

NON-TARGET ANALYSIS OF BIOREMEDIATED SOIL

Zhenyu Tian

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Environmental Sciences and Engineering in the Gillings School of Global Public Health.

Chapel Hill
2018

Approved by:

Michael D. Aitken

Wanda M. Bodnar

Avram Gold

Kun Lu

Jason D. Surratt

© 2018
Zhenyu Tian
ALL RIGHTS RESERVED

ABSTRACT

Zhenyu Tian: Non-target Analysis of Bioremediated Soil
(Under the direction of Michael D. Aitken)

Polycyclic aromatic hydrocarbons (PAHs) are ubiquitous pollutants of environmental concern. Bioremediation, relying on stimulation of natural microbial degradation processes, is a well-established technology to clean up PAH-contaminated soils. However, bioremediation does not necessarily lead to a reduction in soil toxicity. PAH-contaminated sites are affected by extremely complex mixtures, like coal tar or creosote, and biotransformation products or co-occurring compounds can also contribute to the overall toxicological effects of contaminated soil before and after bioremediation. Therefore, the objective of this dissertation was to use non-target analysis workflows to identify the genotoxic transformation products, important co-occurring pollutants, and the unrecognized biotransformation pathways that could contribute to explain the toxicological effects observed beyond parent PAHs.

To identify the source(s) of increased genotoxicity in bioremediated soil, we pursued a non-target analytical approach combining effect-directed analysis (EDA) and metabolite profiling to compare extracts of PAH-contaminated soil before and after bioremediation. A compound with the composition $C_{15}H_8O_2$ and four methylated homologues were shown to accumulate as a result of bioreactor treatment, and the $C_{15}H_8O_2$ compound was determined to be genotoxic. Its structure was established as a heretofore unidentified α,β -unsaturated lactone, 2H-naphtho[2,1,8-*def*]chromen-2-one (NCO), which was confirmed by synthesis. It also accumulated in aerobically incubated soil from two additional PAH-contaminated sites and was

formed from pyrene by two pyrene-degrading bacterial cultures known to be geographically widespread, underscoring its potential environmental significance.

Azaarenes are nitrogen heterocyclic polyaromatic compounds that co-occur with PAHs but have been poorly studied in environmental systems. High resolution mass spectrometry (HRMS) and mass-defect filtering were applied to four PAH-contaminated samples to analyze their diversity, abundance, and biodegradation behavior. The diversity, relatively high concentrations, and persistence of high-molecular-weight azaarenes were highlighted, and isomer-selective biodegradation was observed.

To help elucidate PAH biodegradation pathways and endpoints in contaminated soil, HRMS and stable isotope-assisted metabolomics (SIAM) workflows were tested and applied to a PAH-contaminated soil. Uniformly ^{13}C -labeled fluoranthene, pyrene, or benzo[*a*]anthracene were spiked into the soil and incubated in microcosms. With SIAM, known and unknown metabolites such as ring-cleavage products and conjugates were detected, and the transformation pathways leading to their formation proposed.

ACKNOWLEDGEMENTS

To my advisor, Prof. Michael D. Aitken: thank you for all the guidance, support, and patience. You taught me what is to be a good scientist and a man of integrity.

To my closest collaborator, Dr. Joaquim Vila: thank you for your help and accompany. I miss the coffee breaks and discussions that nourished our research.

To my committee members, Prof. Avram Gold: thank you for teaching me chemistry as well as the healthy life style; Prof. Jason D. Surratt: thank you for those challenging questions and great suggestions for career development; Prof. Wanda Bodnar and Prof. Kun Lu: thank you both for your help on the knowledge and skills of mass spectrometry.

To my wife, Han Xiao: thank you for your love and sacrifice. You offered me the support in my most helpless moments.

To my parents: thank you for the understanding and support, and I am sorry that I am not a good son in respect to the traditional values.

To all the collaborators and lab mates that have helped me in my research: Dr. Jun Nakamura, Dr. Zhenfa Zhang, Leonard Collins, Dr. Vyom Sharma, Dr. Xu Tian, Dr. David Singleton, Dr. Miao Yu, Dr. Alden Adrion, Dr. Eric Staunton, Elizabeth Corteselli, and Elisabeth Rehak. To my friends in the ESE department: Tianqu Cui, Jingbo Wang, Ryan Kingsbury, and Emanuele Sozzi. Thank you all for your help!

P.S.: If I missed out anyone in the list of acknowledgements, I sincerely apologize and I could do a proper *thank you* in person, such as a coffee or chocolate.

TABLE OF CONTENTS

LIST OF TABLES	xi
LIST OF FIGURES	xii
CHAPTER 1 INTRODUCTION.....	1
1.1 Specific aims and rationale	2
1.2 Dissertation organization	4
CHAPTER 2 BACKGROUND.....	5
2.1 PAHs and related compounds in contaminated sites	5
2.2 Toxicity and genotoxicity of polycyclic aromatic compounds.....	8
2.2.1 Toxicities of parent PAHs.....	8
2.2.2 Toxicities of oxy-PAHs	9
2.2.3 Toxicities of azaarenes.....	11
2.3 Biotransformation of PAHs and co-occurring compounds.....	12
2.3.1 Bacterial transformation of PAHs.....	13
2.3.2 Incomplete metabolism of PAHs.....	16
2.3.3 PAH metabolites and increased genotoxicity in contaminated soil.....	17
2.4. Introduction to non-target analysis	17

2.4.1 Workflow for non-target analysis	17
2.4.2 Effect-directed analysis.....	21
2.4.3 Metabolite profiling	22
2.4.4 Stable isotope-assisted metabolomics.....	23
CHAPTER 3 NON-TARGET ANALYSIS REVEALS A BACTERIAL METABOLITE OF PYRENE IMPLICATED IN THE GENOTOXICITY OF CONTAMINATED SOIL AFTER BIOREMEDIATION	26
3.1 Introduction.....	26
3.2 Materials and methods	28
3.2.1 Chemicals.....	28
3.2.2 Soil and bioreactor treatment.....	28
3.2.3 Extraction and Fractionation.....	29
3.2.4 Bioassay	29
3.2.5 Instrumental analysis	30
3.2.6 Purification of NCO and structure identification.....	30
3.2.7 Other biotransformation experiments	31
3.2.8 Data analysis	31
3.3 Results and discussion	32
3.3.1 Fractionation of soil extracts and selection of genotoxic fraction	32
3.3.2 Metabolite profiling	34
3.3.3 Isolation and structural characterization of C ₁₅ H ₈ O ₂	37
3.3.4 Source, putative formation pathway, and potential prevalence.	41

3.3.5 Environmental relevance	44
CHAPTER 4 DIVERSITY AND ABUNDANCE OF HIGH-MOLECULAR-WEIGHT AZAARENES IN PAH-CONTAMINATED ENVIRONMENTAL SAMPLES	46
4.1 Introduction.....	46
4.2 Materials and methods	48
4.2.1 Chemicals.....	48
4.2.2 Contaminated sites, sample collection, and solvent extraction.....	48
4.2.3 Instrumental analysis	49
4.2.4 Nontarget analysis workflow and mass defect filtering.....	49
4.3 Results and discussion	51
4.3.1 PAH quantification and distribution profiles.....	51
4.3.2 Diversity of azaarenes.....	51
4.3.3 Abundances of azaarenes.....	56
4.3.4 Environmental significance	59
CHAPTER 5 ISOMER-SELECTIVE BIODEGRADATION OF HIGH-MOLECULAR- WEIGHT AZAARENES IN CONTAMINATED ENVIRONMENTAL SAMPLES	60
5.1 Introduction.....	60
5.2 Materials and methods	61
5.2.1 Chemicals.....	61
5.2.2 Contaminated samples	61
5.2.3 Microcosm incubations.....	62

5.2.4 Instrumental analysis	63
5.2.5 Statistical analysis	63
5.3 Results and discussion	64
5.3.1 Biodegradability of native azaarenes	65
5.3.2 Isomer-selective biodegradation	69
5.3.3 Effect of exogenous nitrogen addition on azaarene biodegradation	73
5.3.4 Biodegradability of azaarenes relative to their PAH analogues	73
5.3.5 Transformation products from azaarenes.....	74
5.3.6 Environmental significance	76
CHAPTER 6 TRACING BIOTRANSFORMATION OF PAHS IN CONTAMINATED SOIL USING STABLE ISOTOPE-ASSISTED METABOLOMICS	77
6.1 Introduction.....	77
6.2 Materials and methods	79
6.2.1 Chemicals.....	79
6.2.2 Contaminated soil and spiking	79
6.2.3 Microcosm incubations.....	79
6.2.4 Instrumental analysis.	80
6.2.5 Data analysis.	80
6.3 Results and discussions.....	81
6.3.1 Assessing performance of the method.	81
6.3.2 Metabolite identification.....	83

6.3.3 Proposed compound specific degradation pathways in soil	86
6.3.4 Time-course evolution of the identified metabolites.	88
6.3.5 Environmental implications and further applications.....	88
CHAPTER 7 CONCLUSIONS AND RECOMMENDATIONS	90
7.1 Conclusions.....	90
7.2 Recommendations for future work	92
APPENDIX A: SUPPORTING INFORMATION FOR CHAPTER 3	94
APPENDIX B: SUPPORTING INFORMATION FOR CHAPTER 4	111
APPENDIX C: SUPPORTING INFORMATION FOR CHAPTER 5	118
APPENDIX D: SUPPORTING INFORMATION FOR CHAPTER 6	126
REFERENCES	140

LIST OF TABLES

Table 2.1. Chemical structures and selected properties of the 16 USEPA Priority Pollutant PAHs.	7
Table 4.1. Description of the PAH-contaminated samples used in this study	51
Table 4.2 Detailed information on azaarene congeners detected in the samples.....	54
Table 6.1. Detected metabolites from the four-ring PAHs	82
Table A.1. Concentrations ($\mu\text{g}\cdot\text{g}^{-1}$ dry soil) of 14 EPA-regulated PAHs in the soil used to feed the bioreactor and in bioreactor-treated soil.	109
Table A.2. Final list of upregulated ions from metabolite profiling with intensity $\geq 5,000$, fold-change (treated soil:feed soil) ≥ 2 , and $p \leq 0.01$	110
Table A.3. Features from metabolite profiling in APPI+ mode that were “upregulated” in subfraction B2 from extracts of bioreactor-treated soil compared to the same subfraction from feed soil.	110
Table B.1. Initial PAH concentrations ($\mu\text{g}/\text{g}$ dry soil) in the four samples	114
Table B.2. HPLC-MS/MS analytical parameters for quantification.	115
Table B.3. Concentrations of non-alkylated azaarene congeners (mean \pm standard deviation; $n=3$).	116
Table C.1. Concentrations of non-alkylated azaarene congeners ($\mu\text{g}/\text{g}$ dry soil) before (T0) and after (T42) 6 weeks of microcosm incubation after 6 weeks.....	124
Table D.1. Concentrations of monitored PAHs ($\mu\text{g}/\text{g}$ dry soil) in contaminated soil samples (mean \pm standard deviation; $n=3$).	128
Table D.2. Parameter values used for data analysis with XCMS and X13CMS packages.....	129

LIST OF FIGURES

Figure 2.1. Funnel effect in PAH metabolic networks (Kweon et al., 2011). RCP: ring cleavage process; SCP: side chain process; CAP: central aromatic process.....	14
Figure 2.2. Different configurations of benz[<i>a</i>]anthracene binding in an RHD (PhnI) (Jakoncic et al., 2007)	16
Figure 2.3. Non-target screening workflow for environmental samples (Hollender et al., 2017)	18
Figure 2.4. Levels of confidence for the identification of unknowns (Schymanski et al., 2014a).....	21
Figure 3.1. Results of DT40 DNA-damage response bioassays on solvent extracts and fractionated extracts from PAH-contaminated soil used to feed a lab-scale bioreactor (“feed soil”, FS) and bioreactor-treated soil (TS).. ..	34
Figure 3.2. Non-target analysis of extract subfractions B2.....	36
Figure 3.3. DT40 bioassay results on 2 <i>H</i> -naphtho[2,1,8- <i>def</i>]chromen-2-one (NCO, inset) purified from extracts of bioreactor-treated soil.....	39
Figure 3.4. Proposed pathway for the formation of NCO from pyrene by PAH-degrading bacteria.....	42
Figure 4.1 Kendrick mass defect (KMD) plots of azaarene congeners obtained from the four contaminated sites.....	52
Figure 4.2. MS/MS fragmentation patterns of two tentatively identified compounds C ₁₉ H ₁₁ N (F8, above) and C ₂₀ H ₁₃ N (below).....	53
Figure 4.3. Extracted ion chromatograms (EICs) of the SC sample, corresponding to the five major congener groups.....	55
Figure 5.1 Concentrations of the five major azaarene congener groups over time during aerobic microcosm incubations in each of the samples.	66
Figure 5.2 KMD plots over the time series from aerobic microcosm incubations of the SC sample under biostimulated conditions.	68
Figure 5.3. Hierarchical cluster analysis of the quantifiable azaarene congeners according to their biodegradation behavior.	71
Figure 5.4. Overlaid extracted ion chromatograms (EICs) for <i>m/z</i> 204.0800 from KM (top) and SC (middle) samples at day 0 and after 42 days of microcosm incubation, and for <i>m/z</i> 230.0800 from the SC sample (bottom) at day 0 and after 21 days of incubation.	72

Figure 5.5. EICs of azaarene derivatives in the SC sample.	75
Figure 6.1. Mass spectra, extracted ion chromatograms (EIC), and proposed structures of two representative metabolites.	84
Figure 6.2. Proposed degradation pathway for benzo[<i>a</i>]anthracene based on detected metabolites and published literature.	87
Figure A.1. Flow chart of non-target analysis approach (top) and the scheme for effect-directed analysis of fractions (bottom).	100
Figure A.2. (A) LD ₅₀ and (B) relative LD ₅₀ (mutant LD ₅₀ /parental LD ₅₀) values for whole extracts and recombined extracts of feed soil (FS) and bioreactor-treated soil (TS) after the first level of fractionation (n=3).	101
Figure A.3. Relative LD ₅₀ values (LD ₅₀ of mutant cell line/LD ₅₀ of parental DT40 cell line) of soil extract fractions.	101
Figure A.4. Extracted ion chromatograms (<i>a, c, e</i>), MS/MS spectra (<i>b, d, f</i>), and structures of three compounds corresponding to the molecular formula C ₁₅ H ₈ O ₂ synthesized in this study.	102
Figure A.5. Proton NMR spectra (400 MHz, chloroform- <i>d</i>) and numbering of NCO; top: purified metabolite; bottom: synthetic standard.	103
Figure A.6. Heteronuclear multiple bond coherence (HMBC) NMR spectrum for NCO, 500 MHz, chloroform- <i>d</i>	104
Figure A.7. HPLC chromatograms (<i>a,b</i>) and QToF mass spectra (<i>c,d</i>) of NCO purified from bioreactor-treated soil extracts (<i>a,c</i> ; blue lines) and synthetic NCO (<i>b,d</i> ; red lines).	107
Figure A.8. Concentrations of NCO measured after microcosm incubations of bioreactor-treated soil spiked with additional pyrene.	108
Figure B.1. Mass spectra of azaarene isomers analyzed by qTOF in product ion scan mode.	113
Figure C.1. Concentration trend lines of quantifiable azaarene congeners throughout the incubation.	118
Figure C.2. Residual concentrations in each of the five azaarene congener groups quantified at the beginning of the incubation and after 42 days of incubation.	122
Figure C.3. Evolution of the relative concentration (N-PAH/PAH ratio) of the major azaarene congener groups (PANHs) relative to their homocyclic counterparts (PAHs) during the microcosm incubations for each of the four samples.	123

Figure D.1. Time course in unspiked samples of the concentrations of the three PAHs targeted for isotope-assisted metabolomics (left y-axis) and intensities of the detected metabolites (right y-axis).	130
Figure D.2. Mass spectra and extracted ion chromatograms (EIC) of three major metabolites from fluoranthene.....	131
Figure D.3. Mass spectra and EIC of three major metabolites from B[a]A.	133
Figure D.4. Mass spectra and EIC of four major metabolites from pyrene.	135
Figure D.5. Mass spectra of the sulfate conjugate of dihydroxypyrene, under (a) normal (200 V) and (b) low (100 V) source fragmentor voltage.....	137
Figure D.6. MS/MS spectra of pyrene metabolites (a) m/z 369.9945 and (b) m/z 386.9854, fluoranthene metabolite (c) $C_{16}H_{10}O_3$ and (d) $C_{11}H_{10}O_6$, and BaA metabolite (e) $C_{12}H_8O_6$	138
Figure D.7. Proposed degradation pathways for fluoranthene based on identified or putative metabolites.....	139

CHAPTER 1 INTRODUCTION

Polycyclic aromatic hydrocarbons (PAHs) are ubiquitous pollutants of environmental concern, due to their known or suspected toxic, genotoxic, and carcinogenic properties. PAHs have been found in over 700 Superfund sites in the U.S., and they are also the predominant pollutant class at an estimated 45,000 former manufactured gas plant (MGP) sites (US EPA, 2015a). Bioremediation is an established technology to clean up PAH-contaminated soils and sediments by microbial degradation, and it has been proved to remove the US EPA regulated PAHs efficiently (Gillespie and Philp, 2013). According to a recent review, bioremediation is the most frequently (42%) used technology for PAH-contaminated soils (Kuppusamy et al., 2017). For the Superfund sites in the US, bioremediation is also among the most commonly adopted remedial solutions (US EPA, 2015b).

Removal of PAHs during bioremediation does not necessarily lead to a reduced toxicity. A myriad of studies have suggested that bioremediation or biodegradation on PAH-contaminated soil could increase the genotoxicity (Belkin et al., 1994; Hu et al., 2012; Hughes et al., 1998; Lemieux et al., 2009). The source(s) of increased genotoxicity were not elucidated in previous studies, although polar compounds formed during treatment were implicated (Chibwe et al., 2015a). Some microbial metabolites of PAHs, such as oxygenated PAHs (oxy-PAHs), have been suspected because of their known mutagenicity (Lemieux et al., 2008; Zielinska-Park et al., 2004). However, commonly detected oxy-PAHs have been observed to decrease during bioremediation when increases in genotoxicity have been observed, excluding them from the suspected sources (Hu et al., 2014). Therefore, conventional targeted methods for pollutant analysis could not identify the genotoxic compounds in bioremediated soil.

In searching for the genotoxic by-products from bioremediation, I observed the high complexity of chemical composition in contaminated soil by non-target analysis using high resolution mass spectrometry (HRMS). Meanwhile, toxicological studies suggested that regularly monitored pollutants, such as the 16 EPA priority PAHs, only explain small proportions of the observed toxicity in contaminated soil before and after bioremediation (Brinkmann et al., 2014; Lemieux et al., 2015). Therefore, the unregulated and unknown compounds may have significant impact on the health risk of contaminated soil. To better understand the risk of contaminated sites and further improve bioremediation, it is crucial to characterize those unknown compounds.

Non-target analysis (NTA), also known as non-target screening (NTS), is an emerging concept that enables the identification of unknown compounds (Krauss et al., 2010). With the use of HRMS and advanced software or databases, NTA has the power to detect and prioritize the unknown pollutants and transformation products (TPs) from complex mixtures such as sediment and wastewater (Hollender et al., 2017; Schymanski et al., 2014b). The research described herein sought to develop and optimize non-target analysis workflows suitable for contaminated soils in the scenario of bioremediation, and then use these workflows to identify unregulated pollutants, unknown TPs, and unrecognized biodegradation pathways for pollutant transformation.

1.1 Specific aims and rationale

Aim 1: Identification of the unknown genotoxic compound(s) formed in bioremediation.

As indicated above, bioremediation is an accepted technology for cleanup of soil contaminated with polycyclic aromatic hydrocarbons (PAHs), but it can increase the genotoxicity of the soil despite removal of the regulated PAHs (Hu et al., 2012; Hughes et al.,

1998; Lemieux et al., 2009). Although polar biotransformation products have been implicated as causative genotoxic agents, no specific product has been identified (Chibwe et al., 2015a). To identify the unknown genotoxic compounds, in collaboration with Dr. Wanda Bodnar and Mr. Leonard Collins a non-target analytical approach was developed, combining effect-directed analysis (EDA) and metabolite profiling, and was applied to the extracts of PAH contaminated soil from a former manufactured-gas plant site before and after treatment in a laboratory-scale aerobic bioreactor. The DT40 DNA-damage response assay was used for driving the EDA, and HRMS was the major tool for metabolite profiling. The transformation products of interest were prioritized, with their toxicities tested and formation pathways proposed.

Aim 2: Identify and quantify the major azaarenes in samples obtained from PAH-contaminated sites, and observe their susceptibility to biodegradation in microcosm incubations.

In searching for the unknown genotoxic metabolites, a series of abundant features were observed in the whole extracts as well as the semi-polar fractions of contaminated soil. The exact masses and fragmentation patterns from HRMS analysis suggested their identities as N-heterocyclic polyaromatic compounds, also known as azaarenes. Some of these compounds are known to be toxic or carcinogenic (Bleeker et al., 1999; Machala et al., 2001; Wiegman et al., 2002), but there is a lack of knowledge of their fate and environmental behavior (Anyanwu and Semple, 2015a). To fill these gaps of knowledge, non-target analysis was applied on soil samples from four contaminated sites, and the diversity and abundance azaarenes were characterized. In collaboration with Dr. Joaquim Vila, microcosm experiments were set up to observe the biotransformation of azaarenes.

Aim 3: Trace the biotransformation of PAHs in contaminated soil by stable isotope-assisted metabolomics.

The identification of hitherto unknown metabolites under Aim 1 suggested that current knowledge on PAH biodegradation based on isolated microbial strains and single compounds

might be limited. Interactions among soil microbial communities or between components of complex PAH mixtures can result in the formation of unexpected metabolites (Hennessee and Li, 2016; Stringfellow and Aitken, 1995; Zhong et al., 2006). Therefore, to better understand the bioremediation process, it is important to trace the biotransformation of PAHs in real environmental samples such as contaminated soil. We (Dr. Vila and I) assumed that a non-target method called stable isotope-assisted metabolomics (SIAM) would be applicable. We tested the applicability of SIAM on bioremediation of a creosote-contaminated soil, and modified the method to fit the scenario. Using U-¹³C labeled PAHs and automated SIAM workflows, this method allowed us to detect new metabolites originating from specific PAHs, without knowing their identities. Based on these metabolites, biodegradation pathways were proposed.

1.2 Dissertation organization

This dissertation is organized into seven chapters. Chapter 1 provides an introduction and briefly explains the aims of the dissertation. Chapter 2 is a literature review which provides the background knowledge to understand the current work. Chapters 3-6, describe the detailed methods, results and discussions for Aims 1 (Chapter 3), 2 (Chapters 4 and 5), and 3 (Chapter 6). Chapter 3 and Chapter 4 have been published in the journal *Environmental Science and Technology (ES&T)* (Tian et al., 2017a, Tian et al., 2017b). Chapter 5 was originally submitted as part of the manuscript presented in Chapter 4, but at the editor's request was broken out as a separate manuscript; it has not yet been submitted for publication. Chapter 6 has been accepted for publication in the journal *Environmental Science and Technology Letters (ES&T Letters)*. Chapter 7 provides conclusions and recommendations for future research.

CHAPTER 2 BACKGROUND

2.1 PAHs and related compounds in contaminated sites

PAHs are composed of two or more fused aromatic rings, featured by the delocalized conjugated π system. The lower Gibbs free energies of formation make PAHs thermodynamically stable, and the high activation energies for substitution reactions make them less reactive. Their high lipophilicity, low water solubility, and high octanol/water partition coefficients facilitate their accumulation in geosorbents (Johnsen et al., 2005; Wilcke, 2000) and the tissues of organisms (Meador et al., 1995; Nakata et al., 2003). With these properties and their toxicities, PAHs are classified as persistent, bioaccumulative, and toxic (PBT) chemicals (Schaum et al., 2003). The PAHs of environmental concern range from two rings (naphthalene, $C_{10}H_8$) to seven rings (coronene, $C_{24}H_{12}$). For the convenience of monitoring and comparison, 16 representatives of PAHs were prioritized by US EPA according to their prevalence and toxicity (Keith, 2015). Detailed properties of these 16 priority PAHs are listed in **Table 2.1**.

PAHs can be generated from both natural and anthropogenic sources. It is known that forest fires and some biological conversions produce PAHs, and fossil fuels contain considerable amounts of PAHs (Nikolaou et al., 1984). However, the environmental prevalence of PAHs is more closely related to human activities, such as incomplete combustion and the use of fossil fuels (Johnsen et al., 2005). As these processes continue to occur, PAHs will be continuously generated and discharged to the environment. The primary sources of PAHs that contaminate soil and sediment are coal tar, creosote, petroleum, and industrial effluents (Neilson, 1995). Since the nineteenth century, combustible gas manufactured from coal or coke served as a fuel source for urban cooking, heating, and lighting for decades, and the remaining wastes, such as coal tar,

persistently contaminated these manufactured-gas plant (MGP) sites. An estimated 45,000 former MGP sites in the US were contaminated with PAHs, and they have been identified as the predominant pollutants at hundreds of contaminated sites on the Environmental Protection Agency's National Priority List (US EPA, 2015a). In addition, creosote distilled from coal tar was widely used as a wood preservative, and could remain at wood-preserving sites as contaminants (Bleeker et al., 1999).

The contamination sources such as coal tar and creosote are complex mixtures containing thousands of compounds (Koolen et al., 2015; Mueller et al., 1989), and the transformation of these contaminants in the environment may generate other by-products. As a result, many known and unknown compounds, such as oxygenated derivatives of PAHs and N, S, O- heterocyclics, usually co-occur with PAHs in contaminated sites (Lemieux et al., 2009, 2015; Lundstedt et al., 2003; Meyer et al., 1999). With the development of advanced analytical technologies, more recent studies began to focus on N, S, O- heterocyclic aromatic compounds (Anyanwu and Semple, 2015a; Bandowe et al., 2014b; Lundstedt et al., 2014; Manzano et al., 2017). Based on these studies and observations, it has been suggested that a more comprehensive list of polycyclic aromatic compounds (PACs) may serve as an alternative to the 16 EPA priority PAHs (Andersson and Achten, 2015).

Table 2.1. Chemical structures and selected properties of the 16 USEPA Priority Pollutant PAHs ^{a, b}

Compound	Structure	MW	C_{iw}^{sat}	p_i^*	H	$\log K_{ow}$	$\log K_{oc}$
Naphthalene (NAP)		128.18	3.10E+01	1.04E+01	4.30E+01	3.37	3.03
Acenaphthylene (ACY)		152.19	1.61E+01	5.80E-01	5.50E+00	4.00	3.40
Acenaphthene (ACE)		154.20	3.80E+00	3.00E-01	1.22E+01	3.92	3.62
Fluorene (FLU)		166.23	1.90E+00	9.00E-02	7.87E+00	4.18	3.97
Phenanthrene (PHN)		178.24	1.10E+00	2.00E-02	3.24E+00	4.57	4.21
Anthracene (ANT)		178.24	4.50E-02	1.00E-03	3.96E+00	4.54	4.38
Fluoranthene (FLA)		202.26	2.60E-01	1.23E-03	1.04E+00	5.22	4.55
Pyrene (PYR)		202.26	1.32E-01	6.00E-04	9.20E-01	5.18	4.83
Benzo[<i>a</i>]anthracene (BaA)		228.30	1.1E-02	2.80E-05	5.81E-01	5.91	5.08
Chrysene (CHR)		228.30	2.90E-03	5.70E-07	6.50E-02	5.86	5.17
Benzo[<i>b</i>]fluoranthene (BbF)		252.32	2.00E-03	2.18E-07	2.14E-02	5.75	5.72
Benzo[<i>k</i>]fluoranthene (BkF)		252.32	8.00E-04	5.20E-08	1.60E-02	6.00	5.73
Benzo[<i>a</i>]pyrene (BaP)		252.32	3.80E-03	7.00E-07	4.60E-02	6.04	5.99
Benzo[<i>g,h,i</i>]perylene (BgP)		276.34	3.00E-04	1.69E-08	7.50E-02	6.50	6.35
Dibenzo[<i>a,h</i>]anthracene (DBA)		278.36	6.00E-04	3.70E-10	1.28E-03	6.75	5.96
Indeno[<i>1,2,3-c,d</i>]pyrene (INP)		276.34	6.20E-03	2.70E-09	1.00E-02	7.66	6.20

^a abbreviations: MW = molecular weight (g/mol); C_{iw}^{sat} = aqueous solubility (mg/L); p_i^* = vapor pressure (Pa); H = Henry's Law Constant (Pa·m³/mol·K); K_{ow} = octanol/water partitioning coefficient; K_{oc} = organic carbon partitioning coefficient. ^b All data from (Mackay et al., 2006).

2.2 Toxicity and genotoxicity of polycyclic aromatic compounds

2.2.1 Toxicities of parent PAHs

Since the British surgeon Percivall Pott reported the unusually high incidence of scrotal cancer among the London chimney sweeps in 1775, people began to relate PAHs with cancer and other adverse health effects (Baird et al., 2005). The acute toxicities of PAHs were expressed by narcosis as the mode of toxic action, without specific receptor binding or enzymatic activation (Sverdrup et al., 2002). As lipophilic compounds, PAHs could accumulate in the membrane lipid bilayer and damage the normal structural and functional properties of cell membranes (Sikkema et al., 1995). However, the acute toxicities are only observed at high concentrations, and the environmental exposures of PAHs seldom reach such doses due to limited water solubility and bioavailability (Johnsen et al., 2005). Therefore, more studies have focused on the genotoxicity and carcinogenicity of PAHs.

The mutagenicity and genotoxicity of PAHs are caused by their *in vivo* metabolites activated by enzymes. These metabolites could cause direct DNA damage by covalent binding to form stable adducts and indirect DNA damage by inducing oxidative stress in redox cycles (Bolton et al., 2000). Three major pathways have been concluded for the metabolic activation of PAHs: 1) bay region dihydrodiol epoxide pathway by cytochrome P450 enzymes (CYPs), 2) radical cation pathway by one-electron oxidation, and 3) the ortho-quinone pathway via dihydrodiol dehydrogenase (Xue and Warshawsky, 2005). Since the three pathways lead to different reactive intermediates, they cause DNA damage by different mechanisms.

The bay region dihydrodiol epoxide pathway is well known because some of the diol epoxides (e.g. benzo[*a*]pyren-7,8-dihydrodiol-9,10-epoxide, BPDE) have been shown to be potent carcinogens and to be responsible for DNA adduct formation (Meehan et al., 1977; Szeliga and Dipple, 1998). In this pathway, PAHs are transformed into diol epoxides through a

series of reactions catalyzed by CYPs and hydrolases. These diol epoxides are electrophiles that could form a covalent bond with the nucleophilic parts of DNA bases, especially deoxyguanosine (dGuo) and deoxyadenosine (dAdo) (Szeliga and Dipple, 1998). Stable and bulky adducts are generated in such processes, leading to errors in later DNA replication process.

The removal of a single electron in PAH molecules will lead to the formation of highly reactive radical cations (Penning et al., 1999). As electrophiles, the radical cations could form adducts with dGuo and dAdo in a way similar to diol epoxides. The difference in their DNA damage effects is that adducts from radical cations are unstable (Xue and Warshawsky, 2005). While some of these unstable adducts stabilize, others become apurinic sites via a spontaneous depurination process (Bolton et al., 2000).

PAH *o*-quinones are generated when dihydrodiol dehydrogenases compete with CYPs to oxidize the activated PAH dihydrodiols (Jiang et al., 2005). The detailed mechanism for the genotoxicity of *o*-quinones is further discussed in the following chapter, as quinones are common oxygenated PAHs (oxy-PAHs) that are found in the environment (Bandowe et al., 2014b; Wischmann and Steinhart, 1997).

2.2.2 Toxicities of oxy-PAHs

Oxy-PAHs are oxygenated products of PAHs that are formed through biodegradation (Cerniglia, 1992; Lundstedt et al., 2003; Wischmann and Steinhart, 1997), chemical oxidation (Kochany and Maguire, 1994), and phototransformation (Fox and Olive, 1979). As all these processes lead to the oxidation of PAHs, there are actually different compound classes that could be classified as oxy-PAHs, such as quinones, ketones, lactones, and hydroxylated compounds. Compared to parent PAHs, the toxicological behaviors of oxy-PAHs are less understood. It is known that many oxy-PAHs are toxic and genotoxic, and sometimes they could be even more toxic than the parent compounds (McConkey et al., 1997; Shuttleworth and Cerniglia, 1995).

Besides cytotoxicity and genotoxicity, some oxy-PAHs also exhibit aryl hydrocarbon receptor (AhR) activating potency and estrogenic activities (Machala et al., 2001; Wincent et al., 2015). There are several major mechanisms responsible for the toxicological behaviors of oxy-PAHs: adduct formation (Briggs et al., 2003; Lampi et al., 2006), redox production of reactive oxygen species (ROS) (Flowers-Geary et al., 1996; Jarabak et al., 1998), specific receptor binding (Machala et al., 2001), and mitochondrial disruption (Xia et al., 2004; Zhu et al., 1995).

Cytotoxicity of oxy-PAHs is induced by ROS production and mitochondrial disruption. Some PAH quinones and ketones are known to be redox-active, which means they could enter the redox cycle, causing a variety of toxic effects in cells (Jarabak et al., 1998). In the presence of transition metals, redox cycling via the reduction by NADPH to a semiquinone or hydroquinone could initiate a cascade producing superoxide anion and hydrogen peroxide. Important cellular structures like membrane proteins and lipids could be altered by these ROS, resulting in detrimental functional changes (Schweigert et al., 2001). Besides the damage from ROS, PAH quinones could depolarize mitochondrial membrane potential, leading to ATP depletion and cytotoxicity (Henry and Wallace, 1996). It has been shown that these mitochondrial disruption effects were not altered in the presence of superoxide dismutase, suggesting that they were not related to oxidative stress (Zhu et al., 1995).

Genotoxicity of oxy-PAHs is caused by adduct formation and oxidative damage. Some PAH quinones and lactones are known to form stable covalent adducts with DNA bases (Bolton et al., 2000; McCoull et al., 1999). The mechanism for such covalent bonding is that PAH *o*-quinones and unsaturated lactones are Michael addition acceptors, and have the potential to attack the electron-rich parts of DNA bases (Xue et al., 2002). As mentioned previously, redox-active oxy-PAHs could generate ROS such as hydrogen peroxide and hydroxyl radicals. These reactive intermediates are very potent DNA oxidants that could damage DNA by base modification, or lead to strand scission (Penning et al., 1999). The most common oxidative

damage is oxidized dGuo, such as 8-oxo-dG. It has been suggested that PAH *o*-quinones could produce a significant amount of 8-oxo-dG in salmon DNA (Park et al., 2004). If not repaired, they could mismatch with dA in replication, resulting in G→T transversions (Penning et al., 1999). Because of the high potential of introducing oxidative mutations in tumor suppressor p53, PAH *o*-quinones may play an important role in lung cancer (Yu et al., 2002).

2.2.3 Toxicities of azaarenes

Azaarenes, also known as polycyclic aromatic nitrogen heterocycles (PANH), are the heterocyclic analogs of polycyclic aromatic hydrocarbons (PAHs) in which one or more nitrogen atoms are incorporated into the aromatic structure. While sharing some common properties with PAHs, the presence of nitrogen brings special characteristics to azaarenes, such as stronger polarity and higher structural diversity. Many azaarenes are toxic; for example, some of the high-molecular-weight (HMW) compounds (containing 4 or more rings) are mutagenic and carcinogenic to animals (E. A. J. Bleeker et al., 2002; Wiegman et al., 2001) and are major contributors of AhR-mediated activities (Machala et al., 2001).

From the perspective of cytotoxicity, azaarenes have been suspected to be more toxic than their homocyclic counterparts, due to their higher water solubility and thus higher bioavailability (Pearlman et al., 1984). Compared to high-molecular-weight azaarenes, the toxicological behaviors of two- and three-ring congeners have been more extensively studied. The effective concentrations for toxic effects of quinoline and acridine in different testing species are similar, suggesting the mechanism behind their toxicities might be non-specific (E. A. J. Bleeker et al., 2002). Also, lab experiments have shown that these toxicities could be enhanced under radiation by UV light (Wiegman et al., 2002). High-molecular-weight azaarenes tend to be more toxic (Machala et al., 2001), but molecular size is not the only parameter that affects toxicity. It is worth noting that for azaarene isomers that differ only in the position(s) of

the nitrogen atom(s), their toxicities may differ by several orders of magnitude (Bleeker et al., 1999; Kraak et al., 1997).

Many azaarene congeners have been shown to be mutagenic and genotoxic. Low-molecular-weight congeners like quinoline, phenanthridine, and benzo[*h*]quinoline are directly genotoxic, which could damage DNA without metabolic activation (Seixas et al., 1982). The direct genotoxicity might be related to the substituted nitrogen atom (Bleeker et al., 1999). These smaller sized congeners, however, are not carcinogenic. Several high-molecular-weight azaarenes, like benzo[*c*]acridine and 10-aza-benzo[*a*]pyrene, have been shown to be carcinogens (Chang et al., 1984; Yamada et al., 2004). These high-molecular-weight congeners share similar metabolic pathways and indirect carcinogenicity with PAHs, suggesting that the influence of the nitrogen atom decreases in larger molecules (Castagnoli et al., 1997). Like PAHs, the azaarenes could be transformed into highly electrophilic intermediates such as diol epoxides and *o*-quinones, and these reactive intermediates could form adducts with DNA by binding strongly with the bases (Tullis and Banerjee, 1984). The stable oxidized derivatives of azaarenes, like acridone and phenanthridinone, have also been studied for their toxicities. Results from the Mutatox bioassay suggested that although these derivatives presented weaker cytotoxicity, their genotoxicity was stronger than that of the parent compounds (Bleeker et al., 1999).

2.3 Biotransformation of PAHs and co-occurring compounds

It has been known for a long time that microorganisms can degrade and metabolize PAHs and related aromatic compounds, and some species can even mineralize and grow on PAHs (Cerniglia, 1992; Gibson et al., 1975). Bacteria, fungi, and algae have all shown their abilities as PAH degraders, and both aerobic and anaerobic conditions are possible for the biodegradation of PAHs (Mihelcic and Luthy, 1988; Young and Cerniglia, 1995). Since the research of our group on bioremediation has mostly focused on the aerobic biodegradation conducted by bacteria, the

scope of this review is limited to the bacterial degradation of PAHs that occurs under aerobic conditions.

2.3.1 Bacterial transformation of PAHs

Biotransformation of PAHs could serve three functions: assimilative biodegradation, intracellular detoxification, and co-metabolism (Johnsen et al., 2005). When the function is assimilative biodegradation, microbes mineralize a PAH (or part of it), and use the PAH as a source of carbon and energy for growth. Since PAHs are degraded more completely in this way, it is the ideal situation for biodegradation. However, intracellular detoxification and co-metabolism are less understood, as non-specific metabolic pathways and the interactions between multiple compounds may result in unexpected products (Kazunga et al., 2001; van Herwijnen et al., 2003).

The mechanisms of biotransformation of PAHs have been extensively studied and reviewed (Cerniglia, 1992; Kanaly and Harayama, 2000; Vila et al., 2015). In general, the aerobic bacterial degradation of PAHs is initiated by ring activation, followed by ring fission and processing of the side-chains on the opened ring, leading to successive intermediates with fewer rings and eventual channeling of small molecules to the tricarboxylic acid cycle. While the pathways for ring fission and later steps are relatively fixed, ring activation varies among different microbes and different PAHs. In addition, ring activation enzymes typically can act on multiple PAHs, so that many PAHs can be transformed to common intermediates in a single organism. Such diverse initiation and unified later processes result in a funnel-shaped scheme (**Figure 2.1**) for PAH metabolism (Kweon et al., 2011). The bacteria usually activate PAHs to *cis*-dihydrodiols by intracellular dioxygenases, or to *trans*-dihydrodiols by monooxygenases. The formed *cis*-dihydrodiols are oxidized by a dehydrogenase into catechol-like intermediates, then further cleaved by intradiol or extradiol ring-cleavage dioxygenases (*ortho*- or *meta*- cleavage,

respectively). The resulting intermediates are further oxidized and cleaved again, until they are small enough to be further transformed into the tricarboxylic acid (TCA) cycle via the β -keto adipate pathway (Cerniglia, 1992; Kim et al., 2003).

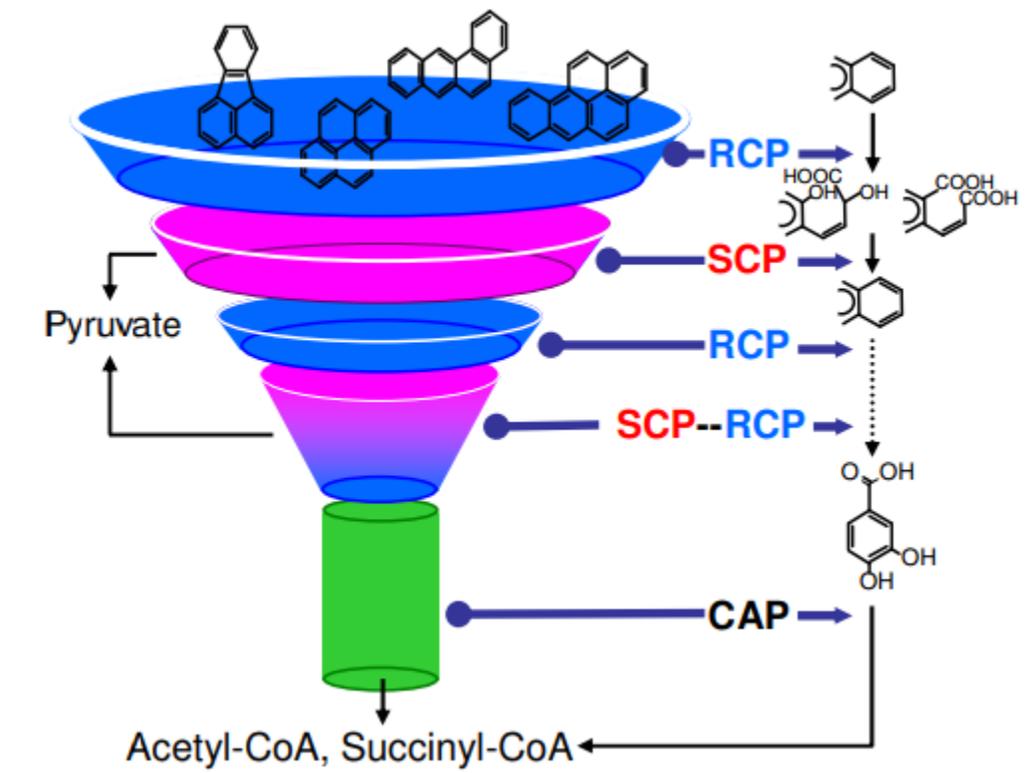


Figure 2.1. Funnel effect in PAH metabolic networks (Kweon et al., 2011). RCP: ring cleavage process; SCP: side chain process; CAP: central aromatic process.

Compared to low-molecular-weight (LMW) PAHs, high-molecular-weight (HMW) PAHs are more persistent in the environment and more difficult to be biodegraded (Cerniglia, 1992). One reason for this is that HMW PAHs are more hydrophobic and less bioavailable, therefore less accessible to the degraders (Johnsen et al., 2005). In addition, the biotransformation pathways required for HMW PAHs are more complex than those for LMW PAHs (Kanaly and Harayama, 2000). Because more aromatic rings need to be broken, the cycle of oxidation and cleavage must be repeated multiple times before smaller intermediates are formed. In fact, although a number of bacterial strains have been shown to grow on two- to four-

ring PAHs (Heitkamp and Cerniglia, 1989), no bacteria have been shown to grow on B[a]P (a 5-ring compound and one of the most carcinogenic PAHs) as a sole carbon and energy source. Studies have suggested that indigenous microbiota can mineralize B[a]P by co-metabolism, if suitable substrates and conditions are provided (Chen and Aitken, 1999; Kanaly and Bartha, 1999).

The biotransformation pathways for PAHs have been suggested as a combination of multiple metabolic networks (Kweon et al., 2011), which could be extremely complex. Ring-hydroxylating dioxygenases (RHD, also called terminal dioxygenase) are the key factor determining if a PAH can be metabolized by a bacterium, because they can oxidize the double bond by introducing two oxygen atoms, forming *cis*-dihydrodiols (Jouanneau et al., 2016; Wammer and Peters, 2006). To activate the aromatic rings of PAHs, an RHD must fit the PAH molecule into its binding cavity, and the bond to be acted on must be in contact with the catalytic center in the cavity. Depending on the PAH, it might fit into a binding cavity in different orientations (Figure 2.2). Since HMW PAHs have larger sizes and diverse shapes, their degraders need to have RHDs with larger binding cavities (Jouanneau et al., 2016). PAH-degrading bacteria usually possess multiple RHDs and can metabolize a myriad of PAHs, so the size and the shape of binding pockets of the enzymes would decide the regioselectivity of the reactions (Jakoncic et al., 2007; Wammer and Peters, 2006). Such diverse initiation of biodegradation leads to complex metabolic networks including numerous pathways, before being narrowed down to the “funnel” in the metabolic network. Furthermore, the interactions between compounds (e.g. co-metabolism) and species (e.g. competition and cooperation) make the big picture even more complicated (Vila et al., 2015).

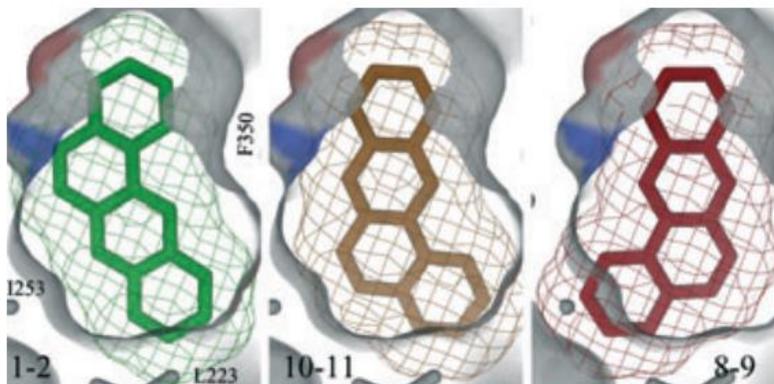


Figure 2.2. Different configurations of benz[*a*]anthracene binding in an RHD (PhnI) (Jakoncic et al., 2007)

2.3.2 Incomplete metabolism of PAHs

As discussed above, the biotransformation of HMW PAHs follows complicated metabolic pathways. Some of the pathways would lead to the complete metabolism of PAHs, which means the PAHs could be mineralized into CO₂ and H₂O. This is the ideal situation for bioremediation because the aim is to reduce the environmental and health risk of a contaminated environment. Some other pathways, however, would lead to the formation of dead-end metabolites instead of the mineralization of PAHs (Andersson et al., 2003; Kazunga and Aitken, 2000; Luan et al., 2007). Two mechanisms might contribute to the incomplete metabolism of PAHs. First, some of the metabolites can originate from co-metabolism with other compounds (Bouchez et al., 1995; van Herwijnen et al., 2003). Since the bacterial enzyme(s) may accidentally metabolize some compounds that do not belong to its own metabolic network, the intermediates might not be accommodated in enzymes further along the pathway (Zhong et al., 2007). Second, some intermediates might be too stable or too toxic for further metabolism. Typical metabolites from incomplete metabolism, like PAH-quinones and lactones, usually show toxicities and inhibitory effects to the microbial community (Bouchez et al., 1999; Kazunga and Aitken, 2000). Although microbial consortia are known to be more efficient at degrading complex contaminants than pure strains, there are still many cases in which dead-end products

accumulated after biodegradation or bioremediation (Lundstedt et al., 2003; Meyer and Steinhart, 2001). One significant problem with this phenomenon is the potential for increased toxicity and genotoxicity in bioremediated soil.

