahydro-5à-hydroxy-4aà,7,7-trimethyl-, acetate	ketone	19.863	1420.9	1.17652	177	765	778	1059
δ-Dodecalactone	ketone	40.146	1792	1.39596	99	913	939	7310
Benzocaine, N-(2-methylpropyl)-	nitrogen	4.2447	2048.81	1.98319	178	836	860	5306
benzenamine, 4,4'-ethylidenebis[N,N-diethyl-3-methyl-	nitrogen	5.3858	2724.1	2.08947	337	721	790	7615
Dodecanamide	nitrogen	19.031	2553.64	1.84704	59	813	879	4604
9-Octadecenamide, (Z)-	nitrogen	27.785	2531.76	1.89163	59	788	788	7276
à-Tocospiro A	unclassified	14.046	2977.1	2.62792	419	782	863	4848
Unknown 2	unclassified	9.1276	2293.39	1.11569	159	NA	NA	NA
Unknown 3	unclassified	26.788	2660.13	2.07447	339	NA	NA	NA
Unknown 1	unclassified	27.147	1561.58	1.19471	121	NA	NA	NA

Although statistically significant correlations of analytes between sites were not found, it does not detract from the value gained by the non-targeted analyses. This study does not focus on the hydrology of the Susquehanna River Basin or examine any environmental matrices. Disease classification was based solely on the gross observations of YOY SMB at the time of collection, and despite these sites being chosen for study due to their history of exhibiting the same disease characteristics for several years, the variation of disease characteristics and rates between collection sites could still be considered large. It is hard to make associations between disease prevalence when using a single time point. This work provides a basis with which to inform future non-targeted investigations of these areas and has tentatively identified previously unknown compounds within the YOY SMB.

5.4 Conclusions

A variety of possible correlations were explored between detected target compounds, non-target compounds, collection sites, fish size, and fish disease state. Although no strong correlations were found, slight correlations between chemical groups and a general subdivision of health anomalies were statistically significant. Individual disease characteristics (e.g., dermal lesions, *Myxobolus* infection, frayed fins, etc.) could not be correlated with any single chemical class. There was a statistically significant, slight positive correlation between the size of the YOY SMB and their classification as Lesioned due to gross observations at the time of collection. This connection along with the lack of additional target chemical analysis data available for YOY SMB, highlights the need for multiple collection time points in future studies.

The reevaluation of the GC×GC data for site-specific significant non-targets resulted in an additional 38 features of interest being identified. Tentative compound identifications were applied for 35 of the 38 features. Of particular importance from this analysis, were the four

compounds found in common with the previous disease state specific non-targeted analysis. As these compounds have now been identified as statistically significant in two different study designs, they should be further investigated. This includes confirming their tentative identifications and determining their source of origin, either externally or internally to SMB.

5.5 References

1. K.A. Hicks, M.L.M. Fuzzen, E.K. McCann, M.J. Arlos, L.M. Bragg, S. Kleywegt, G.R. Tetreault, M.E. McMaster, M.R. Servos, Reduction of Intersex in a Wild Fish Population in Response to Major Municipal Wastewater Treatment Plant Upgrades, Environ. Sci. Technol. 51 (2017) 1811–1819. doi:10.1021/acs.est.6b05370.
2. T.B. Hayes, L.L. Anderson, V.R. Beasley, S.R. De Solla, T. Iguchi, H. Ingraham, P. Kestemont, J. Kniewald, Z. Kniewald, V.S. Langlois, E.H. Luque, K.A. McCoy, M. Muñoz-De-Toro, T. Oka, C.A. Oliveira, F. Orton, S. Ruby, M. Suzawa, L.E. Tavera-Mendoza, V.L. Trudeau, A.B. Victor-Costa, E. Willingham, Demasculinization and feminization of male gonads by atrazine: Consistent effects across vertebrate classes, J. Steroid Biochem. Mol. Biol. 127 (2011) 64–73. doi:10.1016/j.jsbmb.2011.03.015.
3. I.Y. López-Pacheco, A. Silva-Núñez, C. Salinas-Salazar, A. Arévalo-Gallegos, L.A. Lizarazo-Holguin, D. Barceló, H.M.N. Iqbal, R. Parra-Saldívar, Anthropogenic contaminants of high concern: Existence in water resources and their adverse effects, Sci. Total Environ. 690 (2019) 1068–1088. doi:10.1016/j.scitotenv.2019.07.052.
4. C.M. McClure, K.L. Smalling, V.S. Blazer, A.J. Sperry, M.K. Schall, D.W. Kolpin, P.J. Phillips, M.L. Hladik, T. Wagner, Spatiotemporal variation in occurrence and co-occurrence of pesticides, hormones, and other organic contaminants in rivers in the Chesapeake Bay Watershed, United States, Sci. Total Environ. 728 (2020) 138765. doi:10.1016/j.scitotenv.2020.138765.
5. G.D. Smith, V.S. Blazer, H.L. Walsh, L.R. Iwanowicz, C. Starliper, A.J. Sperry, The effects of disease-related mortality of young-of-year smallmouth bass on the population characteristics in the Susquehanna River basin, Pennsylvania and potential implications to conservation of black bass diversity., in: Am. Fish. Soc. Symp., 2015: pp. 82:319-332.
6. H.L. Walsh, V.S. Blazer, G.D. Smith, M. Lookenbill, D.A. Alvarez, K.L. Smalling, Risk