2.3.3 PAH metabolites and increased genotoxicity in contaminated soil

For the biodegradation and bioremediation of PAH-contaminated soil, the main objective is reducing the health and environmental risk. Ideally, soil samples after treatment should be tested for toxicities by bioassays, and compared with their untreated counterparts (Gillespie et al., 2007; Mueller et al., 1991). A number of studies have pointed out that although PAHs were efficiently removed in the bioremediation processes, the genotoxicity could remain or even increase in remediated soil samples (Belkin et al., 1994; Hu et al., 2012; Hughes et al., 1998; Lemieux et al., 2009). The source(s) of increased genotoxicity resulting from bioremediation have not been identified in previous studies (Brooks et al., 1998; Lemieux et al., 2009), although polar compounds formed during treatment have been implicated (Chibwe et al., 2015a). As discussed in the previous section, dead-end products from the incomplete metabolism of PAHs or related compounds are likely to be responsible for the increased toxicity. However, known transformation products from PAHs, such as some commonly observed oxy-PAHs, have been excluded because their concentration decreased after the bioremediation (Chibwe et al., 2015a; Hu et al., 2014). Therefore, to elucidate the source(s) of increased genotoxicity, non-target analysis must be adopted.

2.4. Introduction to non-target analysis

2.4.1 Workflow for non-target analysis

Since complexity is one of the inherent attributes for environmental samples, conventional target analysis of pollutants might be biased by ignoring the potentially high-risk but uncharacterized compounds. The ideal strategy for selecting the contaminants of concern is

to scan everything in the samples of interest, then prioritize and identify the compounds most responsible for adverse effects. Such a strategy is referred to as suspect screening and non-target screening for contaminants (Krauss et al., 2010). Thanks to the developments in soft ionization techniques, high resolution mass spectrometry, and internet-based software and databases, identification of unknown pollutants has become possible in recent years (Hollender et al., 2017). Because a large number of compounds are usually detected in non-selective sampling and analysis, the identification of unknowns could be time- and resource-consuming (Schymanski et al., 2015). Therefore, screening steps like data preprocessing and prioritization, and identification of potential targets are crucial in the workflow, typically between the analysis and the identification. Shown below in Figure 2.3 is an updated workflow for non-target screening of environmental samples.

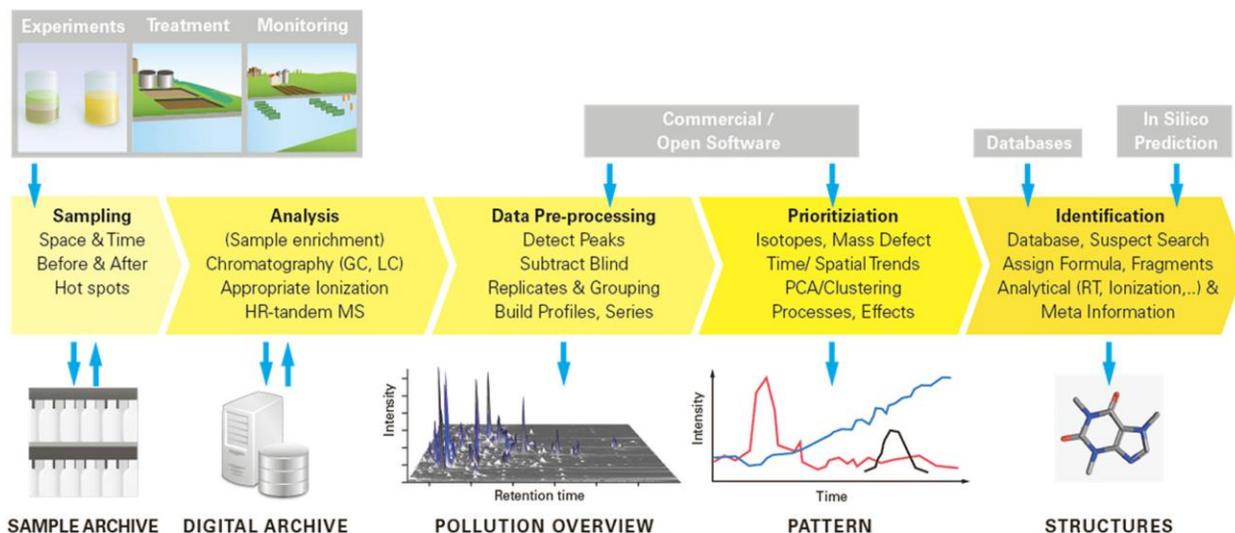


Figure 2.3. Non-target screening workflow for environmental samples (Hollender et al., 2017)

A major principle for the sampling and analysis is to keep the processes non-selective, because the purpose of non-target analysis is to include all possible causes for the adverse effect of interest. For sampling and extraction, the comprehensiveness could be improved by simplifying the extraction steps (Du et al., 2017), or using multifunctional methods such as multi-layer solid phase extraction (SPE)(Kern et al., 2009). It is worth noting that matrix effects

need to be taken into consideration when sample processing is simplified. In the stage of instrumental analysis, it is also important to keep in mind the potential selectivity, as the diversity of compound classes might not be covered by a single technique (Gallampois et al., 2015; Grosse and Letzel, 2007). Combinations of different technologies have been applied to reduce the possible selectivity, such as using both electrospray ionization (ESI) and atmospheric pressure chemical ionization (APCI) in LC-HRMS analysis (Xiao et al., 2016), or analyzing samples by both LC-HRMS and GC-HRMS (Moschet et al., 2017). Other than MS detectors, recent studies on polar contaminants also suggested some new technologies to broaden the range of compounds in LC separation, such as hydrophilic interaction liquid chromatography (HILIC) and mixed-mode liquid chromatography (MMLC) (Bieber et al., 2017; Montes et al., 2017). All these combinations could facilitate more comprehensive sampling, extraction, and analysis on the mixtures, increasing the possibility to identify the responsible toxicants.

Prioritization of non-target pollutants is a challenging task for current workflows. To reduce the number of features from thousands to tens, experiment-driven approaches or data-driven approaches are usually adopted (Hollender et al., 2017). Data-driven approaches include the prioritization based on isotope patterns of hetero atoms (Ballesteros-Gómez et al., 2017; Myers et al., 2014), homologous series (Barzen-Hanson et al., 2017), specific functional groups (Zonja et al., 2015), or difference of intensities among samples (Singh et al., 2015). These are typically faster approaches, but might not be able to narrow down the scope to the specific goals. Experiment-driven approaches, such as effect-directed analysis (Brack et al., 2016), reaction-based screening (Schollée et al., 2015), or isotope labeling experiments (Kolkman et al., 2015), usually consume more time and resources, but they could produce more relevant features.

Another challenge in the development of non-target analysis is the identification and structure elucidation of unknown compounds. Hybrid HRMS instruments like qTOF and qExactive (orbitrap) can provide exact mass, LC retention time, and MS/MS fragmentation

pattern, but such information is not enough for unambiguous assignment of molecular structures. Conventionally, exact structural elucidation requires authentic standards or data from nuclear magnetic resonance (NMR) spectroscopy. However, because of the relatively low concentrations of pollutants and the possibility of “true unknowns”, standards and NMR analysis might not be applicable for all the prioritized compounds (Schymanski et al., 2014a). To deal with the challenge in structural elucidation, various computer-assisted tools have been established, including database search, structure generation, and *in silico* fragmentation (Schymanski et al., 2009). Widely used chemical databases such as PubChem, ChemSpider, and SciFinder contain a large number of compounds that could serve as candidates for a selected list, and specialized databases like Metlin, mzCloud, MassBank, and EPA CompTox Chemistry Dashboard have more specialized information for HRMS data. *In silico* fragmentation could also help with the identification for the “known unknowns” (Ruttkies et al., 2016), while structure generation and pathway prediction could shed light on the structure of “true unknowns” (Huntscha et al., 2014).

In many cases, the structures of unknowns could not be confirmed because of the number of possible isomers and the lack of authentic standards (Krauss et al., 2010; Schymanski et al., 2009). A widely accepted and cited paper (Schymanski et al., 2014a) has summarized the levels of confidence for the identification of unknowns (Figure 2.4). Level 1 is suggested as compounds confirmed with a reference standard, while the lower levels of confidence apply to tentatively identified structures based on different levels of evidence.

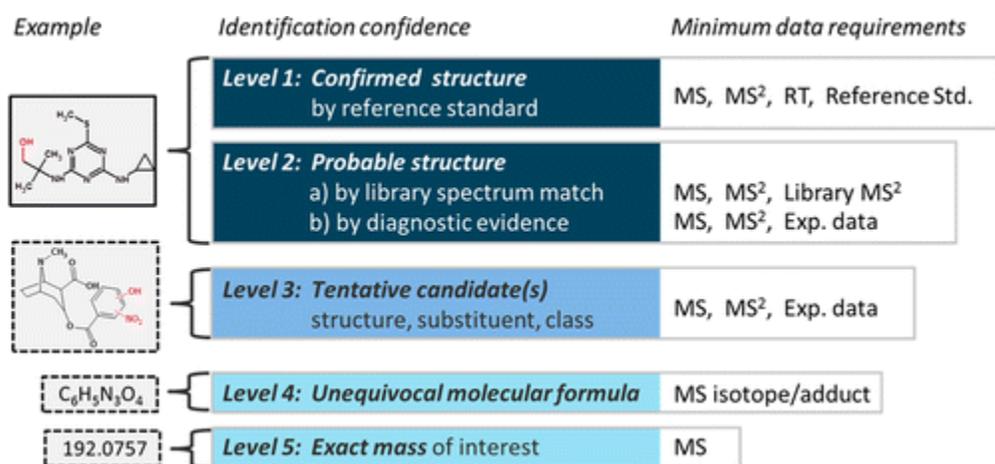


Figure 2.4. Levels of confidence for the identification of unknowns (Schymanski et al., 2014a).

2.4.2 Effect-directed analysis

Effect-directed analysis (EDA) is a well-developed method for identifying the contaminants responsible for observed toxicity in environmental mixtures. EDA can be classified as a subset of non-target analysis, belonging to the experiment-driven prioritization (Hollender et al., 2017). In EDA, a bioassay and chromatographic fractionation are coupled to identify biologically active compounds. The complexity of environmental samples is reduced by fractionation through low- to medium-resolution chromatography, and the resulting fractions are subject to the bioassay of interest to identify those fractions with the highest activity for additional separation by high resolution techniques (Brack et al., 2016; Burgess et al., 2013). In recent years, advanced bioassays allow environmental scientists to trace the origin of different toxicological behaviors, and the development of technology in analytical chemistry, especially HRMS and high performance liquid chromatography (HPLC), enables more comprehensive screening and accurate identification of bioactive compounds. EDA has been successfully applied to various environmental media such as sediment (Brack, 2003; Weiss et al., 2009; Xiao et al., 2016), soil (Legler et al., 2011), house dust (Fang et al., 2015b; Ouyang et al., 2017), and

atmospheric particles (Durant et al., 1998), and a variety of known and previously unknown compounds have been identified as toxicants.

2.4.3 Metabolite profiling

Although EDA can reduce the complexity of environmental samples and prioritize the toxic fractions, it could still be hard to prioritize a compound or compound class in the fractions. Combining EDA with additional statistical analysis or modeling may help identify unknown toxicants (Schollée et al., 2015; Weiss et al., 2011). Metabolite profiling is a data-driven prioritization strategy that is suitable for non-target analysis on processes or paired samples (Singh et al., 2015).

In systems biology, metabolomics represents the study of the complete small molecule profile of an organism. Metabolite profiles of the perturbed and unperturbed system are statistically compared and features that are significantly different are identified in an effort to understand the mechanism producing the change (Fiehn, 2002). MS and NMR are the two major analytical approaches for metabolomics. NMR-based metabolomics could provide more detailed information on the structure of metabolites, and has better reproducibility (Markley et al., 2017). However, MS-based metabolomics is predominantly popular in recent years, attributed to its high sensitivity and high throughput (Dettmer et al., 2007). In MS-based metabolomics, both GC-MS and LC-MS have been widely used, and each of them has special advantages. On one hand, GC-MS is more frequently used in the field of plant metabolomics, and it has better chromatographic resolution and more-comprehensive databases (Jonsson et al., 2005). LC-MS, on the other hand, is more amenable to soft ionization methods (electrospray ionization (ESI), atmospheric pressure chemical ionization (APCI), and atmospheric pressure photoionization (APPI)) that produce molecular ions, and is free of the potential interference from derivatization and thermal degradation associated with GC-MS (Fang et al., 2015a).

In environmental sciences, there are many similar cases where non-target analysis could be applied on pairs of samples, like the mixtures before and after a certain process, or biological samples with and without exposure. A couple of studies have already adopted the concept of metabolite profiling, and achieved inspiring results. Rotander et al. (2015) used metabolite profiling to compare fire fighters and a control group, and found exposure to fluorinated surfactants in firefighters (Rotander et al., 2015). Singh et al. (2015) applied metabolite profiling on the advanced oxidation process, and identified transformation products of iopromide and iopamidol (Singh et al., 2015). Metabolite profiling has also been applied to the biotransformation of pollutants in different processes, and numerous transformation products were identified (LeFevre et al., 2015; Men et al., 2016; Schollée et al., 2015).

Metabolite profiling is a data-driven prioritization strategy, while EDA is an experiment-driven prioritization strategy (Hollender et al., 2010). These two methods can reduce the complexity of the mixture and prioritize the potential toxicants from different dimensions. Therefore, coupling EDA to metabolite profiling can efficiently identify biologically active compounds in complex samples. In fact, such an approach has been used in the field of drug discovery (Kellogg et al., 2016), but until this dissertation research had not been applied in the field of environmental sciences.

2.4.4 Stable isotope-assisted metabolomics

Because of the similarity between non-target analysis and metabolomics, many rationales, methods, and workflows can be learned from the well-developed field of metabolomics. As noted above, PAH metabolites from bioremediation may be toxic or genotoxic (Lemieux et al., 2015; Tian et al., 2017a), and some could be dead-end products that prevent further biodegradation (Kazunga et al., 2001; Schmidt et al., 2010), so it is important to get information on their identity and formation pathway. Although many metabolites and

transformation products of PAHs have been identified, there is a gap between these known transformation products and numerous unknowns. Most of the known metabolites are identified in experiments where pure microbial strains transform pure compounds, and such conditions may be very different from real environmental samples. In real contaminated samples, the PAHs usually exist in complex mixtures, and they are biodegraded or biotransformed by microbial consortia that consist of various species (Boonchan et al., 2000; Luo et al., 2014; Zhong et al., 2011). The cometabolism of some PAHs in the mixture could enhance or inhibit the formation of some metabolites (Bouchez et al., 1995; Zhong et al., 2010), or produce metabolites that would not be present in a single-compound incubation (Hennessee and Li, 2016). Also, the cooperation among microbial strains could increase the diversity of metabolites from PAHs (Zhong et al., 2011). With such complex interactions, it is difficult to identify PAH metabolites in real environmental samples, and even harder to link them to their parent compounds.

Similar challenges exist in the field of metabolomics. HRMS-based metabolomics allows the detection of hundreds or thousands of metabolites in a single run, but it is hard to trace metabolites back to their parent compounds without knowing the metabolic pathways. To solve this issue, methods of stable isotope-assisted metabolomics (SIAM) have been established by different groups (Giavalisco et al., 2008; Guo and Li, 2009; Hegeman et al., 2007), and various automatic workflows have been developed to facilitate these methods (Bueschl et al., 2014; Huang et al., 2014; Leeming et al., 2015; Zhou et al., 2014). The basic idea is that when stable isotope-labeled substrates are metabolized by organisms, their metabolites would also be isotopically labeled, and those could be distinguished in groups with unlabeled metabolites by HPLC-HRMS. These groups confirm the relationship between metabolites and their parent compounds, and provide information on their molecular structure (Neumann et al., 2014). Algorithms enable such pair or group selection by picking up features with 1) the same retention time in chromatography, 2) certain mass differences and 3) diagnostic intensity ratios (Kluger et

al., 2014). The advances of stable isotope-assisted metabolomics in drug development and plant research have been reviewed recently (Fan et al., 2012; Freund and Hegeman, 2017).

The application of SIAM to biodegradation in contaminated soil would provide new insights on the biotransformation of pollutants. Some studies have suggested the use of ^{13}C -labeled PAHs to see their distribution and transformation (Morasch et al., 2011; Richnow et al., 1998, 2000), but due to the availability of only partially labeled compounds commercially and/or the resolution of mass spectrometry, previous attempts at metabolite identification were restricted. In recent studies, isotope labeling combined with HRMS has been applied to track the abiotic transformation products of pollutants (Kolkman et al., 2015; Sun et al., 2016). However, because traditional data analysis and partially labeled pollutants were applied, the screening ranges of products was limited. To my knowledge, SIAM has never been applied in the field of bioremediation.

CHAPTER 3 NON-TARGET ANALYSIS REVEALS A BACTERIAL METABOLITE OF PYRENE IMPLICATED IN THE GENOTOXICITY OF CONTAMINATED SOIL AFTER BIOREMEDIATION¹

3.1 Introduction

Polycyclic aromatic hydrocarbons (PAHs) are a ubiquitous class of soil contaminants found in high concentrations at industrial sites associated with the former production of manufactured gas, wood preservation, and petroleum processing (Neilson, 1995). They are constituents of highly complex mixtures in the primary contamination sources such as coal tar or creosote (Andersson and Achten, 2015; Neilson, 1995), although only 16 PAHs are regulated by the U.S. Environmental Protection Agency (EPA) and many international agencies (Andersson and Achten, 2015). Bioremediation is an accepted technology for PAH-contaminated soil (US EPA, 2015b), but like other remediation methods its efficacy is evaluated only on the basis of removal of the regulated compounds rather than on broader measures of risk. For example, bioremediation of PAH-contaminated soil can have little effect on (Gillespie et al., 2007; Mueller et al., 1991), or actually increase (Belkin et al., 1994; Chibwe et al., 2015a; Hu et al., 2012; Hughes et al., 1998; Lemieux et al., 2009) the soil's genotoxicity, despite extensive removal of the EPA-regulated PAHs.

The source(s) of increased genotoxicity resulting from bioremediation have not been identified in previous studies (Brooks et al., 1998; Chibwe et al., 2015a; Lemieux et al., 2009),

¹Zhenyu Tian , Avram Gold, Jun Nakamura, Zhenfa Zhang, Joaquim Vila, David R. Singleton, Leonard B. Collins, and Michael D. Aitken
Published in Environmental Science & Technology 2017, 51 (12): 7091–7100

JN guided the DT40 bioassay experiments and assisted with data interpretation, ZZ synthesized the reference standards and together with AG helped with structure elucidation, JV and DRS helped with microcosm incubations, and LBC helped with HRMS analysis.

although polar compounds formed during treatment have been implicated (Chibwe et al., 2015a). Prior efforts have focused on possible microbial metabolites, such as oxygenated PAHs (oxy-PAHs), that can cause DNA damage (Zielinska-Park et al., 2004) or exhibit mutagenicity (Lundstedt et al., 2007). However, some commonly observed oxy-PAHs and known bacterial transformation products are not mutagenic (Lundstedt et al., 2007; Park et al., 2008) and have been observed to decrease during bioremediation when increases in genotoxicity have been observed (Chibwe et al., 2015a; Hu et al., 2014).

In this study, we used a non-target approach to investigate extracts of contaminated soil from a former manufactured-gas plant (MGP) site whose genotoxicity has been shown to increase following treatment in a lab-scale, aerobic bioreactor (Chibwe et al., 2015a; Hu et al., 2012). The approach involved a two-tiered separation of soil extracts (pre- and post- treatment) in an effect-directed analysis (EDA) (Brack, 2003; Vughs et al., 2016; Yue et al., 2015) based on the DT40 DNA damage-response assay. A semi-polar fraction of the post-treatment extract exhibiting the greatest cytotoxicity was compared to the corresponding fraction of the pre-treatment extract using metabolite profiling, a method originating in metabolomics (Fernie et al., 2004) and more recently used in environmental analysis (LeFevre et al., 2015; Men et al., 2016; Schollée et al., 2015; Singh et al., 2015; Strynar et al., 2015), which combines high-resolution analytical techniques with rigorous statistical analysis. Coupling EDA to metabolite profiling can efficiently identify biologically active compounds in complex samples, as has been shown in the field of drug discovery (Kellogg et al., 2016), but it has had limited application in environmental studies.

3.2 Materials and methods

3.2.1 Chemicals

PAH standards (EPA 610 PAH mixture), silica gel (high purity, 70-230 mesh), pyrene, methyl methanesulfonate (MMS), dimethylsulfoxide (DMSO), and phosphate buffer solution (PBS) were obtained from Sigma-Aldrich (St. Louis, MO, U.S.A.). Anhydrous sodium sulfate and HPLC-grade solvents, including n-hexane, dichloromethane (DCM), acetone, ethyl acetate (EA), and methanol, were purchased from Fisher Scientific (Pittsburgh, PA, U.S.A.).

5*H*-naphtho[8,1,2-*cde*]chromen-5-one (4-oxapyrene-5-one) and 4*H*-cyclopenta[*def*]phenanthrene-8,9-dione were synthesized according to previously published methods (Gillis and Porter, 1989; Trost and Kinson, 1975). The authentic standard of 2*H*-naphtho[2,1,8-*def*]chromen-2-one (NCO) was synthesized *de novo* and characterized by high-resolution mass spectrometry (HRMS), ¹H nuclear magnetic resonance (NMR) spectroscopy, and 2-dimensional NMR methods as described in Appendix A. Details of the synthetic scheme and characterization of intermediates will be published in a forthcoming manuscript.

3.2.2 Soil and bioreactor treatment

PAH-contaminated soil treated in the bioreactor was collected from a former manufactured gas plant site in Salisbury, North Carolina, USA. Detailed information on collection, processing and characterization of the soil can be found in a previous publication (Hu et al., 2012). The concentrations of target PAHs (14 of the 16 EPA-regulated PAHs, excluding acenaphthylene and indeno-[1,2,3-*cd*]pyrene) were quantified by HPLC with fluorescence detection (Richardson et al., 2011) and are shown in Table A1.

The bioremediation process was achieved in a laboratory-scale (2 L), semi-continuous, aerated bioreactor containing a slurry of soil in an aqueous buffer (15% w/w). Every week, 20% of the treated slurry was removed from the bioreactor and replaced with new untreated (feed) soil

suspended in fresh buffer (pH 7.5) containing 5 mM phosphate and 2.5 mM ammonium nitrate (“bioreactor buffer”). Bioreactor-treated soil was recovered from the slurry by centrifugation at 3900 rpm for 20 min.

3.2.3 Extraction and Fractionation

Soil samples were extracted according to the method described in SI. Extracts from both feed and bioreactor-treated soils were fractionated using the two-level EDA method described in Figure A1. In the first level of fractionation, extracts were separated into broad compound classes by column chromatography using silica gel, according to the method of Bandowe et al (Bandowe et al., 2010). The second level of fractionation was achieved with solid-phase extraction (SPE) cartridges containing cyanopropyl-bonded silica (SiliCycle, QC, Canada). Details on the fractionation are provided Appendix A.

3.2.4 Bioassay

To avoid toxicity effects related to the extraction or elution solvents, extracts were dried and re-dissolved in DMSO prior to the bioassay. Aliquots of fractions and subfractions were evaporated in pre-weighed vials, and dry masses of residues were determined on an analytical balance. Mass balances of residues in the fractions were calculated according to the dry masses and the proportions relative to the mass of residue in the whole extract. The residues were then dissolved in DMSO under sterile conditions and stored at -80 °C until analysis.

Toxicity and genotoxicity of the soil extracts and fractions were evaluated using a 96-well plate-based DT40 chicken lymphocyte DNA-damage response assay (Hu et al., 2012; Ji et al., 2009). The DT40 isogenic cell line, and its mutants knocked out in specific DNA repair pathways, were exposed to residues dissolved in DMSO and serially diluted with PBS. The concentration of DMSO was adjusted to 0.3% for all cell exposures. To verify the consistency of

cells across different batches, methyl methanesulfonate (MMS) was used as positive control while the vehicle blank (DMSO diluted in PBS) was used as negative control.

LD₅₀ values were calculated by fitting the dose-response curve (survival percentage versus log concentration) using GraphPad Prism 6.05 for Windows. The mass of residue in each dose was converted to the equivalent mass of soil from which that mass would have been obtained (Hu et al., 2012) for the purpose of comparing doses between samples. The LD₅₀ for the DT40 parental cell line represents the cytotoxicity, and the relative LD₅₀ (LD₅₀ of mutant cell line divided by LD₅₀ of parental cell line) was calculated as the measure of genotoxicity (Adrion et al., 2016; Hu et al., 2012).

3.2.5 Instrumental analysis

For the selected fractions, analysis was performed using an Agilent 1100 series HPLC coupled to a 6520 series high-resolution quadrupole time-of-flight MS (HR-QToF-MS) equipped with an electrospray ionization (ESI) source (Agilent Technologies, Palo Alto, CA, USA). HPLC-MS analysis using an atmospheric pressure photoionization (APPI) source was performed on a Thermo TSQ Quantum Ultra triple quadrupole mass spectrometer (QqQ MS; Thermo Fisher Scientific). Details of HPLC-MS methods are described in Appendix A.

Quantification of NCO was accomplished on the triple quadrupole mass spectrometer using the synthetic compound as an analytical standard. The samples were analyzed in selected reaction monitoring (SRM) mode using the transition m/z 221.06 ([M+H]⁺) to m/z 165.06 ([M+H-2(CO)]⁺) and the optimized collision energy, 29 V.

3.2.6 Purification of NCO and structure identification

To obtain enough material for NMR analysis, 150 g (dry weight) of bioreactor-treated soil samples were extracted and combined for purification. NCO formed during bioreactor treatment was purified by a three-level separation method: low-pressure silica gel-column

chromatography, medium-pressure column chromatography, and HPLC. Details of the purification and structural identification are provided in Appendix A.

3.2.7 Other biotransformation experiments

To demonstrate that the isolated product was derived from pyrene, microcosm incubations were performed with the bioreactor-treated slurry in the presence of pyrene. To confirm that NCO is a bacterial metabolite, transformation of pyrene by selected bacterial strains was evaluated in resting-cell incubations. To test the potential prevalence of NCO formation in PAH-contaminated soils, microcosm incubations were performed with soil samples from two independent and geographically distant PAH-contaminated sites: creosote-contaminated soil from a wood-treatment facility in Andalucía, Spain, which had a 100-year history of soil pollution (Tejeda-Agredano et al., 2013) and soil from the Holcomb Creosote Superfund Site in Yadkinville, North Carolina, USA. Incubation details are provided in Appendix A.

3.2.8 Data analysis

The HPLC-QToF profiles of extract fractions from feed soil and bioreactor-treated soil were compared with the metabolomics platform XCMS Online (<https://xcmsonline.scripps.edu/>). Data for the treated and feed soil samples were set as the test and control groups, respectively. Data files from three replicate samples were obtained for each group, and the input parameters were based on the default method optimized for HPLC-QToF (details in Appendix A). It has been shown that three replicates are sufficient for metabolite profiling of complex samples that have been simplified through EDA (Kellogg et al., 2016). Independent *t*-tests (Welch's *t*-tests) were used to test for statistically significant differences of features between the two groups (feed and treated soils).

The R language and environment for statistical computing (version 3.2.3) was used for analysis and hypothesis testing on toxicological data. Paired *t*-tests were employed to test for

statistically significant differences of LD₅₀ or relative LD₅₀ values between samples. Significance level was set at $\alpha = 0.05$.

3.3 Results and discussion

3.3.1 Fractionation of soil extracts and selection of genotoxic fraction

Two-level fractionation of soil extracts was directed by cytotoxicity and genotoxicity measured with the DT40 DNA damage-response bioassay, which is based on DT40 isogenic chicken lymphocytes and mutants that are each knocked out in a specific DNA repair mechanism (Ji et al., 2009). The *Rad54*^{-/-} and *Rev1*^{-/-} mutants were selected for this study because of their sensitivity to the increased genotoxicity of bioreactor-treated soil (Hu et al., 2012). The *Rad54*^{-/-} mutant is deficient in Rad54 proteins and therefore is sensitive to genotoxicants that cause double-strand breaks and replication-fork blockage. The *Rev1*^{-/-} mutant is deficient in Rev1 proteins, resulting in vulnerability to chemicals causing DNA depurination or stable covalent adducts.

The first level of fractionation using deactivated silica gel is a classical adsorption-based method to separate the extracts into broad classes (Bandowe et al., 2010), in this case a nonpolar fraction (containing compounds such as PAHs and alkylated PAHs), a semi-polar fraction (containing compounds such as oxy-PAHs and azaarenes), and a polar fraction (containing acids and phenols) designated A, B, and C, respectively. As indicated in Figure A1, mass balance was achieved in the separation, with the greatest mass in Fraction B. Complete recovery of the toxic compounds was demonstrated by comparison of the cytotoxicity and genotoxicity of recombined fractions and whole-soil extracts (Figures A2 a and b, respectively). Consistent with earlier work (Chibwe et al., 2015a; Hu et al., 2012), whole-soil extracts indicated that the bioreactor-treated soil was more genotoxic than the untreated feed soil (Figure A2b).

Fraction C did not cause toxicity and is not discussed further. Nonpolar fraction A from the treated soil had a significantly higher LD₅₀ than for feed soil for all three cell lines (Figure 3.1a), suggesting detoxification of the nonpolar compounds; this would be consistent with removal of the PAHs in the bioreactor (Table A1). Semi-polar fraction B showed the opposite trend (Figure 3.1b); the LD₅₀ for all three cell lines was significantly lower for treated soil than for feed soil, indicating greater toxicity of treated soil. The relative LD₅₀ values of fraction B, however, were not significantly different between treated soil and feed soil (Figure A3a). Given that fractionation did not appear to influence the recovery of cytotoxic or genotoxic compounds when the fractions were recombined (Figures A2 a and b), we inferred that the high cytotoxicity of the semi-polar compounds in fraction B for all three cell lines could have masked a genotoxic effect. On this basis, we selected fraction B for further fractionation.

The second level of fractionation on fraction B was achieved by cyanopropyl solid-phase extraction (SPE), which retains semi-polar to polar compounds based on partitioning and has been used previously to separate mutagens without affecting activity (Durant et al., 1998). Three subfractions (B1, B2, and B3) were obtained for each soil (Figure A1). Subfractions B1 from each sample (feed and treated soil) accounted for the greatest dry mass of residue, but the cytotoxicity in treated soil was lower than in feed soil (Figure 3.1c). Subfractions B3 did not present any detectable toxicity for either soil. For subfraction B2, treated soil had significantly lower LD₅₀ than feed soil (Figure 3.1d), implying greater cytotoxicity. The relative LD₅₀ values for treated soil and feed soil were not significantly different at $\alpha = 0.05$ (Figure A3b); nevertheless the relative LD₅₀ values indicate genotoxicity. As a result, subfractions B2 were selected for metabolite profiling.

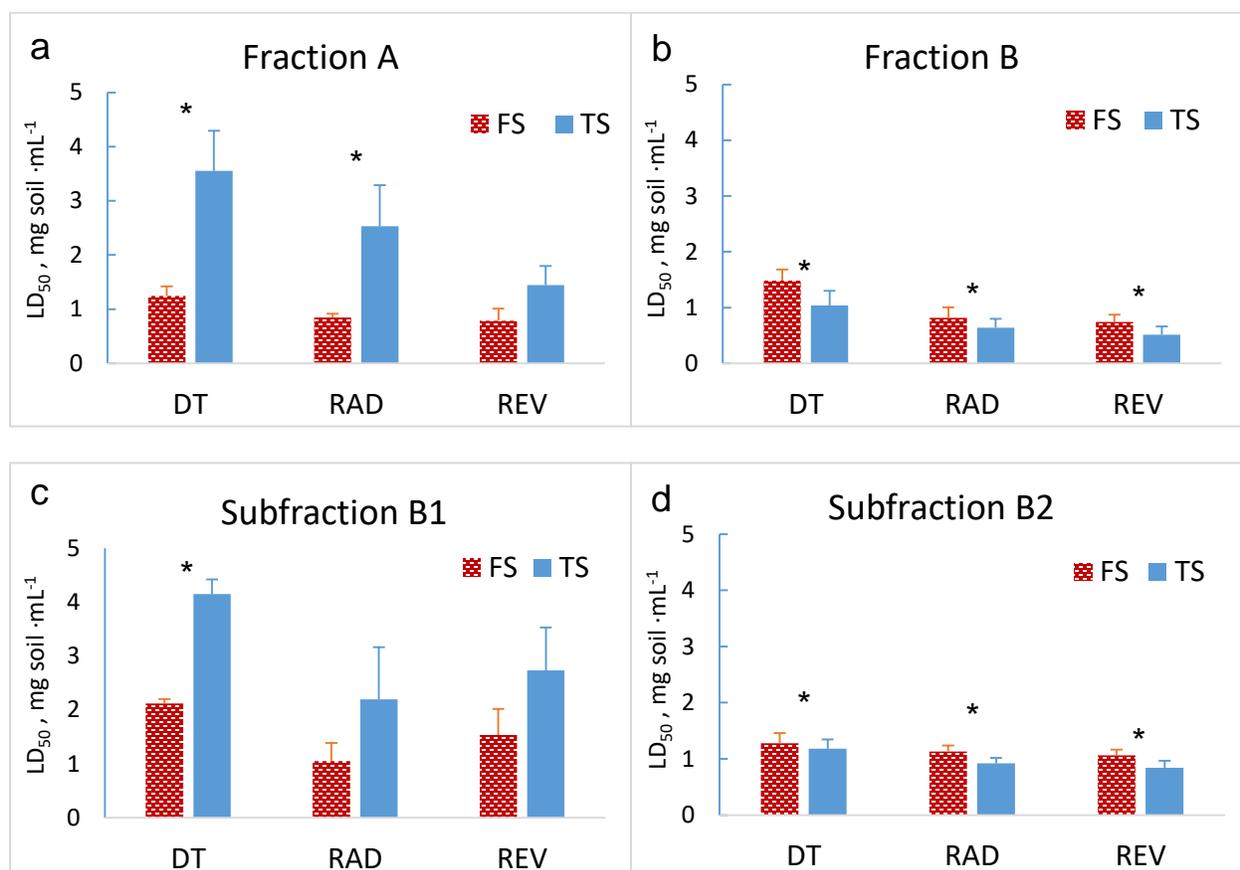


Figure 3.1. Results of DT40 DNA-damage response bioassays on solvent extracts and fractionated extracts from PAH-contaminated soil used to feed a lab-scale bioreactor (“feed soil”, FS) and bioreactor-treated soil (TS). Assays were conducted with the parental DT40 cell line and two mutant cell lines deficient in different DNA-repair mechanisms, *Rad54*^{-/-} and *Rev1*^{-/-}. (a and b) LD₅₀ values from the two fractions exhibiting measurable cytotoxicity (over the dose range tested) after the first level of extract fractionation. (c and d) LD₅₀ values from the two subfractions of primary fraction B that exhibited cytotoxicity. Error bars represent standard deviations from replicates at each dose (n = 5). Asterisks indicate a significant difference between feed soil and bioreactor-treated soil for a given cell line. Plots of relative LD₅₀ values corresponding to data in panels a-d are in Figure A2.

3.3.2 Metabolite profiling

Metabolite profiling was achieved by HPLC-MS and data analysis with XCMS Online, a widely used metabolomics platform (Tautenhahn et al., 2012). Assuming that the compounds responsible for increased genotoxicity would be significantly more abundant in bioreactor-treated soil than in feed soil, we focused on what the software categorizes as “upregulated” analytes and infer that they were formed in the bioreactor from constituents of feed soil.

Although these analytes were likely to be transformation products of soil constituents, they would not necessarily be microbial metabolites.

Because the sensitivity of different ionization sources in mass spectrometry varies for different compound classes and functional groups (Grosse and Letzel, 2007; O'Connell et al., 2013), we employed electrospray ionization in positive (ESI+) and negative (ESI-) ionization modes, and atmospheric pressure photoionization in positive (APPI+) and negative (APPI-) ionization modes. The composition of ions of interest was determined from exact mass measurements on the HR-QToF-MS instrument. Few upregulated ions, none with notable intensities, were detected by ESI operated in the negative ionization mode or by APPI in negative ionization mode. The greatest number of upregulated features was observed in the ESI+ mode, a small subset of which were also detected by APPI+ (Table A3). As a result, we focused on the ions detected in ESI+ mode for further investigation.

Of the ions detected in ESI+ mode, XCMS Online prioritized 146 with absolute intensity ≥ 1000 and fold-change ≥ 1.5 , with $p \leq 0.01$ (Figure 3.2a); 54 of them had higher concentrations in bioreactor-treated soil ("upregulated"). Application of more-restrictive criteria (absolute intensity ≥ 5000 and fold-change ≥ 2 , $p \leq 0.01$) yielded 14 significantly upregulated ions (Table A2). Two of the upregulated ions can be attributed to the neutral composition $C_{15}H_8O_2$: m/z 221.0595, $[M+H]^+$ (Figure 3.2b) and m/z 243.0416, $[M+Na]^+$. A third ion, at m/z 463.0934, can be assigned as the sodium adduct of the dimer, $[2M+Na]^+$. Intensity at m/z 235.0753, corresponding to the composition $C_{16}H_{10}O_2$, was resolved into three ions in the extracted ion chromatogram (EIC; Figure 3.2c); because two of the three ions co-eluted, only two were distinguished by the XCMS Online software (Table A2). The composition $C_{16}H_{10}O_2$ corresponds to addition of a methyl group to the parent $C_{15}H_8O_2$ framework, suggesting the presence of three methyl congeners of the $C_{15}H_8O_2$ parent. Similarly, an upregulated ion at m/z 249.0912, with the composition $C_{17}H_{12}O_2$, may represent either a dimethyl or ethyl congener of $C_{15}H_8O_2$.

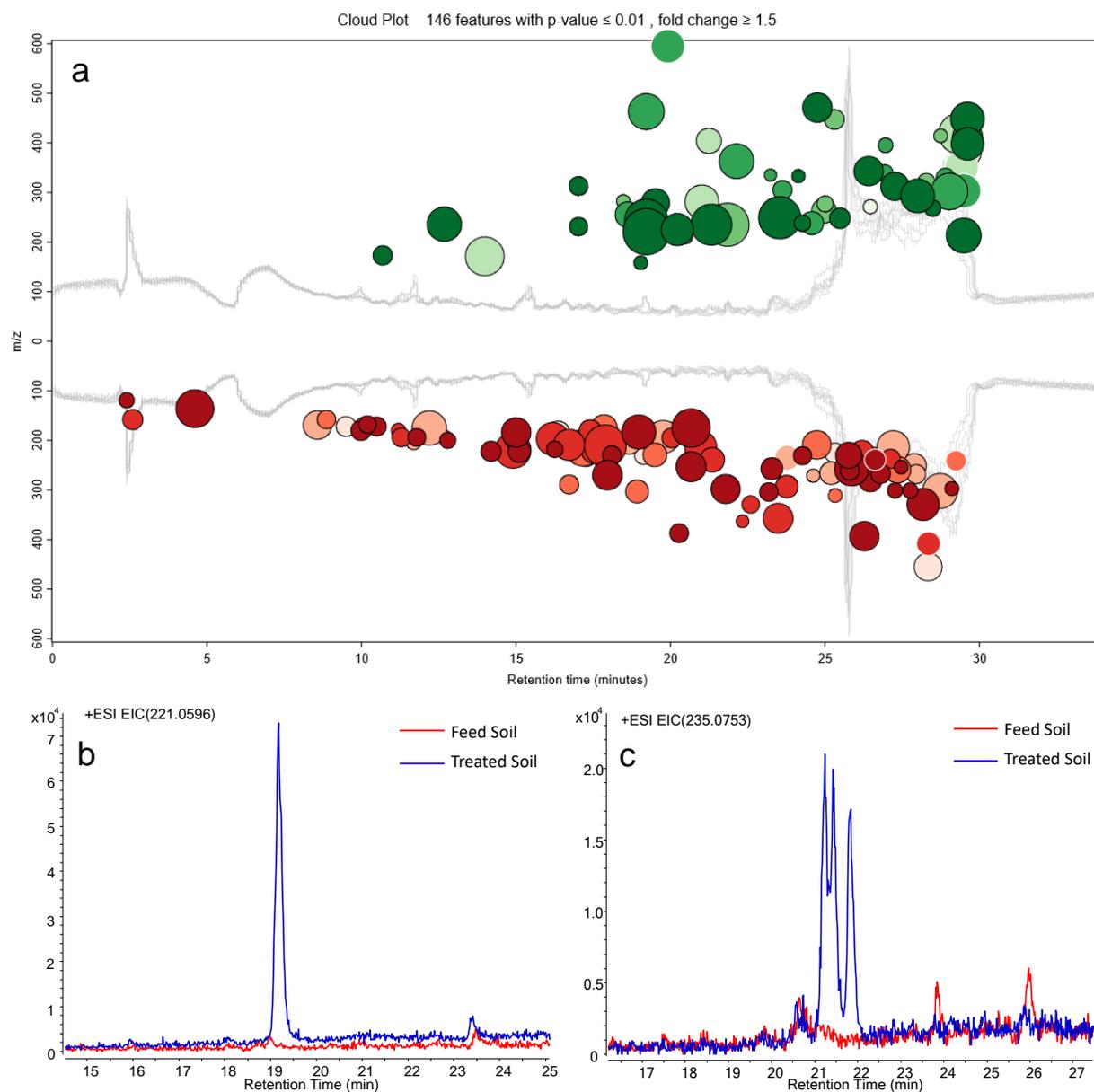


Figure 3.2. Non-target analysis of extract subfractions B2. (a) Cloud plot from metabolite profiling comparing the high-resolution mass spectra (ESI+ mode) of subfraction B2 from bioreactor-treated soil to subfraction B2 from feed soil. Mass spectral features with significantly higher intensities in the treated soil than in the feed soil (upregulated) are in green and those significantly lower in the treated soil are in red. Bubble size is proportional to peak intensity; the darker the shading, the higher fold change. All 146 features with intensity $\geq 1,000$, fold-change ≥ 1.5 in either direction, and $p \leq 0.01$ are shown. (b) and (c) Extracted ion chromatograms for m/z 221.0596 and m/z 235.0753, respectively.

Based on the summed intensities of the protonated and sodiated adduct ions, $C_{15}H_8O_2$ was the most abundant of the upregulated compounds in the bioreactor-treated soil extract, with the mono- and dimethylated (or ethylated) congeners also present at high levels. This group of

compounds showed exclusively high values in fold-change between feed soil and bioreactor-treated soil, 861 for C₁₅H₈O₂ and 196-347 for the remaining compounds (Table A2). Among the limited number of upregulated ions detected in APPI+ mode that met the input criteria, *m/z* 221.06 and *m/z* 235.08 were also prominent (Table A3), confirming that these compounds were not artifacts of the ionization source. In a preliminary EDA study on our bioreactor-treated soil using one level of fractionation (Chibwe et al., 2015a), genotoxic fractions subjected to non-target analysis by HPLC/(ESI+)-MS were also observed to contain analytes corresponding to C₁₅H₈O₂ and C₁₆H₁₀O₂, although the compounds were not identified (Chibwe et al., 2017).

The relevance of the putative biotransformation product C₁₅H₈O₂ to our previous studies on increased genotoxicity of treated soil (Hu et al., 2012) was determined by analysis of eight archived extracts of bioreactor-treated soil samples collected over a four-year period. The compound and its methylated homologs were detected in all samples at levels comparable to those in this study. Consequently the most abundant product, C₁₅H₈O₂, was considered the most likely source of increased genotoxicity resulting from bioreactor treatment of the soil and was targeted for structural characterization.

3.3.3 Isolation and structural characterization of C₁₅H₈O₂

The fragmentation pattern obtained from MS/MS of the parent ion at *m/z* 221.0595 was characterized by a major product ion at *m/z* 165.0691 (rel. intensity 100), corresponding to [M+H-2(CO)]⁺, and a second product ion at *m/z* 193.0640 (rel. intensity ~0.3) corresponding to [M+H-(CO)]⁺ (Figure A4f). Such a fragmentation pattern is typical of PAH-derived quinones and lactones (Gallampo et al., 2015; O'Connell et al., 2013).

Based on current knowledge of PAH biotransformation (Kim et al., 2007; Lundstedt et al., 2003; Vila et al., 2001) and database searches (SciFinder and ChemSpider), two structures were deduced as candidates for C₁₅H₈O₂. 5*H*-naphtho[8,1,2-*cde*]chromen-5-one (also called 4-

oxapyrene-5-one; Figure A4a) has been found in particulate matter from diesel exhaust and verified by comparison with an authentic standard (Pitts et al., 1982). Lundstedt et al. (Lundstedt et al., 2003) reported that 4-oxapyrene-5-one accumulated in an aerobic bioreactor treating a creosote-contaminated soil; to our knowledge, however, this assignment was based solely on mass spectrometry, which does not allow a definitive structural assignment in the absence of an authentic standard. We synthesized 5*H*-naphtho[8,1,2-*cde*]chromen-5-one and definitively ruled it out as the C₁₅H₈O₂ product observed in this study by HPLC/HR-QToF-MS (Figure A4).

A second known structure with the composition C₁₅H₈O₂ is 4*H*-cyclopenta[*def*]phenanthrene-8,9-dione, a quinone derived from 4*H*-cyclopenta[*def*]phenanthrene. We also synthesized this compound and while the MS/MS spectrum was similar to that of the target product in this study (Figure A4d), it was ruled out based on a different HPLC retention time. Since the target product did not correspond to any known structure, structural assignment required isolation and characterization by an unambiguous technique. Hence, we purified a sufficient quantity for NMR studies from extracts of bioreactor-treated soil using bulk extraction and three-level chromatographic separation (SI).