Factors Associated with Mortality of Age-0 Smallmouth Bass in the Susquehanna River Basin, Pennsylvania, J. Aquat. Anim. Health. 30 (2018) 65–80. doi:10.1002/aah.10009.

7. C. Starliper, V.S. Blazer, L.R. Iwanowicz, H.L. Walsh, Microbial isolates in diseased fishes, primarily smallmouth bass (Micropterus dolomieu), within the Chesapeake Bay drainage in 2009-2011, in: Proc. West VIrginia Acad. Sci. Submitt. Pap. Eight-Eighth Annu. Sess. 2013, 2013.
8. J.J. Chaplin, J. Kent Crawford, R.A. Brightbill, Water-Quality Monitoring in Response to Young-of-the- Year Smallmouth Bass (Micropterus dolomieu) Mortality in the Susquehanna River and Major Tributaries, Pennsylvania: 2008, 2009. https://pubs.usgs.gov/of/2009/1216/ofr2009-1216.pdf.
9. V.S. Blazer, D.D. Iwanowicz, H.L. Walsh, A.J. Sperry, L.R. Iwanowicz, D.A. Alvarez, R.A. Brightbill, G. Smith, W.T. Foreman, R. Manning, Reproductive health indicators of fishes from Pennsylvania watersheds: association with chemicals of emerging concern, Environ. Monit. Assess. 186 (2014) 6471–6491. doi:10.1007/s10661-014-3868-5.
10. L.R. Iwanowicz, A.E. Pinkney, C.P. Guy, A.M. Major, K. Munney, V.S. Blazer, D.A. Alvarez, H.L. Walsh, A. Sperry, R. Braham, L.R. Sanders, D.R. Smith, Temporal evaluation of estrogenic endocrine disruption markers in smallmouth bass (Micropterus dolomieu) reveals seasonal variability in intersex, Sci. Total Environ. 646 (2019) 245–256. doi:10.1016/j.scitotenv.2018.07.167.
11. V.S. Blazer, S.E. Gordon, H.L. Walsh, C.R. Smith, Perfluoroalkyl Substances in Plasma of Smallmouth Bass from the Chesapeake Bay Watershed, Int. J. Environ. Res. Public Health. 18 (2021) 5881. doi:10.3390/ijerph18115881.
12. Luzio, S.M. Monteiro, S. Garcia-Santos, E. Rocha, A.A. Fontaínhas-Fernandes, A.M. Coimbra, Zebrafish sex differentiation and gonad development after exposure to 17α-ethinylestradiol, fadrozole and their binary mixture: A stereological study, Aquat. Toxicol. 166 (2015) 83–95. doi:10.1016/j.aquatox.2015.07.015.
13. T. Ye, M. Kang, Q. Huang, C. Fang, Y. Chen, H. Shen, S. Dong, Exposure to DEHP and MEHP from hatching to adulthood causes reproductive dysfunction and endocrine disruption in marine medaka (Oryzias melastigma), Aquat. Toxicol. 146 (2014) 115–126. doi:10.1016/j.aquatox.2013.10.025.
14. V.R.T. Zanotelli, S.C.F. Neuhauss, M.U. Ehrengruber, Long-term exposure to bis(2-ethylhexyl)phthalate (DEHP) inhibits growth of guppy fish (Poecilia reticulata), J. Appl. Toxicol. 30 (2010) 29–33. doi:10.1002/jat.1468.