The 1D ¹H NMR spectrum (Figure A5) along with 2D homonuclear coherence (COSY; Figure A6c), heteronuclear multiple bond coherence (HMBC, Figure A6a), and heteronuclear single quantum coherence (HSQC; Figure A6b) spectra of the purified compound were acquired. The expected eight proton signals observed in the ¹H NMR spectrum appeared as two pairs of doublets and two doublets coupling into a triplet in the aromatic region along with a singlet at higher field than generally observed for aromatic protons. As inferred from the ¹H NMR spectrum, the ¹³C NMR spectrum indicated 8 carbon signals with intensities expected for H-bearing carbons and 7 low-intensity signals characteristic of quaternary carbons. Importantly, only one quaternary carbon signal appeared at a chemical shift associated with carbonyl carbon (162 ppm), ruling out a quinone structure. A second quaternary carbon signal at 154 ppm

appeared in a region characteristic of an aromatic carbon bearing an oxygen atom. These two observations were consistent with a PAH-derived lactone structure. The COSY spectrum confirmed two sets of coupled proton signals with two additional protons coupled into the triplet. This pattern is typical for the K-region and apical ring signals of pyrene. The pattern of proton signals suggested *2H*-naphtho[2,1,8-*def*]chromen-2-one (NCO; Figure 3.3 inset), an apparent pyrene-derived structure in which an apical ring had been oxidized to a lactone.

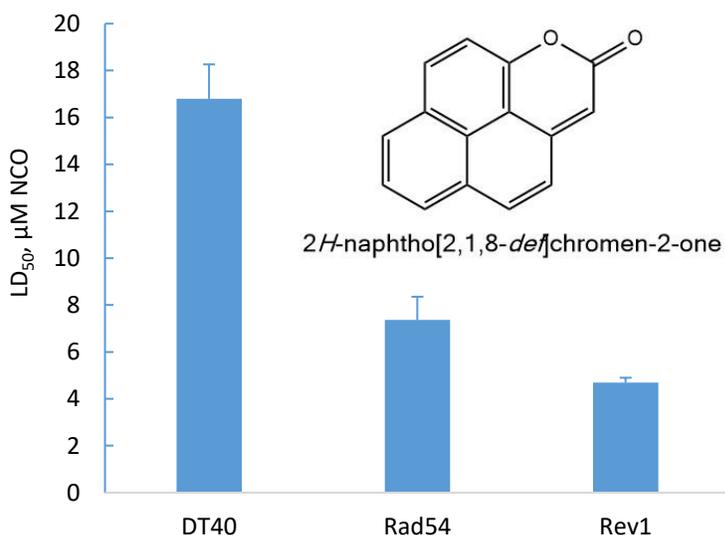


Figure 3.3. DT40 bioassay results on *2H*-naphtho[2,1,8-*def*]chromen-2-one (NCO, inset) purified from extracts of bioreactor-treated soil.

Connectivities observed in the HSQC and HMBC spectra were key to the structural determination. A high-field carbon signal (106 ppm) in the HSQC spectrum was coupled to the high-field proton, confirming the high-field proton signal as non-aromatic. No C-H coupling was detected for the carbon signals at 162 and 154 ppm, in line with expectations for carbonyl and quaternary aromatic oxygen-substituted lactone carbons, respectively. In the HMBC spectrum, couplings predicted for NCO were observed between the non-aromatic singlet and the carbonyl carbon, and between a K-region doublet and the oxygen-substituted carbon assigned as part of the lactone structure. Thus, the NMR data are uniquely satisfied by the NCO structure.

Finally, NCO synthesized by an unambiguous route was identical in all respects to the product isolated from extracts of bioreactor-treated soil (Figures A5 and A7), establishing the NCO structure at the highest level of confidence for identifying unknown compounds (Schymanski et al., 2014a). The synthetic standard was subsequently used as an analytical standard for quantification of NCO in samples.

The cytotoxicity and genotoxicity of NCO were confirmed by the DT40 bioassay. The LD₅₀ of NCO for the parental cell line was 16.8 μ M (Figure 3.3), lower than that of the positive control methyl methanesulfonate (MMS; 58.5 μ M), suggesting stronger cytotoxicity. The LD₅₀ was 7.36 μ M for the *Rad54*^{-/-} mutant and 6.49 μ M for the *Rev1*^{-/-} mutant, corresponding to relative LD₅₀ values of 0.44 and 0.28, respectively, clearly demonstrating the genotoxicity of this compound. For comparison, the relative LD₅₀ values for MMS were 0.31 and 0.07 for the *Rad54*^{-/-} and *Rev1*^{-/-} mutants, respectively. In a recently published study (Sharma et al., 2016), the relative LD₅₀ values for hydrogen peroxide for these same two mutants were 0.54 and 0.32, respectively, comparable to that of NCO.

One possible mechanism for the observed genotoxicity of NCO might be through DNA alkylation, as α,β -unsaturated lactones have been demonstrated to modify DNA bases (Pfenning et al., 2016) and formation of bulky adducts is consistent with the type of damage detected by both DT40 mutant cell lines used in this study. While biotransformation products other than NCO might contribute to the increased genotoxicity resulting from bioreactor treatment of the soil, the high concentration of NCO and its potent genotoxicity support its significant contribution to the overall genotoxicity of the treated soil. As a newly discovered biotransformation product in an environmentally relevant system, further studies on its toxicological behavior would be warranted.

3.3.4 Source, putative formation pathway, and potential prevalence.

The structure of NCO is consistent with its proposed origin as a transformation product of pyrene. This hypothesis was verified in microcosm incubations of the bioreactor-treated soil with or without added pyrene. After 3 days of incubation, the average concentration of NCO in the pyrene-spiked microcosms increased from $0.27 \mu\text{g}\cdot\text{mL}^{-1}$ to $2.05 \mu\text{g}\cdot\text{mL}^{-1}$, while the concentration in the control group was essentially unchanged (Figure A8). This experiment unambiguously demonstrated that NCO was a transformation product of pyrene.

The concentration of pyrene in the untreated soil used to feed the bioreactor was $108 \pm 8.5 \mu\text{g}\cdot\text{g}^{-1}$ dry soil, second highest among the EPA-regulated PAHs monitored (Table A1). The estimated yield of NCO in the bioreactor was $\sim 13\%$, based on its concentration in treated soil samples ($11.2 \pm 2.5 \mu\text{g}\cdot\text{g}^{-1}$ dry soil, $n=8$) compared to the removal of pyrene in the bioreactor ($84.0 \pm 9.0 \mu\text{g}\cdot\text{g}^{-1}$ dry soil). We assume that the detected alkylated congeners of NCO ($\text{C}_{16}\text{H}_{10}\text{O}_2$, $\text{C}_{17}\text{H}_{12}\text{O}_2$) are likewise biotransformation products from the corresponding methylated and dimethylated pyrene isomers also present at relatively high concentrations ($53.5 \pm 2.0 \mu\text{g}\cdot\text{g}^{-1}$ and $21.4 \pm 1.0 \mu\text{g}\cdot\text{g}^{-1}$ dry soil, respectively) in the soil used to feed the bioreactor (Adrion et al., 2016).

To further confirm NCO as a bacterial metabolite of pyrene, the potential biotransformation to NCO was screened with three bacteria known to grow on pyrene as a sole carbon and energy source: *Mycobacterium vanbaalenii* strain PYR-1^T, *Rugisobacter aromaticivorans* strain Ca6^T (Corteselli et al., 2017a), and *Immundisolibacter cernigliae* strain TR3.2^T (Corteselli et al., 2017b). *M. vanbaalenii* PYR-1^T is the most comprehensively studied (Kim et al., 2007; Kweon et al., 2011) bacterial isolate capable of growing on pyrene but is not known to be present in our bioreactor. *R. aromaticivorans* strain Ca6^T and *I. cernigliae* strain TR3.2^T (previously referred to as PG1 and PG2, respectively) were identified by stable-isotope

probing as major contributors to pyrene degradation in several PAH-contaminated soils (Jones et al., 2008, 2011; Singleton et al., 2006) and possess 16S rRNA gene sequences highly similar to those recovered from geographically widespread contaminated soils (Corteselli et al., 2017b, 2017a). *Rugosibacter* sequences are present in the soil that was treated in the bioreactor in this study (Richardson et al., 2011), but are not present in high relative abundance in the bioreactor. Strain TR3.2^T, recently isolated from bioreactor-treated soil (Corteselli et al., 2017a; Singleton et al., 2016), and uncultivated *Immundisolibacter* strains are significant members of the bioreactor microbial community (Singleton et al., 2011). Both strains Ca6^T and TR3.2^T accumulated NCO after 24 hours of resting-cell incubation in the presence of pyrene, whereas strain PYR-1 did not. The yield of NCO from pyrene was $1.3 \pm 0.4\%$ for strain Ca6^T and $0.3 \pm 0.1\%$ for strain TR3.2^T. These results confirmed that NCO is a bacterial metabolite of pyrene. The low yields of NCO with the pure cultures compared to the yield observed in the bioreactor could have resulted from significant differences between bioreactor conditions and incubations of the pure cultures; for example, inocula were uninduced for pyrene metabolism, the incubation was short-term, and there were no other substrates that would be present in the bioreactor that could influence metabolic flux or differential transcription of the multiple ring-hydroxylating dioxygenases in both organisms. It is also possible that other, undetermined PAH-degrading bacteria in the bioreactor contributed to NCO accumulation.

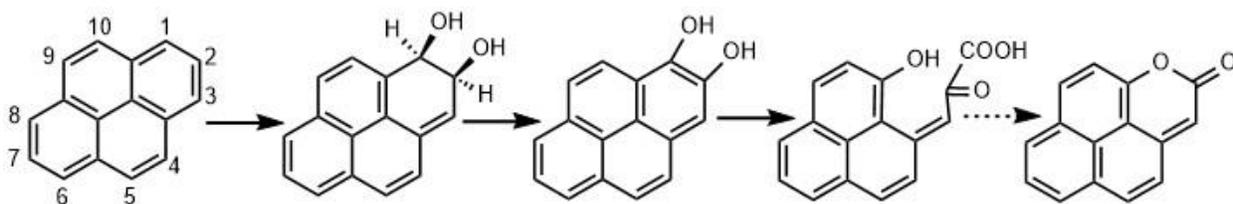


Figure 3.4. Proposed pathway for the formation of NCO from pyrene by PAH-degrading bacteria.

The proposed pathway for formation of NCO is shown in Figure 3.4. The first three steps are consistent with the canonical steps for aerobic bacterial metabolism of PAHs (Kweon et al.,

2011): *cis*-dihydroxylation by a ring-hydroxylating dioxygenase (RHD), dehydrogenation, and ring cleavage by an extradiol dioxygenase. Lactonization of the α -ketohydroxycarboxylic acid resulting from extradiol ring cleavage has been reported for the bacterial oxidation of naphthalene (Davies and Evans, 1964), phenanthrene (Evans et al., 1965) and anthracene (Herwijnen et al., 2003; Moody et al., 2001); in most cases, however, the lactone was considered to be a minor product relative to further metabolism of the ring-cleavage product by an aldolase-catalyzed release of pyruvate.

Initial ring hydroxylation would have to occur at the C1-C2 position to form NCO from pyrene, whereas complete (growth-related) metabolism of pyrene involves initial attack at the C4-C5 position, or K-region (Kim et al., 2007). *M. vanbaalenii* PYR-1 is known to initiate C1,2 oxidation of pyrene (Heitkamp et al., 1988; Kim et al., 2007; Kweon et al., 2014) via Nid-type dioxygenases (Kim et al., 2007), although as a minor side reaction relative to dihydroxylation at C4-C5 (Kweon et al., 2014). Dead-end products retaining three (Heitkamp et al., 1988) or four (Kim et al., 2007) rings have been found to accumulate from initial C1,2-dioxygenation of pyrene by strain PYR-1, but NCO was not reported as a product in either case. Given the low yield of NCO from the two pyrene-degrading pure cultures shown to produce it, it may be a dead-end metabolite initiated by a dioxygenase whose dominant regiospecificity is for C4-C5. Alternatively, given the differential substrate specificity of RHDs in strain Ca6^T (Singleton et al., 2012), it is possible that pyrene was oxidized in the pure cultures via an RHD with higher specificity for a different PAH substrate in which pyrene would align differently at the active site than in an RHD highly specific for pyrene. After ring hydroxylation, the next two steps would have to be carried out by enzymes that can accommodate the substitutions at C1-C2. It appears that the hydratase-aldolase that would normally be involved in releasing pyruvate from an extradiol ring-cleavage product (Kweon et al., 2011) of other PAHs is either not present in the

organism(s) that produce NCO or is unable to accommodate the product resulting from oxidation of pyrene at C1-C2.

To verify that NCO could be produced in PAH-contaminated soils other than the soil we treated in our bioreactor, we prepared aerobic microcosms of creosote-contaminated soil samples from two other sites, to which no additional pyrene was added: the Holcomb Superfund site in North Carolina, USA, and a wood treatment facility in Andalucía, Spain (Tejeda-Agredano et al., 2013). After six weeks of incubation, NCO was detected in both samples (no NCO was present in either sample initially). In the Holcomb sample, $0.5 \pm 0.02 \mu\text{g NCO} \cdot \text{g}^{-1}$ dry soil was produced, and in the Andalucía sample $10.3 \pm 0.54 \mu\text{g NCO} \cdot \text{g}^{-1}$ dry soil was produced.

3.3.5 Environmental relevance

We have identified a genotoxic bacterial metabolite of pyrene, one of the most abundant of the EPA-regulated PAHs in contaminated systems (Haeseler et al., 1999), produced during lab-scale bioremediation of field-contaminated soil from a former MGP site. Its yield and concentration in the lab-scale bioreactor are high enough to be of potential concern, because oxy-PAHs such as NCO and quinones are likely to be more mobile in the environment than their parent PAHs (Lundstedt et al., 2007). Furthermore, we have shown that NCO formation from pyrene in our bioreactor is not likely to be an isolated observation. We observed its formation in soils from three different sites, two in North Carolina (USA) and one in Europe, with different contamination sources (coal tar and creosote). It is also produced by two major pyrene-degrading bacteria that are closely related phylogenetically to uncultivated bacteria found in geographically diverse contaminated environments (Corteselli et al., 2017b, 2017a). In addition, it is possible that NCO has been observed in previous studies but not recognized as such. A metabolite of pyrene from two different *Mycobacterium* strains, including *M. vanbaalenii* PYR-1, was reported (Kim et al., 2005; Vila et al., 2001) to have an electron-ionization mass spectrum with a

molecular ion at m/z 220 and fragment ions at m/z 192 [$M^+ - (CO)$] and 164 [$M^+ - 2(CO)$] or 163 [$M^+ - COOCH$], consistent with the composition $C_{15}H_8O_2$. However, the metabolite was not further characterized in either study; in the case of strain PYR-1, the authors assumed that the metabolite was 4-oxapyrene-5-one (Kim et al., 2005). Although we did not observe NCO production (or any other product with m/z 220) by strain PYR-1 in the present study, this may have been due to differences in incubation conditions. As noted above, Lundstedt et al (Andersson et al., 2003) reported the accumulation of a compound assumed to be 4-oxapyrene-5-one in an aerobic bioreactor treating a PAH-contaminated soil. In these earlier studies it would have been reasonable to assume that an analyte with the same molecular formula as NCO was 4-oxapyrene-5-one; it had already been identified definitively in particulate matter from diesel exhaust (Pitts et al., 1982), and it would be reasonable to predict a bacterial metabolite of pyrene via oxidation at C4-C5, the predominant site of attack in aerobic bacterial metabolism.

Because of its cytotoxicity, genotoxicity, and potential occurrence in other PAH-contaminated environmental systems, we believe that attention should be paid to NCO during active bioremediation efforts or as a result of natural attenuation *in situ*. This study adds to the increasing awareness of the potential toxicity and persistence of transformation products of environmental contaminants (Cwiertny et al., 2014; Fenner et al., 2013; Mohler et al., 2013; Qu et al., 2013) and reinforces the need to include such products in risk analysis. Traditional environmental analysis has focused on target compounds that could be identified and quantified by commercial standards, overlooking many possibly important contaminants (Krauss et al., 2010). As illustrated in this study, combining non-target analysis with EDA methods allows greater insights into the chemicals that can pose substantial threats to human health.

CHAPTER 4 DIVERSITY AND ABUNDANCE OF HIGH-MOLECULAR-WEIGHT AZAARENES IN PAH-CONTAMINATED ENVIRONMENTAL SAMPLES²

4.1 Introduction

Azaarenes, also known as polycyclic aromatic nitrogen heterocycles (PANH), are the heterocyclic analogues of polycyclic aromatic hydrocarbons (PAHs) in which one or more nitrogen atoms are incorporated into the aromatic structure. While sharing some common properties with PAHs, the presence of nitrogen brings special characteristics to azaarenes, such as greater polarity and higher structural diversity. Some azaarenes are produced for the manufacture of dyes, paints, pharmaceuticals, and pesticides, but they also originate from petrogenic and pyrogenic processes, and therefore they exist in derivatives of coal (tar, creosote) and crude oil (Anyanwu and Semple, 2015a; Manzano et al., 2017). As a consequence they can potentially co-occur at relatively high concentrations in PAH-contaminated sites such as those associated with coke works, manufactured-gas plants, and wood-treatment facilities (Neilson, 1995). However, current practices in risk assessment and remediation overlook the presence and potential risks of contaminants that co-occur with the regulated PAHs, such as oxygenated PAHs and heterocyclic aromatics (Andersson and Achten, 2015; Brinkmann et al., 2014). Many azaarenes are toxic; for example, some of the high-molecular-weight (HMW) compounds (containing 4 or more rings) are mutagenic and carcinogenic to animals (Bleeker et al., 1999; Wiegman et al., 2001), being major contributors of AhR-mediated activities (Machala et al.,

² Zhenyu Tian, Joaquim Vila, Hanyan Wang, Wanda Bodnar and Michael D. Aitken
Published in *Environmental Science & Technology* 2017, 51 (24): 14047–14054

JV (co-first author) conducted the experiments and helped with writing, HW helped with the data processing and statistical analysis, and WB assisted with the HRMS data interpretation.

2001). In addition, azaarenes are more water-soluble and mobile than PAHs, and thus have been found in groundwater at sites contaminated with creosote (Pereira et al., 1987) and coal tar (Blum et al., 2011). Their toxicity and environmental mobility suggest that azaarenes have implications for human and environmental health.

Azaarenes in contaminated sites have rarely been analyzed, and there is a general lack of knowledge on their environmental behavior and fate (Anyanwu and Semple, 2015a). While recent studies have focused on targeted analysis of low-molecular-weight (LMW) azaarenes such as quinoline and acridine (Bandowe et al., 2014a; Lundstedt et al., 2014; Mahler et al., 2015; Wei et al., 2015), limited attention has been paid to the occurrence of isomers of these compounds, methylated derivatives and HMW homologs in contaminated soils (Meyer et al., 1999; Švábenský et al., 2009). As a result the overall contribution of azaarenes to toxicity may be significantly underestimated, especially for the HMW compounds. Considering their potential persistence and the risk to public health, it is necessary to fill this gap of knowledge by elucidating the existence and diversity of these pollutants.

In this study, we applied high-resolution mass spectrometry (HRMS) and Kendrick mass defect (KMD) analysis to samples from four different PAH-contaminated sites to reveal the diversity and distribution of azaarene congeners containing one nitrogen atom. Mass-defect filtering is a nontarget analytical method based on HRMS that has been used to group and illustrate the occurrence of compound classes featuring hetero atoms in complex samples (Ballesteros-Gómez et al., 2017; Myers et al., 2014). Compound classes are distinguished by transforming the exact mass according to a specific mass scale and then calculating the mass defects (Hughey et al., 2001; Sleno, 2012); KMD is the most commonly used scale and is based on the monoisotopic and nominal mass of a CH₂ group (described in Supporting Information). The concept has been widely used to identify different compound classes in complex mixtures

and environmental samples, such as sulfur-containing (Müller et al., 2005; Tao et al., 2014) and perfluorinated compounds (Barzen-Hanson et al., 2017; Newton et al., 2017).

4.2 Materials and methods

4.2.1 Chemicals

Anhydrous sodium sulfate and high-performance liquid chromatography (HPLC)-grade solvents, including dichloromethane (DCM), acetone, and methanol, were purchased from Fisher Scientific (Pittsburgh, PA, U.S.A.).

Acridine, phenanthridine, benzo[*h*]quinoline, benzo[*c*]acridine, and the PAH standard solution (EPA 610 PAH mixture) were obtained from Sigma-Aldrich (St. Louis, MO, U.S.A.). Benzo[*a*]acridine was purchased from LGC standards (Teddington, Middlesex, UK). Acridine-*d*₉ and dibenzo[*a,j*]acridine were from Cambridge Isotope Laboratories (Tewksbury, MA, U.S.A.).

4.2.2 Contaminated sites, sample collection, and solvent extraction.

Four PAH-contaminated samples (three soils and one sediment) were collected from three sites located in North Carolina, U.S.A (Hu et al., 2012; Tian et al., 2017a), and one site in Andalucía, Spain (Tejeda-Agredano et al., 2013). Further details on the origin, associated industrial activity, and total PAH concentration and degree of weathering are provided in Table 4.1. Concentrations of individual PAHs in each sample are listed in Table B1.

Before chemical analysis, all samples were air-dried to constant weight for 48 hours, sieved through 2 mm, and homogenized. All glassware used in the process was clean and autoclaved at 121 °C to avoid contamination. Triplicate 1-g aliquots for each sample were extracted by two successive overnight extractions using 20 mL of a DCM: acetone mixture (1:1, v/v) with agitation on a wrist-action shaker. Prior to extraction, deuterated compounds (100 µg acridine-*d*₉ and 200 µg anthracene-*d*₁₀) were spiked into the samples in acetone solution as recovery standards. Sodium sulfate (10 g) was added to dry the sample, and glass beads were

added for better extraction efficiency. The collected extracts were filtered through a 0.2 μm pore-size nylon membrane, brought up to 50 mL with DCM and acetone (1:1, v/v), and stored at 4 °C until further analysis.

4.2.3 Instrumental analysis

Nontarget analysis was performed using an Agilent 1200 series HPLC interfaced with a 6520 accurate mass quadrupole time-of-flight mass spectrometer (qTOF MS, Agilent Technologies, Santa Clara, CA.). Quantitative target analysis on azaarenes was performed using a TSQ Quantum Ultra triple quadrupole mass spectrometer (QqQ MS, Thermo Fisher Scientific, Waltham, MA). Electrospray ionization in positive (ESI+) mode was adopted for both nontarget and target analysis, and the detailed parameters for MS and HPLC are provided in Appendix B (Table B2).

The concentrations of the USEPA-regulated PAHs in the samples were quantified using HPLC with fluorescence detection as described elsewhere (Hu et al., 2012).

4.2.4 Nontarget analysis workflow and mass defect filtering

High resolution exact mass data were converted to .mzData format, and peak detection and alignment were achieved using the publicly available Scripps XCMS Online platform (<https://xcmsonline.scripps.edu/>). The resulting feature tables were imported and analyzed with an in-house script written in R language (R 3.3.3).

Since accurate, monoisotopic atomic masses are defined in relation to ^{12}C , there are small differences between the nominal mass and the accurate mass of other elements (“mass defects”) (Hughey et al., 2001; Sleno, 2012). Molecules with different numbers of hydrogens and other heteroatoms each have unique mass defects. Therefore, it is feasible to distinguish between compound classes by transforming the exact mass according to a specific mass scale (in this case

the Kendrick scale based on CH₂) and then calculating the mass defects. The following equations were used as the basis for mass-defect filtering (where X represents a halogen atom if present):

$$\text{Kendrick mass (KM)} = \text{exact mass} \times (14 / 14.01565)$$

$$\text{Kendrick mass defect (KMD)} = \text{nominal mass} - \text{KM}$$

$$\text{Double-bond equivalents (DBE)} = \text{C} - \text{H}/2 - \text{X}/2 + \text{N}/2 + 1$$

According to previous knowledge on mass-defect filtering, the CH₂-based KMD difference between neighboring series of analytes of interest (here, those containing one N atom) is 0.0134.²⁰ Because the Kendrick mass defect was defined by the difference of CH₂ units, the neighboring series on the Kendrick scale should differ in their DBE by 1. By calculation and accurate mass measurements on standards, we determined that the KMD of acridine and phenanthridine (C₁₃H₉N, DBE 10) is 0.1205, and the KMD of benzo[*a*]acridine (C₁₇H₁₁N, DBE 13) is 0.1607, which is equal to 0.1205 + 3* 0.0134. Another frequently detected signal, probably an azapyrene (C₁₅H₉N, DBE 12), featured a KMD of 0.1473, equal to 0.1205 + 2*0.0134. Therefore, the KMD differences of these detected compounds were in agreement with the theoretical pattern. Accordingly, within the mass range of interest (100 < *m/z* < 500), we selected data that fit in the series by having KMD values that differed with those of known compounds by increments of 0.0134. The resulting sets of series represented KMDs from 0.1071 to 0.2009, corresponding to DBE from 9 to 16. The diversity and intensity of all the detected azaarenes are illustrated by plotting the KMD versus Kendrick mass, known as a Kendrick mass defect plot. Partial transparent bubble plots (transparency = 1/4) based on peak area data were created using the ggplot2 package in R.

4.3 Results and discussion

4.3.1 PAH quantification and distribution profiles.

Samples from the four contaminated sites varied greatly in the concentrations of the EPA-regulated PAHs, with total PAH concentrations ranging between 147 and 7,430 $\mu\text{g/g}$ dry weight (Table 4.1). The individual PAH concentration profiles (Table B1) revealed that KM and HS samples were highly enriched in HMW PAHs relative to the LMW compounds, with HMW/LMW PAH concentration ratios of 42.6 and 22.9, respectively (Table 4.1). These results indicated a higher degree of weathering compared to FS and SC samples (HMW/LMW PAH ratios of 3.1 and 0.63, respectively) (Banger et al., 2010; Yang et al., 2014), which contained high concentrations of LMW PAHs.

Table 4.1. Description of the PAH-contaminated samples used in this study

Site ID	Origin	Industrial activity	PAH conc. ^a	Weathering ratio ^b
FS	Salisbury, NC	Manufactured gas plant	421 \pm 27.8	3.1
HS	Holcomb Superfund Site, NC	Wood treatment facility	147 \pm 19.9	22.9
KM	Kerr McGee Superfund Site, NC	Wood treatment facility	2,490 \pm 64	42.6
SC	Andalucía, Spain	Wood treatment facility	7,430 \pm 315	0.6

4.3.2 Diversity of azaarenes.

The solvent extracts of the samples from the four sites were analyzed by HRMS followed by KMD filtering. Eight homologous series of azaarenes, with DBE values ranging from 9 to 16, were tentatively identified and are presented in KMD plots (Figure 4.1). The identities of eight detected azaarenes were confirmed by authentic standards, or from their spectral characteristics by MS/MS fragmentation when the standards were not available. In general, the non-alkylated compounds featured a loss of m/z 28, consistent with a fragment of CH_2N (Figure 4.2 and Figure B1), in accordance with a previously reported method for azaarene analysis (Lintelmann et al., 2010). The possible mechanism is that the nitrogen atom was protonated in electrospray

ionization, and the loss of CH_2N corresponds to the protonated CHN group. Methylated azaarenes were also confirmed by the consecutive loss of alkyl groups (m/z 15, $-\text{CH}_3$) and CH_2N (Figure 4.2). Therefore, we concluded that the series we observed from KMD plots were azaarene homologs and their alkylated derivatives. According to the widely accepted levels of confidence in nontarget analysis (Schymanski et al., 2014a), the compounds confirmed with authentic standards would be among level 1 (highest), while others identified by MS/MS and KMD analysis would fit in level 3 (tentative candidates). The formulas and the respective number of isomers detected for each congener group in the 4 samples are listed in Table 4.2.

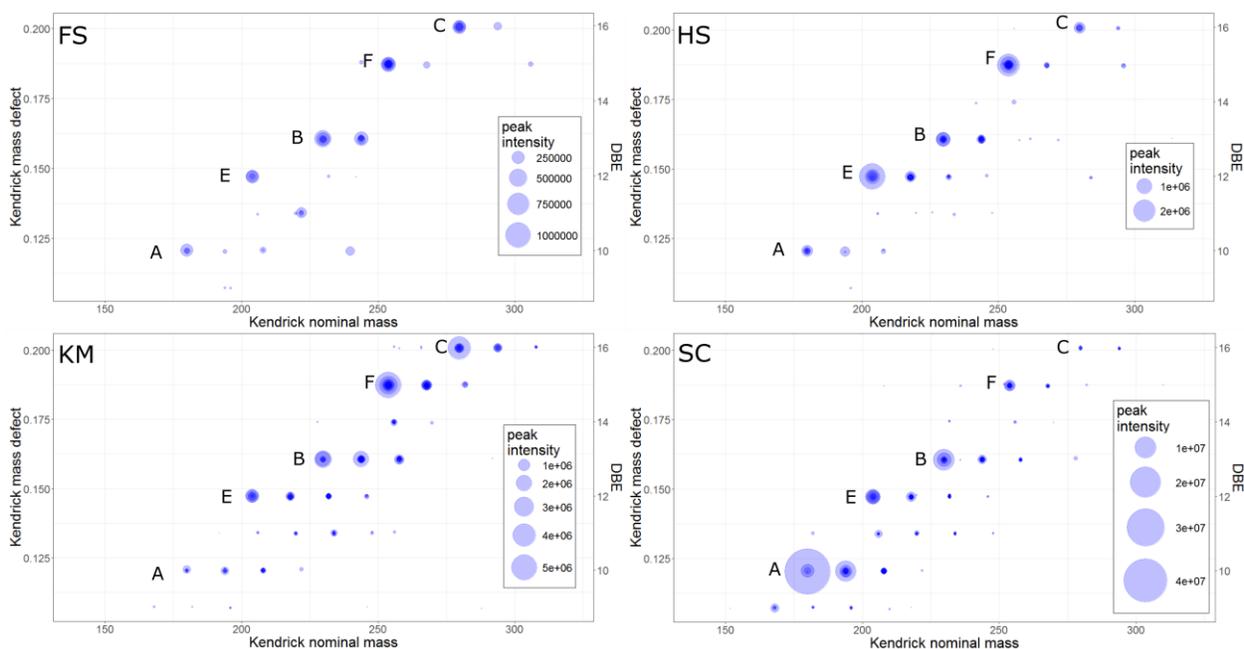


Figure 4.1 Kendrick mass defect (KMD) plots of azaarene congeners obtained from the four contaminated sites. (a) Soil sample from a manufactured-gas plant site in Salisbury, NC (FS sample); (b) Soil sample from the Holcomb Creosote Superfund site, NC (HS sample); (c) Sediment sample from the Kerr-McGee Superfund site, NC (KM sample). (d) Soil sample from a wood-treatment facility in Andalucía, Spain (SC sample). The five groups of non-alkylated azaarenes of greatest interest (most abundant congeners) are labeled as follows: A ($\text{C}_{13}\text{H}_9\text{N}$), E ($\text{C}_{15}\text{H}_9\text{N}$), B ($\text{C}_{17}\text{H}_{11}\text{N}$), F ($\text{C}_{19}\text{H}_{11}\text{N}$), and C ($\text{C}_{21}\text{H}_{13}\text{N}$). Each horizontal series of bubbles corresponds to a homologous series of congeners that share the same basic structure, while the overlapping bubbles are isomers with identical molecular formula and position on the KMD plot. The color darkened by overlapping suggests the emergence of multiple isomers, and the areas of the bubbles are proportional to the intensities.

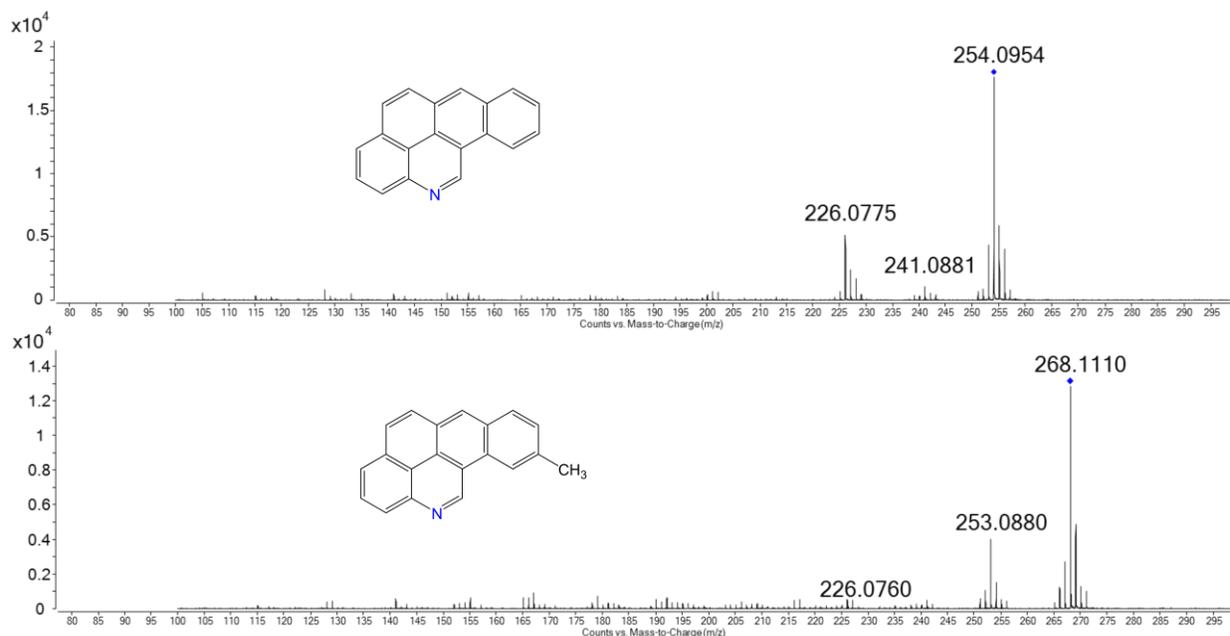


Figure 4.2. MS/MS fragmentation patterns of two tentatively identified compounds C₁₉H₁₁N (F8, above) and C₂₀H₁₃N (below). The structures were inferred from MS/MS and literature data.

In total, 232 congeners were detected and tentatively identified (except for the location of the N atom in a given congener for which there were no standards) within these 8 homologous series. Each series containing as many as 4 congener groups, corresponding to the non-alkylated and the mono-, di-, and tri-methylated (C1-C3) derivatives. The deeper color in the first three positions of each line in the KMD plots implied higher isomer diversities for the non-alkylated and the C1 and C2 azaarenes. Of the eight homologous series listed in Table 4.2, six of them could be related to known azaarene structures previously detected in atmospheric particulate matter (Lintelmann et al., 2010) or during the thermal degradation of nylon polymers (Wilhelm et al., 2000). The other two homolog series (DBE = 11 and 14), tentatively corresponding to phenylquinoline isomers (C₁₅H₁₁N) and phenylacridine isomers (C₁₉H₁₃N), were tentatively identified here for the first time, with no data being available regarding their detection in other complex matrices. We note that, except for the standards, the names used here are example chemicals that represent the congener groups with a certain molecular formula and ring number, but do not indicate identified structures. For instance, ‘azapyrene isomers’ represent the formula

C₁₅H₉N, which can also include azafluoranthene. Because these congener groups were only tentatively identified and not confirmed with authentic standards, we have used their common names (e.g. azapyrene, azabenz[*a*]pyrene) instead of the more specific IUPAC nomenclatures.

Table 4.2 Detailed information on azaarene congeners detected in the samples

example chemical	formula	DBE	exact mass (Da)		# of isomers
			M	[M+H] ⁺	
azafluorene	C ₁₂ H ₉ N	9	167.0735	168.0806	4
C1-azafluorene	C ₁₃ H ₁₁ N	9	181.8910	182.0962	4
C2-azafluorene	C ₁₄ H ₁₃ N	9	195.1048	196.1118	6
acridine	C ₁₃ H ₉ N	10	179.0735	180.0806	6
C1-acridine	C ₁₄ H ₁₁ N	10	193.0891	194.0965	11
C2-acridine	C ₁₅ H ₁₃ N	10	207.1048	208.1120	19
C3-acridine	C ₁₆ H ₁₅ N	10	221.1204	222.1275	3
phenylquinoline	C ₁₅ H ₁₁ N	11	205.0891	206.0965	7
C1-phenylquinoline	C ₁₆ H ₁₃ N	11	219.1048	220.1117	7
C2-phenylquinoline	C ₁₇ H ₁₅ N	11	233.1204	234.1277	8
azapyrene	C ₁₅ H ₉ N	12	203.0735	204.0807	7
C1-azapyrene	C ₁₆ H ₁₁ N	12	217.0891	218.0965	18
C2-azapyrene	C ₁₇ H ₁₃ N	12	231.1048	232.1120	15
C3-azapyrene	C ₁₈ H ₁₅ N	12	245.1204	246.1275	4
benzo[<i>a</i>]acridine	C ₁₇ H ₁₁ N	13	229.0891	230.0965	13
C1-benzo[<i>a</i>]acridine	C ₁₈ H ₁₃ N	13	243.1048	244.1120	16
C2-benzo[<i>a</i>]acridine	C ₁₉ H ₁₅ N	13	257.1204	258.1278	13
C3-benzo[<i>a</i>]acridine	C ₂₀ H ₁₇ N	13	271.1361	272.1433	1
phenylacridine	C ₁₉ H ₁₃ N	14	255.1048	256.1118	5
C1-phenylacridine	C ₂₀ H ₁₅ N	14	269.1204	270.1278	1
azabenz[<i>a</i>]pyrene	C ₁₉ H ₁₁ N	15	253.0891	254.0965	17
C1-azabenz[<i>a</i>]pyrene	C ₂₀ H ₁₃ N	15	267.1048	268.1122	22
C2-azabenz[<i>a</i>]pyrene	C ₂₁ H ₁₅ N	15	281.1204	282.1276	3
dibenzo[<i>a,j</i>]acridine	C ₂₁ H ₁₃ N	16	279.1048	280.1120	10
C1-dibenzo[<i>a,j</i>]acridine	C ₂₂ H ₁₅ N	16	293.1204	294.1277	7
C2-dibenzo[<i>a,j</i>]acridine	C ₂₃ H ₁₇ N	16	307.1361	308.1433	5
Total					232

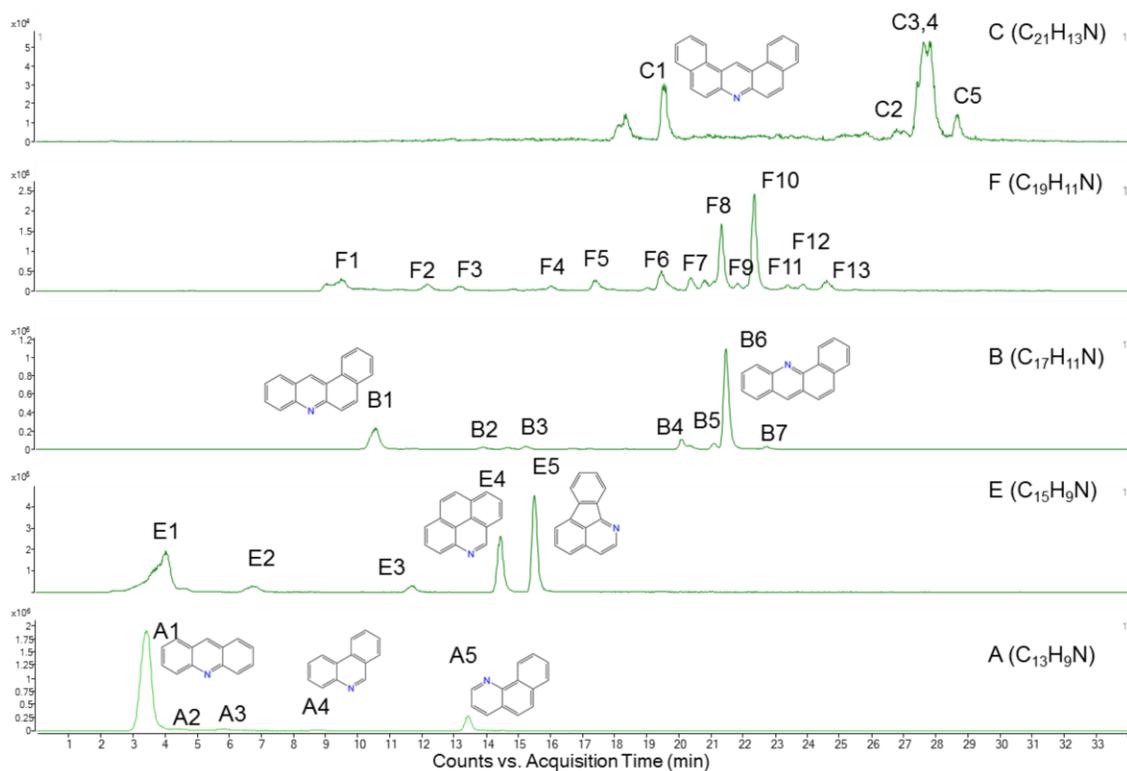


Figure 4.3. Extracted ion chromatograms (EICs) of the SC sample, corresponding to the five major congener groups. Structures are shown for peaks that corresponded to available standards: A1, acridine; A4, phenanthridine; A5, benzo[*h*]quinoline; B1, benzo[*a*]acridine; B6, benzo[*c*]acridine; C1, dibenzo[*a,j*]acridine. Structures E4 (azapyrene) and E5 (azafluoranthene) were inferred from literature data.

Figure 4.3. depicts the extracted ion chromatograms (EIC) for the five major non-alkylated congener groups in the SC sample labeled A, B and C for the cata-condensed groups $C_{13}H_9N$ (DBE 10, 3-ring), $C_{17}H_{11}N$ (DBE 13, 4-ring), and $C_{21}H_{13}N$ (DBE 16 5-ring), respectively, and E and F for the peri-condensed groups $C_{15}H_9N$ (DBE 12, 4-ring) and $C_{19}H_{11}N$ (DBE 15, 5-ring), respectively. For each EIC, multiple chromatographic peaks were observed, corresponding to the different isomers. This high diversity of azaarene isomers derives from both their distinctive carbon skeleton as well as the position of the substituted nitrogen atom. The substitution of a given isomer with methyl groups amplifies the structural diversity, resulting in a large number of isomers detected for some methylated congener groups. Such diversity of azaarene isomers has never been reported for environmental samples, and suggests that

conventional target analysis of azaarenes based on available standards might result in a biased and very limited estimation of their abundance.

In general, these five congener groups presented the largest bubbles in KMD plots (Figures 4.1), indicating their predominance among azaarene congener groups in the four samples. The intensities (abundances) of methylated azaarenes decreased with increasing level of methylation, indicating that non-alkylated and mono-methylated derivatives accounted for the major signal intensities among these azaarene congeners. However, because the relative response factors (RRF) might differ for various compounds, the estimation of concentrations exclusively based on peak intensities might not be accurate (Huba et al., 2016). More-accurate quantification was achieved using commercially available standards; each standard was used to quantify all isomers of the standard itself (*i.e.*, response factors were assumed to be equivalent for each isomer).

4.3.3 Abundances of azaarenes.

Quantification of non-alkylated azaarenes was based on the MS/MS transition from $[M+H]^+$ to $[M+H-28]^+$, corresponding to the loss of a CH_2N group, and further confirming the identities of azaarene isomers. Authentic standards for acridine, phenanthridine, benzo[*h*]quinoline, benzo[*a*]acridine, benzo[*c*]acridine, and dibenzo[*a,j*]acridine were used to quantify their series of isomers, and semi-quantification of azapyrene and azabenzopyrene isomers was based on the RRF of benzo[*a*]acridine and dibenzo[*a,j*]acridine, respectively. Concentrations of quantifiable isomers in each series are shown in Table B3.

There were obvious differences in the abundance of congener groups among samples (Figure 4.4). Consistent with the degree of weathering observed for PAHs, the KM sample was particularly enriched in the HMW azaarenes, with congener groups B, C and F at the highest concentrations (20.5, 21.4 and 16.9 $\mu\text{g/g}$, respectively). In contrast, the relatively unweathered

SC sample was highly enriched in 3-ring azaarenes, with a concentration (320 µg/g) at least four times higher than that of any of the 4- and 5-ring congener families (ranging between 7.5 and 72 µg/g). The less-contaminated FS and HS samples presented a more balanced distribution among the different azaarene congener groups, ranging between 1.0 and 2.8 µg/g for FS, and between 2.0 and 9.7 µg/g for HS. In general, except for the SC sample, the most abundant azaarene congener groups were those with four or five rings. The varied profiles in congener groups demonstrated that the occurrence and environmental prevalence of azaarenes might be site-specific and affected by weathering processes. Consequently, the exclusive analysis of compounds from one congener group, such as the most frequently targeted 3-ring compounds, will likely lead to underestimation of the environmental impact of azaarenes.

Within each of the congener groups, we analyzed the isomer distribution among the four samples. In congener group A (C₁₃H₉N), acridine (A1) was the most abundant among the five quantifiable isomers at all four sites; this contrasts with results from a study (Lundstedt et al., 2014) on three PAH-contaminated soil samples in which benzo[*h*]quinoline was at higher concentration than acridine. The average concentration of acridine was as high as 281 µg/g (SC sample), dramatically higher than those previously reported for other PAH-contaminated soils (Lundstedt et al., 2014; Meyer et al., 1999) (ranging from 0.18 to 10 µg/g). For the FS, HS, and KM samples, the concentrations of acridine were 0.9, 1.0, and 3.2 µg/g, respectively, similar to concentrations reported in other studies.

In congener group B (C₁₇H₁₁N), seven isomers were quantified; in general, benzo[*a*]acridine (B1) and benzo[*c*]acridine (B6) were the predominant isomers. In accordance with their lower degree of contamination, the FS and HS samples presented the lowest concentrations among the four sites (respectively 1.0 and 2.8 µg/g for benzo[*a*]acridine, and 0.90 and 1.0 µg/g for benzo[*c*]acridine). The heavily contaminated KM and SC samples had higher concentrations (12.2 and 13.1 µg/g for benzo[*a*]acridine, and 4.2 and 33.9 µg/g for

benzo[*c*]acridine, respectively). Compared to the data from the only published study on PAH-contaminated soil that considered these compounds (Meyer et al., 1999), the concentrations in the KM sample were comparable, while the concentrations in the SC sample were substantially higher. Consistent with results in the earlier study, the concentration of benzo[*c*]acridine was higher than that of benzo[*a*]acridine in the FS, HS and SC samples.

Congener group C (C₂₁H₁₃N) comprised five quantified isomers. Isomers C3 and C4, which co-eluted in a single peak, were the most abundant in all of the samples. According to the chromatographic profiles previously described by Lintelmann et al during the analysis of azaarenes in air particulate matter (Lintelmann et al., 2010), these compounds were tentatively identified as dibenzo[*a,h*]acridine and dibenzo[*a,c*]acridine. The latter is the only 5-ring azaarene for which there is previously published data for contaminated soil (Meyer et al., 1999), with concentrations in two soils (0.4 and 0.7 µg/g) comparable to those detected here in the FS and HS samples.

Regarding the peri-condensed groups of congeners, 5 congeners in group E (C₁₅H₉N) and 13 congeners in group F (C₁₉H₁₁N) were semi-quantified. Within group E, isomers E1 (0.55-31 µg/g), E4 (0.4-12 µg/g), and E5 (0.13-14 µg/g) had relatively higher concentrations, the last two tentatively corresponding to 4-azapyrene (E4) and 1-azafluoranthene (E5) based on their LC elution profiles (Lintelmann et al., 2010). Within the diverse congener group F (C₁₉H₁₁N), F1 (0.17-2.2 µg/g), F6 (0.17-3.8 µg/g), F8 (0.39-4.6 µg/g) and F10 (0.26-4.6 µg/g) were the major isomers. However, due to a lack of authentic standards, we could not establish their identities. To our knowledge, this is the first report of peri-condensed 5-ring azaarenes in contaminated soil. Several peri-condensed, unsubstituted 5-ring PAHs such as benzo[*a*]pyrene, benzo[*b*]fluoranthene, and benzo[*k*]fluoranthene are known human carcinogens; therefore, the presence of their azaarene analogues in contaminated soils should not be ignored.