15. Z. Pang, J. Chong, G. Zhou, D.A. de Lima Morais, L. Chang, M. Barrette, C. Gauthier, P.-É. Jacques, S. Li, J. Xia, MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights, Nucleic Acids Res. 49 (2021) W388–W396. doi:10.1093/NAR/GKAB382.
16. H.L. Walsh, V.S. Blazer, L.R. Iwanowicz, G. Smith, A Redescription of Myxobolus inornatus from Young-of-the-Year Smallmouth Bass (Micropterus dolomieu), J. Parasitol. 98 (2013) 1236–1242. doi:10.1645/ge-3081.1.
17. W.-K. Yang, L.-F. Chiang, S.-W. Tan, P.-J. Chen, Environmentally relevant concentrations of di(2-ethylhexyl)phthalate exposure alter larval growth and locomotion in medaka fish via multiple pathways, Sci. Total Environ. 640–641 (2018) 512–522. doi:10.1016/j.scitotenv.2018.05.312.
18. M. Bayley, M. Junge, E. Baatrup, Exposure of juvenile guppies to three antiandrogens causes demasculinization and a reduced sperm count in adult males, Aquat. Toxicol. 56 (2002) 227–239. doi:10.1016/S0166-445X(01)00210-7.
19. Y. Lor, A. Revak, J. Weigand, E. Hicks, D.R. Howard, T.C. King-Heiden, Juvenile exposure to vinclozolin shifts sex ratios and impairs reproductive capacity of zebrafish, Reprod. Toxicol. 58 (2015) 111–118. doi:10.1016/j.reprotox.2015.09.003.
20. S. Ullah, M.J. Zorriehzahra, Ecotoxicology: A Review of Pesticides Induced Toxicity in Fish, Adv. Anim. Vet. Sci. 3 (2015) 40–57. doi:10.14737/journal.aavs/2015/3.1.40.57.
21. D. Ruiz-Perez, H. Guan, P. Madhivanan, K. Mathee, G. Narasimhan, So you think you can PLS-DA?, BMC Bioinforma. 2020 211. 21 (2020) 1–10. doi:10.1186/S12859-019-3310-7.

Chapter 6

Summary and Conclusions

The goal of this study was to develop a new methodology for a non-traditional approach to examining the role of environmental contaminants in disease in SMB. This was accomplished by applying an exposomics based approach to evaluate YOY SMB for target and non-target compounds of interest using both gas and liquid chromatographic platforms for the analysis of a single sample extraction. This study developed a new application for the QuEChERS sample extraction methodology, providing future researchers a simple, quick, and efficient extraction method that can be analyzed by multiple chromatographic platforms without repeated sample preparation steps. It has helped to fill the gap in knowledge surrounding exposures in YOY SMB using targeted analyses. Non-targeted analyses have provided new information about compounds not routinely monitored in fish and identified compounds that are statistically significant to specific classes of YOY SMB. The work also highlights challenges and solutions for analyzing complex biological matrices as well as for simultaneously characterizing datasets for target and non-target compounds. The complexity associated with non-target sample and data analyses is a major hurdle to overcome before non-targeted analyses become routine for environmental monitoring efforts. However, as the effects of unknown anthropogenic contaminants on both wildlife and humans is constantly growing, there is a need to establish robust and routine non-targeted analyses. This study provides a framework for creating high-throughput, non-targeted methodologies that can be transitioned into routine settings in the future.

In Chapter 2, the sample preparation, analysis, and data processing methodology used throughout the entire study was established. A new modification to the QuEChERS extraction was optimized specifically for use on YOY fish. Due to their small size and high fat content,

existing EPA protocols would require sample pooling and extensive cleanup that would detract from if not completely prevent a non-targeted analysis. The methodology was validated by spiking salmon tissue with analytical standards of compounds routinely monitored in fish and analyzing them by GC×GC-TOFMS. The methodology was applied to a subset of 21 YOY SMB from two collection sites for proof-of-concept in the study organism. In total, 34 out of 233 target compounds were detected within the 21 YOY SMB. The non-targeted analysis identified an additional 10 significant features that were unique to either the Normal or Lesioned sample classes. This work resulted in a valid methodology for the analysis of the entire YOY SMB sample set.

Chapter 3 presents the application of the optimized methodology to 146 YOY SMB collected from 14 sampling sites. This set included a reevaluation of the subset of 21 samples on a new column combination. Fish were evaluated for the presence of 127 target compounds. A total of 15 non-PCB target compounds were detected in samples. PCBs were not reported as individual congeners, but as total PCB loads. The plasticizer DEHP and the insect repellant DEET were found at the highest concentrations within samples. An additional 65 features were identified as statistically significant in the non-targeted analysis. These features were tentatively identified and grouped by chemical family for comparison between sample classes. Three chemical families, esters, ketones, and nitrogen-containing compounds were found at higher percent abundances in Lesioned class fish than Normal class fish. This finding is important, as a leading theory for the cause in the decline of SMB health relates to increased nitrogen and phosphorus agricultural run-off entering waterways. Esters and ketones have also been associated with oxidative stress in organisms and these compounds may be possible biomarkers of disease characteristics. These findings provide new information for future studies investigating biomarkers and metabolites associated with the observed signs of disease, an area of research that has not yet been explored in SMB.

Chapter 4 utilized the same sample extracts evaluated in Chapter 3 and analyzed them by UPLC-QTOF-MS as a complement to the GC×GC-TOFMS data. This showed the versatility of the modified QuEChERS extraction methodology, as only a simple solvent exchange was needed to analyze the samples by a completely different chromatographic technique. The target list consisted of 33 pesticide and 5 PFAS compounds. Eight pesticides and one PFAS were quantified within the YOY SMB. Three additional PFAS were detected. PFAS have only been reported once in adult SMB and never before in YOY SMB from the Susquehanna River Basin. Additional compounds, not always included on routine monitoring target lists, were detected in the non-targeted analysis. However, no compounds were determined to be statistically significant to a specific sample class as with the GC×GC data. While the sample preparation provided an extract that could be used for both platforms, the data processing pipelines for these two analytical methods are vastly different. GC×GC data benefits from an alignment software that allows for feature reduction during alignment and focuses on identifying compounds that significantly contribute to the variation between sample classes. A similar feature reduction method was not available for the UPLC data. Statistical tools designed for metabolomics data were employed to mine for significant features, but without success. Though class specific significant features were not revealed by this analysis, other significant non-target compounds including never before reported pesticides, pharmaceuticals, and steroids were.

Chapter 5 examined the relationships between the target compounds and the YOY SMB. Correlation tests were used to determine if any correlations could be made between the detected targets and the observed disease characteristics present on the YOY SMB in this study. Weak, but statistically significant, correlations between two chemical groups, PFAS and pesticides, and broadly defined lesion subtype categories were established. However, specific lesion types could not be correlated with any single target chemical class. An additional slight positive correlation between the size of the YOY SMB and their classification as Lesioned was found to be

statistically significant, though no correlations between size and target chemical body burdens were found. An additional evaluation of the GC×GC data was performed to focus on identifying non-target compounds that were statistically significant among different sampling sites. A total of 38 features were found and 35 of these were tentatively identified. Comparisons between the two non-targeted GC×GC analyses revealed four compounds that were deemed statistically significant to both as well as having matching tentative identifications, retention times, and mass spectra. While all the non-target compounds identified warrant further investigation, these four should be a priority in future analyses to determine what role they play in the health of SMB.