Overall, the detection of azaarene isomers at relatively high concentrations in four independent and geographically distant samples suggests that they might be prevalent and abundant in sites contaminated with coal tar and creosote. Congener groups B, C, and F encompass some known or suspected human carcinogens, such as benzo[*c*]acridine and 10-azaB[*a*]P (Bleeker et al., 1999; Chang et al., 1984; Yamada et al., 2004). The high concentrations and isomer diversity of these congener groups therefore suggests that they could significantly contribute to the overall toxicity and/or genotoxicity of PAH-contaminated soils. We suggest that future analysis and monitoring of azaarenes should also include HMW compounds, such as benz[*c*]acridine, azapyrenes, and/or dibenzo[*a,h*]acridine.

4.3.4 Environmental significance

This work highlights the application of nontarget analytical methods such as mass defect filtering that can help select and prioritize important compound classes from complex environmental samples. By applying these nontarget analytical techniques we have demonstrated the occurrence of a wide range of N-heterocyclic aromatics in soil and sediment samples contaminated with coal tar and creosote, many at relatively high concentrations. Azaarenes have received far less attention than unsubstituted PAHs in contaminated environmental systems, and they currently are not among the regulated contaminants in these systems. Their biotransformation, biodegradation, and fate in the environment were not well understood, and further research would be warranted. In a companion manuscript, we illustrate that many of these compounds are recalcitrant to biodegradation (Chapter 5). Given the known toxicological properties of several azaarenes, the diversity, abundance and, in some cases, persistence in environmental samples of a wide range of previously unknown compounds suggests that further attention be paid to this overlooked class of contaminants.

CHAPTER 5 ISOMER-SELECTIVE BIODEGRADATION OF HIGH-MOLECULAR-WEIGHT AZAARENES IN CONTAMINATED ENVIRONMENTAL SAMPLES³

5.1 Introduction

Polycyclic aromatic nitrogen heterocycles (PANHs), or azaarenes, are known to be common co-contaminants at sites contaminated with polycyclic aromatic hydrocarbons (PAHs). They tend to be more polar than their PAH counterparts, and thus have been found at significant concentrations in groundwater at sites associated with contamination by creosote (Pereira et al., 1988) or coal tar (Blum et al., 2011). However, the range of azaarene congeners that have been studied in environmental systems has been limited (Anyanwu and Semple, 2015a). Given the known toxicity of some of these compounds (Bleeker et al., 1999; Machala et al., 2001; Wiegman et al., 2001), their fate in the environment is important to consider.

Relatively limited research has been conducted on the biodegradation of PANHs in soil and sediment systems, particularly for alkyl-substituted PANHs or those of higher molecular weight (HMW, more than three rings). A few studies have reported that bacteria isolated from soil can degrade or transform some two-ring (quinoline, isoquinoline) and three-ring (phenanthridine, acridine) PANHs (Parshikov et al., 2012; Pereira et al., 1988; Sutherland et al., 2009; Van Herwijnen et al., 2004; Willumsen et al., 2005). However, little is still known about their fate in the environment, and to our knowledge only two reports focused on azaarene degradation in soil. Anyanwu and Semple analyzed the fate of PANH analogues of phenanthrene spiked into a pristine agricultural soil (Anyanwu and Semple, 2015b), and recently Biache and

³ Joaquim Vila, Zhenyu Tian, Hanyan Wang, Wanda Bodnar, and Michael D. Aitken (manuscript in preparation) JV (co-first author) conducted the experiments and helped with writing, HW helped with the data processing and statistical analysis, and WB assisted with HRMS data interpretation.

colleagues (Biache et al., 2017) demonstrated the removal of four target PANHs during the slurry incubation of three PAH-contaminated soil samples inoculated with a suspension of soil from a gasification plant. These works have been limited to a small number of target compounds; the capabilities of autochthonous microbial communities or the effects of the N position, degree of methylation or aromatic structure on the susceptibility of azaarenes to microbial transformation remain to be elucidated.

In Chapter 4 (Tian et al., 2017b), we report the detection of a wide range of azaarenes and methylated azaarenes in soil and sediment samples from four different PAH-contaminated sites, using high-resolution mass spectrometry (HRMS) combined with Kendrick mass defect filtering. Based on the observation of a large number of azaarene congeners in all four samples using these methods, we set up microcosm incubations to provide insight into congener biodegradation patterns, isomer-selective degradation, and potential formation of transformation products.

5.2 Materials and methods

5.2.1 Chemicals

Anhydrous sodium sulfate, monobasic and dibasic sodium phosphates, and high-performance liquid chromatography (HPLC)-grade solvents, including dichloromethane (DCM), acetone, and methanol, were purchased from Fisher Scientific (Pittsburgh, PA, U.S.A.).

Acridine, phenanthridine, benzo[*h*]quinoline, benzo[*c*]acridine, and the PAH standard solution (EPA 610 PAH mixture) were obtained from Sigma-Aldrich (St. Louis, MO, U.S.A.).

Benzo[*a*]acridine was purchased from LGC standards (Teddington, Middlesex, UK). Acridine-*d*₉ and dibenzo[*a,j*]acridine were from Cambridge Isotope Laboratories (Tewksbury, MA, U.S.A.).

5.2.2 Contaminated samples

PAH-contaminated samples (three soils and one sediment) were collected from three sites located in North Carolina (NC), U.S.A (Hu et al., 2012; Tian et al., 2017a), and a site in

Andalucía, Spain (Tejeda-Agredano et al., 2013). Details on the origin and characteristics of these four samples are provided in Table 4.1 (previous chapter). The samples are referred to below as FS (coal tar-contaminated soil from a former manufactured-gas plant in NC that had been used to feed a lab-scale bioreactor); HS (creosote-contaminated soil from the Holcomb Superfund site wood-treatment facility in NC); KM (creosote-contaminated sediment from the Kerr McGee Superfund site wood-treatment facility in NC); and SC (creosote-contaminated soil from a wood-treatment facility in Andalucía). The KM and SC samples were the most highly contaminated, and the HS and KM samples demonstrated the highest degree of weathering as defined by the ratio of HMW PAHs to low-molecular-weight PAHs (Table 4.1). The SC sample was both the most highly contaminated and the least weathered of the four samples.

5.2.3 Microcosm incubations

The biodegradation patterns of azaarene congener groups were evaluated during aerobic incubations of the four samples. Aliquots of 1 g of air-dried, sieved (2 mm) and homogenized samples were placed in sterile 125-mL Erlenmeyer flasks with PTFE-coated screw caps, containing 30 mL of 10 mM phosphate buffer (pH 7.5) supplemented with urea as nitrogen source. A higher concentration of urea (1 mM) was used for the more highly contaminated samples (SC and KM), while a lower concentration (0.2 mM) was applied to the FS and HS samples, to maintain an approximate C:N ratio of 10:1 in accordance with their respective total PAH concentrations. Triplicate flasks were sacrificed initially (time 0) and after 7, 21, and 42 days of incubation. To evaluate the effect of urea addition on azaarene removal, identical triplicate microcosms except for the addition of urea were set up for the FS and SC samples; these microcosms were sacrificed after 42 days for azaarene analysis. For each sample, triplicate killed controls were prepared in the same manner as for biostimulated microcosms, but with the addition of phosphoric acid to pH 2, and sacrificed after 42 days of incubation. Before extraction,

killed controls were neutralized with a concentrated NaOH solution to facilitate azaarene extraction. All flasks were incubated at 25 °C and under agitation at 150 rpm. Every two days, the caps of flasks were loosened for 5 minutes to avoid oxygen limitations.

At the end of incubation, the entire contents from each flask were transferred to a 30-mL centrifuge tube with PTFE-lined septum. The slurry was then centrifuged at 3900 rpm, and the pellets were stored at -20 °C until extraction. Solvent extractions on the pelleted solid residues were conducted as described in Chapter 4.

5.2.4 Instrumental analysis

Nontarget analysis and quantitative target analysis on selected PANHs were conducted as described in Chapter 4. Briefly, nontarget analysis was conducted by HPLC-HRMS using electrospray ionization in positive (ESI+) mode on an Agilent 6520 accurate mass quadrupole time-of-flight MS. Target analysis was performed on a TSQ Quantum Ultra triple quadrupole MS in ESI+ mode. Concentrations of selected PAHs were determined by HPLC with fluorescence detection. Data from nontarget analysis were processed as described in Chapter 4.

The compounds for which we had standards (acridine, phenanthridine, benzo[*h*]quinoline, benzo[*c*]acridine, benzo[*a*]acridine, and dibenzo[*a,j*]acridine) were used to quantify all isomers of the standard itself, assuming that the response factors were equivalent for each isomer. Overall, there were 34 quantifiable homologs.

5.2.5 Statistical analysis

The susceptibility to degradation of azaarene isomer groups was analyzed by hierarchical cluster analysis using Ward's minimum variance method, and the distances between clusters were measured by Euclidean distance. The number of groups was set at 2 ("biodegradable" and

“recalcitrant”). The differences between clustered groups were assessed by Mann-Whitney-Wilcoxon test. All data analyses were conducted in R. Statistically significant differences ($p < 0.05$) between pairs of samples were evaluated by t -tests using Microsoft Excel 2010.

Statistical Analysis. The susceptibility to degradation of azaarene isomer groups was analyzed by hierarchical cluster analysis using Ward’s minimum variance method, and the distances between clusters were measured by Euclidean distance. The number of groups was set at 2 (“readily biodegradable” and “recalcitrant”). The differences between clustered groups were assessed by Mann-Whitney-Wilcoxon test. All data analyses were conducted in R. Statistically significant differences ($p < 0.05$) between samples amended with urea vs. unamended samples were evaluated by t -tests using Microsoft Excel 2010.

5.3 Results and discussion

In our companion publication (Tian et al., 2017b), we assessed the diversity and abundance of azaarenes containing a single N hetero-atom in the four PAH-contaminated samples. Collectively across the four samples, the detected azaarenes represented 8 homologous series of congeners ranging from three rings to five rings. Each series contained congener groups corresponding to the unsubstituted azaarenes and their monomethyl-, dimethyl- and/or trimethyl-derivatives; the greatest intensities (which we infer as abundance, although we recognize that this is not always accurate, Huba et al., 2016) were observed for the unsubstituted and monomethyl derivatives. Overall, 232 congeners were detected, as summarized in Table 4.2. There were five major congener groups, designated A, B, C, E, and F. Groups A, B and C are cata-condensed compounds with molecular formulas $C_{13}H_9N$ (3-ring), $C_{17}H_{11}N$ (4-ring) and $C_{21}H_{13}N$ (5-ring), respectively; groups E and F are peri-condensed compounds with molecular formulas $C_{15}H_9N$ (4-ring) and $C_{19}H_{11}N$ (5-ring), respectively.

5.3.1 Biodegradability of native azaarenes

Each of the four PAH-contaminated samples was incubated for six weeks under biostimulated conditions to evaluate the biodegradability of the native azaarenes. The concentrations of the 34 quantifiable non-methylated congeners were monitored throughout the incubation, and the data are summarized in Appendix C; the time-course evolution of each congener in each sample is shown in Figure C1, and initial (T0) and final (T42) concentrations in microcosm incubations are provided in Table C1. Their concentrations in killed controls after six weeks of incubation did not show any significant differences with respect to initial (T0) concentrations, indicating that potential abiotic losses during incubation were negligible (Table C1). The extent of total azaarene degradation was moderate for FS and HS samples (36% and 43%, respectively). However, there were big differences between the more-highly contaminated KM and SC samples; in agreement with its higher degree of weathering, in the KM sample the five quantified azaarene groups were degraded only 15% after six weeks. In contrast, the much less-weathered SC sample had the highest extent of degradation, with a total azaarene removal of 85%.

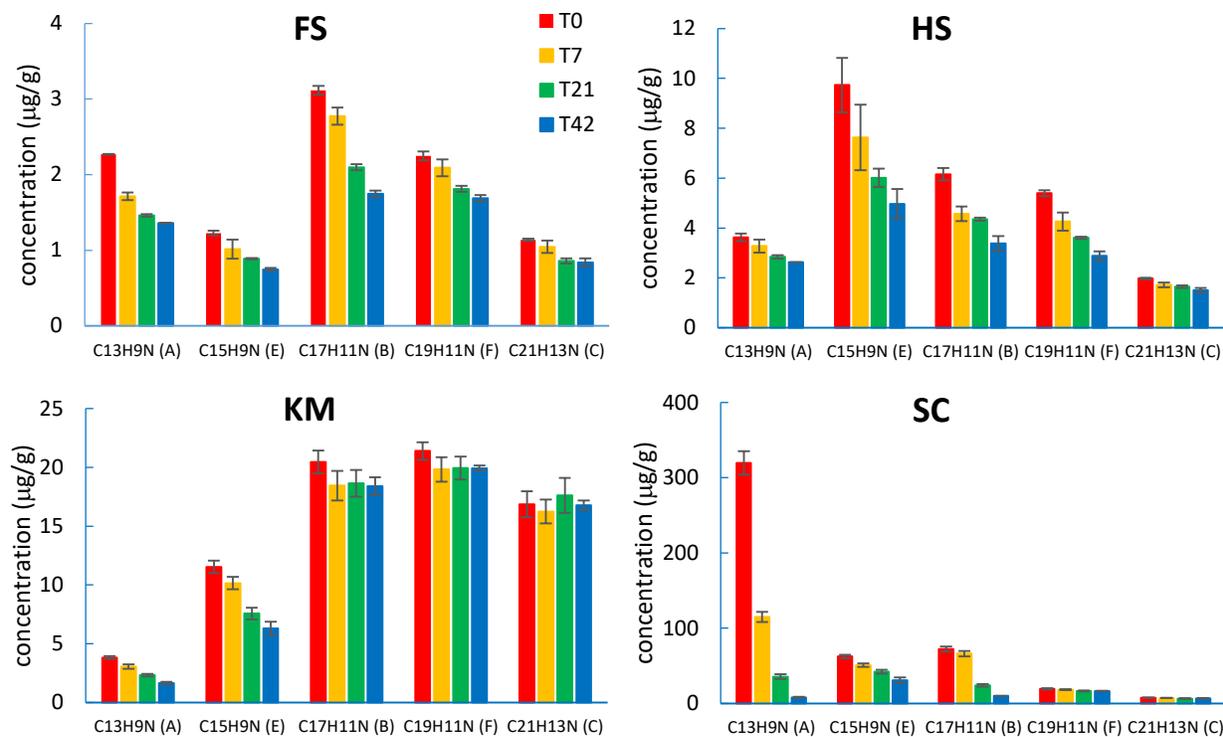


Figure 5.1 Concentrations of the five major azaarene congener groups over time during aerobic microcosm incubations in each of the samples. Note the differences in the y-axis scales.

The extent of degradation after six weeks was also compared across the different azaarene congener groups (Figure 5.1). In general, the LMW congener group A (3-ring PANHs) was the most extensively degraded (40-97%) except for the HS sample (27%), which might be attributed to the low initial concentrations in this sample (3.6 µg/g total group A compounds). The higher extent of degradation for congener group A is in agreement with the general trend of PAH biodegradation observed in many studies, for which LMW compounds are generally more easily degraded than 4- or 5-ring PAHs. Indeed, across samples the 5- ring congener groups F (15-46% removal) and C (0-26% removal) were the most recalcitrant.

For the more-weathered samples HS and KM, the azaarene analogues of fluoranthene and pyrene (group E) were more extensively degraded (49% and 45%) than the benzo[*a*]anthracene and chrysene analogues (group B, 45% and 10% removal, respectively). In contrast, for the less-weathered samples FS and SC, congener group B was more substantially degraded (44% and 86%, respectively) than group E (39% and 50%, respectively). It is not possible to conclude

definitively from this evidence that the degree of weathering (ratio of HMW to LMW PAHs) was the primary factor influencing differences in the extent of biodegradation among the samples. However, it is plausible that weathering would decrease the bioavailability of residual HMW contaminants and also would remove the more-readily biodegradable fractions of unsubstituted PAHs (and/or PANHs) if cometabolism is important for azaarene biodegradation. It is also of interest that for the less-weathered SC sample, the order of azaarene removal was A>B>E>F>C (Figure 5.1). This result is inconsistent with the typical pattern for PAH biodegradation in which fluoranthene and pyrene (azaarene analogues in group E) are more biodegradable than benzo[*a*]anthracene and chrysene (azaarene analogues in group B, Bossert and Bartha, 1986).

The biodegradability of the methylated congener groups was also evaluated qualitatively. Figure 5.2 shows the time course evolution of Kendrick Mass-Defect (KMD) plots for the SC sample; KMD plots are an efficient and convenient way of summarizing the congener groups as explained in Chapter 4. In general, the removal of the non-methylated compounds was greater than the removal of the mono- and di-methylated homologs, while tri-methylated families were not significantly removed (Figure 5.2). Methylated derivatives of 5-ring compounds (groups C and F) were generally resistant to microbial attack. Greater removal of the non-methylated congeners than mono- and di-methylated derivatives is in agreement with the degradation trends generally observed during microbial degradation of alkylated PAHs within petrogenic mixtures or in crude oil-polluted samples (Vila and Grifoll, 2009; Wang et al., 1998).

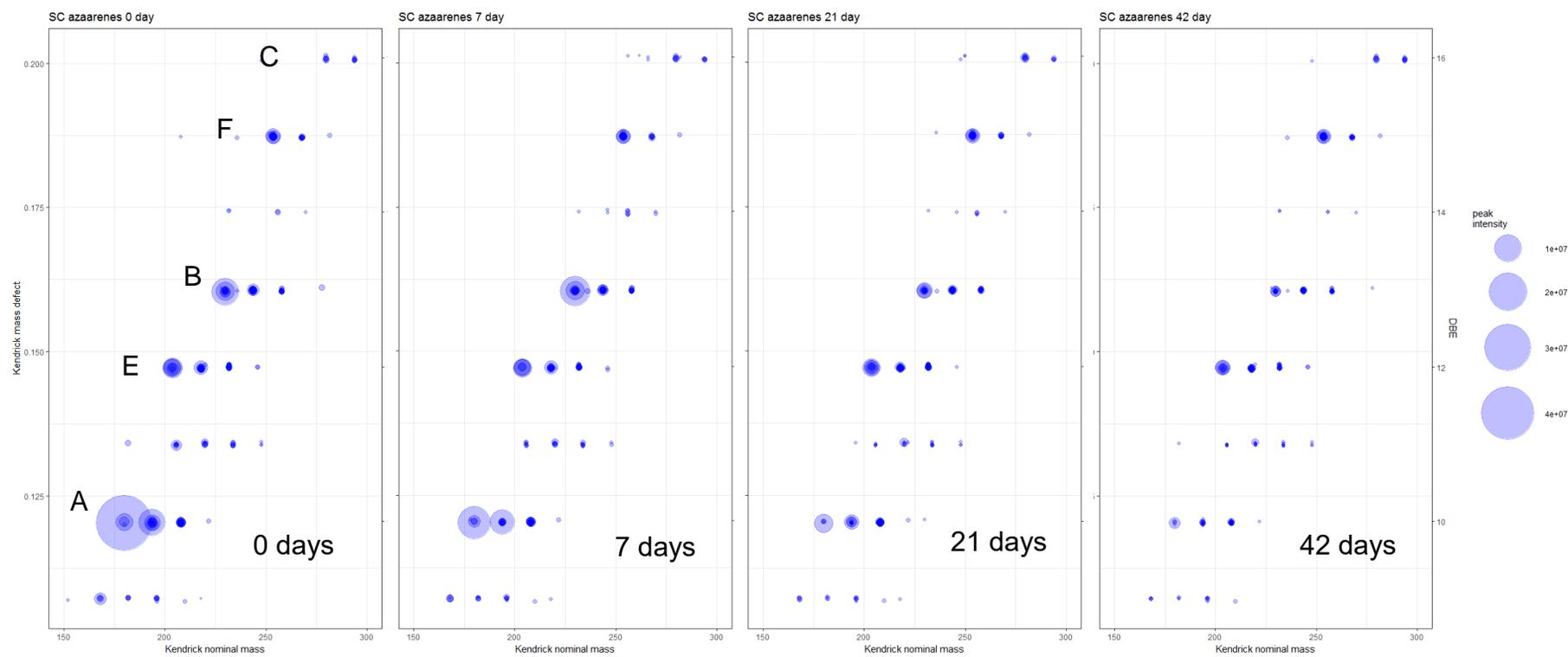


Figure 5.2 KMD plots over the time series from aerobic microcosm incubations of the SC sample under biostimulated conditions. The KMD plot describes the abundance and diversity of the detected azaarenes in eight homologous series of congener groups; the five groups of greatest interest (most abundant congeners) are labeled as follows: A (C₁₃H₉N), E (C₁₅H₉N), B (C₁₇H₁₁N), F (C₁₉H₁₁N), and C (C₂₁H₁₃N). Each horizontal series of bubbles corresponds to a homologous series of congeners that share the same basic structure, while the overlapping bubbles are isomers with identical molecular formula and position on the KMD plot. The color darkened by overlapping suggests the emergence of multiple isomers, and the areas of the bubbles are proportional to the intensities. Within a given series, the degree of methylation increases from left to right.

5.3.2 *Isomer-selective biodegradation*

Within a given congener group we found significant differences in the extent of biodegradation among isomers, suggesting a selective microbial transformation of certain isomers. Such isomer-selective biodegradation has previously been observed in structurally diverse pollutants such as nonylphenols (Lu and Gan, 2014) and alkylated PAHs (Bayona et al., 1986; Lamberts et al., 2008). To quantitatively assess the biodegradability of individual non-methylated isomers over the course of the incubations, the concentration data from the time series for all samples were normalized by averaging the ratio of concentrations between neighboring time points for each isomer; this approach accounts for both the rate and extent of biodegradation over the six-week incubations. Isomers were then categorized by cluster analysis into two significantly different ($p < 0.01$) groups: recalcitrant compounds and readily biodegradable compounds (Figure 5.3). It is important to note that “recalcitrant” does not imply non-biodegradable. Here it means a statistically significant, lesser removal of a compound relative to one categorized as readily biodegradable when considering all of the time-series data for all four samples. Absolute biodegradability of a particular congener (significantly lower concentration at 42 days compared to the initial concentration) in a given sample is illustrated in Table C1.

As expected from the analysis of congener groups described above, all the isomers in groups C and F (5-ring compounds) clustered in the recalcitrant group, while isomers in group A (3-ring compounds) fell into the readily biodegradable group (Figure 5.3). However, for the 4-ring congener groups B and E, significant differences were observed in the biodegradability among isomers. While most of the B isomers clustered in the readily biodegradable group, isomer B4 clustered in the recalcitrant group. Two E group isomers (E3 and E5) were classified

among the readily biodegradable compounds, whereas three isomers (E1, E2 and E4) were classified as recalcitrant.

Examples of differences in biodegradation between isomers are shown in Figure 5.4. Of particular interest is the difference between isomers B1 and B6 in the SC sample (Figure 5.4c), which differ only in the position of the N atom in the structure. The great difference between E4 (azapyrene) and E5 (azafluoranthene) is striking (Figures 5.3, 5.4a and 5.4b), as their PAH counterparts are both relatively biodegradable. Among the B- and E-group isomers, the relationship between structure and biodegradation potential is still unclear due to the lack of authentic standards. However, the HPLC elution order of isomers (Figure 4.3) generally corresponded to isomer biodegradability (Figure 5.3), such that later-eluting isomers were generally more biodegradable (illustrated for two of the samples in Figure 5.4). This observation suggests that one of the parameters controlling the elution (polarity or basicity) could be inversely related to biodegradability.

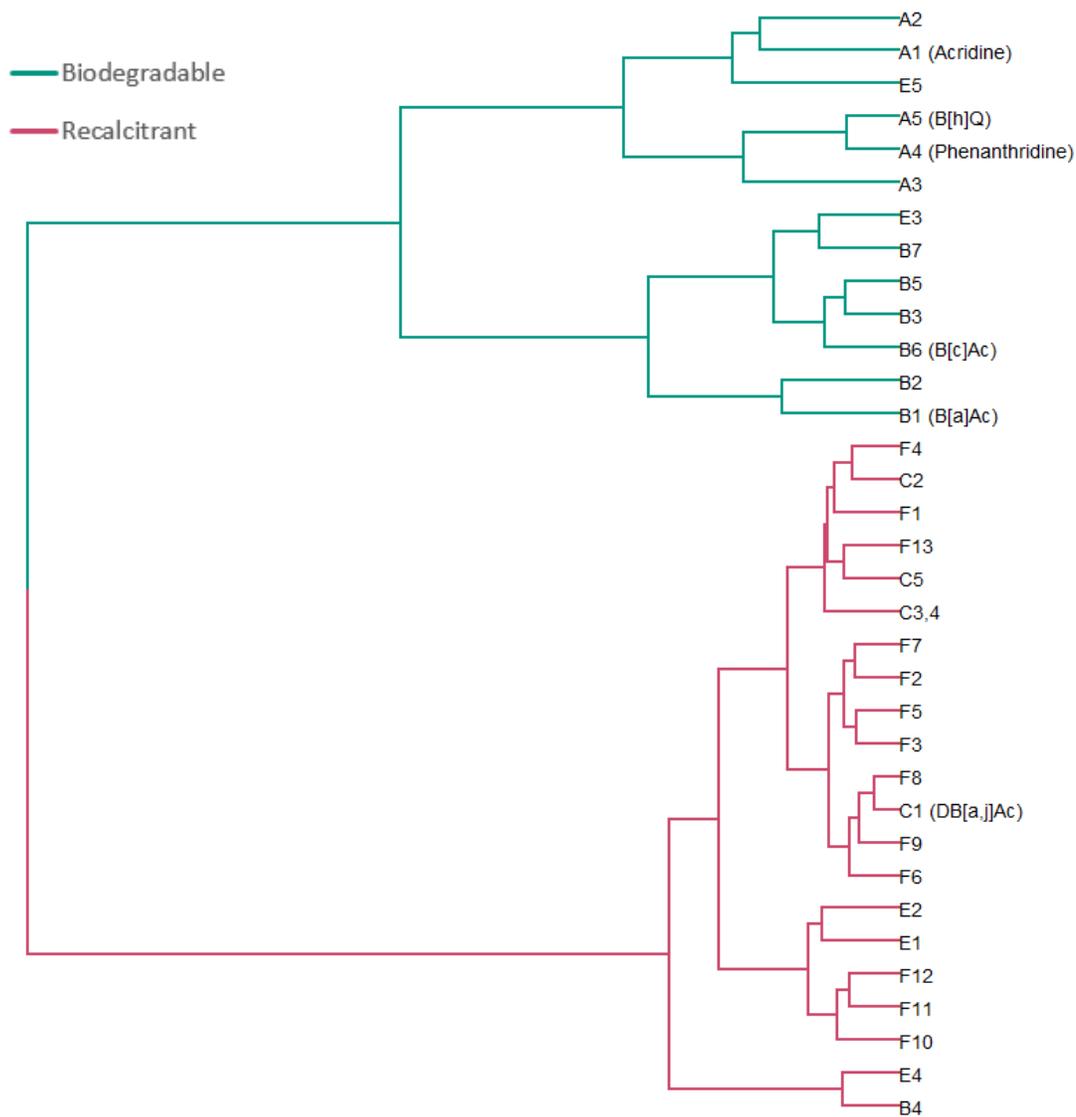


Figure 5.3. Hierarchical cluster analysis of the quantifiable azaarene congeners according to their biodegradation behavior (using a time-series measure that combines the rate and extent of removal) during microcosm incubations. Ward's minimum variance method was used, and the congeners were categorized into the two clearly separated groups.

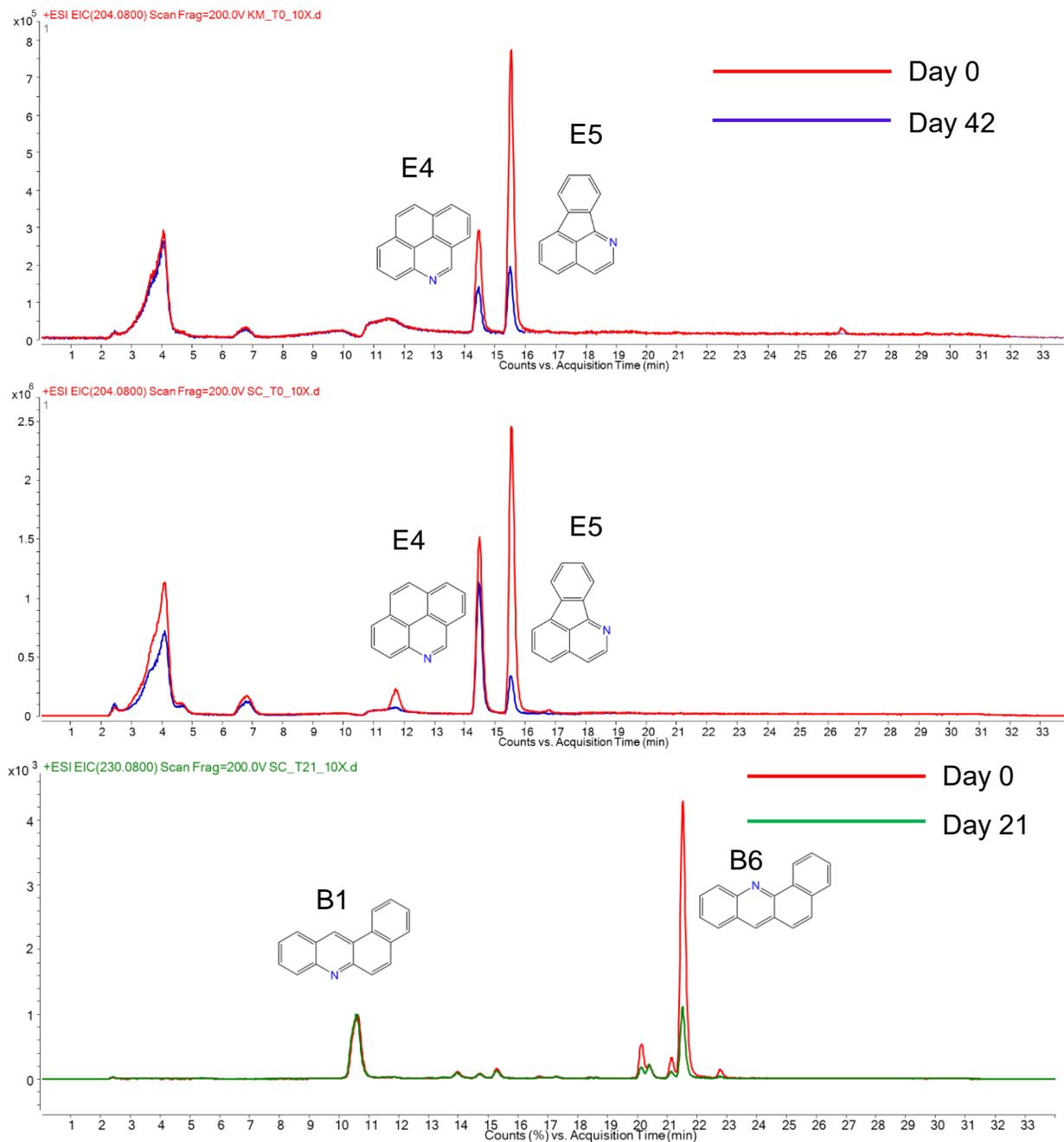


Figure 5.4. Overlaid extracted ion chromatograms (EICs) for m/z 204.0800 from KM (top) and SC (middle) samples at day 0 and after 42 days of microcosm incubation, and for m/z 230.0800 from the SC sample (bottom) at day 0 and after 21 days of incubation. Each panel depicts the isomer-specific extent of degradation within an azaarene congener group. Data in the bottom panel are shown for 21 days of incubation because several of the analytes were completely removed by Day 42.

5.3.3 *Effect of exogenous nitrogen addition on azaarene biodegradation*

The presence of a nitrogen atom in the azaarene molecule suggested its potential to be used as a nitrogen source for microbial growth, facilitating the assimilation of the carbon skeleton. To evaluate whether the addition of an external nitrogen source was essential for azaarene biodegradation, we conducted 42-day incubations of the FS and SC samples in the presence and absence of added urea (Figure C2). In general, in the absence of urea there was a significantly lower degradation of all congener groups, and in no case did the omission of urea lead to improved degradation of any isomer. This finding does not exclude the possibility that the azaarenes could serve as a nitrogen source for the microbial community, but neither does it provide support for such a possibility (we do not know the background concentrations of other possible nitrogen sources in these samples). Conversely, the addition of urea clearly led to significantly greater removal of a number of the isomers in both samples, particularly of some of those concluded to be recalcitrant to biodegradation (*e.g.*, E1, E2, E4, F5 or F9 for the FS sample, and E1, E2, F5 or F9 for the SC sample). Overall, these results underscore the benefit of adding an external nitrogen source to maximize biostimulation.

5.3.4 *Biodegradability of azaarenes relative to their PAH analogues*

We compared the concentrations and biodegradation behaviors of azaarene congener groups with respect to their corresponding homocyclic counterparts. It has been estimated that NSO-heterocyclic polycyclic aromatic hydrocarbons are on the order of 1% to 10% of the total mass of homocyclic PAHs in environmental mixtures (Eric A. J. Bleeker et al., 2002; Neilson, 1995). For two of the samples in our case (HS and SC), ratios of azaarene congeners to their respective PAH analogues were substantially higher than that (Figure C3). During the microcosm incubations, the proportions of azaarenes to analogue PAHs remained relatively constant for the weathered samples HS and KM. However, they generally increased over time for the HMW congeners in

the highly contaminated SC sample (Figure C3), suggesting the greater recalcitrance of HMW azaarene congeners relative to their homocyclic counterparts. Depending on the toxicity of azaarenes relative to their PAH counterparts, these observations may have implications for residual risk following bioremediation.

5.3.5 Transformation products from azaarenes

The nontarget analysis approach was applied to detect the formation or fate of putative transformation products from azaarenes during the microcosm incubations. Considering that incubations were performed under aerobic conditions, primary attention was paid to oxygenated derivatives of azaarenes that included one or two oxygen atoms. Four major features (groups of related analytes) were detected, whose occurrence and degradation were specific to samples and time points (Figure 5.5). The first group corresponded to the formula $C_{13}H_9NO$ (m/z 196.0760), which includes acridone, phenanthridinone, and their isomers (Figure 5.5a). These compounds have been detected in estuarine environments (de Voogt and Laane, 2009), and similar ketones derived from quinolines were suggested as indicators of natural attenuation in groundwater at coal tar-contaminated sites (Reineke et al., 2007). Their concentrations decreased during the incubation process, suggesting that they were already present in the initial samples and were readily biodegradable. Other groups of compounds possessed the formulas $C_{13}H_9NO_2$ (m/z 212.0706), $C_{14}H_{11}NO_2$ (m/z 226.0864), and $C_{17}H_{11}NO_2$ (m/z 262.0862), corresponding to dioxygenated azaarenes. Although these formulas are consistent with dihydroxy-compounds, analogous HMW dihydroxy-PAHs are highly unstable and easily autoxidized to quinones (Jouanneau et al., 2007; Penning et al., 1996). Other possibilities include the monohydroxy derivatives of acridone-like structures (Figure 5.5 b and c), which are more-stable tautomers of dihydroxy compounds (Fetzner, 1998). In the SC samples the concentrations of these analytes peaked at 7 d (for $C_{14}H_{11}NO_2$) or 21 d (for $C_{17}H_{11}NO_2$) and decreased thereafter (Figure 5.5 b

and c), suggesting their transient formation and subsequent removal during the biodegradation process. In addition, their emergence was concomitant with the removal of the corresponding parent azaarene congener groups, further suggesting that they were oxidized products from azaarenes. Except for the acridine derivatives ($C_{13}H_9NO$, Bobeldijk et al., 2002; Brinkmann et al., 2014), the other oxidized azaarenes are reported here for the first time.

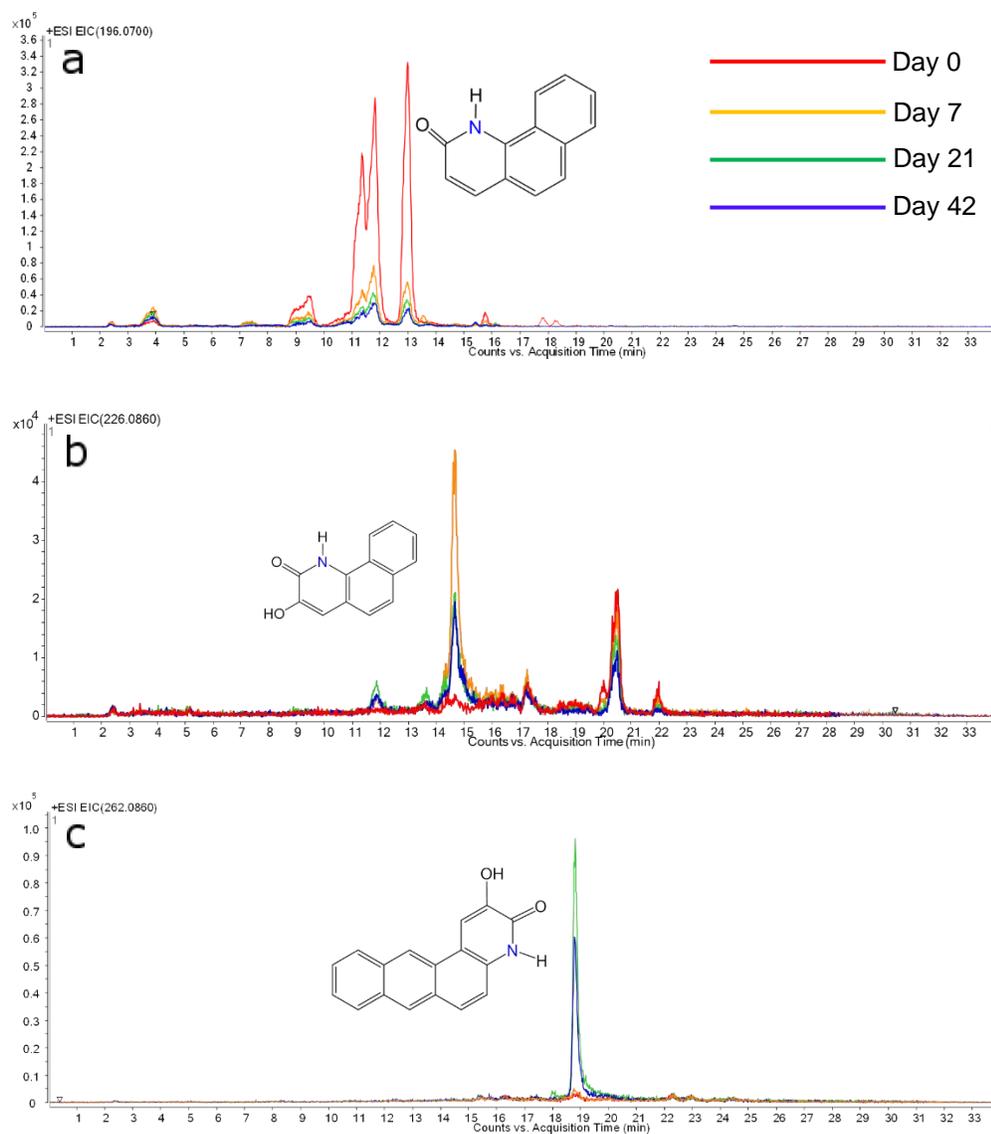


Figure 5.5. EICs of azaarene derivatives in the SC sample. (a) $C_{13}H_9NO$ (m/z 196.0760); (b) $C_{14}H_{11}NO_2$ (m/z 226.0864); (c) $C_{17}H_{11}NO_2$ (m/z 262.0862). Structures are hypothetical, and positions of the N atom and the substituents in the structures shown are illustrative only.

5.3.6 *Environmental significance*

Current risk assessment approaches during remediation of PAH-contaminated soils generally consider reference levels for only 16 regulated PAHs. However, we have demonstrated the co-occurrence of a wide range of PANHs, including HMW compounds and methylated derivatives. Considering the known toxicological properties of several azaarenes, here we assessed their potential environmental fate during bioremediation. Like their PAH counterparts, there are substantial differences in their susceptibility to biodegradation as a function of molecular weight (number of rings) and degree of methylation. However, the position of the N atom and other parameters associated with their physicochemical properties (such as basicity or polarity) may affect the susceptibility to biodegradation between isomers of the same congener group. We have also demonstrated the potential for accumulation of azaarene transformation products under aerobic conditions. These findings on the biodegradation of azaarenes will lend support to future research in environmental monitoring as well as bioremediation for this overlooked class of contaminants.

CHAPTER 6 TRACING BIOTRANSFORMATION OF PAHS IN CONTAMINATED SOIL USING STABLE ISOTOPE-ASSISTED METABOLOMICS⁴

6.1 Introduction

Bioremediation is a well-accepted technology for the treatment of PAH-contaminated soil, but it might not be sufficient to reduce soil genotoxicity due to the formation of oxidation products (Chibwe et al., 2015b; Hu et al., 2012; Lemieux et al., 2015). In a recent study we identified a previously unknown bacterial metabolite of pyrene in contaminated soil after bioremediation (Tian et al., 2017a), illustrating that current knowledge on PAH biodegradation pathways gathered from the study of isolated strains on single compounds can be limited. Interactions among organisms in soil microbial communities or between components of complex PAH-mixtures (Grifoll et al., 1995; Zhong et al., 2006) can result in cometabolism, (Bouchez et al., 1995) competitive metabolism (Hennessee and Li, 2016; Knightes and Peters, 2006; Stringfellow and Aitken, 1995), or cooperation (Luo et al., 2014; Zhong et al., 2011); such phenomena can enhance, inhibit, or modify the extent and/or profiles of metabolite formation.

An effective strategy to trace the biotransformation of specific pollutants within environmental samples is using stable or radioactive isotope-labeled compounds (Liu et al., 2013; Richnow et al., 1998, 2000; Wang et al., 2017). For radioactive isotopes, their special safety requirements can limit wider application. Conversely, ¹³C-labeled substrates could be a suitable tool to shed light on the biodegradation pathways of organic pollutants (Fischer et al.,

⁴ Zhenyu Tian, Joaquim Vila, Miao Yu, Wanda Bodnar, and Michael D. Aitken
Published in Environmental Science & Technology Letters 2018, 5 (2), 103–109

JV (co-first author) designed and conducted the experiments, and helped with writing, MY helped with the data analysis, and WB assisted with the HRMS data interpretation.

2016; Morasch et al., 2011). However, the identification of labeled metabolites has been restricted so far by the resolution of mass spectrometry and by data analysis methods. Recently, isotope labeling combined with high-resolution mass spectrometry (HRMS) has been applied to track the abiotic transformation of pollutants in aqueous mixtures (Kolkman et al., 2015; Sun et al., 2016). Adapting this combination of techniques to understanding biotransformation of the components of complex environmental mixtures, such as coal or oil derivatives, requires more powerful data analysis of HRMS data.

The application of stable isotope-assisted metabolomics (SIAM) could provide new insights on the biotransformation of pollutants and unleash the inherent power of isotope labeling. Such methods have already been established in the field of metabolomics (Giavalisco et al., 2008; Guo and Li, 2009; Hegeman et al., 2007; Hiller et al., 2010), and various automatic workflows are available (Bueschl et al., 2017; Huang et al., 2014; Leeming et al., 2015; Zhou et al., 2014). The basic principle of SIAM is that when ^{13}C labeled substrates are added to mixtures containing natural-abundance (unlabeled) counterparts and are biotransformed into labeled metabolites, they can be detected in groups with unlabeled metabolites by high performance liquid chromatography (HPLC) coupled with HRMS. Algorithms enable such pair or group selection by picking up features with (1) the same chromatographic retention time; (2) certain exact mass differences; and (3) diagnostic intensity ratios (Kluger et al., 2014; Neumann et al., 2014). The SIAM approach differs in this respect from previous applications of HRMS with isotope-labeled compounds in environmental science cited above. Among other differences, SIAM cannot be used with deuterated compounds because they elute at different retention times than their unlabeled counterparts. Here, we applied SIAM to trace the biotransformation of PAHs in contaminated soil samples. Our objectives were (1) to check the feasibility of this method for organic pollutants in complex mixture using uniformly ^{13}C ($\text{U-}^{13}\text{C}$) labeled PAHs as a model, and (2) to apply this method to identify the compound-specific metabolite profiles and

degradation pathways for three four-ring PAHs in real contaminated soil samples during an active, laboratory-scale bioremediation process.

6.2 Materials and methods

6.2.1 Chemicals.

Anhydrous sodium sulfate and HPLC-grade solvents were purchased from Fisher Scientific (Pittsburgh, PA, U.S.A.). U-¹³C labeled fluoranthene, pyrene, and benzo[*a*]anthracene (B[*a*]A) were synthesized according to previous publications (Singleton et al., 2006; Zhang et al., 2011).

6.2.2 Contaminated soil and spiking.

Creosote-contaminated soil was collected from a wood-treatment facility in Andalucía, Spain with a 100-year history of pollution (Tejeda-Agredano et al., 2013). PAHs concentrations are listed in Table D1. For the preliminary method assessment, treated soil from a different site was collected from a lab-scale aerobic bioreactor (Tian et al., 2017a).

For each U¹³C-labeled PAH (fluoranthene, pyrene, or B[*a*]A), the desired mass in acetone solution was spiked into 0.4 g (20% of a 2 g sample) of dry contaminated soil. After solvent evaporation overnight, the spiked soil was mixed with the remaining 1.6 g of unspiked soil in successive 0.4 g increments (Jones et al., 2008). The amount of spiked U-¹³C PAH was adjusted to maintain a 1:3 (mass:mass) proportion with respect to the corresponding native PAH concentration in the soil. Unspiked samples were amended with the same level of acetone as the spiked samples.

6.2.3 Microcosm incubations.

Aliquots of 1 g of soil (dry wt) were placed in sterile 125 mL Erlenmeyer flasks containing 30 mL of 10 mM phosphate buffer (pH 7.5) supplemented with urea (1 mM) as nitrogen source. All flasks were incubated at 25 °C under agitation at 150 rpm. Triplicate

unspiked flasks were sacrificed initially (time 0) and after 7, 14, and 21 d of incubation. Triplicate killed controls were prepared by acidifying (pH 2) using phosphoric acid, and sacrificed after 21 d. Duplicate flasks containing U-¹³C fluoranthene, pyrene or B[a]A were sacrificed after 7 or 14 d. Optimal incubation times for each target PAH were determined in a preliminary incubation using unspiked soil as described in Figure D1.

After incubation, the flask contents were centrifuged to recover the solid and aqueous phases. The solid phase was processed as described previously (Hu et al., 2012). The aqueous phase was acidified (pH 2) and extracted with ethyl acetate. Further details on sample extraction and processing are provided in Appendix D.