Overall, this study has provided new information about unknown compounds of interest as well as providing a point of reference for future investigations into chemical contaminant exposure in YOY SMB. The exposomics based approach led to the development of a broad range sample extraction and minimalized cleanup suitable for analysis by both GC and LC platforms. This sample preparation methodology can be used for future studies of YOY SMB but can also be used to examine other aquatic populations undergoing similar health crises throughout the world. The use of simultaneous targeted and non-targeted data processing methodologies provides the most comprehensive dataset possible for each sample. This allows researchers to quantify known contaminants and act immediately in situations where remediation is warranted as well as qualitatively identifying previously unknown compounds of interest. These unknowns may include other priority contaminants that were not selected for targeted analysis, degradation products or metabolites of the selected target compounds, or previously undiscovered contaminants of emerging concern.

While this study provides new information about both target and non-target compounds within YOY SMB, it does not provide a complete overview of all the factors that can impact their health. Further examination of environmental matrices (e.g., water, sediment, food sources) is needed to understand sources of exposure for the YOY SMB. Additional collection time points of

either the same spawning classes throughout their maturation or collections from the same populations across multiple consecutive years are needed to properly evaluate the relationship between chemical contaminants and disease characteristics. Sampling of YOY SMB from out-of-basin reference sites with no disease characteristics during the same collection periods is needed for all additional studies to ensure there are proper control samples to use as references. And finally, the non-target compounds identified in this study should be further investigated to determine their role in SMB health.

Appendix

Portions of this dissertation, including text and figures, were from the following peer-reviewed publication:

Teehan, P., Schall, M. K., Blazer, V. S., Gruber, B., and Dorman, F. L. (2020). Modified QuEChERS Extraction for the Analysis of Young-of-Year Smallmouth Bass Using GCxGC-TOFMS. Anal. Methods 12, 3697-3704

Teehan, P., Schall, M. K., Blazer, V.S., and Dorman, F.L. (2022). Targeted and Non-targeted Analysis of Young-of-Year Smallmouth Bass Using Comprehensive Two-Dimensional Gas Chromatography Coupled with Time-of-Flight Mass Spectrometry. Sci. Total Environ. 806, 150378

VITA

Paige D. Teehan

Education

Ph.D. (2021) Biochemistry, Microbiology, and Molecular Biology
The Pennsylvania State University

B.S. (2012) Biological Sciences, University of South Carolina

Additional Publications

1. Pérez-Morales, D., Nava-Galeana, J., Rosales-Reyes, R., **Teehan, P**., Yakhnin, H., De la Cruz, M. A., Babitzke, P., and Bustamante, V. H. (2021). An incoherent feedforward loop formed by SirA/BarA, HilE and HilD controls the growth cost of virulence factors expression by Salmonella Typhimurium. PLoS Pathog. 17, 5

2. Murrell, K.A., **Teehan, P**., and Dorman, F. L. (2021). Determination of Contaminants of Emerging Concern and Their Transformation Products in Treated-Wastewater Irrigated Soil and Corn. Chemosphere 218, 130735.

3. Stultz, C., Jaramillo, R., **Teehan, P**., and Dorman, F. L. (2020). Comprehensive Two-Dimensional Gas Chromatography Thermodynamic Modeling and Method Optimization of Polychlorinated Dibenzo-p-dioxins and Dibenzofurans Separations in Fish Tissue Matrix. J. Chromatogr. A 1626, 461311

4. **Teehan, P**., Murrell, K. A., Jaramillo, R., Wicker, A. P., Parette, R., Schug, K. A., and Dorman, F. L. (2020). Environmental Analysis of Soil, Water, and Air. In M.N. Dunkle & W. L. Winniford (Eds.) Analytical Techniques in the Oil and Gas Industry for Environmental Impact Monitoring (pp. 381- 416). Wiley.

5. Weggler, B. A., Gruber, B., **Teehan, P**., Jaramillo, R., and Dorman, F. L. (2020). Inlets and Sampling. In N. Snow (Ed.). Basic Multidimensional Gas Chromatography (pp. 141-203). Elsevier Inc.

Presentations

2021 Multidimensional Chromatography Workshop (Liège, Belgium*Virtual)
2020 Multidimensional Chromatography Workshop (Honolulu, HI)
2019 International Symposium on Capillary Chromatography (Ft. Worth, TX)
Environmental Chemistry and Microbiology Student Symposium (State College, PA)