6.2.4 Instrumental analysis.

Extracts were analyzed using an Agilent 1200 series HPLC interfaced with a 6520 accurate mass quadrupole time-of-flight mass spectrometer (qTOF-MS, Agilent Technologies, Santa Clara, CA.). Full scan (m/z 100-1000) with electrospray ionization (ESI) was used for SIAM analysis, and product ion scanning (MS/MS) was applied to elucidate the putative structures of metabolites. Details and parameters are provided in Appendix D.

6.2.5 Data analysis.

Raw data from qTOF-MS were converted into .mzData or .mzXML format for in-depth analysis. Peak detection and alignment were achieved using the R package XCMS (Patti et al., 2012) and the publicly available Scripps XCMS Online platform (<https://xcmsonline.scripps.edu/>) (Tautenhahn et al., 2012). Isotope group picking was conducted in R using the package X13CMS (Huang et al., 2014). Detailed processing methods and parameter settings are provided in Appendix D.

6.3 Results and discussions

6.3.1 Assessing performance of the method.

Considering the previous identification of a genotoxic metabolite (2H-naphtho[2,1,8-*def*]chromen-2-one; NCO) during the incubation of bioreactor-treated soil with pyrene (Tian et al., 2017a), we first assayed three different ratios of U-¹³C to natural abundance (native) pyrene in the bioreactor slurry to test the feasibility of the method. Ratios of 1:1 and 1:3 produced ¹³C-NCO in sufficient quantity for its analytical detection, while 1:9 could not. Considering the cost and scarcity of U-¹³C compounds, and to minimize their impact on the microbial community as an enrichment substrate, we chose the 1:3 ratio to perform our experiments on the creosote-contaminated soil.

To detect the natural abundance and ¹³C-enriched pairs of metabolites, we applied the “getIsoDiffReport” function of the X13CMS package to select the peak groups that shared the same retention time but presented *m/z* values that differed in $N \times 1.00335$, where N is the number of ¹³C atoms (N varied depending on the metabolite from a given PAH) (Huang et al., 2014). With this approach a smaller number of isotope groups were identified in the solid-phase compared to the liquid-phase extracts. However, with this approach most of the selected groups were false positives that only contained ions corresponding to [M-H]⁻ and [M-H+1]⁻, which could be assigned to the background of natural abundance ¹³C.

To restrict the number of selected groups, we first restricted the ratio of labeled and unlabeled metabolites according to the 1:3 ratio between the spiked ¹³C PAHs and the native compounds, using the “enrichmentLvsU” parameter in the X13CMS package. We then compared spiked and unspiked samples obtained from the same incubation time using XCMS Online. These two samples differed only in the presence of isotope labeling, and the features significantly higher in the spiked samples should correspond to labeled metabolites. By

integrating the results from both approaches we generated the final list of prioritized features (Table 6.1).

Table 6.1. Detected metabolites from the four-ring PAHs

RT ^a	Phase ^b	Ionization	<i>m/z</i> ¹² C	<i>m/z</i> ¹³ C	formula	Confidence level ^c	tentative identification	PubChem Identifier
Fluoranthene								
23.1	W	ESI-	217.0652	233.1192	C ₁₆ H ₁₀ O	3	Hydroxy-fluoranthene	3014645
20.1	S	ESI+	233.0567	249.1094	C ₁₆ H ₈ O ₂	3	Fluoranthene <i>o</i> -quinone	21481
14.2	W	ESI-	249.0562	265.1068	C ₁₆ H ₁₀ O ₃	4		
14.1	W	ESI-	195.0459	208.0874	C ₁₃ H ₈ O ₂	3	Hydroxy-fluorenone	80663
7.3	W	ESI-	237.0398	248.0695	C ₁₁ H ₁₀ O ₆	4	(2-Carboxybenzyl) malonic acid	
Pyrene								
20.0	S	ESI+	233.0567	249.1094	C ₁₆ H ₈ O ₂	3	Pyrene <i>o</i> -quinone	160814
18.1	W	ESI-	313.0179	329.0716	C ₁₆ H ₁₀ O ₂ ·SO ₃	3	Dihydroxy-pyrene sulfate conjugate	
20.4	W	ESI-	358.0043	374.0572	C ₁₆ H ₉ NO ₄ ·S O ₃	4		
21.9	W	ESI-	369.9945	385.0452	C ₁₅ H ₅ N ₃ O ₉	4		
22.3								
21.2	W	ESI-	342.0005	356.0448	C ₁₄ H ₅ N ₃ O ₈	4		
21.4	W	ESI-	386.9854	401.0332	C ₁₄ H ₄ O ₄ ·2N ₂ O ₃	4		
Benzo[<i>a</i>]anthracene								
19.9	W	ESI-	339.0331	357.0932	C ₁₈ H ₁₂ O ₂ ·SO ₃	3	Dihydroxy benzo(<i>a</i>)anthracene	
15.9	W	ESI-	289.0178	303.0645	C ₁₄ H ₁₀ O ₂ ·SO ₃	3	Dihydroxy-anthracene; dihydroxy-phenanthrene	
11.1	W	ESI-	271.0614	286.1104	C ₁₅ H ₁₂ O ₅	4	1-Hydroxy-2-carboxy-anthracene dihydrodiol	
2.6, 5.9	W	ESI-	247.0246	259.0640	C ₁₂ H ₈ O ₆	3	Dihydroxynaphthalene-dicarboxylic acid (2 isomers)	69414683

^a RT, retention time; ^b W, aqueous (water) phase; S, solid phase

6.3.2 Metabolite identification

Because we tested only ESI in positive and negative ionization modes, it is possible that some metabolites escaped analysis. Metabolites present in low abundance, with very low ionization potential, or which would be better detected with a different ionization source could have been missed. However, our primary objective was to establish proof-of-concept of SIAM for applications in complex environmental systems, rather than an exhaustive search for all possible metabolites.

For each target PAH, several labeled features were detected. Tentative structures of metabolites (Table 6.1) were based on their exact mass, MS/MS fragmentation patterns, deduced number of labeled carbons, and further comparison with data available in the literature and pathway prediction systems such as EAWAG-BBD (Wicker et al., 2010) and enviPath (Latino et al., 2017). Standards were not available for any of the proposed metabolites. With a formula and MS/MS pattern, each structure was processed with the MetFrag (Ruttkies et al., 2016) platform and searched in the PubChem database (<https://pubchem.ncbi.nlm.nih.gov/>), and five of them matched the record.

In the solid phase, labeled metabolites were detected only for fluoranthene and pyrene. The identified products with the formula $C_{16}H_8O_2$ corresponded to the respective *o*-quinones (see Figure D2 for the fluoranthene quinone), which were previously found to accumulate during incubation with pure bacterial strains (Kazunga et al., 2001; Kazunga and Aitken, 2000). The low number of products detected in the solid phase suggested that most of the metabolites were formed in or transferred to the aqueous phase of the slurry due to their increased polarity relative to the parent compounds.

Two major groups of compounds were detected in the liquid phase: highly oxygenated ring-cleavage products and PAH conjugates. Within the first group, we detected several formulas originating from fluoranthene ($C_{11}H_{10}O_6$; Figure 6.1) and B[a]A ($C_{12}H_8O_6$, Figure D3; and

$C_{15}H_{12}O_5$) corresponding to ring-cleavage products containing one or more carboxyl group (mass spectra consistent with loss of at least one fragment of m/z 44). Regarding the fluoranthene products, $C_{11}H_{10}O_6$ did not correspond to any previously identified metabolite, and is proposed as (2-carboxybenzyl) malonic acid (Figure 6.1A and Figure D6d). The B[a]A metabolite $C_{15}H_{12}O_5$ has not been previously described in the literature, and is proposed as a dihydrodiol of the 1-hydroxy-2-anthracene carboxylic acid. For the formula $C_{12}H_8O_6$, the MS/MS spectral characteristics (Figure D6e) are identical to those of 1,6-dihydroxynaphthalene-2,7-dicarboxylic acid, a known ring-fission product of B[a]A from *Sphingobium* sp. strain KK22 (Kunihiro et al., 2013).

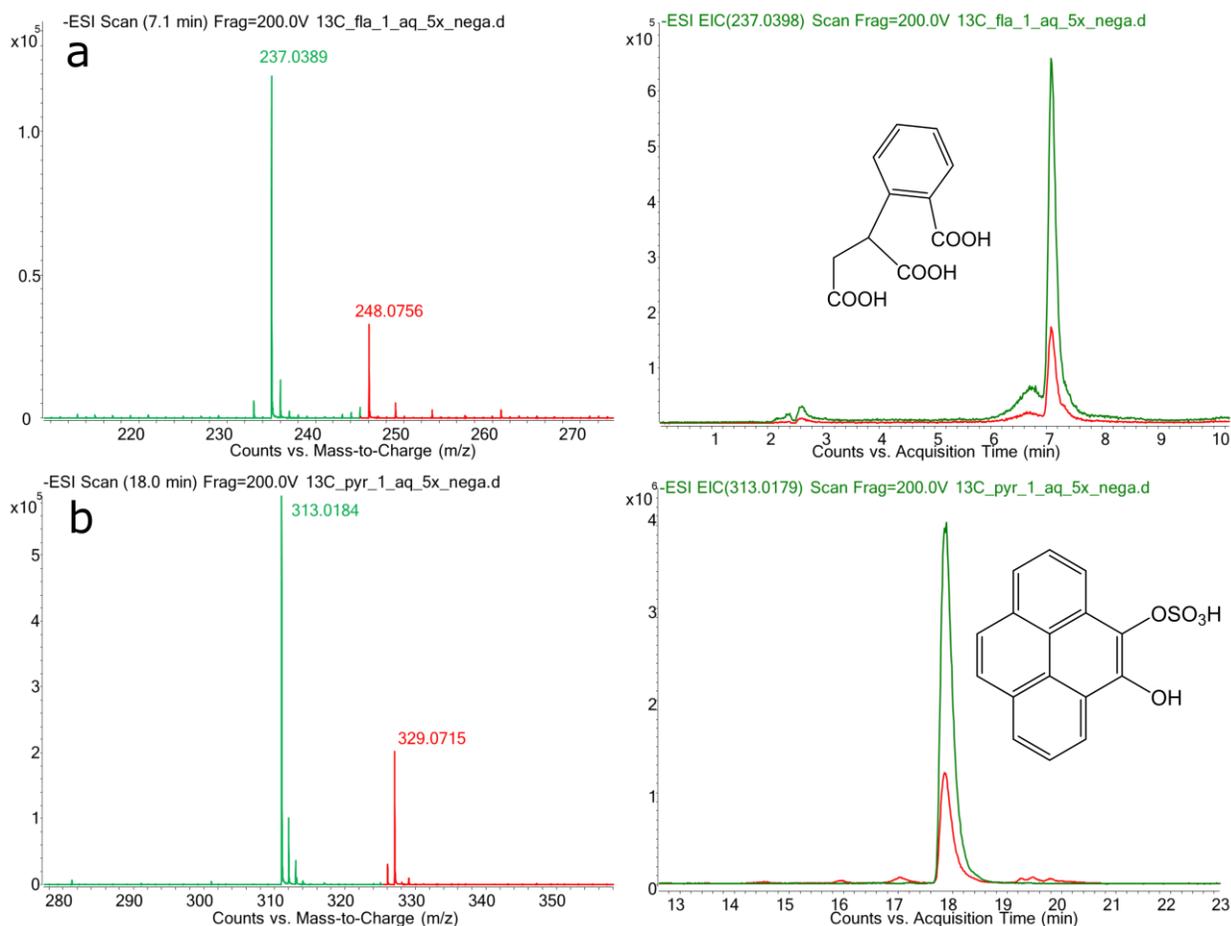


Figure 6.1. Mass spectra, extracted ion chromatograms (EIC), and proposed structures of two representative metabolites. **a:** $C_{11}H_{10}O_6$ from fluoranthene; **b:** $C_{16}H_{10}O_2 \cdot SO_3$ from pyrene. The green lines represent ^{12}C metabolites, and the red lines represent $U-^{13}C$ labeled metabolites.

A second group of products were conjugates of PAH metabolites, mainly sulfate conjugates of dihydroxylated PAHs. We detected several peaks corresponding to the sulfate conjugates of dihydroxy-pyrene (Figure 6.1b), dihydroxy-B[a]A, and either dihydroxy-anthracene or -phenanthrene derived from B[a]A (Figure D3). However, the accumulation of dihydroxylated PAHs was unexpected, as they are highly unstable in the environment (Jouanneau et al., 2007; Kazunga and Aitken, 2000). Each of these dihydroxy-compounds was always associated with a co-eluting peak with m/z corresponding to $[M-H+80]^-$ (Figure D5). According to the exact mass (m/z 79.9570), this fragment was assigned to an SO_3 group, indicative of a sulfate conjugate (Riva et al., 2015). To further confirm this observation, we repeated the analysis decreasing the fragmentor voltage to 100V, which increased the sulfate conjugate signal while that of the dihydroxy compounds disappeared, confirming the sulfate conjugates as the actual metabolites. Such sulfate conjugation of dihydroxy-PAHs is common in eukaryotic organisms, including fungi (Pothuluri et al., 1996; Schmidt et al., 2010), but has never been observed in bacteria. These results suggest that fungi might have played a significant role in PAH biotransformation and detoxification in the contaminated soil.

Besides the sulfate conjugates, we observed a number of other candidate conjugation products, especially in the incubations containing U- ^{13}C pyrene (m/z 342.0005, 358.0003, 369.9945, and 386.9854, Figure D4). Although their formulas could not be unambiguously assigned, all of these features had an m/z over 300, and some of the deprotonated ions presented nominal masses in even numbers, suggesting the existence of nitrogen atoms. Their MS/MS spectra revealed fragments with exact masses compatible with the loss of different nitrogen oxides (NO or NO_2) but especially of m/z 75.99, corresponding to the formula N_2O_3 (Figure D6). The formulas of the conjugates were therefore assigned according to their fragmentation patterns. PAH metabolites with such fragments, as well as conjugation reactions with nitrogen oxides, have not previously been described. Based on current evidence, it is difficult to elucidate

the identities of these unknown pyrene metabolites. Considering their novelty and diversity, these nitrogen-containing metabolites should be interesting targets for future research.

6.3.3 Proposed compound specific degradation pathways in soil

Based on the detected ^{13}C -labeled metabolites, the pathways for B[a]A and fluoranthene degradation in the PAH-contaminated soil are proposed (Figure 6.2 and Figure D, respectively). It is important to note that these pathways do not necessarily result from metabolism of the parent compound by an individual organism, and therefore the specific metabolites and their accumulation over time are less meaningful than for incubation of a single substrate with a pure culture. B[a]A biodegradation mainly followed the preferential degradation route previously described for *Sphingobium yanoikuyae* B1 (formerly *Beijerinckia*) (Jerina et al., 1984; Mahaffey et al., 1988), which is initiated by dioxygenation of the molecule in positions 1,2. Further *meta*-cleavage and pyruvate release would lead to the formation of 1-hydroxy-2-anthracenoic acid, an intermediate also observed in other *Sphingobium* strains (Kunihiro et al., 2013), followed by oxidation to a dihydrodiol ($\text{C}_{15}\text{H}_{12}\text{O}_5$) that is reported here for the first time. The identification of two isomers of dihydroxynaphthalene dicarboxylic acid (Table 6.1) indicated the dehydrogenation and ring cleavage of the latter (Kunihiro et al., 2013). Several fluoranthene metabolites are not consistent with established pathways, but the identification of hydroxy-fluorenone suggests the processing of one of the two fused aromatic rings (Kweon et al., 2007; Rehmann et al., 2001). We observed evidence for both initial mono- ($\text{C}_{16}\text{H}_{10}\text{O}$) and dioxygenation ($\text{C}_{16}\text{H}_8\text{O}_2$) of the fluoranthene molecule (Figure D1), but none of the previously described ring-cleavage products was detected. The MS/MS analysis of the formula $\text{C}_{16}\text{H}_{10}\text{O}_3$, corresponding to a metabolite not previously identified, presented a fragment of m/z 44, indicative of a carboxyl group (Figure D5c) and suggesting its formation as a product from the initial ring fission. Because this metabolite contained 16 C atoms and only three oxygen atoms,

instead of the expected four for a common *meta*- or *ortho*-cleavage product, it could not be assigned to any previously described ring fission metabolites (Kweon et al., 2007; van Herwijnen et al., 2003). Finally, the product with formula $C_{11}H_{10}O_6$ was identified here for the first time, and could be consistent with a metabolite resulting from the cleavage and further degradation of the biphenylene derivative originating from angular attack on fluorenone (Kweon et al., 2007).

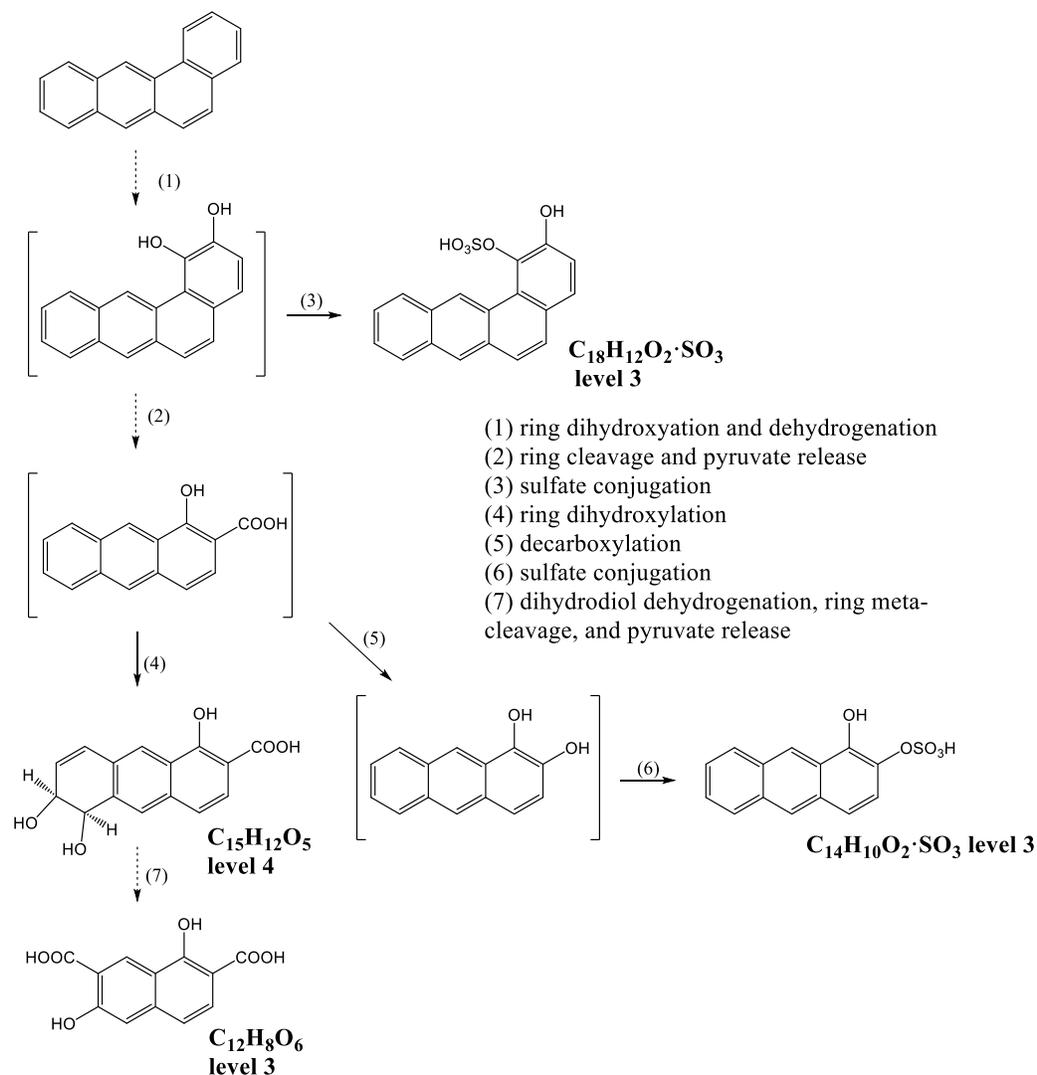


Figure 6.2. Proposed degradation pathway for benzo[*a*]anthracene based on detected metabolites and published literature (Kunihiro et al., 2013). Note that substituent positions are either inferred from known mass spectra or are hypothesized. Dashed arrows indicate more than one successive reaction. Structures in brackets are putative and were not observed directly. Reactions follow the canonical reactions in aerobic fungal and bacterial metabolism of PAHs summarized in Cerniglia et al. (Cerniglia, 1997) and Kweon et al. (Kweon et al., 2011), respectively. The levels indicated next to structures are the level of confidence as defined in Schymanski et al. (Schymanski et al., 2014a). Level 3: tentative candidates (compounds identified by molecular formula and substructures), based on accurate mass, MS/MS pattern, and database search. Level 4: unequivocal molecular formula based on accurate mass, but no other supporting data.

6.3.4 Time-course evolution of the identified metabolites.

We traced the detected metabolites through the course of the incubations according to their m/z and retention time (Figure D1). Due to the lack of standards, metabolites were quantified relative to their peak areas. Consistent with PAH removal (Figure D1), fluoranthene metabolites peaked at 7 d (except $C_{16}H_8O_2$, which peaked at 14 d), then decreased in abundance, clearly indicative of their subsequent metabolism. B[a]A metabolites, including two sulfate conjugates, appeared at 14 d and increased in abundance after 21 d incubation. Sulfate conjugates from pyrene also accumulated over the course of incubation, suggesting the potential persistence of such conjugates even in biologically active soil. The recalcitrance of analogous PAH conjugates accumulated by the fungus *Cunninghamella elegans* to soil microbial communities was previously demonstrated (Schmidt et al., 2010). Further studies are needed to understand their influence on the biotransformation process.

6.3.5 Environmental implications and further applications.

In this study, we demonstrated that SIAM can be successfully applied to the biotransformation of pollutants. This method allowed us to trace the biotransformation of individual pollutants in the context of a real contaminated sample. SIAM in combination with MS/MS enabled the detection of novel metabolites and preferential metabolic pathways that could result from interactions within complex microbial communities and contaminant mixtures. Despite the analytical limitations associated with the use of a single ionization source, our method allowed us to highlight major mechanisms actually implicated in PAH biodegradation during biostimulation of a contaminated soil. The SIAM flow could be applied to a wider range of studies in environmental sciences, such as the biotransformation of pesticides and other micropollutants in aquatic systems and in water treatment processes. In addition, the combination

of SIAM with compound specific isotope analysis (CSIA) may become a powerful tool for field-based bioremediation studies.

CHAPTER 7 CONCLUSIONS AND RECOMMENDATIONS

7.1 Conclusions

The overall goal of this work was to identify unknown transformation products, unregulated pollutants, and unrecognized metabolic pathways in bioremediated soil, by using multiple non-target analytical approaches. Conclusions for the specific aims of this dissertation are outlined below:

Aim 1: Identification of the unknown genotoxic compound(s) formed in bioremediation.

A non-target analytical approach combining effect-directed analysis (EDA) and metabolite profiling was developed, and applied to compare extracts of PAH-contaminated soil before and after treatment in the aerobic bioreactor. A compound with the composition $C_{15}H_8O_2$ and four methylated homologs were shown to accumulate as a result of bioreactor treatment, and the $C_{15}H_8O_2$ compound purified from soil extracts was determined to be genotoxic. Its structure was established by nuclear magnetic resonance and mass spectroscopy as a heretofore unidentified α,β -unsaturated lactone derived from dioxygenation of pyrene at an apical ring, 2H-naphtho[2,1,8-def]chromen-2-one (NCO), which was confirmed by synthesis. The concentration of NCO in the bioreactor was $11 \mu\text{g g}^{-1}$ dry soil, corresponding to 13% of the pyrene removed. It also accumulated in aerobically incubated soil from two additional PAH-contaminated sites and was formed from pyrene by two pyrene-degrading bacterial cultures known to be geographically widespread, underscoring its potential environmental significance.

Aim 2: Identify and quantify the major azaarenes in samples obtained from PAH-contaminated sites, and observe their susceptibility to biodegradation in microcosm incubations.

We applied high-resolution mass spectrometry and mass-defect filtering to four PAH-contaminated samples from geographically distant sites and detected 232 azaarene congeners distributed in eight homologous series, including alkylated derivatives and hitherto unknown series. Four- and five-ring azaarenes were detected among these series, and the most abundant non-alkylated congeners ($C_{13}H_9N$, $C_{15}H_9N$, $C_{17}H_{11}N$, $C_{19}H_{11}N$, and $C_{21}H_{13}N$) were quantified. Their biodegradability by autochthonous microbial communities in each of the four samples was evaluated in aerobic microcosms under biostimulated conditions. The removal of total quantifiable azaarenes ranged from 15-85%, and was related to the initial degree of weathering for each sample. While three-ring azaarenes were readily biodegradable, the five-ring congeners were the most recalcitrant. Isomer-selective biodegradation of four-ring azaarenes was observed, which might be attributed to the position of the nitrogen atom that also influences the physicochemical properties of azaarenes.

Aim 3: Trace the biotransformation of PAHs in contaminated soil by stable isotope-assisted metabolomics.

We applied stable isotope-assisted metabolomics (SIAM) to PAH-contaminated soil collected from a wood treatment facility. Soil samples were separately spiked with uniformly ^{13}C -labeled fluoranthene, pyrene, or benzo[*a*]anthracene at a level below that of the native contaminant, and incubated for 1 or 2 weeks under aerobic biostimulated conditions. Combining high-resolution mass spectrometry and automated SIAM workflows, chemical structures of metabolites and biodegradation pathways in the soil were proposed. Ring-cleavage products, including previously unreported intermediates such as $C_{11}H_{10}O_6$ and $C_{15}H_{12}O_5$, were detected originating from fluoranthene and benzo[*a*]anthracene, respectively. Sulfate conjugates of dihydroxy compounds were found as major metabolites of pyrene and benzo[*a*]anthracene, suggesting the potential role of fungi in their biotransformation in soils. A series of unknown N-containing metabolites were identified from pyrene incubation, but their structural elucidation

requires further investigation. The results suggest that SIAM can be successfully applied to understand the fate of organic pollutants in environmental samples, opening lines of evidence for novel mechanisms of microbial transformation within such complex matrices.

7.2 Recommendations for future work

Along with the path of current research, there are several potential lines of research that could be explored in the future.

First, further investigation of the environmental and toxicological behavior of 2*H*-naphtho[2,1,8-*def*]chromen-2-one (NCO) would be necessary. In Aim 1, we identified this toxic and genotoxic metabolite from bioremediated soil, and traced its sources to pyrene and certain bacterial strains. The formation pathway for NCO was proposed, but there are interesting questions to be answered about its formation mechanism. For instance, the production of NCO in pure strains and in microcosm incubations was much slower than in the bioreactor, suggesting that co-metabolism or kinetic factors might contribute to the formation of NCO. Alternatively, the bioreactor might be enriched in bacteria that inherently transform pyrene to NCO in high yield. Understanding such mechanisms may help avoid the formation of NCO in bioremediation. Meanwhile, we do not exactly know how NCO causes DNA damage, and further studies on toxicology would be warranted. In addition, it would be interesting to see if NCO really exists in the samples in other earlier studies, in which increased genotoxicity was found in bioremediated soils. It would also be important to check the existence of NCO in real sites that have been cleaned up by bioremediation. Finally, as PAHs are prevalent pollutants found in various environmental media, it is possible that NCO and its analogues might occur in other systems like water or air particulates.

Second, application of ultra-high resolution mass spectrometry (UHRMS) on contaminated and bioremediated soil would be promising. In Aim 2, we have observed the diversity of azaarenes in the coal tar- and creosote-contaminated environmental samples, and it is natural to consider that similar diversities could be observed for other compound classes, such as alkylated-PAHs, oxy-PAHs, and sulfur-containing PAHs. All these different compound classes contribute to the complex nature of coal tar and creosote contamination. In addition, biodegradation or bioremediation would have an impact on the overall chemical composition of contaminated soil, which was not investigated. Our samples collected in the bioremediation process would be good materials for characterizing the chemical compositions in contaminated soil, and tracking their changes during the course of bioremediation. UHRMS-based “petroleomics” (Marshall and Rodgers, 2008) analysis would help achieve such a bigger picture, and multiple ionization modes would be necessary for a comprehensive profiling.

Third, the unknown metabolites detected in SIAM experiments might shed light on the complex biodegradation behavior in real soil samples. Most interestingly, we observed multiple sulfate conjugates of dihydroxylated PAHs, as well as unknown nitrogen-containing conjugates from pyrene. The sulfate conjugates, known as the phase II metabolites from eukaryotes, might suggested the fungal activities in bioremediation. The interactions between bacteria and fungi in bioremediation would be an interesting topic for further exploration. Furthermore, as most of these conjugates were identified from the aqueous phase, the transport of pollutant metabolites among phases could be a direction that is worth investigating.

APPENDIX A: SUPPORTING INFORMATION FOR CHAPTER 3

Soil sample extraction

Replicates of feed and bioreactor-treated soil samples (3 g wet weight aliquots) were each dried with 10 g of sodium sulfate and transferred to 30-mL glass centrifuge tubes containing glass beads and capped with PTFE-lined septa; moisture content was determined in separate replicate aliquots. Extracts were obtained by two successive overnight extractions using 20 mL of a DCM: acetone mixture (1:1, v/v) with agitation on a wrist-action shaker. The collected extracts were filtered through a 0.2 μm pore-size nylon membrane. For each sample, the extracts from eight replicate aliquots (24 g wet weight of soil total) were combined in a final volume of 500 mL and stored at 4°C for chemical and biological analysis.

Fractionation

The extracts from feed and bioreactor-treated soil were fractionated using the two-level EDA method described in Figure A1. In the first level of fractionation, extracts were separated into broad compound classes by column chromatography using silica gel, according to the method of Bandowe et al (Bandowe et al., 2010). The silica gel was pre-cleaned with methanol and dichloromethane (DCM). A glass column (0.8 cm inner diameter) loaded with 5 g of 10% (w/w) deactivated silica gel was topped with 1g of sodium sulfate and conditioned with 10 mL n-hexane before adding samples. Glass wool extracted with DCM was placed at the bottom of the column to prevent the break-through of silica particles. Aliquots of extracts corresponding to 1.2 g dry weight of soil were concentrated to 1 mL under a gentle nitrogen flow, and transferred onto the column. Three fractions – nonpolar (A), semi-polar (B), and polar (C) – were collected by eluting the column with 15 mL n-hexane/DCM (5:1, v/v), 16 mL DCM/acetone (5:3, v/v) and 15 mL methanol, respectively. The fractions were used for bioassay testing or further fractionation.

The second level of fractionation was achieved with solid phase extraction (SPE) cartridges containing 1 g of cyanopropyl-bonded silica (SiliCycle, QC, Canada). Fractions from the first level were concentrated to 1 mL under a gentle nitrogen flow, and applied to the SPE cartridges. Three subfractions from the semi-polar fraction (B) of the first level were collected by eluting the cartridge with 10 mL hexane (B1), 10 mL DCM (B2), and 10 mL methanol (B3).

For each level of fractionation, replicates were obtained on the same day, and the stored fractions were used for chemical analysis, bioassay, or further fractionation. Method blanks were prepared in parallel with samples by adding pure solvents to the columns, and then eluting the columns in the same manner as for samples. No toxicity was observed for these blanks.

Instrumental analysis

For the selected B2 fractions, LCMS analysis was performed at high mass resolution on an Agilent 1100 series HPLC interfaced with a 6520 series quadrupole time-of-flight mass spectrometer (QToF-MS; Agilent Technologies, Palo Alto, CA, USA) equipped with an electrospray ionization (ESI) source operated in both positive and negative ion (ESI+ and ESI-) modes. The ToF mass selector was tuned to a resolution $> 15,000$ at m/z 322. During analysis, data were acquired in full scan mode (m/z 100-1000), and mass accuracy was continuously corrected by two references, purine (m/z 121.050873) and HP-921 (m/z 922.009798). HPLC separation was accomplished using a Poroshell 120 EC C18 column (3.0×100 mm, $2.7 \mu\text{m}$, Agilent). The mobile phase consisted of 0.1% formic acid in water (v/v, A) and methanol (B), and was delivered at a flow rate of $0.2 \text{ mL}\cdot\text{min}^{-1}$ applying the following gradient: 0-8 min, 10-50% B; 8-25 min, 50-90% B; 25-26 min 90% B; 26-33 min, 90-10% B.

LCMS analysis at unit mass resolution was performed on a Thermo TSQ Quantum Ultra triple quadrupole mass spectrometer (QqQ MS, Thermo Fisher Scientific) equipped with an atmospheric pressure ionization (APPI) source in both positive and negative ionization (APPI+

and APPI-) modes. Full-scan mass spectra were acquired over the range m/z 100-1000. A Waters Acquity BEH C18 column (2.1 mm \times 100 mm, 1.7 μ m, Waters, Milford, MA, USA) was used for HPLC separation. Flow rate and mobile phases were the same as those used with QToF, and the following gradient was applied: 0-3.6 min, 10-30 % B; 3.6-22.6 min, 30-90% B; 22.6-24 min 90% B; 24-30 min, 90-10% B.

Data analysis for metabolite profiling

Data files acquired by LCMS were processed by the metabolomics platform XCMS Online (<https://xcmsonline.scripps.edu/>). In brief, the centWave algorithm was used for peak detection, with a peak width range from 5 to 30 seconds. The mass error tolerance was set at 10 ppm. Final annotation and molecular formula assignment were achieved by XCMS Online and Agilent MassHunter. Annotation for features allowed the search for isotopic features and multiple adducts ($[M+H]^+$, $[M+Na]^+$, $[M+K]^+$, $[M+NH_4]^+$). The differences of features between groups (feed vs. bioreactor- treated soil) were evaluated by Welch t -test with a threshold p -value set at 0.01, and a minimum fold-change set at 1.5.

Purification of NCO

NCO formed during bioreactor treatment was purified by a three-level separation method consisting of sequential low-pressure silica gel-column chromatography, medium-pressure column chromatography, and HPLC.

For the first level of separation, glass chromatography columns were filled with 20 g of activated silica gel. Extracts corresponding to 15 g (dry weight) of treated soil were concentrated and added onto the column. The column was eluted with 60 mL hexane/DCM (1:1, v/v), 30 mL pure DCM, and 100 mL DCM/ethyl acetate (9:1, v/v). The latter, containing the target compound, was collected for the next stage of separation.

The second level of separation was by flash-chromatography (CombiFlash Rf, Teledyne Isco, NE, USA) on a pre-filled silica column using gradient elution (0-100% ethyl acetate/hexane

over 12 min). The target compound was eluted in the fractions containing 50-65% ethyl acetate, which were dried and re-dissolved in acetone for the next stage of separation.

The fractions reconstituted in acetone were separated on an HPLC system (Agilent 1100) equipped with a fraction collector using a C18 reverse phase column (Atlantis dC18, 10 × 100 mm, 5 μm, Waters) with a deionized water/methanol gradient program (50-90% methanol, 0-10 min). The purified target compound was collected in fractions composed of 75% methanol.

NMR spectrometry

NMR data were acquired at 25 °C in CDCl₃ on a Varian INOVA 400 at 400 MHz for proton spectra and 100 MHz for carbon spectra or on a Varian INOVA 500 NMR spectrometer at 500 MHz for proton spectra and 125 MHz for carbon spectra. Chemical shifts are reported relative to tetramethylsilane protons or carbons as appropriate. ¹H, COSY, HMBC and HSQC spectra were acquired. NMR data are given in Figures A5-A7.

Incubations of bioreactor-treated soil with pyrene

Microcosm incubations of the bioreactor-treated slurry in the presence of pyrene were prepared in 30-mL glass centrifuge tubes. Empty tubes were first autoclaved, and half were spiked with 200 μg of pyrene in acetone solution. After solvent evaporation under a gentle nitrogen flow, 2.5 mL of phosphate buffer and 5 mL bioreactor-treated slurry were added to all tubes. Samples were incubated on an orbital shaker at 200 rpm at room temperature. At time 0 and after 3 and 7 days of incubation, duplicate tubes spiked with pyrene and duplicate controls were centrifuged and the soil pellets were extracted by 30 minutes of agitation on a wrist-action shaker with a DCM: acetone mixture (1:1, v:v). The organic extracts were filtered (0.2 μm pore size) and made up to a volume of 50 mL. The concentrations of NCO were quantified by HPLC-MS, and pyrene concentrations were measured by HPLC with fluorescence detection. Results are shown in Figure A8.

Resting cell incubations of pure strains with pyrene

The potential formation of NCO by three bacterial strains was tested in resting cell incubations in the presence of pyrene. Pure cultures of *Rugosibacter aromaticivorans* Ca6^T and *Immundisolibacter cernigliae* strain TR3.2^T were first grown to turbidity in sRB2 medium amended with sodium pyruvate (2 g L⁻¹). *Mycobacterium vanbaalenii* PYR-1^T was grown to turbidity in MM2 medium containing sodium pyruvate. All cultures were incubated at 30 °C with shaking at 225 rpm. Cells were pelleted by centrifugation and washed three times in bioreactor buffer. The washed pellets were resuspended in 1.5 mL of bioreactor buffer and 500 µL of cell suspension were aliquoted in triplicate into 5-mL glass, screw-cap tubes containing 25 µg of pyrene and 500 µL bioreactor buffer. Uninoculated controls were prepared in triplicate in the same manner. All tubes were incubated at 30°C on an orbital shaker at 225 rpm for 24 hours in the dark. After incubation, the samples were subjected to three successive liquid-liquid extractions with 2 mL of DCM, and made up to 10 mL for quantification.

Incubations with PAH-contaminated soil from other sites

Microcosm incubations were performed on soil samples from two independent and geographically distant PAH-contaminated sites, which were not subject to bioremediation conditions before microcosm incubation. A sample of creosote-contaminated soil was obtained from a wood-treatment facility in Andalucía, Spain, that had a 100-year history of soil pollution (Tejeda-Agredano et al., 2013), although for this work a more highly contaminated sample was collected than that used in the cited study; the sample was collected with a clean shovel from the soil surface. The second sample was obtained from the Holcomb Creosote Superfund Site in Yadkinville, North Carolina, USA. A clean, sterile hand auger was used to obtain a core approximately 50 cm below the ground surface. The core was placed in a sterile glass dish and mixed with a sterile metal spoon before the soil was placed into sterile glass jars for transfer to

the laboratory. Initial concentrations of pyrene were 1,130 $\mu\text{g/g}$ for the soil from Spain and 26.9 $\mu\text{g/g}$ for the Holcomb sample. Each soil sample was air-dried, sieved, and homogenized. Triplicate 1-g aliquots were then transferred to 150-mL sterile Erlenmeyer flasks, topped with PTFE-lined screw caps, containing 10 mM phosphate buffer (pH 7.5) and urea as the nitrogen source (1 mM for the Spanish soil and 0.2 mM for the Holcomb soil). Concentrations of urea were adjusted to maintain a C:N ratio of 30:1 with respect to the total PAH concentration. Killed controls were prepared in the same manner, but with the addition of phosphoric acid (pH 2). All flasks were incubated with agitation at 150 rpm on an orbital shaker, at 25 °C for 42 days. The caps of flasks were loosened every 2 days to maintain an aerobic environment. After 42 days of incubation, the slurry was centrifuged and extracted, as described above for the bioreactor-treated slurry samples, to quantify the presence of NCO.

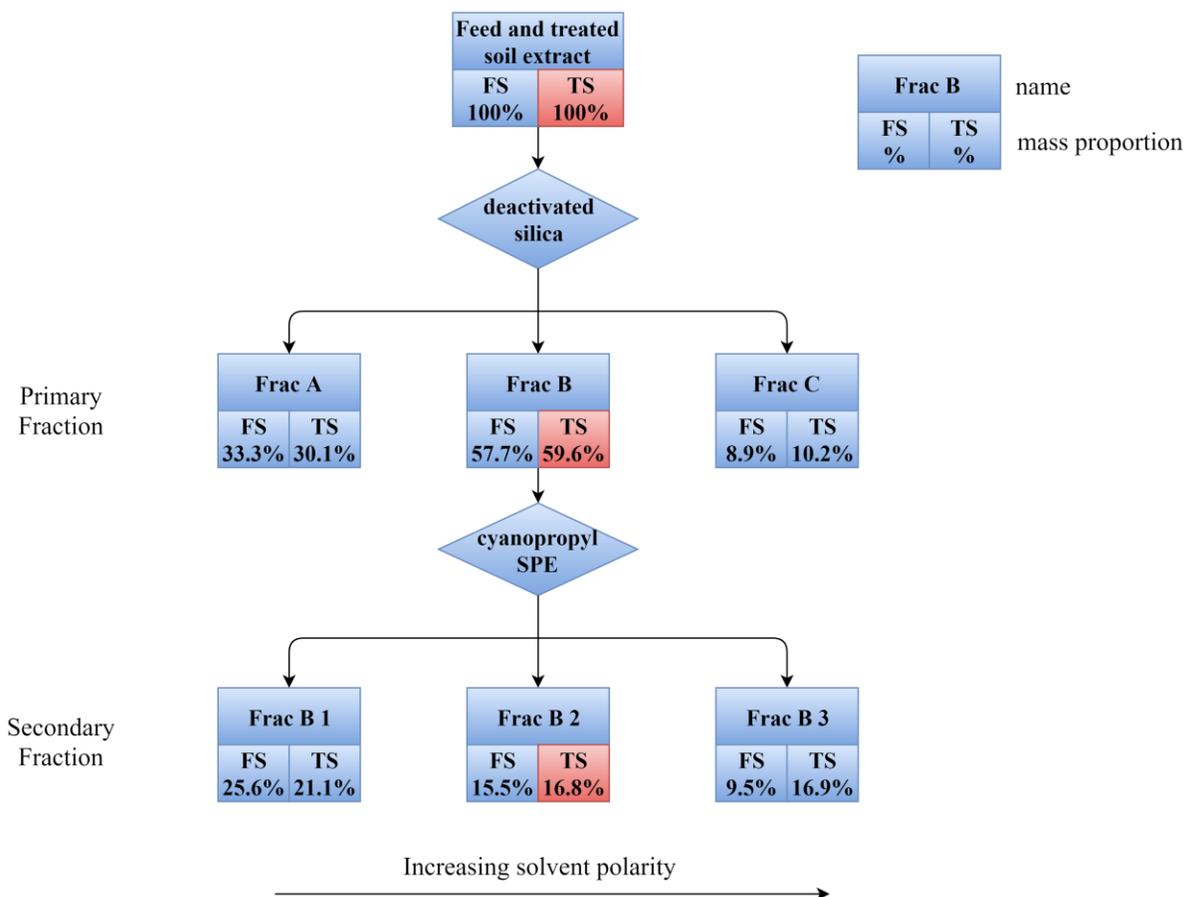
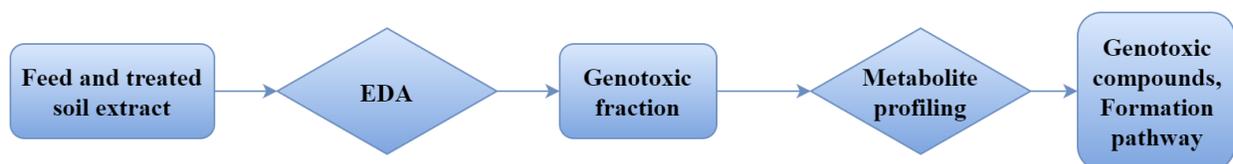


Figure A.1. Flow chart of non-target analysis approach (top) and the scheme for effect-directed analysis of fractions (bottom). The lower half of the boxes for each fraction represent the percent of total residue mass from the initial whole extract that was recovered in the fraction, for both feed soil (FS) and bioreactor-treated soil (TS). The fractions that were prioritized for further chemical analysis according to bioassay data are highlighted in red.

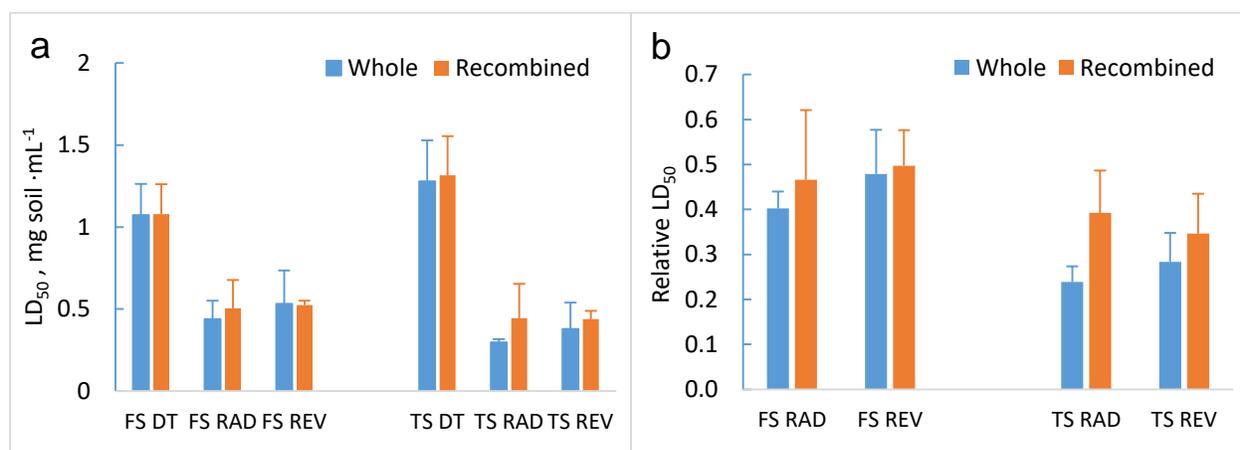


Figure A.2. (A) LD₅₀ and (B) relative LD₅₀ (mutant LD₅₀/parental LD₅₀) values for whole extracts and recombined extracts of feed soil (FS) and bioreactor-treated soil (TS) after the first level of fractionation (n=3); a relative LD₅₀ < 1.0 indicates genotoxicity; asterisks indicate a significant difference between feed soil and bioreactor-treated soil for a given cell line. DT, parental DT40 cells; RAD, *Rad54*^{-/-} mutant cell line; REV, *Rev1*^{-/-} mutant cell line.

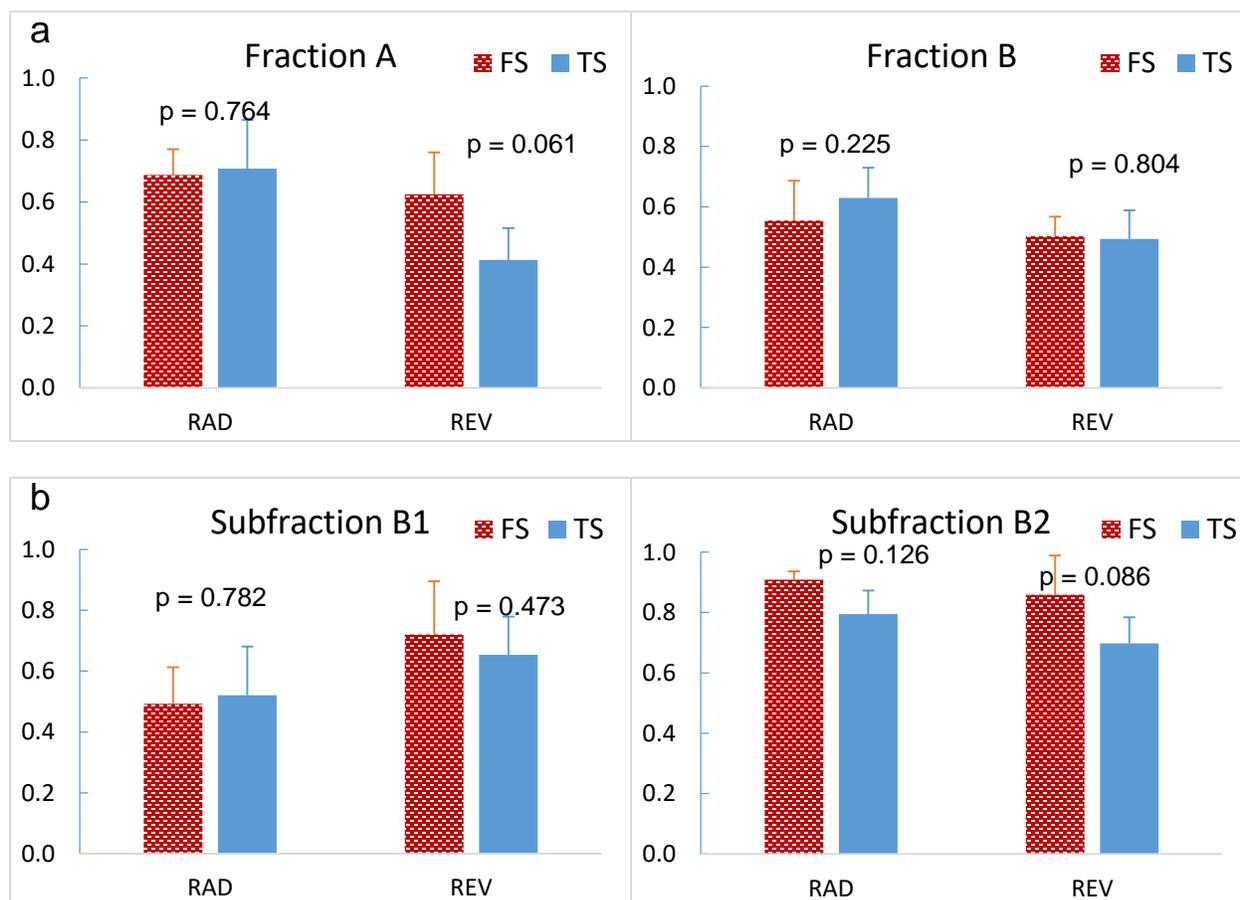


Figure A.3. Relative LD₅₀ values (LD₅₀ of mutant cell line/LD₅₀ of parental DT40 cell line) of soil extract fractions. (a) Fractions A and B from the first level of fractionation. (b) Subfractions B1 and B2 from the second level of fractionation of primary fraction B. FS, feed soil; TS, bioreactor-treated soil; RAD, *Rad54*^{-/-} mutant cell line; REV, *Rev1*^{-/-} mutant cell line. Error bars are standard deviations of replicates (n = 5), and the p-values from t-tests are indicated above each pair of bars.

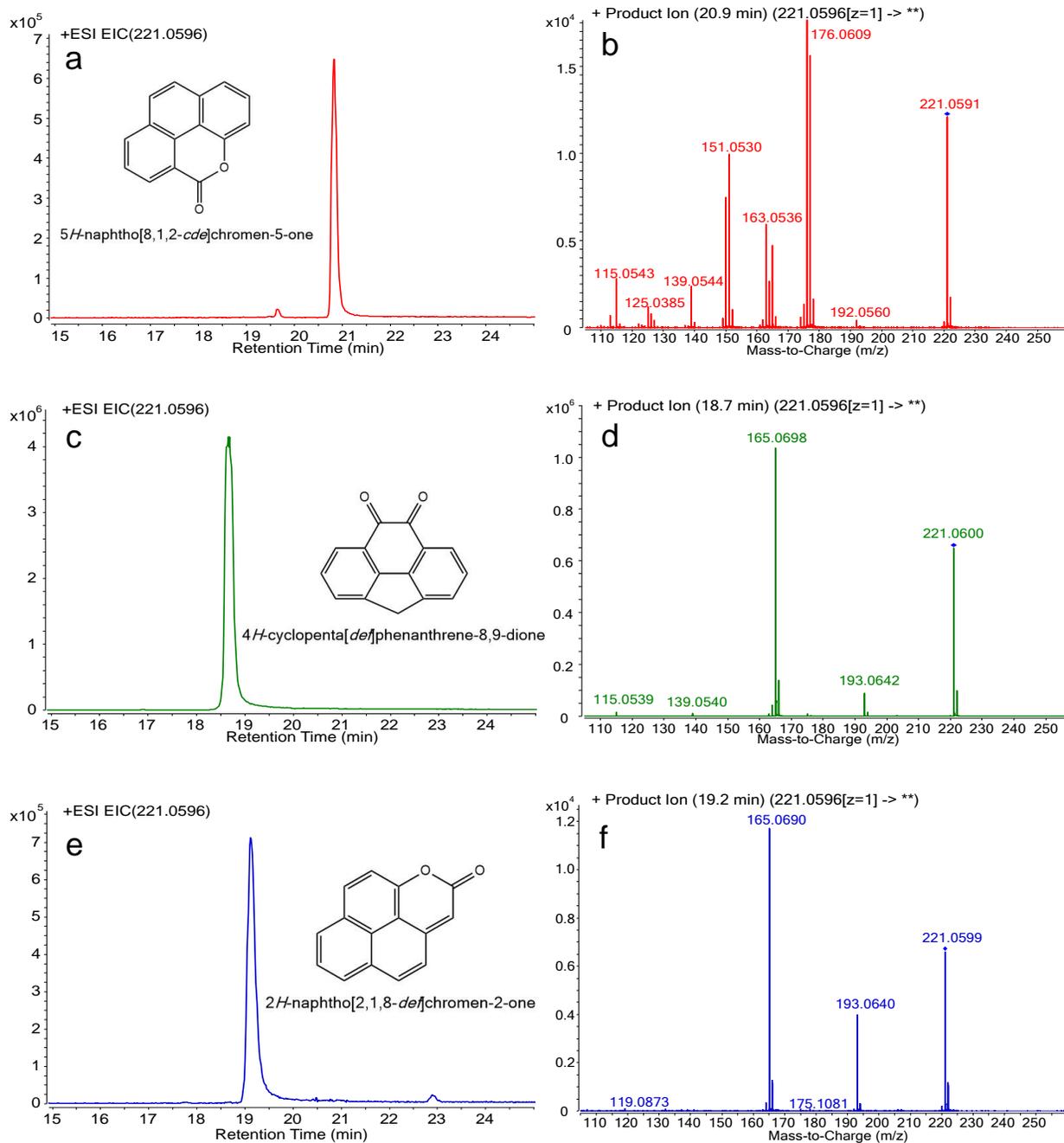


Figure A.4. Extracted ion chromatograms (*a*, *c*, *e*), MS/MS spectra (*b*, *d*, *f*), and structures of three compounds corresponding to the molecular formula $C_{15}H_8O_2$ synthesized in this study. (*a,b*) 5*H*-naphtho[8,1,2-*cde*]chromen-5-one (4-oxaypyrene-5-one). (*c,d*) 4*H*-cyclopenta[*def*]phenanthrene-8,9-dione. (*e,f*) 2*H*-naphtho[2,1,8-*def*]chromen-2-one (NCO).

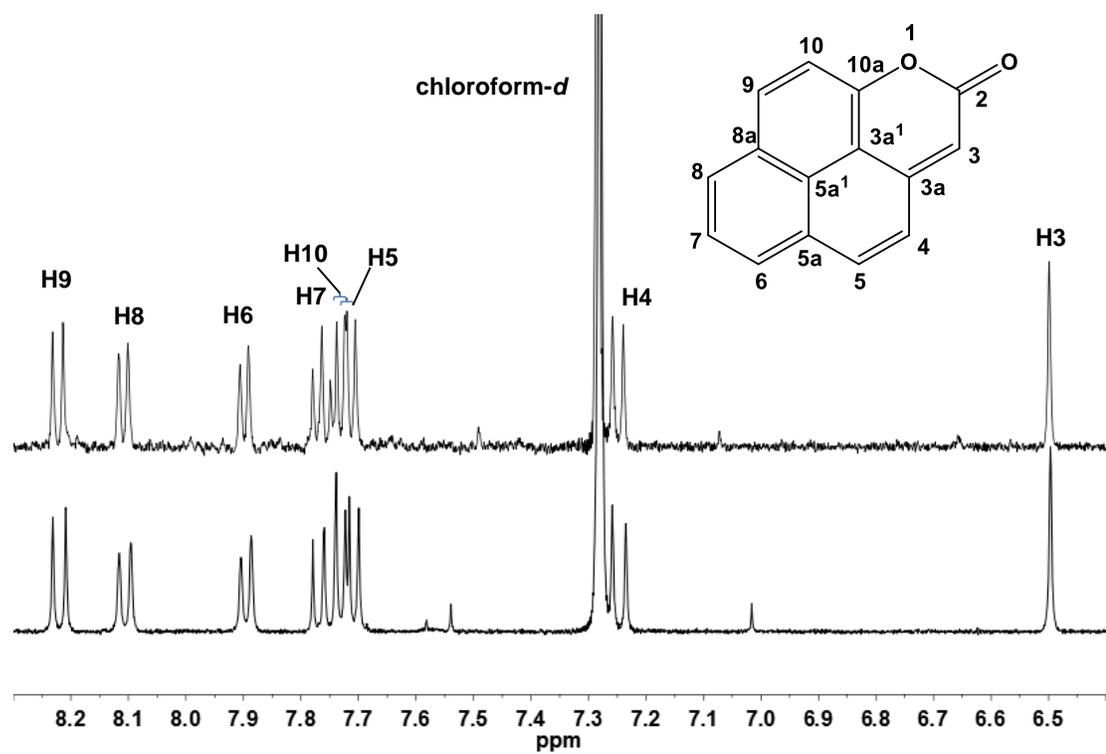


Figure A.5. Proton NMR spectra (400 MHz, chloroform-*d*) and numbering of NCO; top: purified metabolite; bottom: synthetic standard.

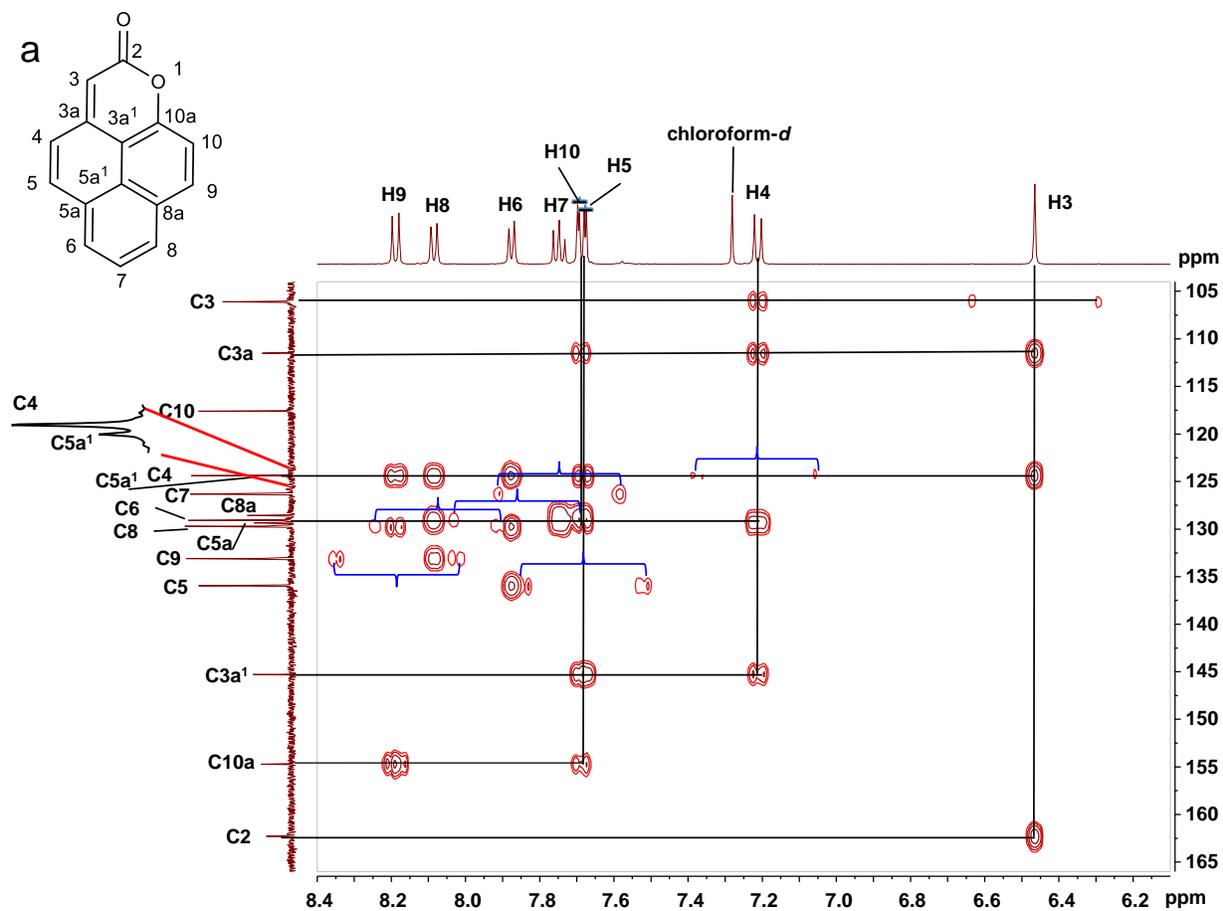


Figure A.6. Heteronuclear multiple bond coherence (HMBC) NMR spectrum for NCO, 500 MHz, chloroform-*d*. One-bond couplings are indicated by blue brackets.

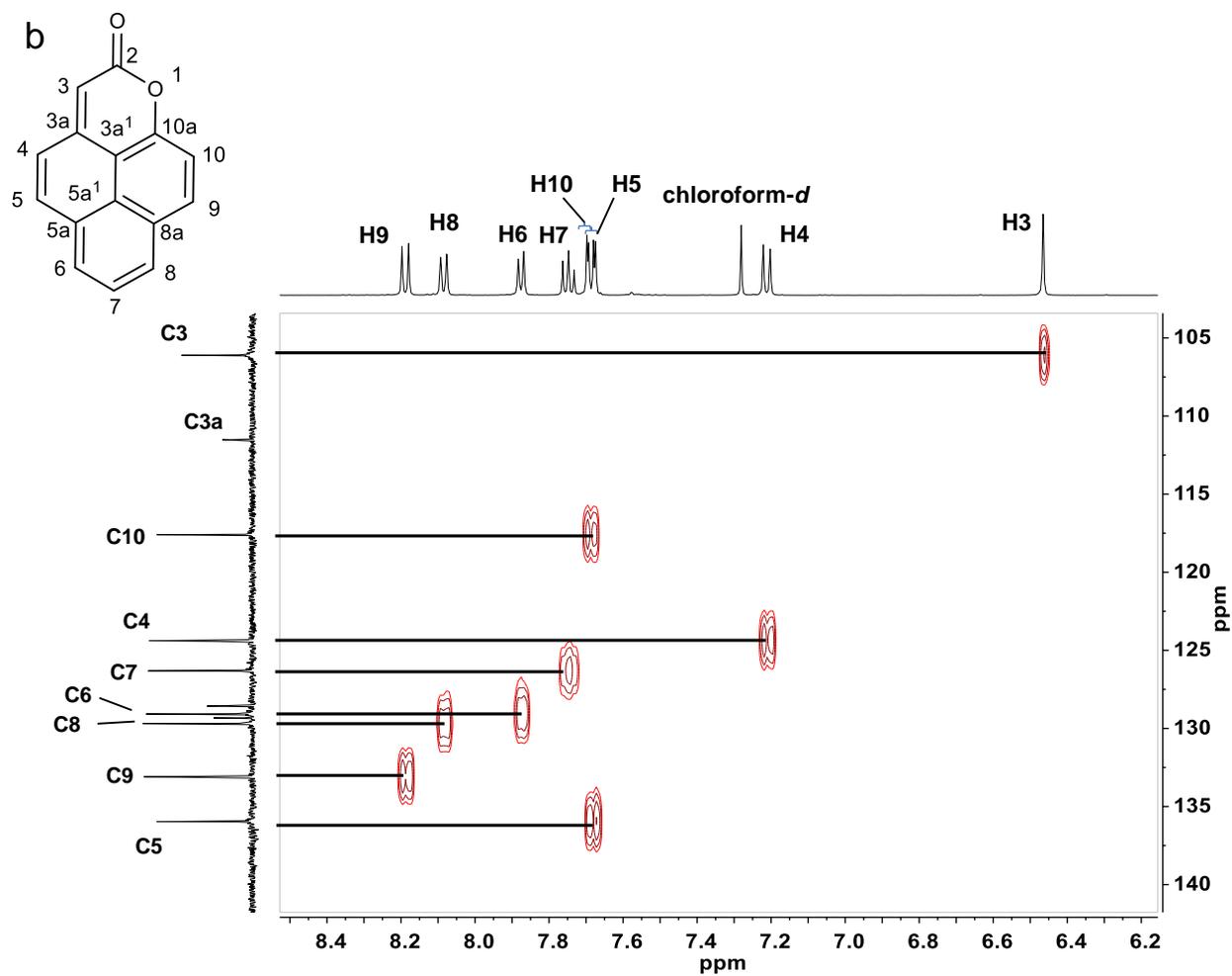


Figure A6b. Heteronuclear Single Quantum Coherence (HSQC) NMR spectrum for NCO, 500 MHz, chloroform-*d*.

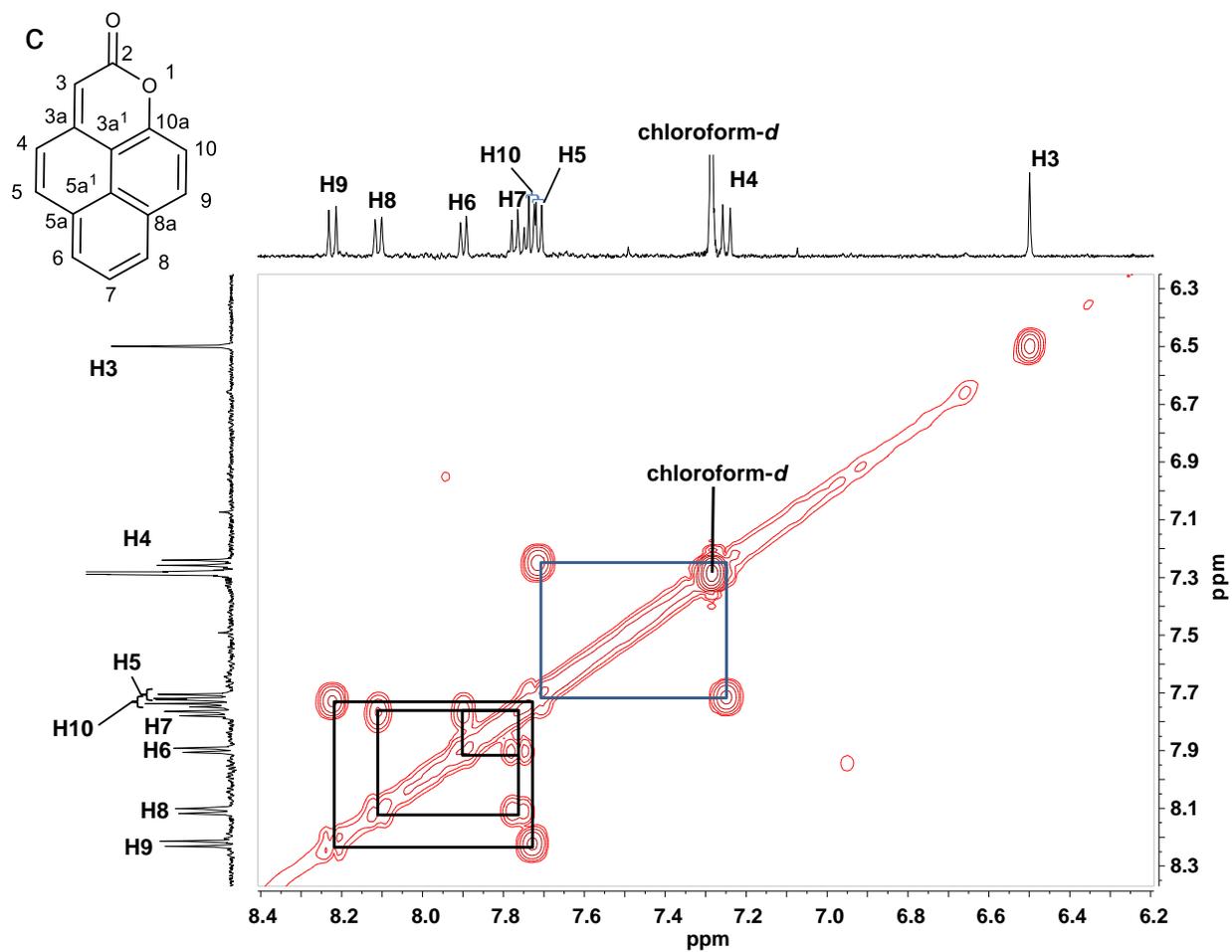


Figure A6c. Homonuclear correlation spectroscopy (COSY) NMR spectrum for NCO, 400 MHz, chloroform-*d*. Proton connectivities are indicated by boxes.

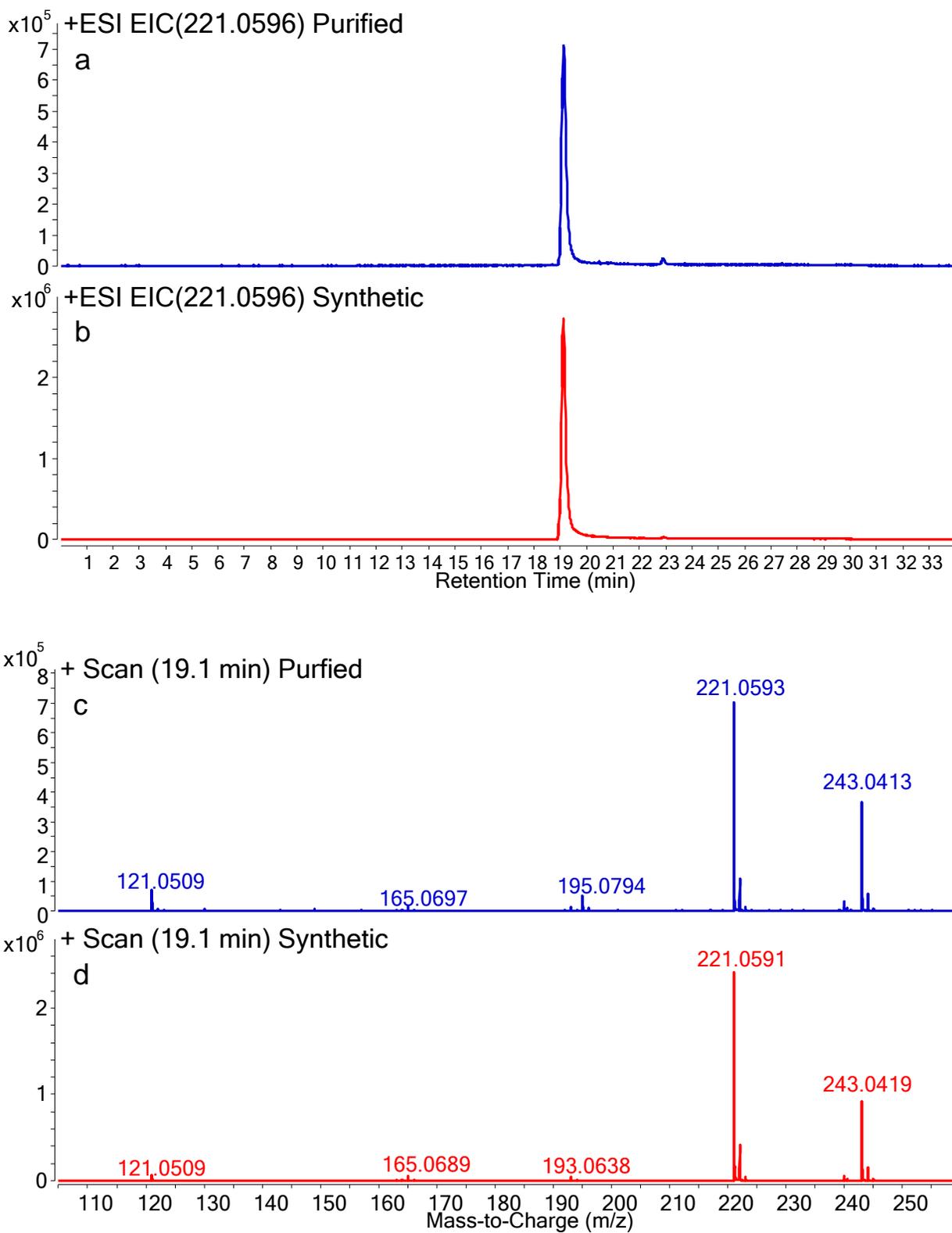


Figure A.7. HPLC chromatograms (*a,b*) and QToF mass spectra (*c,d*) of NCO purified from bioreactor-treated soil extracts (*a,c*; blue lines) and synthetic NCO (*b,d*; red lines).

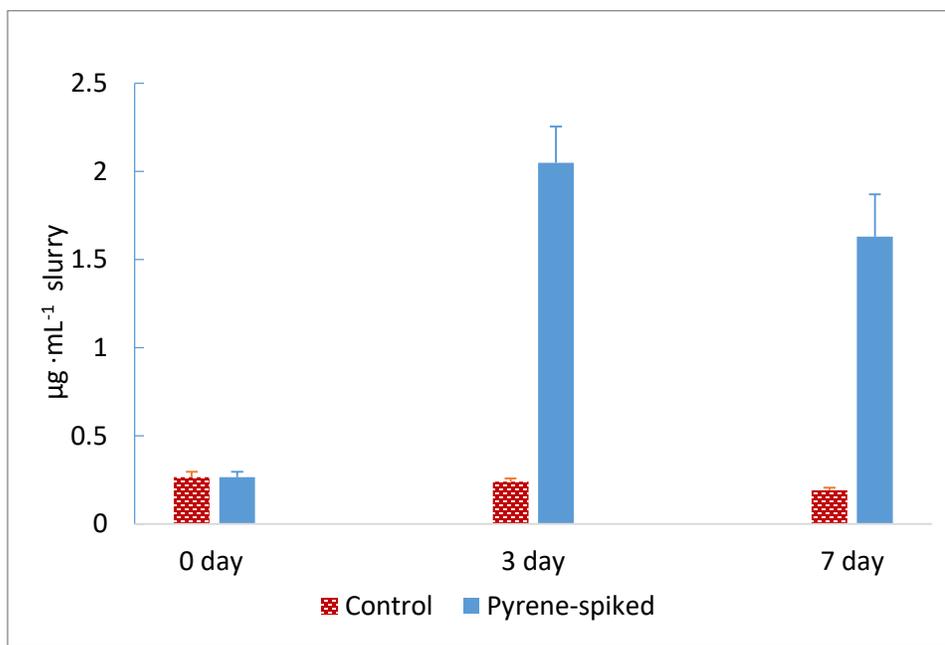


Figure A.8. Concentrations of NCO measured after microcosm incubations of bioreactor-treated soil spiked with additional pyrene. Error bars represent standard deviations from duplicate microcosms and asterisks indicate significant differences between conditions ($\alpha = 0.05$).

Table A.1. Concentrations ($\mu\text{g}\cdot\text{g}^{-1}$ dry soil) of 14 EPA-regulated PAHs in the soil used to feed the bioreactor and in bioreactor-treated soil.^a

	Feed soil	Treated Soil	LOQ
NAP	18.4 ± 1.8	19.5 ± 2.8	1.5
ACE	7.73 ± 1.34	0.85 ± 0.37	0.3
FLU	5.61 ± 1.44	1.40 ± 0.19	0.6
PHN	114 ± 15	18.7 ± 2.2	1.6
ANT	9.39 ± 1.94	2.29 ± 0.22	0.5
FLA	66.2 ± 5.7	13.7 ± 1.6	1.1
PYR	108 ± 8.5	24.2 ± 3.1	1.4
BaA	36.0 ± 3.3	12.7 ± 1.6	0.6
CHR	44.9 ± 3.5	13.1 ± 1.9	1.4
BbF	23.0 ± 3.2	11.9 ± 1.6	0.9
BkF	12.5 ± 1.8	6.27 ± 0.76	0.2
BaP	27.2 ± 4.4	14.7 ± 1.9	1.0
DBA	1.74 ± 0.70	1.32 ± 0.24	1.1
BgP	21.9 ± 1.8	24.3 ± 4.0	1.1

^a Data are means and standard deviations from triplicate aliquots of soil. LOQ: Limit of quantification. NAP, naphthalene; ACE, acenaphthene; FLU, fluorene; PHN, phenanthrene; ANT, anthracene; FLA, fluoranthene; PYR, pyrene; BaA, benz[*a*]anthracene; CHR, chrysene; BbF, benzo[*b*]fluoranthene; BkF, benzo[*k*]fluoranthene; BaP, benzo[*a*]pyrene; DbA, dibenz[*a,h*]anthracene; BgP, benzo[*g,h,i*]perylene.

Table A.2. Final list of upregulated ions from metabolite profiling with intensity $\geq 5,000$, fold-change (treated soil:feed soil) ≥ 2 , and $p \leq 0.01$.^a

fold change	p-value	<i>m/z</i>	RT	molecular formula
861	0.0002	221.0595	19.22	C ₁₅ H ₈ O ₂
363	0.0004	243.0416	19.22	C ₁₅ H ₈ O ₂
347	0.0058	235.0753	21.85	C ₁₆ H ₁₀ O ₂
273	0.0020	249.0912	23.54	C ₁₇ H ₁₂ O ₂
196	0.0002	235.0753	21.32	C ₁₆ H ₁₀ O ₂
66.4	0.0021	302.0968	29.03	C ₂₃ H ₁₁ N
58.0	0.0034	463.0934	19.22	C ₁₅ H ₈ O ₂
8.2	0.0021	255.0809	18.63	C ₁₉ H ₁₀ O
6.0	0.0046	297.24	29.61	C ₁₆ H ₃₄ O ₃
3.8	0.0009	248.1071	25.48	C ₁₇ H ₁₃ NO
3.2	0.0010	173.1078	10.69	C ₁₁ H ₁₂ N
3.0	0.0034	330.128	28.9	C ₂₅ H ₁₅ N
2.8	0.0012	231.0804	17.02	C ₁₇ H ₁₁ O
2.3	0.0008	268.112	28.49	C ₂₀ H ₁₃ N

^a RT, retention time. *m/z* and RT are median values from triplicates.

Table A.3. Features from metabolite profiling in APPI+ mode that were “upregulated” in subfraction B2 from extracts of bioreactor-treated soil compared to the same subfraction from feed soil.^a

fold change	p-value	<i>m/z</i>	RT	molecular formula
12.3	0.0119	221.06	15.21	C ₁₅ H ₈ O ₂
9.1	0.0048	222.06	15.21	C ₁₅ H ₈ O ₂
3.1	0.0205	235.09	16.77	C ₁₆ H ₁₀ O ₂

^a RT, retention time. *m/z* and RT are median values from triplicates.

APPENDIX B: SUPPORTING INFORMATION FOR CHAPTER 4

Information on sampling sites.

The FS soil was collected from a former manufactured-gas plant (MGP) site in Salisbury, NC, USA in the vicinity of the former tar well at a depth of 1.2 m below ground surface (Richardson, 2010). The soil was processed by mixing and screening through 10 mm wire mesh, and then stored at 4 °C before further screening at the time of use to remove small stones. Processed soil contained 83% sand, 14% silt, and 3% clay, with total organic matter of 8.3% and extractable organic matter of 0.64%. PAH concentrations of the processed soil are in Table S1.

Creosote-contaminated HS soil was collected from a former wood-treating facility site in Yadkinville, NC, USA, which is now a Superfund site. The sample was obtained using a clean, sterile hand auger to obtain a core approximately 50 cm below the ground surface. The core was placed in a sterile glass dish and mixed with a sterile metal spoon before the soil was placed into sterile glass jars for transfer to the laboratory. More detailed information about the site can be found at: <https://cumulis.epa.gov/supercpad/cursites/csinfo.cfm?id=0402724>

The creosote-contaminated KM sediment sample was collected from a former wood-treating facility site in Navassa, NC, USA, also now a Superfund site. The sample was collected from a marsh area impacted by contaminated wastewater, and was contained in sterile glass jars for transfer to the laboratory. More detailed information about the site can be found at: <https://cumulis.epa.gov/supercpad/cursites/csinfo.cfm?id=0403028>

The SC soil sample was collected from a former wood-treatment facility in Andalucía, Spain, that had a 100-year history of creosote pollution (Tejeda-Agredano et al., 2013). The sample was collected with a clean shovel from the soil surface.

Instrumental analysis.

Nontarget analysis was performed using an Agilent 1200 series high-performance liquid chromatography (HPLC) system interfaced with a 6520 accurate mass quadrupole time-of-flight mass spectrometer (qTOF MS, Agilent Technologies, Santa Clara, CA). Electrospray ionization in positive mode (ESI+) was adopted, and the TOF mass selector was tuned in high resolution mode (4 GHz, resolution > 15,000 at m/z 322). The TOF detector was set to acquire spectra over the range m/z 100-1000, and the mass accuracy was continuously corrected in real time using two references, purine (m/z 121.050873) and HP-921 (m/z 922.009798). HPLC separation was accomplished using a Poroshell 120 EC C18 column (3.0 × 100 mm, 2.7 μm, Agilent). The mobile phase consisted of 0.1% (v/v) formic acid in water (A) and methanol (B), and was delivered at a flow rate of 0.2 mL·min⁻¹ applying the following gradient: 0-8 min, 10-50 % B; 8-25 min, 50-90% B; 25-26 min 90% B; 26-33 min, 90-10% B. The identities of azaarene isomers were confirmed using product ion scan mode at collision energy of 30 eV to obtain fragmentation spectra for structure elucidation.

Quantitative targeted analysis of azaarenes was performed using a TSQ Quantum Ultra triple quadrupole mass spectrometer (QqQ MS, Thermo Fisher Scientific, Waltham, MA). Samples were analyzed in selected reaction monitoring (SRM) mode focusing on 6 sets of m/z transitions, corresponding to the 5 non-alkylated azaarene formulas and that of the recovery standard acridine-d₉ (Table B2). The MS/MS fragmentation pattern was selected according to previous studies (Lintelmann et al., 2010; Švábenský et al., 2009) and the collision energies were optimized by infusion of standard compounds. In order to make results comparable in target and nontarget analyses, the same HPLC column and parameters were used for chromatographic separation. Quantitative data were processed with Xcalibur software (Thermo Scientific), and concentrations were calculated according to external standard curves. The recoveries calculated

from acridine-d9 ranged from 53% to 102%. Data on limits of quantification are provided in Table B2.

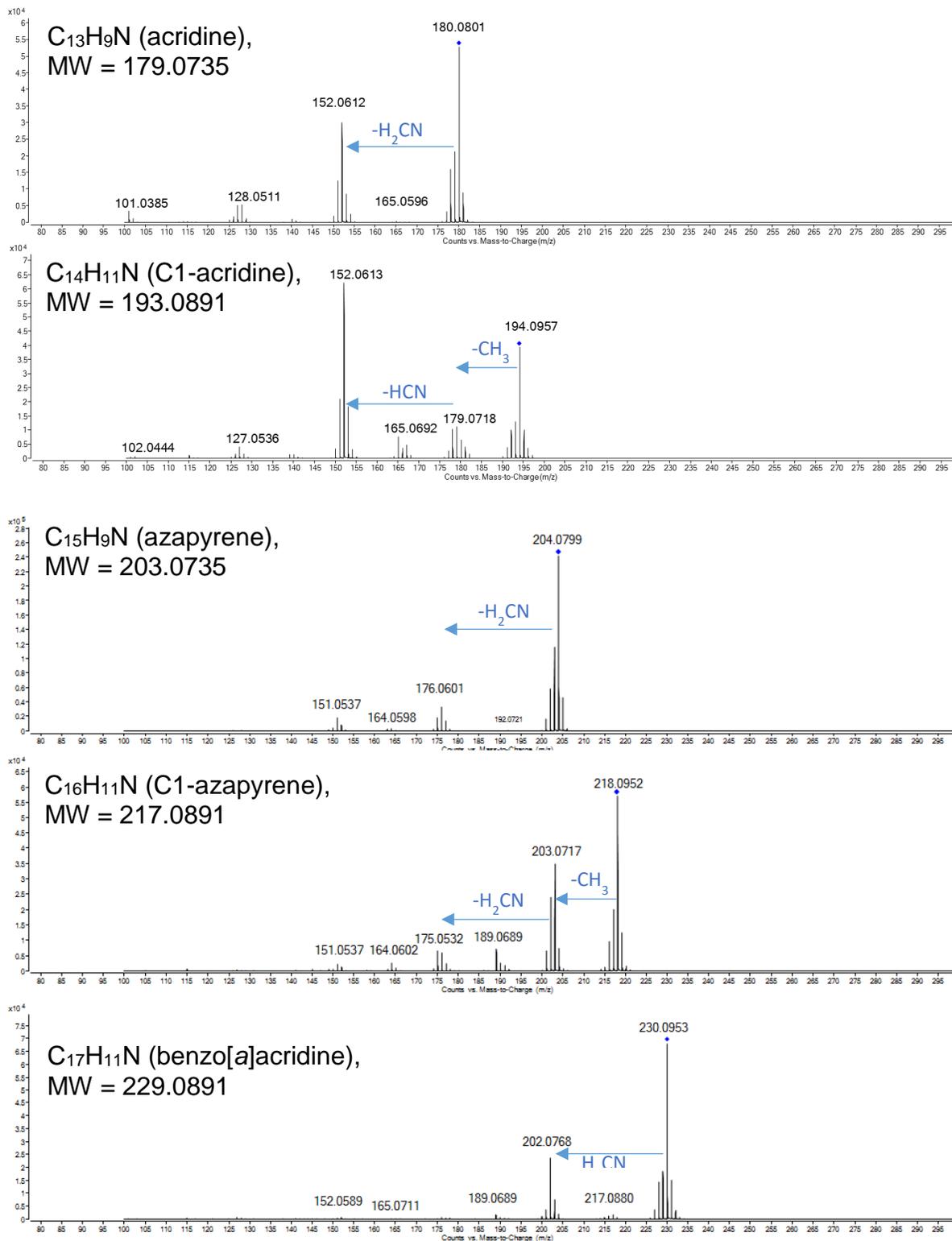


Figure B.1. Mass spectra of azarene isomers analyzed by qTOF in product ion scan mode.

Table B.1. Initial PAH concentrations ($\mu\text{g/g}$ dry soil) in the four samples (mean \pm standard deviation; $n=3$)

PAH	FS	HS	KM	SC
NAP	17.6 \pm 3.37	ND	ND	ND
ACE	7.18 \pm 0.78	ND	0.00 \pm	425 \pm 22.9
FLU	3.56 \pm 0.65	ND	0.00 \pm	479 \pm 24.6
PHN	68.1 \pm 6.14	4.45 \pm 0.30	25.5 \pm 2.47	2,680 \pm 108
ANT	6.58 \pm 0.78	1.71 \pm 0.68	31.6 \pm 4.44	963 \pm 122
FLA	53.5 \pm 0.86	17.6 \pm 3.63	1004 \pm 15.6	1,390 \pm 25.5
PYR	87.0 \pm 1.44	26.9 \pm 4.48	438 \pm 26.1	865 \pm 37.2
B[a]A	33.9 \pm 7.95	20.7 \pm 1.80	195 \pm 13.4	250 \pm 6.87
CHR	70.2 \pm 9.07	20.9 \pm 1.90	188 \pm 13.2	178 \pm 4.45
B[b]F	19.9 \pm 0.42	25.1 \pm 1.58	221 \pm 2.60	97.5 \pm 22.0
B[k]F	10.2 \pm 0.19	12.2 \pm 0.68	121 \pm 1.55	43.8 \pm 1.69
B[a]P	20.0 \pm 0.45	10.8 \pm 1.26	179 \pm 5.53	57.6 \pm 7.23
DBA	1.83 \pm 0.23	0.93 \pm 0.00	9.16 \pm 1.00	ND
B[g]P	21.7 \pm 0.43	6.13 \pm 0.59	80.0 \pm 4.85	ND
Sum	421 \pm 27.8	147 \pm 20.0	2,490 \pm 65	7,430 \pm 315

NAP, naphthalene; ACE, acenaphthene; FLU, fluorene; PHN, phenanthrene; ANT, anthracene; FLA, fluoranthene; PYR, pyrene; B[a]A, benz[a]anthracene; CHR, chrysene; B[b]F, benzo[b]fluoranthene; B[k]F, benzo[k]fluoranthene; B[a]P, benzo[a]pyrene; DBA, dibenz[a,h]anthracene; B[g]P, benzo[g,h,i]perylene; ND, not detected.

Table B.2. HPLC-MS/MS analytical parameters for quantification.

Standard	formula	mass transition	CE (eV)	LOQ	Method
				Instru ment (µg/L)	(µg/g dry soil)
acridine-d9	C13D9N	189 → 161	25	0.125	0.062
acridine	C13H9N	180 → 152	25	0.119	0.059
phenanthridine	C13H9N	180 → 152	30	0.095	0.047
B[<i>h</i>]Q	C13H9N	180 → 152	30	0.511	0.250
B[<i>a</i>]Ac	C17H11N	230 → 202	51	0.704	0.352
B[<i>c</i>]Ac	C17H11N	230 → 202	51	0.327	0.163
DB[<i>a,j</i>]Ac	C21H13N	280 → 252	61	0.223	0.112

CE, collision energy; LOQ, limit of quantification; B[*h*]Q, benzo[*h*]quinoline; B[*a*]Ac, benz[*a*]acridine; B[*c*]Ac, benz[*c*]acridine; DB[*a,j*]Ac, dibenz[*a,j*]acridine.

Table B.3. Concentrations of non-alkylated azaarene congeners (mean \pm standard deviation; n=3). ND: not detected.

ID	FS	HS	KM	SC
A1 Acridine	0.913 \pm 0.008	1.04 \pm 0.054	3.20 \pm 0.137	281 \pm 12.6
A2	0.334 \pm 0.006	1.06 \pm 0.059	ND	7.99 \pm 0.145
A3	0.398 \pm 0.004	0.480 \pm 0.037	0.060 \pm 0.023	10.3 \pm 1.14
A4 Phenanthridine	0.266 \pm 0.004	0.509 \pm 0.002	ND	0.220 \pm 0.107
A5 Benzo[<i>h</i>]quinoline	0.357 \pm 0.005	0.537 \pm 0.006	0.561 \pm 0.045	20.0 \pm 1.777
B1 Benzo[<i>a</i>]acridine	1.04 \pm 0.012	2.79 \pm 0.115	12.3 \pm 0.676	13.1 \pm 0.320
B2	0.172 \pm 0.004	0.708 \pm 0.063	0.917 \pm 0.101	4.62 \pm 0.220
B3	0.360 \pm 0.013	0.370 \pm 0.027	1.11 \pm 0.081	7.21 \pm 0.434
B4	0.234 \pm 0.019	0.759 \pm 0.026	ND	4.76 \pm 0.228
B5	0.287 \pm 0.009	0.391 \pm 0.009	0.416 \pm 0.024	7.74 \pm 0.477
B6 Benzo[<i>c</i>]acridine	0.903 \pm 0.020	0.996 \pm 0.054	4.23 \pm 0.125	33.9 \pm 1.683
B7	0.122 \pm 0.002	0.136 \pm 0.002	1.53 \pm 0.032	0.902 \pm 0.068
E1	0.554 \pm 0.007	6.53 \pm 0.916	6.36 \pm 0.231	31.1 \pm 0.802
E2	0.078 \pm 0.005	0.640 \pm 0.042	0.343 \pm 0.034	3.89 \pm 0.198
E3	0.058 \pm 0.003	0.191 \pm 0.009	ND	1.43 \pm 0.093
E4 Azapyrene	0.403 \pm 0.021	1.78 \pm 0.114	2.33 \pm 0.146	11.9 \pm 0.878
E5 Azafluoranthene	0.133 \pm 0.003	0.594 \pm 0.044	2.50 \pm 0.135	14.1 \pm 0.622
C1 Dibenzo[<i>a,j</i>]acridine	0.264 \pm 0.007	0.368 \pm 0.018	2.46 \pm 0.047	1.87 \pm 0.073
C2	0.211 \pm 0.002	0.290 \pm 0.006	2.39 \pm 0.134	0.898 \pm 0.066
C3,4	0.362 \pm 0.004	0.652 \pm 0.018	11.6 \pm 0.948	4.60 \pm 0.454
C5	0.301 \pm 0.007	0.664 \pm 0.008	0.403 \pm 0.011	0.217 \pm 0.013
F1	0.172 \pm 0.006	0.373 \pm 0.026	2.20 \pm 0.086	2.13 \pm 0.098
F2	0.104 \pm 0.000	0.260 \pm 0.005	0.831 \pm 0.038	1.11 \pm 0.021
F3	0.146 \pm 0.004	0.172 \pm 0.002	0.398 \pm 0.014	0.740 \pm 0.036
F4	0.085 \pm 0.003	0.179 \pm 0.002	0.706 \pm 0.028	1.03 \pm 0.035

Table B.3., continued.

ID	FS	HS	KM	SC
F5	0.122 ± 0.011	0.399 ± 0.010	0.174 ± 0.008	0.665 ± 0.026
F6	0.173 ± 0.009	0.354 ± 0.012	3.83 ± 0.194	2.17 ± 0.131
F7	0.165 ± 0.006	0.354 ± 0.007	1.55 ± 0.040	1.17 ± 0.046
F8	0.391 ± 0.010	0.892 ± 0.018	4.58 ± 0.200	3.63 ± 0.089
F9	0.094 ± 0.003	0.263 ± 0.003	ND	0.762 ± 0.021
F10	0.260 ± 0.011	1.07 ± 0.028	5.26 ± 0.174	3.17 ± 0.150
F11	0.237 ± 0.006	0.450 ± 0.027	0.326 ± 0.014	1.07 ± 0.098
F12	0.148 ± 0.003	0.348 ± 0.006	0.752 ± 0.046	0.897 ± 0.033
F13	0.152 ± 0.005	0.279 ± 0.010	0.787 ± 0.014	0.901 ± 0.039

APPENDIX C: SUPPORTING INFORMATION FOR CHAPTER 5

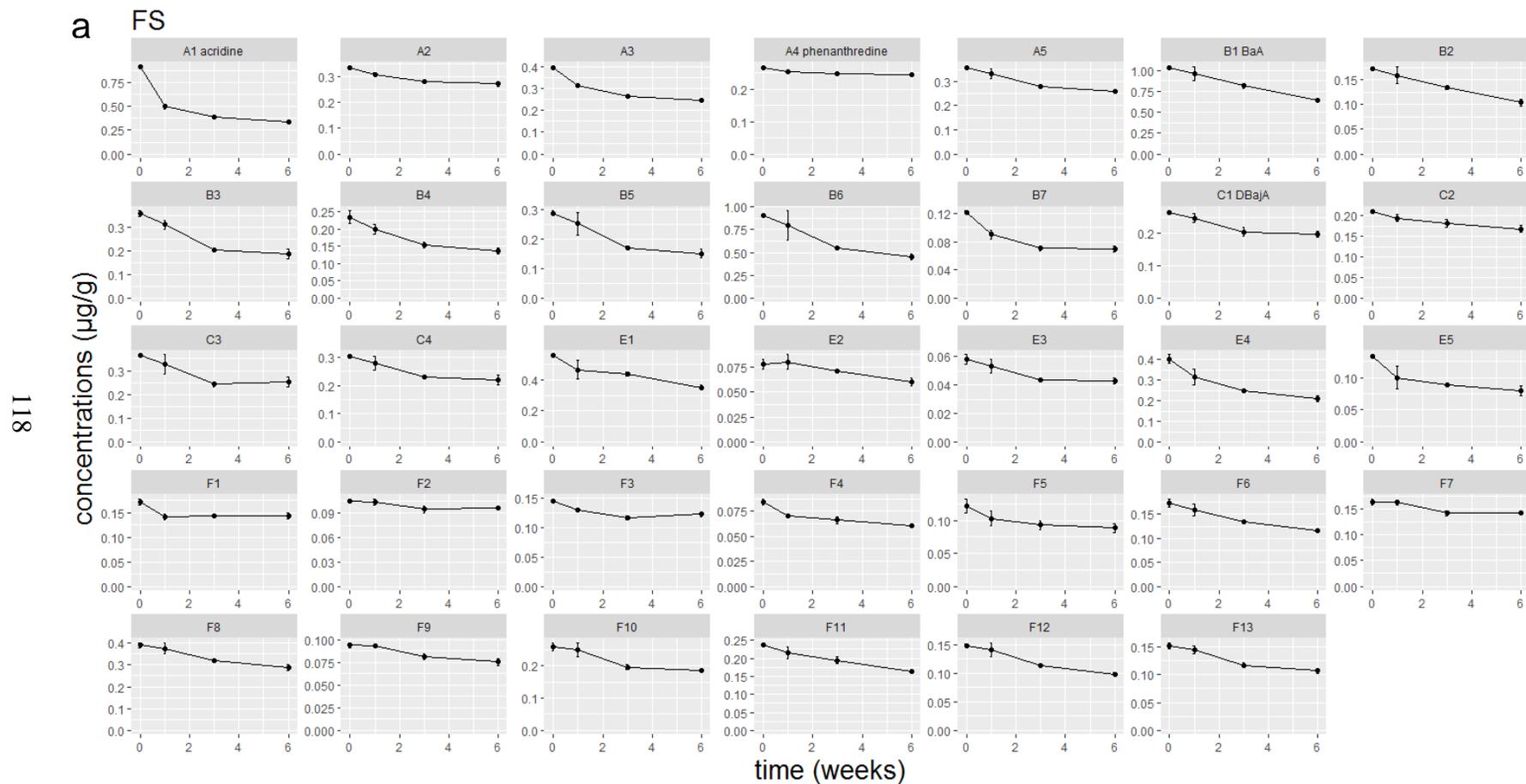


Figure C.1. Concentration trend lines of quantifiable azaarene congeners throughout the incubation. The data points and error bars represent the mean and standard deviation of triplicates. **(a)** FS, **(b)** HS, **(c)** KM, **(d)** SC.

b HS

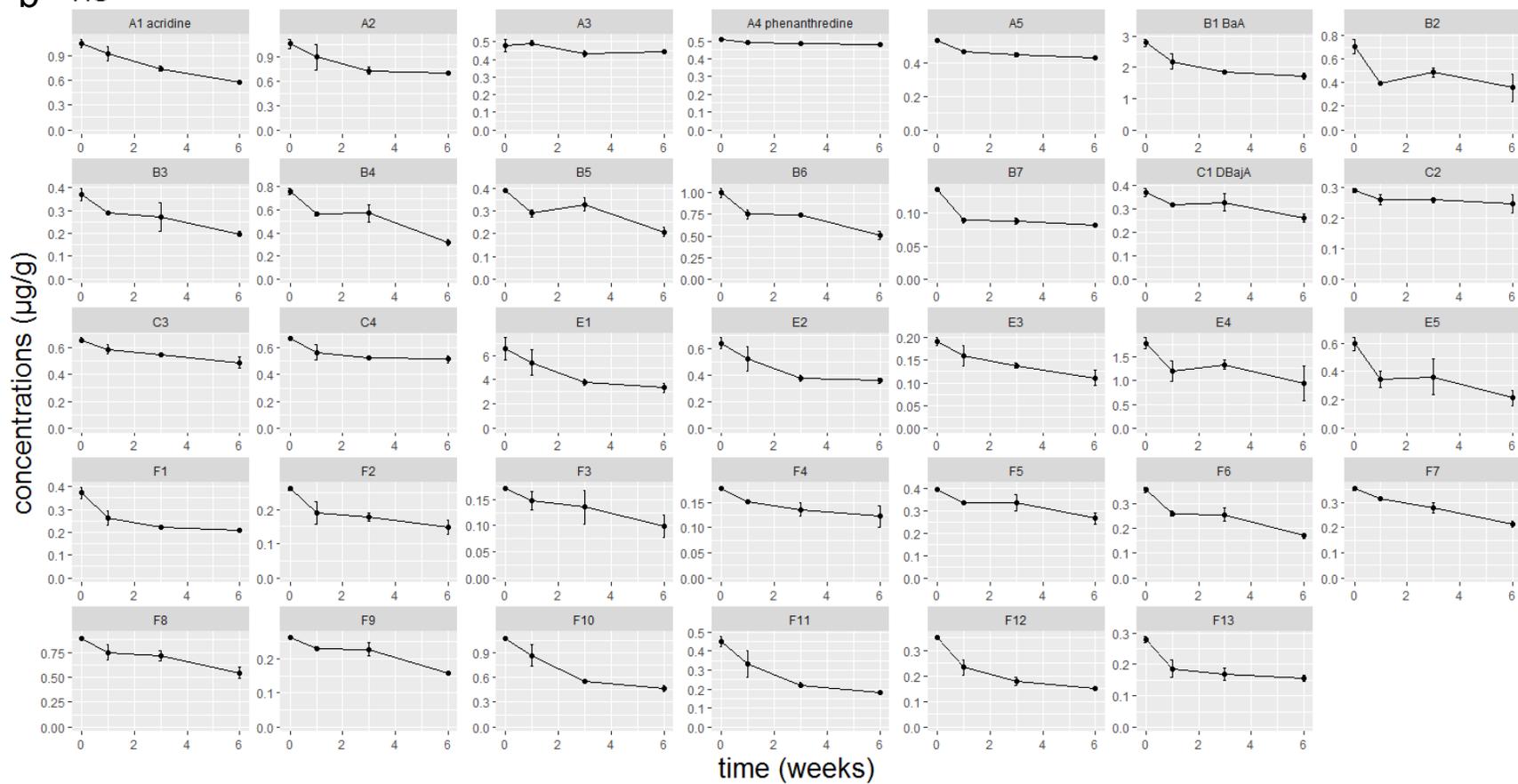


Figure C1, Continued.

C KM

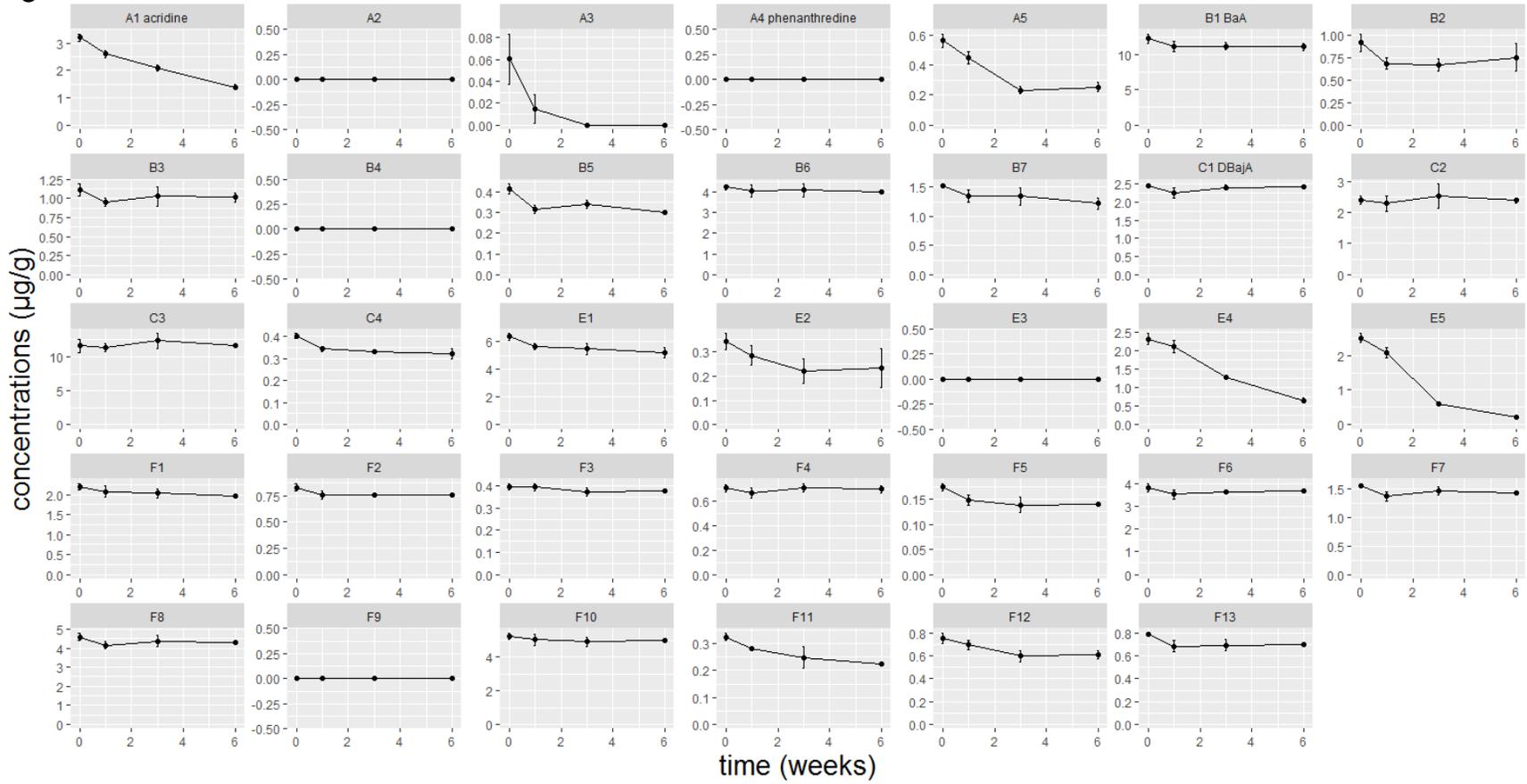


Figure C1, Continued.

d SC

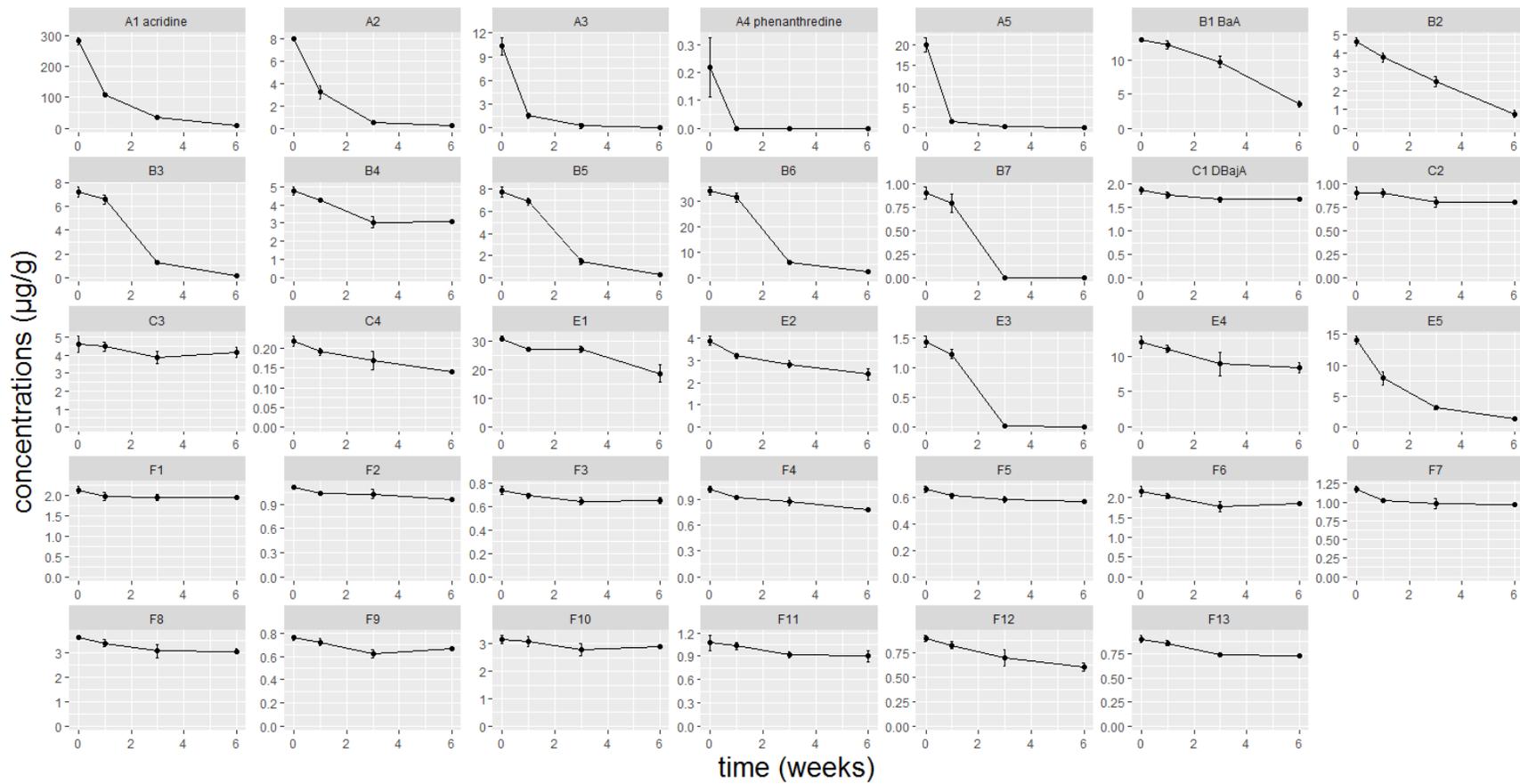


Figure C1, Continued.

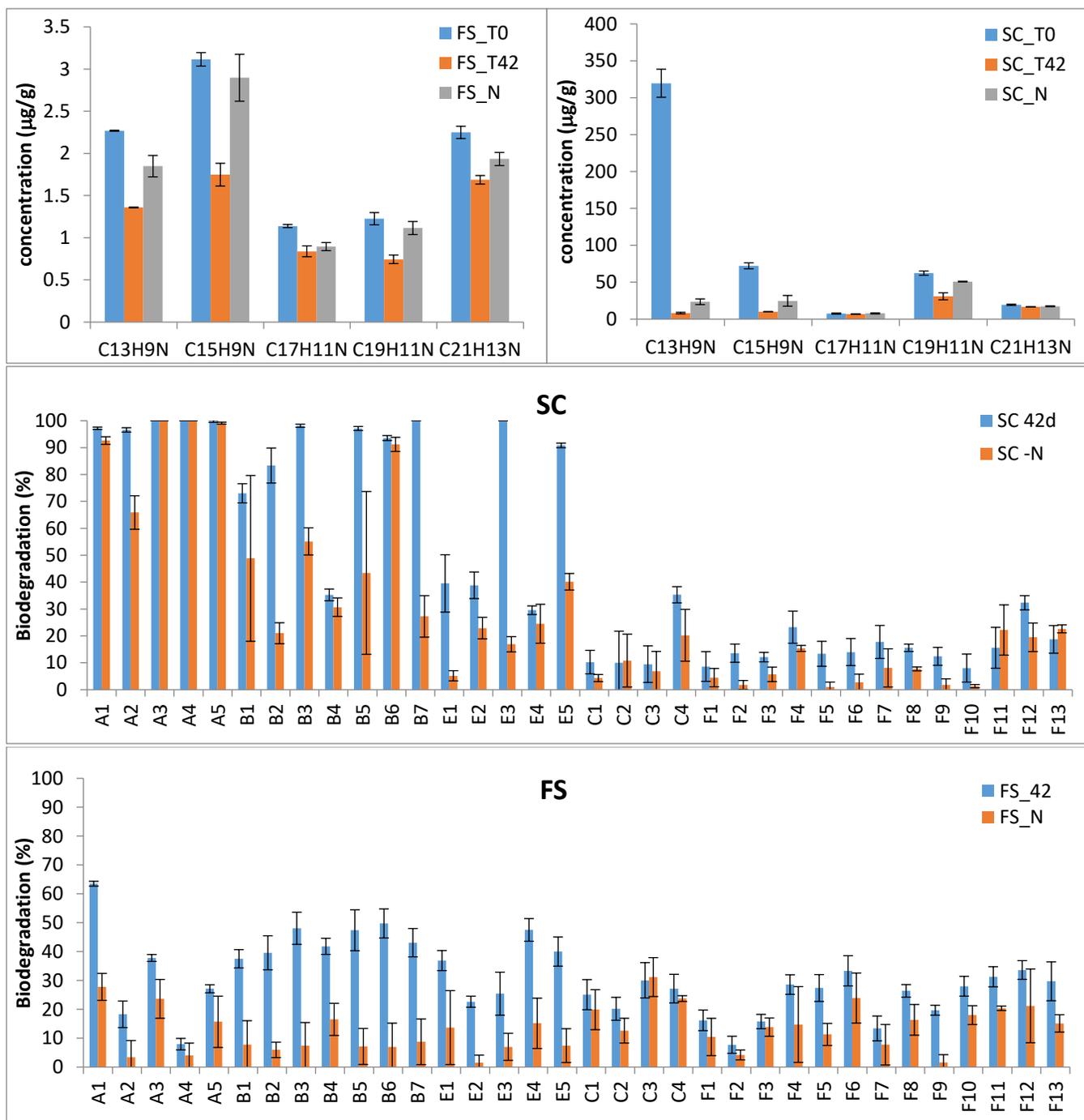


Figure C.2. Top, bar plots showing the sum of residual concentrations in each of the five azaarene congener groups quantified at the beginning of the incubation (time 0, blue), and after 42 days of incubation in the presence (orange) or absence (grey) of urea in microcosms of FS and SC soils. Middle and bottom, bar plots showing the removal of individual quantified azaarene isomers in the presence (blue) or absence (orange) of urea in SC and FS soil microcosms. Each bar represents the mean and standard deviation of triplicates.

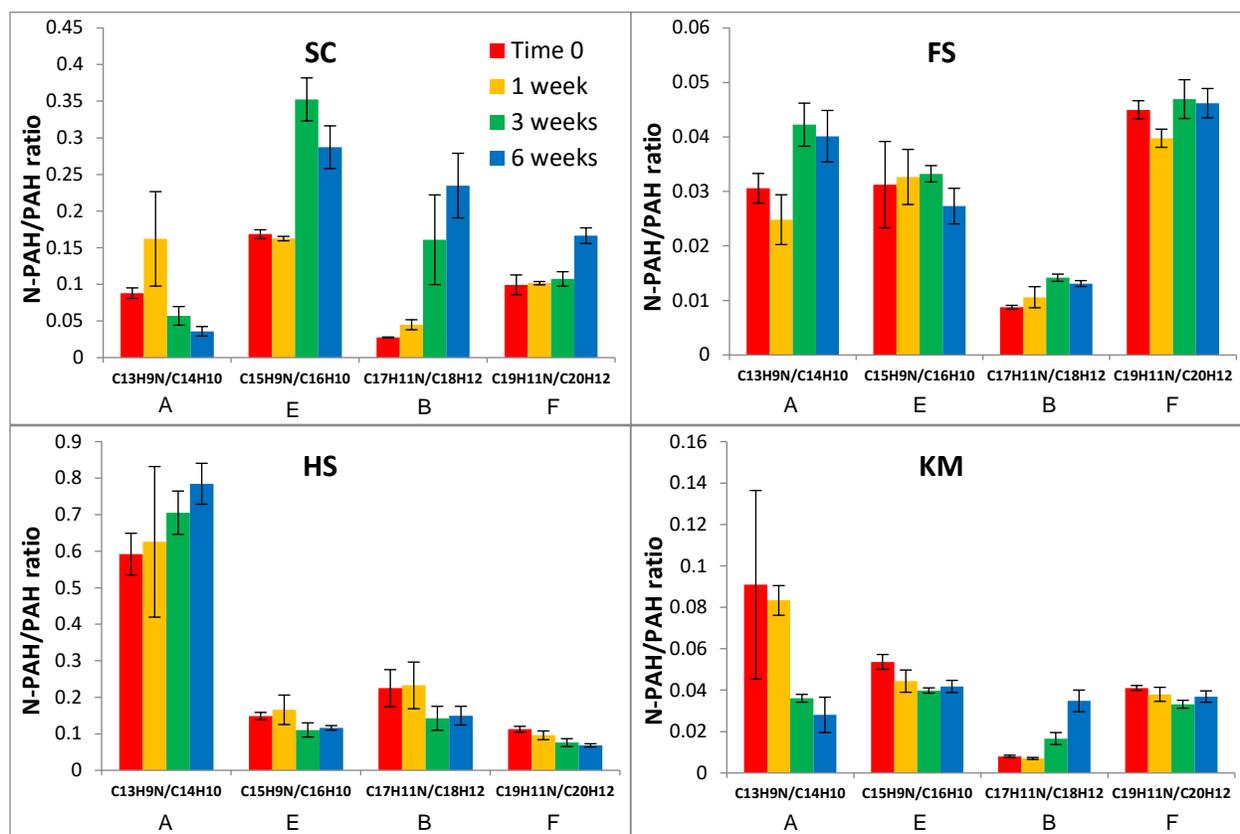


Figure C.3. Evolution of the relative concentration (N-PAH/PAH ratio) of the major azaarene congener groups (PANHs) relative to their homocyclic counterparts (PAHs) during the microcosm incubations for each of the four samples. Note differences in the y-axis scales. The PAHs included were anthracene and phenanthrene ($C_{14}H_{10}$); pyrene and fluoranthene ($C_{16}H_{10}$); benz[*a*]anthracene and chrysene ($C_{18}H_{12}$); and benzo[*a*]pyrene, benzo[*b*]fluoranthene and benzo[*k*]fluoranthene ($C_{20}H_{12}$). Group C azaarenes are not shown because the only PAH analogue we measured was dibenz[*a,h*]anthracene, which was present at low concentration in all the samples.

Table C.1. Concentrations of non-alkylated azaarene congeners ($\mu\text{g/g}$ dry soil) before (T0) and after (T42) 6 weeks of microcosm incubation (mean \pm standard deviation; n=3; KC=kill control). In bold, values significantly reduced ($p<0.05$) after 6 weeks.

ID	FS			HS			KM			SC		
	T0	T42	KC									
A1 Acridine	0.913 \pm 0.008	0.333 \pm 0.004	0.565 \pm 0.024	1.04 \pm 0.054	0.575 \pm 0.024	0.772 \pm 0.015	3.20 \pm 0.137	1.38 \pm 0.099	3.368 \pm 0.102	281 \pm 12.6	7.85 \pm 0.819	267 \pm 11.4
A2	0.334 \pm 0.006	0.273 \pm 0.009	0.304 \pm 0.056	1.06 \pm 0.059	0.699 \pm 0.009	0.958 \pm 0.060	ND	ND	ND	7.99 \pm 0.145	0.272 \pm 0.049	8.59 \pm 0.464
A3	0.398 \pm 0.004	0.248 \pm 0.006	0.351 \pm 0.027	0.480 \pm 0.037	0.446 \pm 0.003	0.528 \pm 0.030	0.060 \pm 0.023	ND	0.070 \pm 0.027	10.3 \pm 1.14	ND	9.10 \pm 0.354
A4 phenanthridine	0.266 \pm 0.004	0.245 \pm 0.001	0.258 \pm 0.002	0.509 \pm 0.002	0.482 \pm 0.001	0.510 \pm 0.002	ND	ND	ND	0.220 \pm 0.107	ND	0.123 \pm 0.018
A5 Benzo[h]quinoline	0.357 \pm 0.005	0.260 \pm 0.002	0.366 \pm 0.004	0.537 \pm 0.006	0.431 \pm 0.014	0.446 \pm 0.001	0.561 \pm 0.045	0.251 \pm 0.030	0.606 \pm 0.128	20.0 \pm 1.777	0.044 \pm 0.062	18.0 \pm 0.268
B1 Benzo[a]acridine	1.04 \pm 0.012	0.648 \pm 0.024	0.929 \pm 0.024	2.79 \pm 0.115	1.71 \pm 0.102	2.393 \pm 0.246	12.3 \pm 0.676	11.1 \pm 0.474	11.8 \pm 0.363	13.1 \pm 0.320	3.55 \pm 0.458	13.3 \pm 0.540
B2	0.172 \pm 0.004	0.104 \pm 0.008	0.142 \pm 0.020	0.708 \pm 0.063	0.358 \pm 0.118	0.535 \pm 0.081	0.917 \pm 0.101	0.754 \pm 0.153	0.917 \pm 0.076	4.62 \pm 0.220	0.760 \pm 0.207	4.54 \pm 0.256
B3	0.360 \pm 0.013	0.187 \pm 0.022	0.363 \pm 0.001	0.370 \pm 0.027	0.197 \pm 0.011	0.304 \pm 0.029	1.11 \pm 0.081	1.02 \pm 0.058	1.098 \pm 0.022	7.21 \pm 0.434	0.133 \pm 0.028	7.04 \pm 0.403
B4	0.234 \pm 0.019	0.136 \pm 0.008	0.219 \pm 0.003	0.759 \pm 0.026	0.318 \pm 0.027	0.555 \pm 0.052	ND	ND	ND	4.76 \pm 0.228	3.08 \pm 0.086	4.40 \pm 0.227
B5	0.287 \pm 0.009	0.151 \pm 0.015	0.293 \pm 0.016	0.391 \pm 0.009	0.208 \pm 0.020	0.324 \pm 0.038	0.416 \pm 0.024	0.303 \pm 0.007	0.434 \pm 0.004	7.74 \pm 0.477	0.221 \pm 0.034	7.27 \pm 0.280
B6 Benzo[c]acridine	0.903 \pm 0.020	0.454 \pm 0.034	0.901 \pm 0.044	0.996 \pm 0.054	0.510 \pm 0.045	0.766 \pm 0.088	4.23 \pm 0.125	4.00 \pm 0.110	4.355 \pm 0.203	33.9 \pm 1.683	2.18 \pm 0.149	32.5 \pm 1.445
B7	0.122 \pm 0.002	0.069 \pm 0.004	0.120 \pm 0.004	0.136 \pm 0.002	0.082 \pm 0.002	0.114 \pm 0.013	1.53 \pm 0.032	1.22 \pm 0.093	1.369 \pm 0.039	0.902 \pm 0.068	ND	0.899 \pm 0.037
E1	0.554 \pm 0.007	0.349 \pm 0.015	0.257 \pm 0.178	6.53 \pm 0.916	3.34 \pm 0.355	6.276 \pm 1.072	6.36 \pm 0.231	5.21 \pm 0.383	6.093 \pm 0.338	31.1 \pm 0.802	18.9 \pm 3.13	29.7 \pm 0.596
E2	0.078 \pm 0.005	0.060 \pm 0.004	0.056 \pm 0.015	0.640 \pm 0.042	0.360 \pm 0.021	0.569 \pm 0.079	0.343 \pm 0.034	0.235 \pm 0.080	0.373 \pm 0.016	3.89 \pm 0.198	2.39 \pm 0.260	3.78 \pm 0.139
E3	0.058 \pm 0.003	0.043 \pm 0.002	0.048 \pm 0.007	0.191 \pm 0.009	0.111 \pm 0.016	0.159 \pm 0.042	ND	ND	ND	1.43 \pm 0.093	ND	1.35 \pm 0.035
E4 Azapyrene	0.403 \pm 0.021	0.212 \pm 0.016	0.318 \pm 0.031	1.78 \pm 0.114	0.945 \pm 0.357	1.198 \pm 0.311	2.33 \pm 0.146	0.650 \pm 0.067	1.876 \pm 0.160	11.9 \pm 0.878	8.39 \pm 0.740	11.5 \pm 0.537
E5 Azafluoranthene	0.133 \pm 0.003	0.080 \pm 0.007	0.124 \pm 0.002	0.594 \pm 0.044	0.214 \pm 0.053	0.418 \pm 0.125	2.50 \pm 0.135	0.213 \pm 0.062	1.891 \pm 0.053	14.1 \pm 0.622	1.29 \pm 0.115	13.6 \pm 0.433
C1 Dibenzo[a,j]acridine	0.264 \pm 0.007	0.198 \pm 0.009	0.184 \pm 0.078	0.368 \pm 0.018	0.260 \pm 0.019	0.305 \pm 0.011	2.46 \pm 0.047	2.42 \pm 0.030	2.466 \pm 0.062	1.87 \pm 0.073	1.68 \pm 0.008	1.74 \pm 0.050
C2	0.211 \pm 0.002	0.168 \pm 0.008	0.208 \pm 0.007	0.290 \pm 0.006	0.244 \pm 0.030	0.292 \pm 0.013	2.39 \pm 0.134	2.39 \pm 0.074	2.443 \pm 0.162	0.898 \pm 0.066	0.802 \pm 0.024	0.888 \pm 0.043
C3,4	0.362 \pm 0.004	0.254 \pm 0.021	0.327 \pm 0.009	0.652 \pm 0.018	0.486 \pm 0.040	0.649 \pm 0.025	11.6 \pm 0.948	11.7 \pm 0.310	11.9 \pm 0.765	4.60 \pm 0.454	4.14 \pm 0.293	4.96 \pm 0.503

Table C.1., continued.

ID	FS			HS			KM			SC		
	T0	T42	KC									
C5	0.301 ± 0.007	0.220 ± 0.017	0.296 ± 0.005	0.664 ± 0.008	0.512 ± 0.029	0.581 ± 0.012	0.403 ± 0.011	0.323 ± 0.022	0.339 ± 0.003	0.217 ± 0.013	0.140 ± 0.004	0.194 ± 0.007
F1	0.172 ± 0.006	0.144 ± 0.006	0.161 ± 0.062	0.373 ± 0.026	0.210 ± 0.007	0.347 ± 0.038	2.20 ± 0.086	1.97 ± 0.017	2.11 ± 0.080	2.13 ± 0.098	1.94 ± 0.023	2.04 ± 0.090
F2	0.104 ± 0.000	0.096 ± 0.003	0.070 ± 0.033	0.260 ± 0.005	0.148 ± 0.020	0.206 ± 0.016	0.831 ± 0.038	0.763 ± 0.019	0.840 ± 0.028	1.11 ± 0.021	0.958 ± 0.014	1.08 ± 0.047
F3	0.146 ± 0.004	0.123 ± 0.004	0.100 ± 0.049	0.172 ± 0.002	0.099 ± 0.021	0.137 ± 0.021	0.398 ± 0.014	0.376 ± 0.003	0.426 ± 0.015	0.740 ± 0.036	0.650 ± 0.024	0.595 ± 0.016
F4	0.085 ± 0.003	0.060 ± 0.001	0.065 ± 0.014	0.179 ± 0.002	0.122 ± 0.021	0.153 ± 0.006	0.706 ± 0.028	0.698 ± 0.032	0.723 ± 0.028	1.03 ± 0.035	0.788 ± 0.024	0.950 ± 0.019
F5	0.122 ± 0.011	0.088 ± 0.007	0.082 ± 0.029	0.399 ± 0.010	0.268 ± 0.025	0.335 ± 0.042	0.174 ± 0.008	0.140 ± 0.003	0.181 ± 0.011	0.665 ± 0.026	0.575 ± 0.003	0.667 ± 0.003
F6	0.173 ± 0.009	0.115 ± 0.003	0.147 ± 0.031	0.354 ± 0.012	0.169 ± 0.011	0.261 ± 0.028	3.83 ± 0.194	3.70 ± 0.040	3.99 ± 0.088	2.17 ± 0.131	1.86 ± 0.041	1.87 ± 0.009
F7	0.165 ± 0.006	0.142 ± 0.003	0.149 ± 0.030	0.354 ± 0.007	0.212 ± 0.012	0.301 ± 0.025	1.55 ± 0.040	1.43 ± 0.035	1.56 ± 0.019	1.17 ± 0.046	0.963 ± 0.026	1.17 ± 0.021
F8	0.391 ± 0.010	0.288 ± 0.012	0.358 ± 0.016	0.892 ± 0.018	0.545 ± 0.053	0.812 ± 0.075	4.58 ± 0.200	4.23 ± 0.063	4.40 ± 0.196	3.63 ± 0.089	3.06 ± 0.115	3.52 ± 0.039
F9	0.094 ± 0.003	0.075 ± 0.003	0.096 ± 0.001	0.263 ± 0.003	0.158 ± 0.007	0.205 ± 0.045	ND	ND	ND	0.762 ± 0.021	0.667 ± 0.003	0.751 ± 0.031
F10	0.260 ± 0.011	0.187 ± 0.001	0.237 ± 0.005	1.07 ± 0.028	0.464 ± 0.034	0.848 ± 0.117	5.26 ± 0.174	5.02 ± 0.120	5.29 ± 0.119	3.17 ± 0.150	2.91 ± 0.015	3.08 ± 0.214
F11	0.237 ± 0.006	0.163 ± 0.003	0.228 ± 0.004	0.450 ± 0.027	0.184 ± 0.005	0.339 ± 0.048	0.326 ± 0.014	0.224 ± 0.001	0.311 ± 0.030	1.07 ± 0.098	0.901 ± 0.068	0.941 ± 0.008
F12	0.148 ± 0.003	0.098 ± 0.003	0.149 ± 0.002	0.348 ± 0.006	0.150 ± 0.006	0.271 ± 0.021	0.752 ± 0.046	0.607 ± 0.034	0.644 ± 0.012	0.897 ± 0.033	0.607 ± 0.042	0.904 ± 0.040
F13	0.152 ± 0.005	0.107 ± 0.005	0.145 ± 0.002	0.279 ± 0.010	0.156 ± 0.010	0.239 ± 0.011	0.787 ± 0.014	0.698 ± 0.018	0.696 ± 0.007	0.901 ± 0.039	0.731 ± 0.006	0.866 ± 0.014

APPENDIX D: SUPPORTING INFORMATION FOR CHAPTER 6

Determination of optimal incubation time for U-¹³C tracer experiments

The optimal incubation time for each of the target compounds was determined in a preliminary incubation using unspiked soil. To be consistent with later experiments using U-¹³C labeled PAHs, 20% of the contaminated soil was amended with the same volume of acetone used for spiked samples. After solvent evaporation overnight, the acetone-amended soil was gradually mixed with the remaining 80% of unspiked soil. Aliquots of 1 g of soil (dry wt) were placed in sterile 125 mL Erlenmeyer flasks with PTFE screw caps containing 30 mL of 10 mM phosphate buffer (pH 7.5) supplemented with urea (1 mM) as nitrogen source. All flasks were incubated at 25 °C under agitation at 150 rpm, and their caps were loosened for 5 mins every 2d to avoid oxygen limitation. Triplicate samples were sacrificed at each of four time points: 0 (initial), 7 days, 14 days, and 21 days.

Extraction methods

After incubation for the selected time point, the slurry was collected and centrifuged to recover the solid and aqueous phases separately. The solid phase was processed by two successive overnight extractions using 20 mL of a DCM: acetone mixture (1:1, v/v) with agitation on a wrist-action shaker. Prior to extraction, 200 µg of anthracene-d₁₀ was spiked into the samples in acetone solution as a recovery standard. Sodium sulfate (10 g) was added to dry the sample, and glass beads were added for better extraction efficiency. The collected extracts were filtered through a 0.2 µm pore-size nylon membrane, and brought up to 50 mL with DCM and acetone (1:1, v/v). The aqueous phase (30 mL) was acidified (pH 2) with 1M HCl, extracted three times with 10 mL of ethyl acetate each time, and the pooled extracts concentrated to a final volume of 5 mL. All extracts were stored at 4 °C until further analysis.

Instrumental analysis

Extracts were analyzed using an Agilent 1200 series high-performance liquid chromatography (HPLC) system interfaced with a 6520 accurate mass quadrupole time-of-flight mass spectrometer (qTOF-MS, Agilent Technologies, Santa Clara, CA). Preliminary analysis suggested that most signals in extracts of the solid phase were detected in electrospray ionization positive (ESI+) mode, while the negative (ESI-) mode detected more signals in extracts of the aqueous phase. Therefore, different polarities were applied for each ionization mode and extract source. The TOF mass selector was tuned in high resolution mode (4 GHz, resolution > 15,000 at m/z 322), and the mass accuracy was continuously corrected in real time using two references, purine (m/z 121.050873) and HP-921 (m/z 922.009798). In full-scan mode, the TOF detector was set to acquire spectra over the range m/z 100-1000. In product ion scan mode, the collision energy was set at 20 eV to obtain fragmentation spectra for structure elucidation.

Chromatographic separation was accomplished using a Poroshell 120 EC C18 column (3.0 × 100 mm, 2.7 μm, Agilent). The mobile phase consisted of deionized water (A) and methanol (B), and was delivered at a flow rate of 0.2 mL·min⁻¹ applying the following gradient: 0-8 min, 10-50 % B; 8-25 min, 50-90% B; 25-26 min 90% B; 26-33 min, 90-10% B. To enhance the ionization efficiency, 0.1% formic acid was added for both A and B in the analysis of ESI+ mode, while 5 mM ammonium formate was added in the analysis of ESI- mode.

Data analysis

To compare the samples spiked with a ¹³C-labeled PAH and the unspiked samples from the same incubation time, we used the metabolomics platform XCMS Online (<https://xcmsonline.scripps.edu/>). In brief, we used the centWave algorithm for peak detection, with a peak width range from 10 to 60 seconds. The mass error tolerance was set at 10 ppm.

Retention time correction was performed with the peak groups method. Final annotation and molecular formula assignment were achieved by XCMS Online and Agilent MassHunter.

Annotation for features allowed the search for isotopic features and multiple adducts ($[M+H]^+$, $[M+Na]^+$, $[M+K]^+$, $[M+NH_4]^+$). The differences of features between groups (U¹³-C PAH spiked vs. unspiked soil) were evaluated by Welch *t*-test with a threshold *p*-value set at 0.01, and a minimum fold-change set at 3.

In parallel, feature detection and alignment were processed by R package XCMS (Patti et al., 2012), and isotope group picking was finished by R package X13CMS (Huang et al., 2014).

The parameters were optimized according to the conditions of analysis as listed in Table C2.

Table D.1. Concentrations of monitored PAHs ($\mu\text{g/g}$ dry soil) in contaminated soil samples (mean \pm standard deviation; $n=3$). KC: killed control after 21 days. n.d.: not detected.

PAHs	t0	t7	t14	t21	KC
naphthalene	n.d.	n.d.	n.d.	n.d.	n.d.
acenaphthene	324 \pm 32.1	n.d.	n.d.	n.d.	364 \pm 7.51
fluorene	384 \pm 16.5	90.1 \pm 14.5	33.9 \pm 14.0	n.d.	413 \pm 54.8
phenanthrene	2490 \pm 57	179 \pm 14.9	103 \pm 10.6	55.6 \pm 28.2	2420 \pm 65
anthracene	847 \pm 112	623 \pm 70.5	292 \pm 26.8	164 \pm 53.4	827 \pm 3.50
fluoranthene	1250 \pm 36	389 \pm 34.5	63.3 \pm 19.3	6.61 \pm 1.86	1280 \pm 71
pyrene	761 \pm 10.4	748 \pm 4.13	317 \pm 59.9	79.2 \pm 12.1	819 \pm 40.4
benzo[<i>a</i>]antherencene	191 \pm 14.1	221 \pm 10.5	24.9 \pm 4.25	14.6 \pm 11.4	222 \pm 13.0
chrysene	197 \pm 22.5	166 \pm 15.2	n.d.	n.d.	178 \pm 12.1
BbF	52.1 \pm 7.59	49.5 \pm 3.46	44.0 \pm 13.0	31.4 \pm 12.8	72.7 \pm 9.77
BkF	37.3 \pm 2.13	35.4 \pm 3.09	30.7 \pm 1.60	21.2 \pm 3.02	44.3 \pm 1.19
BaP	63.5 \pm 3.77	52.4 \pm 3.84	53.2 \pm 3.95	55.6 \pm 2.38	59.8 \pm 3.32
DBA	n.d.	n.d.	n.d.	n.d.	n.d.
BgP	n.d.	n.d.	n.d.	n.d.	n.d.

Table D.2. Parameter values used for data analysis with XCMS and X13CMS packages.

package	function	parameter values
XCMS	xcmsSet()	method = centWave ppm= 20 peakwidth= c(10,200) bw=5 mzwid=0.025
X13CMS	getIsoLabelReport()	RTwin = 10 ppm = 20 noiseNoise = 10000 alpha = 0.05 enrichTol = 0.1

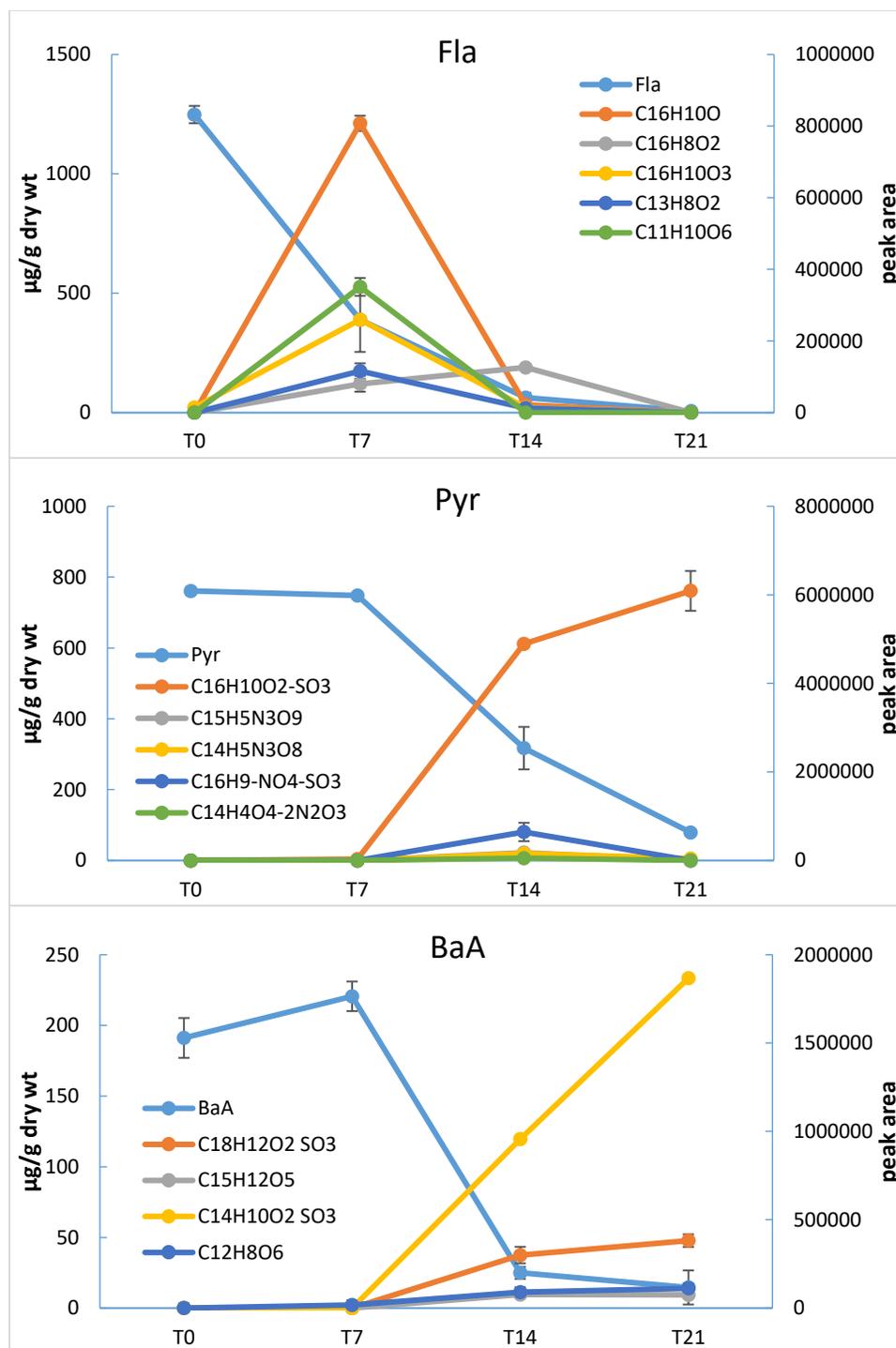


Figure D.1. Time course in unspiked samples of the concentrations of the three PAHs targeted for isotope-assisted metabolomics (left y-axis) and intensities of the detected metabolites (right y-axis). Error bars, where visible, represent standard deviations of triplicate samples.

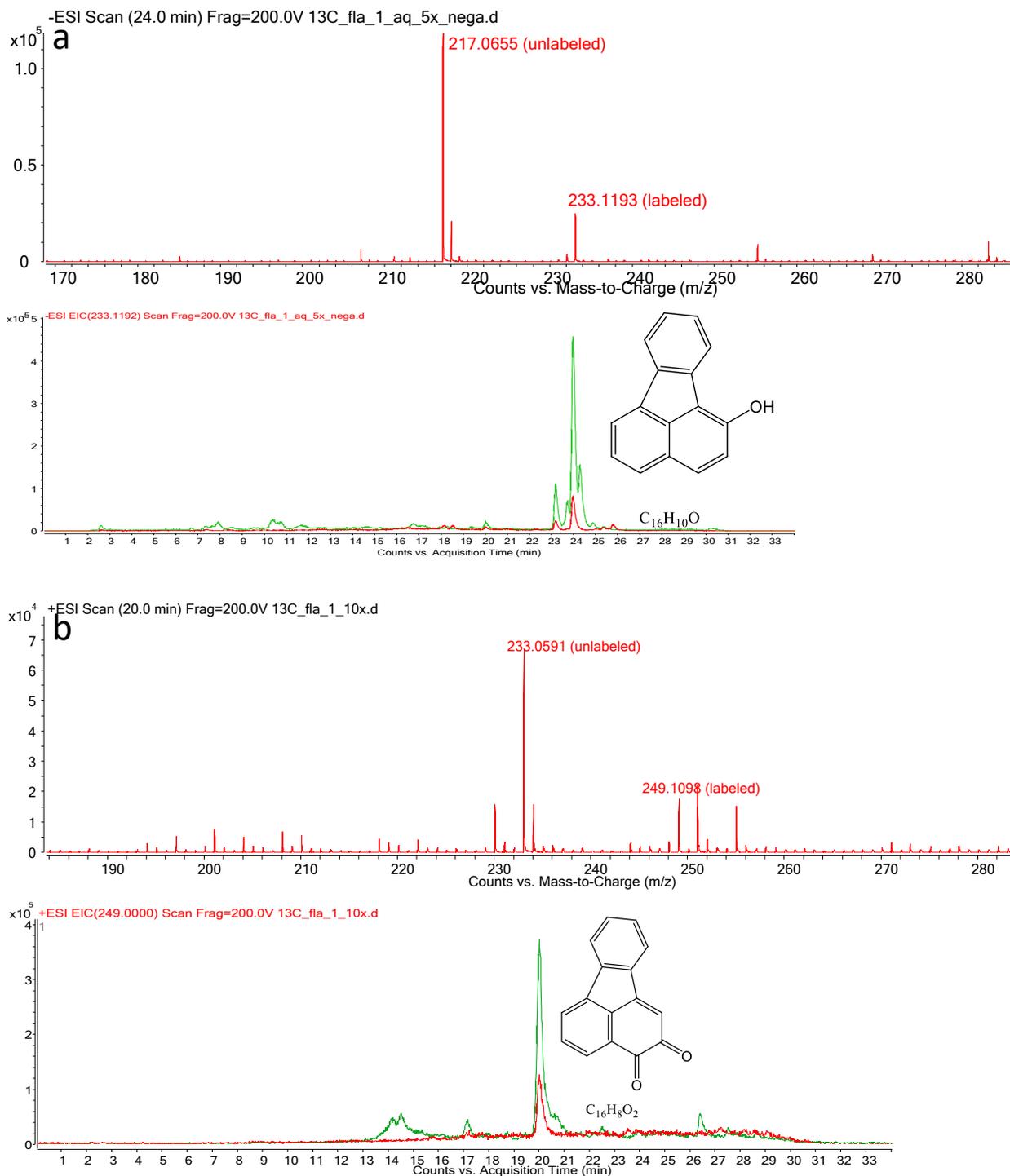


Figure D.2. Mass spectra and extracted ion chromatograms (EIC) of three major metabolites from fluoranthene. **(a)** $C_{16}H_{10}O$, **(b)** $C_{16}H_8O_2$, **(c)** $C_{16}H_{10}O_3$. The structures drawn in EIC panels show substituents in hypothesized locations in the absence of standards or known mass spectral properties.

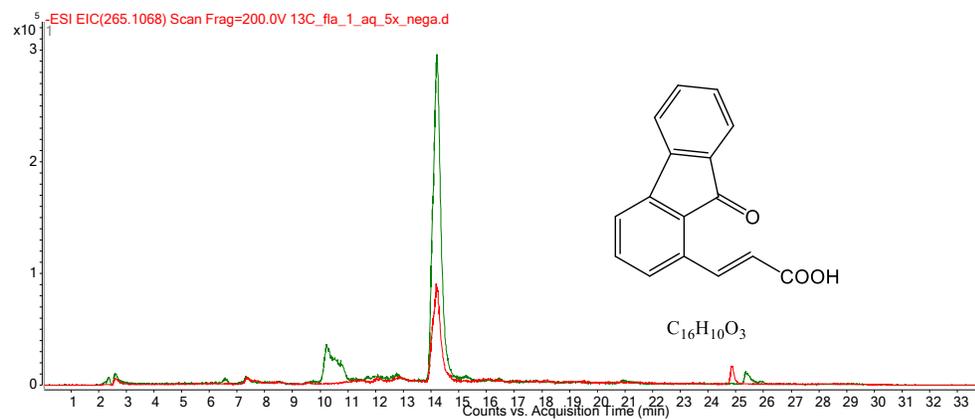
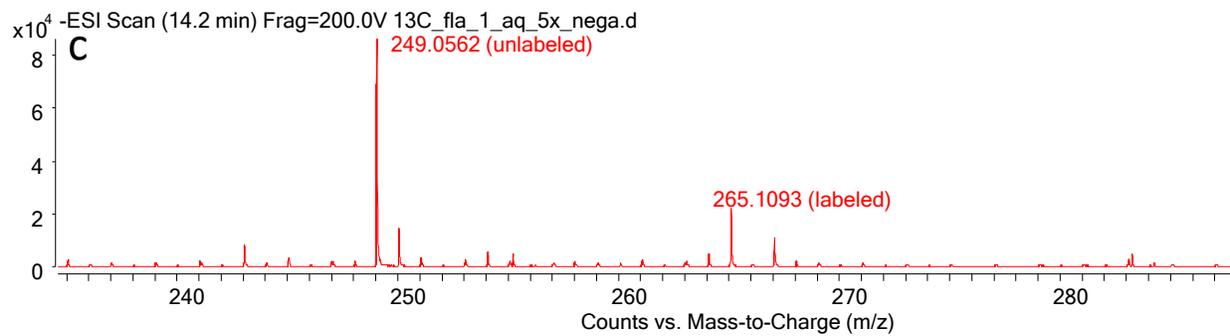


Figure D.2. (continued)

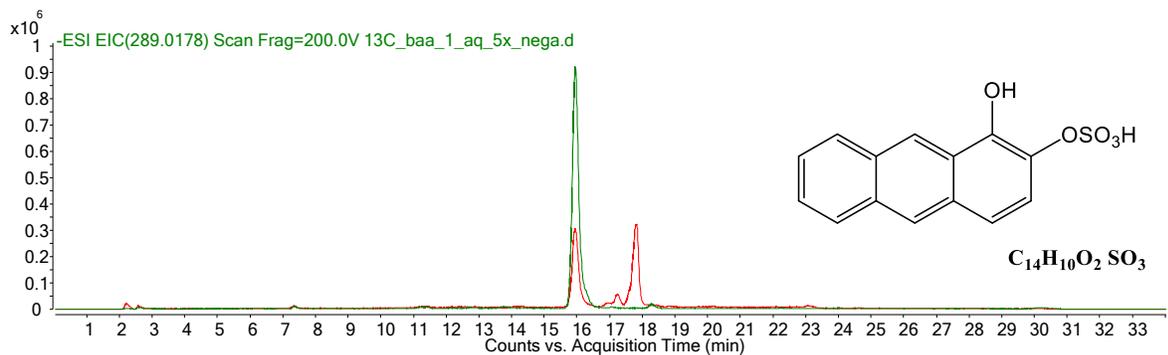
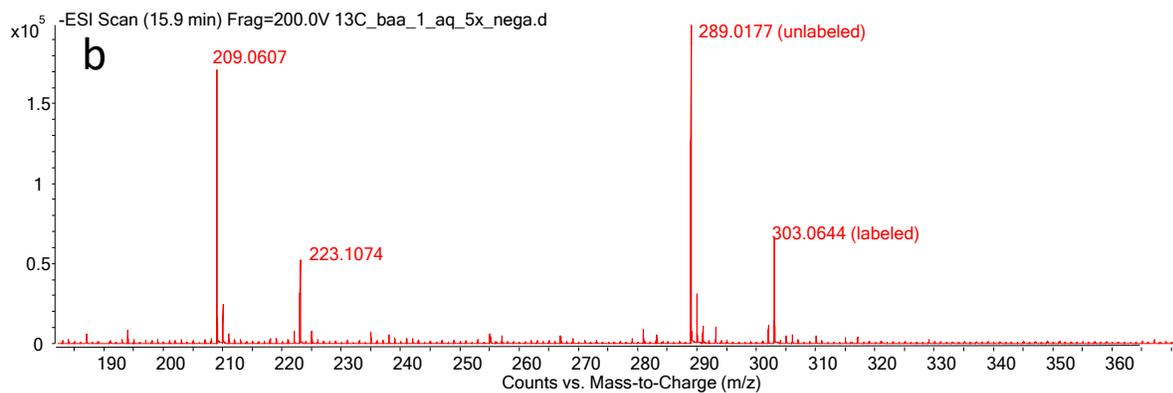
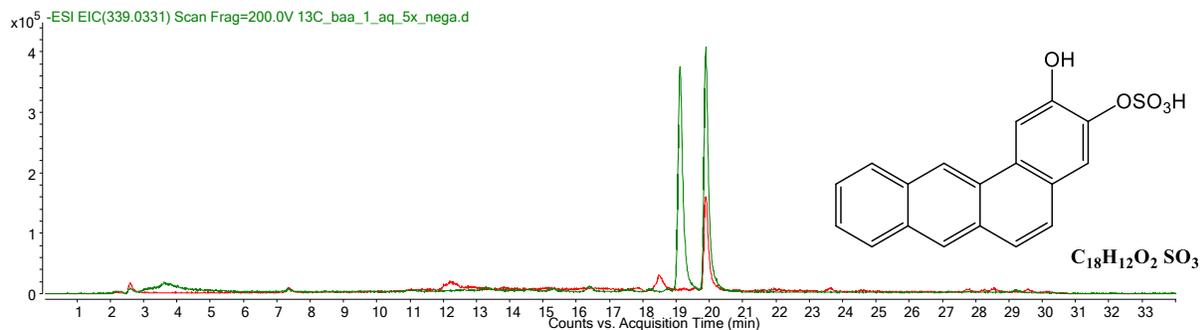
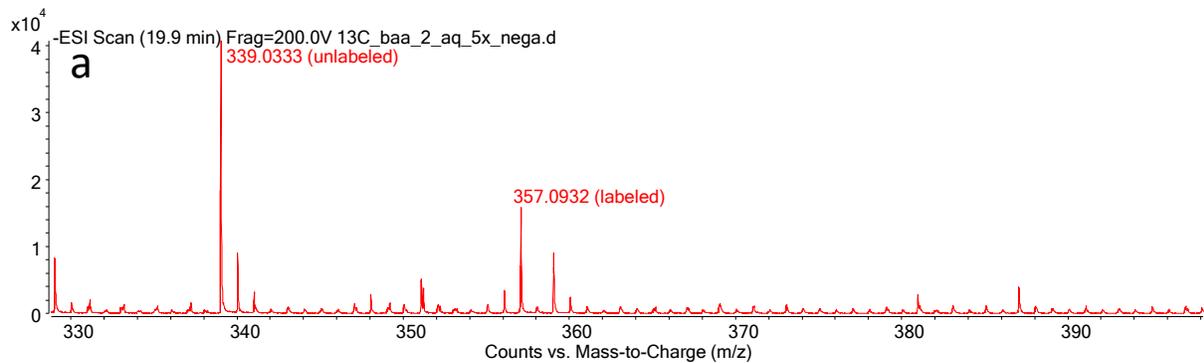


Figure D.3. Mass spectra and EIC of three major metabolites from B[a]A. (a) $C_{18}H_{12}O_2 \cdot SO_3$, (b) $C_{14}H_{10}O_2 \cdot SO_3$, (c) $C_{12}H_8O_6$. Note that unmatched peaks are noise peaks from EIC extraction windows.

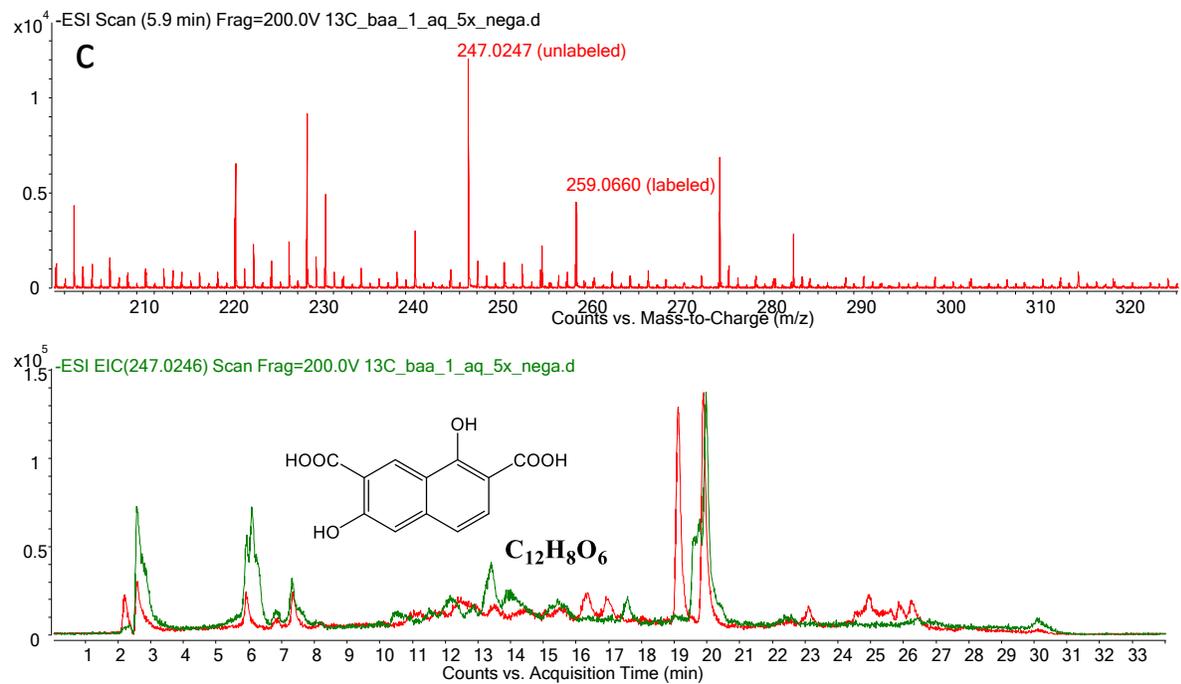


Figure D.3. (continued)

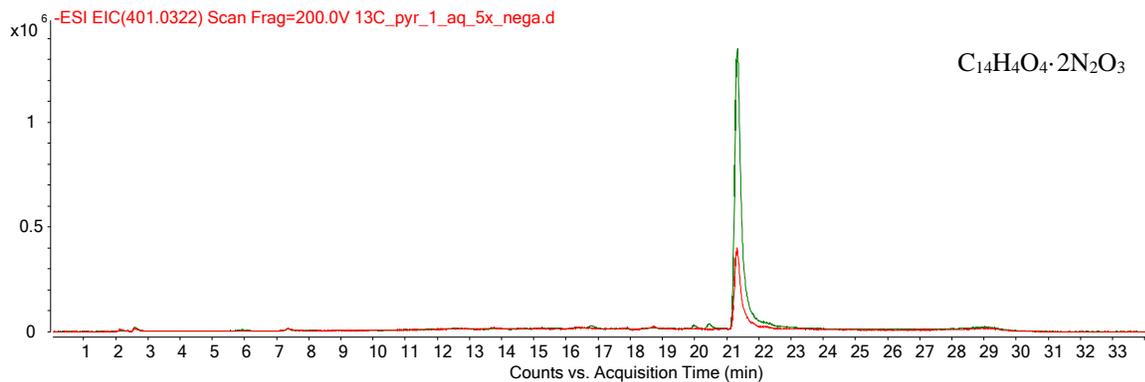
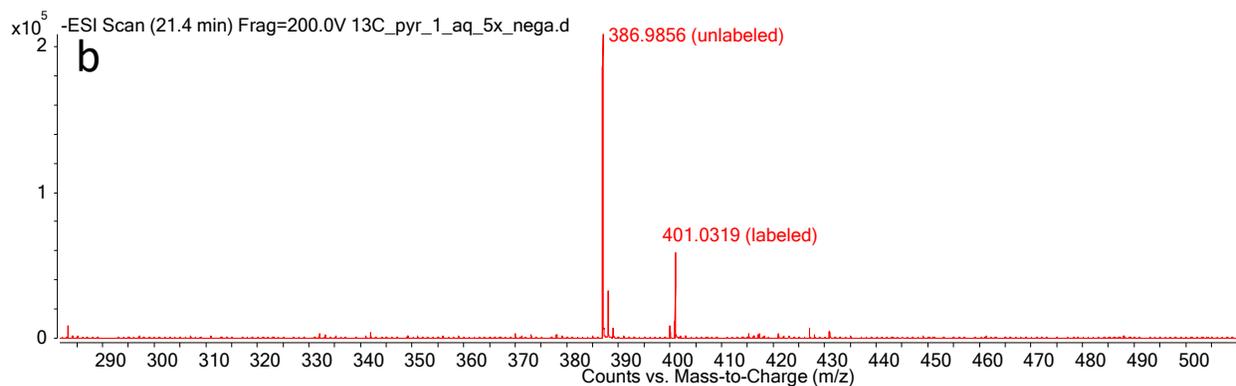
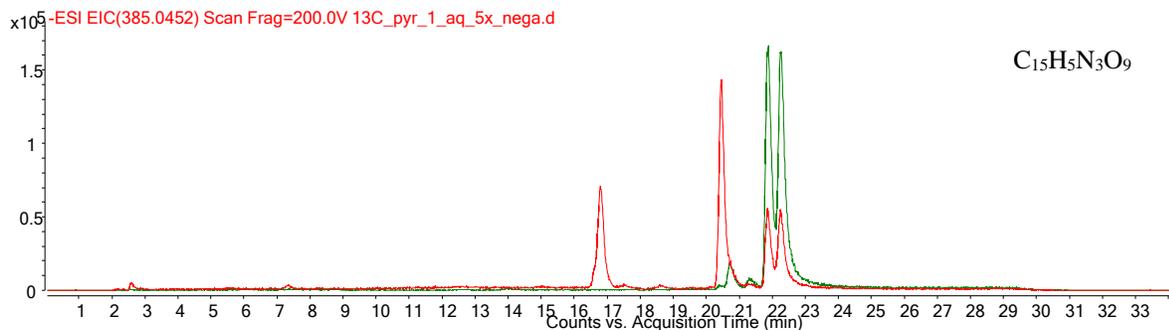
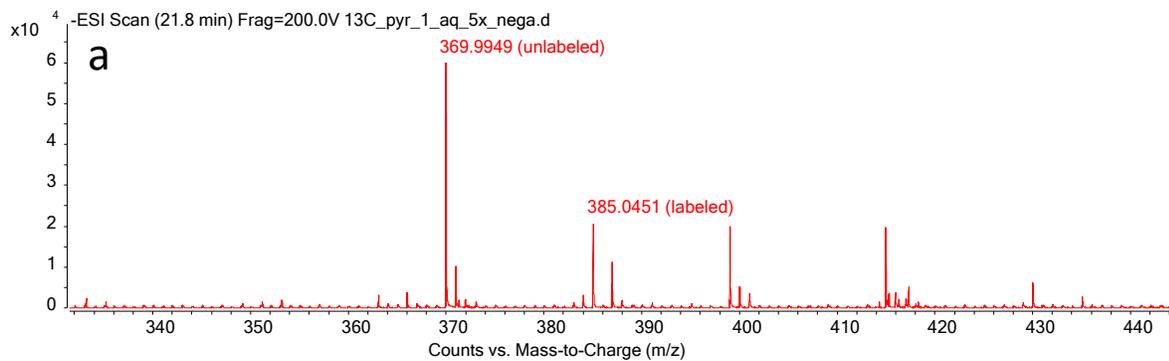


Figure D.4. Mass spectra and EIC of four major metabolites from pyrene. **(a)** $C_{15}H_5N_3O_9$, **(b)** $C_{14}H_4O_4 \cdot 2N_2O_3$, **(c)** $C_{14}H_5N_3O_8$, **(d)** $C_{16}H_9NO_4 \cdot SO_3$. Note that unmatched peaks are noise peaks from EIC extraction windows.

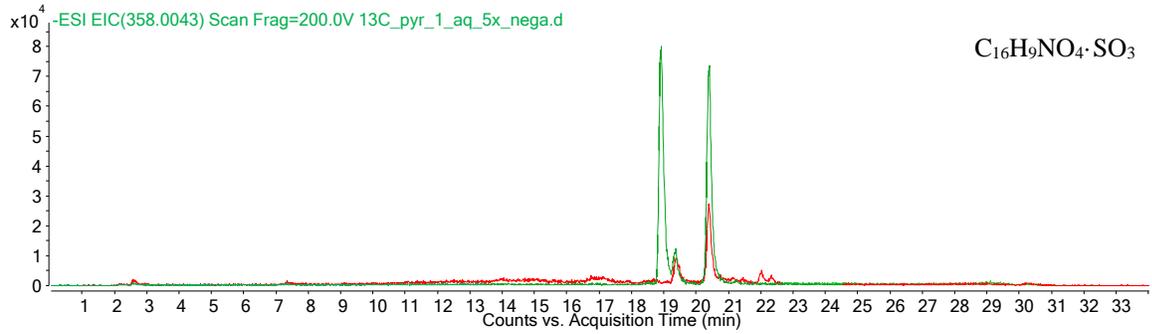
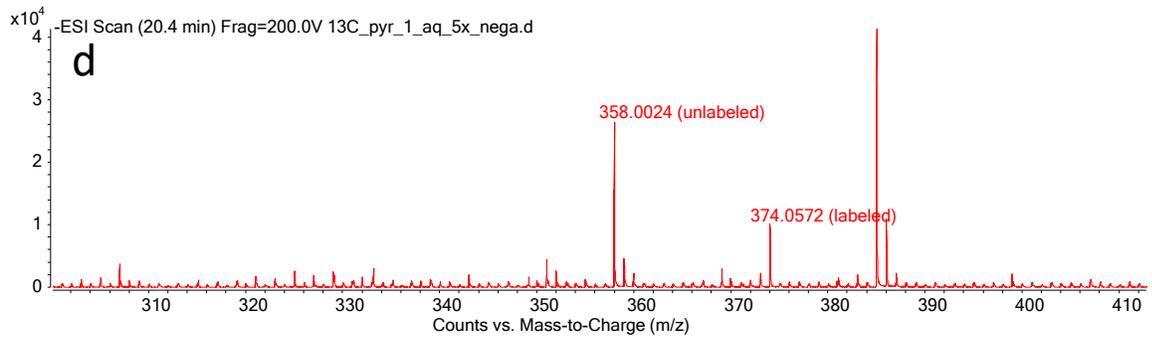
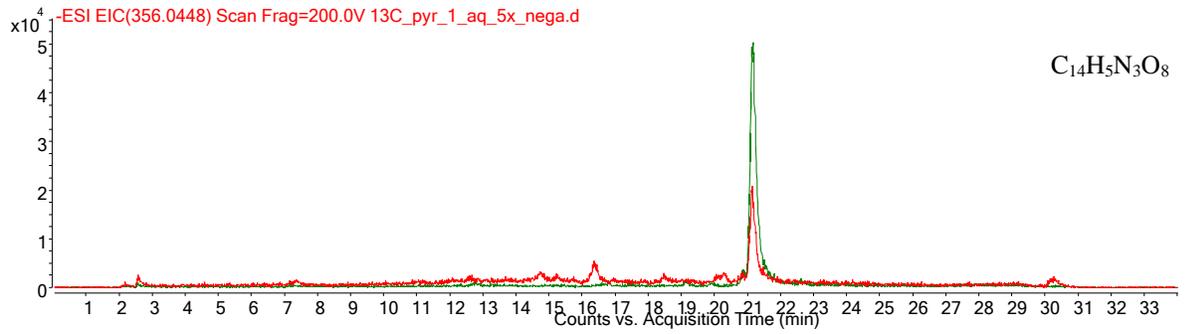
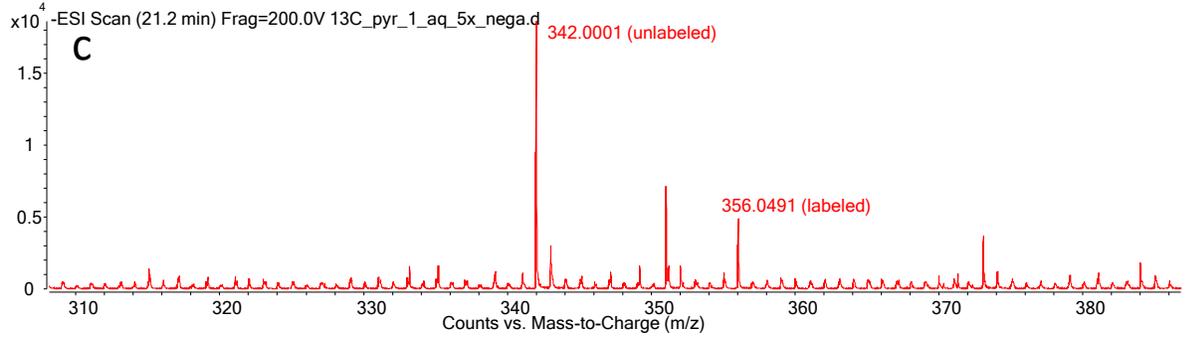


Figure D.4. (continued)

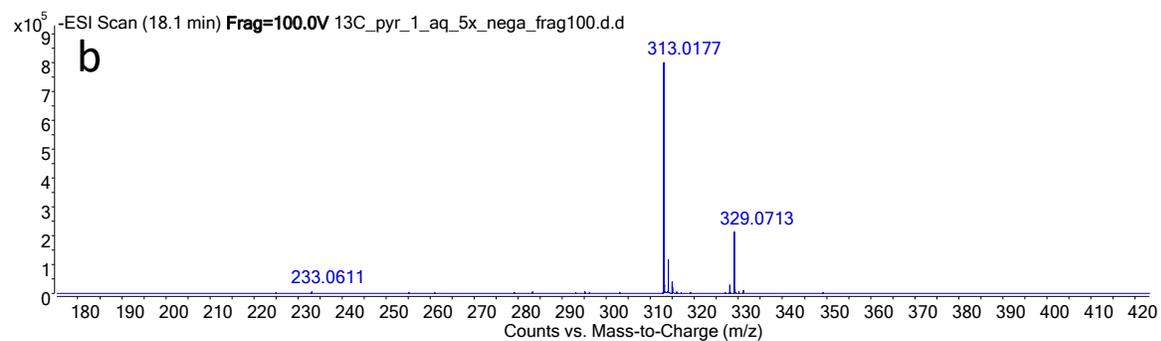
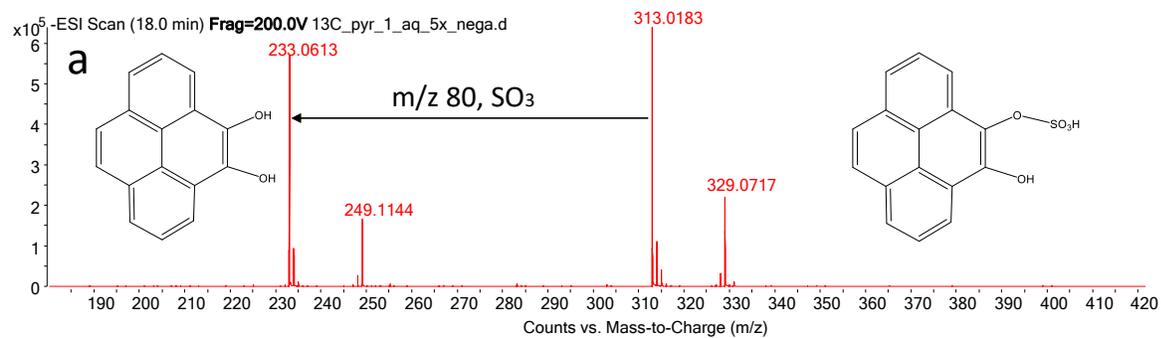


Figure D.5. Mass spectra of the sulfate conjugate of dihydroxypyrene, under **(a)** normal (200 V) and **(b)** low (100 V) source fragmentor voltage.

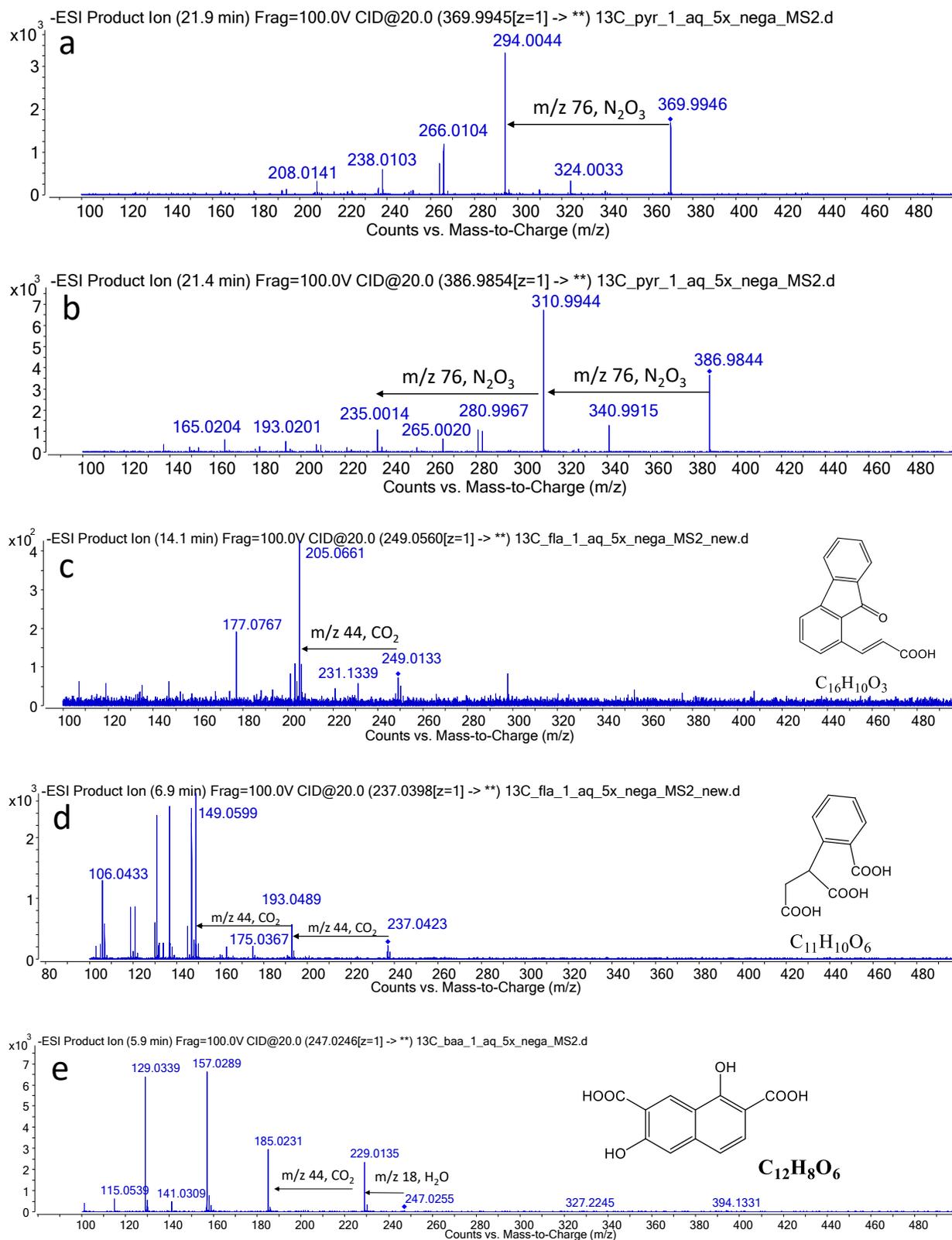


Figure D.6. MS/MS spectra of pyrene metabolites (a) m/z 369.9945 and (b) m/z 386.9854, fluoranthene metabolite (c) $C_{16}H_{10}O_3$ and (d) $C_{11}H_{10}O_6$, and BaA metabolite (e) $C_{12}H_8O_6$.

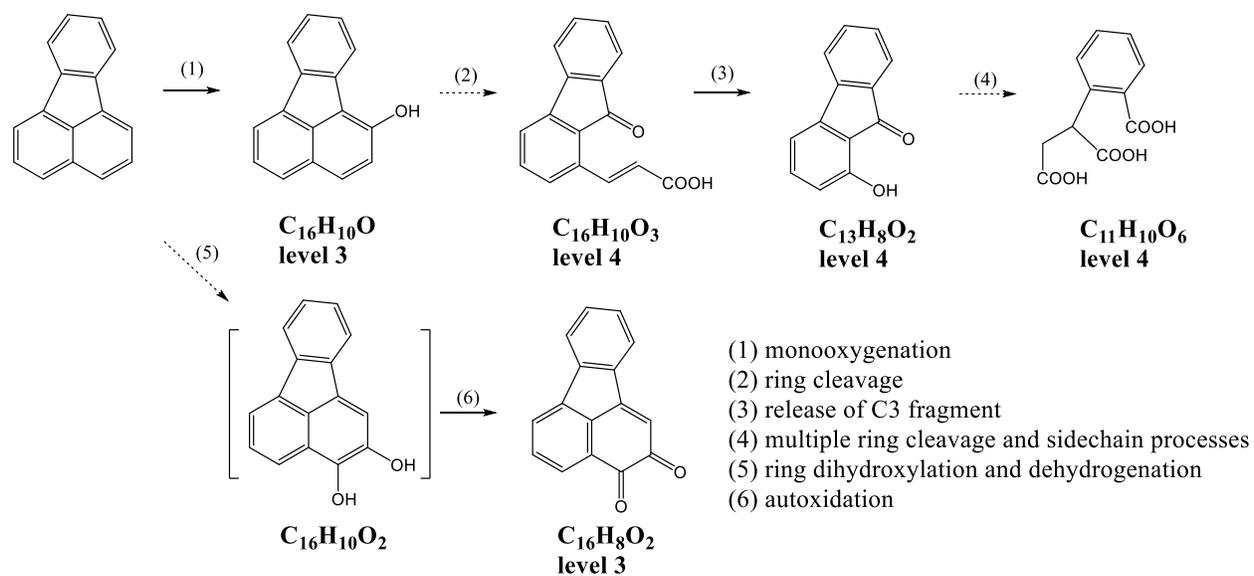


Figure D.7. Proposed degradation pathways for fluoranthene based on identified or putative metabolites. Dashed arrows indicate more than one successive reaction. Proposed reactions generally follow canonical reactions for aerobic bacterial metabolism of PAHs as summarized in Kweon et al (Kweon et al., 2011).

REFERENCES

- Adrion, A.C., Nakamura, J., Shea, D., Aitken, M.D., 2016. Screening Nonionic Surfactants for Enhanced Biodegradation of Polycyclic Aromatic Hydrocarbons Remaining in Soil After Conventional Biological Treatment. Environ. Sci. Technol. 50, 3838–3845.
- Andersson, B.E., Lundstedt, S., Tornberg, K., Schnürer, Y., Öberg, L.G., Mattiasson, B., 2003. Incomplete degradation of polycyclic aromatic hydrocarbons in soil inoculated with wood-rotting fungi and their effect on the indigenous soil bacteria. Environ. Toxicol. Chem. 22, 1238–1243.
- Andersson, J.T., Achten, C., 2015. Time to Say Goodbye to the 16 EPA PAHs? Toward an Up-to-Date Use of PACs for Environmental Purposes. Polycycl. Aromat. Compd. 35, 330–354.
- Anyanwu, I.N., Semple, K.T., 2015a. Fate and behaviour of nitrogen-containing polycyclic aromatic hydrocarbons in soil. Environ. Technol. Innov. 3, 108–120.
- Anyanwu, I.N., Semple, K.T., 2015b. Biodegradation of Phenanthrene-Nitrogen-Containing Analogues in Soil. Water. Air. Soil Pollut. 226, 252.
- Baird, W.M., Hooven, L.A., Mahadevan, B., 2005. Carcinogenic polycyclic aromatic hydrocarbon-DNA adducts and mechanism of action. Environ. Mol. Mutagen. 45, 106–114.
- Ballesteros-Gómez, A., Ballesteros, J., Ortiz, X., Jonker, W., Helmus, R., Jobst, K.J., Parsons, J.R., Reiner, E.J., 2017. Identification of Novel Brominated Compounds in Flame Retarded Plastics Containing TBBPA by Combining Isotope Pattern and Mass Defect Cluster Analysis. Environ. Sci. Technol. 51, 1518–1526.
- Bandowe, B.A.M., Lueso, M.G., Wilcke, W., 2014a. Oxygenated polycyclic aromatic hydrocarbons and azaarenes in urban soils: A comparison of a tropical city (Bangkok) with two temperate cities (Bratislava and Gothenburg). Chemosphere 107, 407–414.
- Bandowe, B.A.M., Meusel, H., Huang, R., Ho, K., Cao, J., Hoffmann, T., Wilcke, W., 2014b. PM_{2.5}-bound oxygenated PAHs, nitro-PAHs and parent-PAHs from the atmosphere of a Chinese megacity: Seasonal variation, sources and cancer risk assessment. Sci. Total Environ. 473, 77–87.
- Bandowe, B.A.M., Shukurov, N., Kersten, M., Wilcke, W., 2010. Polycyclic aromatic hydrocarbons (PAHs) and their oxygen-containing derivatives (OPAHs) in soils from the Angren industrial area, Uzbekistan. Environ. Pollut. 158, 2888–2899.
- Banger, K., Toor, G.S., Chirenje, T., Ma, L., 2010. Polycyclic Aromatic Hydrocarbons in Urban Soils of Different Land Uses in Miami, Florida. Soil Sediment Contam. Int. J. 19, 231–243.

- Barzen-Hanson, K.A., Roberts, S.C., Choyke, S., Oetjen, K., McAlees, A., Riddell, N., McCrindle, R., Ferguson, P.L., Higgins, C.P., Field, J.A., 2017. Discovery of 40 Classes of Per- and Polyfluoroalkyl Substances in Historical Aqueous Film-Forming Foams (AFFFs) and AFFF-Impacted Groundwater. Environ. Sci. Technol. 51, 2047–2057.
- Bayona, J.M., Albaigés, J., Solanas, A.M., Pares, R., Garrigues, P., Ewald, M., 1986. Selective Aerobic Degradation of Methyl-Substituted Polycyclic Aromatic Hydrocarbons in Petroleum by Pure Microbial Culture. Int. J. Environ. Anal. Chem. 23, 289–303.
- Belkin, S., Stieber, M., Tiehm, A., Frimmel, F.H., Abeliovich, A., Werner, P., Ulitzur, S., 1994. Toxicity and genotoxicity enhancement during polycyclic aromatic hydrocarbons' biodegradation. Environ. Toxicol. Water Qual. 9, 303–309.
- Biache, C., Ouali, S., Cébron, A., Lorgeoux, C., Colombano, S., Faure, P., 2017. Bioremediation of PAH-contaminated soils: Consequences on formation and degradation of polar-polycyclic aromatic compounds and microbial community abundance. J. Hazard. Mater. 329, 1–10.
- Bieber, S., Greco, G., Grosse, S., Letzel, T., 2017. RPLC-HILIC and SFC with Mass Spectrometry: Polarity-Extended Organic Molecule Screening in Environmental (Water) Samples. Anal. Chem. 89, 7907–7914.
- Bleeker, E.A.J., Geest, H.G.V.D., Klamer, H.J.C., Voogt, P.D., Wind, E., Kraak, M.H.S., 1999. Toxic and Genotoxic Effects of Azaarenes: Isomers and Metabolites. Polycycl. Aromat. Compd. 13, 191–203.
- Bleeker, E.A.J., Pieters, B.J., Wiegman, S., Kraak, M.H.S., 2002. Comparative (Photoenhanced) Toxicity of Homocyclic and Heterocyclic PACs. Polycycl. Aromat. Compd. 22, 601–610.
- Bleeker, E.A.J., Wiegman, S., de Voogt, P., Kraak, M., Leslie, H.A., de Haas, E., Admiraal, W., 2002. Toxicity of azaarenes. Rev. Environ. Contam. Toxicol. 173, 39–83.
- Blum, P., Sagner, A., Tiehm, A., Martus, P., Wendel, T., Grathwohl, P., 2011. Importance of heterocyclic aromatic compounds in monitored natural attenuation for coal tar contaminated aquifers: A review. J. Contam. Hydrol. 126, 181–194.
- Bobeldijk, I., Stoks, P.G.M., Vissers, J.P.C., Emke, E., van Leerdam, J.A., Muilwijk, B., Berbee, R., Noij, T.H.M., 2002. Surface and wastewater quality monitoring: combination of liquid chromatography with (geno)toxicity detection, diode array detection and tandem mass spectrometry for identification of pollutants. J. Chromatogr. A, Mass Spectrometry: Innovation and Application. Part I 970, 167–181.
- Bolton, J.L., Trush, M.A., Penning, T.M., Dryhurst, G., Monks, T.J., 2000. Role of Quinones in Toxicology. Chem. Res. Toxicol. 13, 135–160.
- Boonchan, S., Britz, M.L., Stanley, G.A., 2000. Degradation and Mineralization of High-Molecular-Weight Polycyclic Aromatic Hydrocarbons by Defined Fungal-Bacterial Cocultures. Appl. Environ. Microbiol. 66, 1007–1019.

- Bossert, I.D., Bartha, R., 1986. Structure-biodegradability relationships of polycyclic aromatic hydrocarbons in soil. Bull. Environ. Contam. Toxicol. 37, 490–495.
- Bouchez, M., Blanchet, D., Bardin, V., Haeseler, F., Vandecasteele, J.-P., 1999. Efficiency of defined strains and of soil consortia in the biodegradation of polycyclic aromatic hydrocarbon (PAH) mixtures. Biodegradation 10, 429–435.
- Bouchez, M., Blanchet, D., Vandecasteele, J.-P., 1995. Degradation of polycyclic aromatic hydrocarbons by pure strains and by defined strain associations: inhibition phenomena and cometabolism. Appl. Microbiol. Biotechnol. 43, 156–164.
- Brack, W., 2003. Effect-directed analysis: a promising tool for the identification of organic toxicants in complex mixtures? Anal. Bioanal. Chem. 377, 397–407.
- Brack, W., Ait-Aissa, S., Burgess, R.M., Busch, W., Creusot, N., Di Paolo, C., Escher, B.I., Mark Hewitt, L., Hilscherova, K., Hollender, J., Hollert, H., Jonker, W., Kool, J., Lamoree, M., Muschket, M., Neumann, S., Rostkowski, P., Ruttkies, C., Schollee, J., Schymanski, E.L., Schulze, T., Seiler, T.-B., Tindall, A.J., De Aragão Umbuzeiro, G., Vrana, B., Krauss, M., 2016. Effect-directed analysis supporting monitoring of aquatic environments — An in-depth overview. Sci. Total Environ. 544, 1073–1118.
- Briggs, M.K., Desavis, E., Mazzer, P.A., Sunoj, R.B., Hatcher, S.A., Hadad, C.M., Hatcher, P.G., 2003. A New Approach to Evaluating the Extent of Michael Adduct Formation to PAH Quinones: Tetramethylammonium Hydroxide (TMAH) Thermochemolysis with GC/MS. Chem. Res. Toxicol. 16, 1484–1492.
- Brinkmann, M., Maletz, S., Krauss, M., Bluhm, K., Schiwy, S., Kuckelkorn, J., Tiehm, A., Brack, W., Hollert, H., 2014. Heterocyclic Aromatic Hydrocarbons Show Estrogenic Activity upon Metabolization in a Recombinant Transactivation Assay. Environ. Sci. Technol. 48, 5892–5901.
- Brooks, L.R., Hughes, T.J., Claxton, L.D., Austern, B., Brenner, R., Kremer, F., 1998. Bioassay-directed fractionation and chemical identification of mutagens in bioremediated soils. Environ. Health Perspect. 106, 1435–1440.
- Bueschl, C., Kluger, B., Lemmens, M., Adam, G., Wiesenberger, G., Maschietto, V., Marocco, A., Strauss, J., Bödi, S., Thallinger, G.G., Krska, R., Schuhmacher, R., 2014. A novel stable isotope labelling assisted workflow for improved untargeted LC-HRMS based metabolomics research. Metabolomics Off. J. Metabolomic Soc. 10, 754–769.
- Bueschl, C., Kluger, B., Neumann, N.K.N., Doppler, M., Maschietto, V., Thallinger, G.G., Meng-Reiterer, J., Krska, R., Schuhmacher, R., 2017. MetExtract II: A software suite for stable Isotope-assisted untargeted metabolomics. Anal. Chem. 89, 9518–9526.
- Burgess, R.M., Ho, K.T., Brack, W., Lamoree, M., 2013. Effects-directed analysis (EDA) and toxicity identification evaluation (TIE): Complementary but different approaches for diagnosing causes of environmental toxicity. Environ. Toxicol. Chem. 32, 1935–1945.

- Castagnoli, Rimoldi, J.M., Bloomquist, J., Castagnoli, K.P., 1997. Potential Metabolic Bioactivation Pathways Involving Cyclic Tertiary Amines and Azaarenes. Chem. Res. Toxicol. 10, 924–940.
- Cerniglia, C.E., 1997. Fungal metabolism of polycyclic aromatic hydrocarbons: past, present and future applications in bioremediation. J. Ind. Microbiol. Biotechnol. 19, 324–333.
- Cerniglia, C.E., 1992. Biodegradation of polycyclic aromatic hydrocarbons. Biodegradation 3, 351–368.
- Chang, R.L., Levin, W., Wood, A.W., Kumar, S., Yagi, H., Jerina, D.M., Lehr, R.E., Conney, A.H., 1984. Tumorigenicity of Dihydrodiols and Diol-Epoxides of Benz[*c*]acridine in Newborn Mice. Cancer Res. 44, 5161–5164.
- Chen, S.-H., Aitken, M.D., 1999. Salicylate Stimulates the Degradation of High-Molecular Weight Polycyclic Aromatic Hydrocarbons by *Pseudomonas saccharophila* P15. Environ. Sci. Technol. 33, 435–439.
- Chibwe, L., Davie-Martin, C.L., Aitken, M.D., Hoh, E., Massey Simonich, S.L., 2017. Identification of polar transformation products and high molecular weight polycyclic aromatic hydrocarbons (PAHs) in contaminated soil following bioremediation. Sci. Total Environ. 599–600, 1099–1107.
- Chibwe, L., Geier, M.C., Nakamura, J., Tanguay, R.L., Aitken, M.D., Simonich, S.L.M., 2015a. Aerobic bioremediation of PAH contaminated soil results in increased genotoxicity and developmental toxicity. Environ. Sci. Technol.
- Chibwe, L., Geier, M.C., Nakamura, J., Tanguay, R.L., Aitken, M.D., Simonich, S.L.M., 2015b. Aerobic bioremediation of PAH contaminated soil results in increased genotoxicity and developmental toxicity. Environ. Sci. Technol. 49, 13889–13898.
- Corteselli, E.M., Aitken, M.D., Singleton, D.R., 2017a. *Rugosibacter aromaticivorans* gen. nov., sp. nov., a bacterium within the family Rhodocyclaceae, isolated from contaminated soil, capable of degrading aromatic compounds. Int. J. Syst. Evol. Microbiol. 67, 311–318.
- Corteselli, E.M., Aitken, M.D., Singleton, D.R., 2017b. Description of *Immundisolibacter cernigliae* gen. nov., sp. nov., a high-molecular-weight polycyclic aromatic hydrocarbon-degrading bacterium within the class Gammaproteobacteria, and proposal of *Immundisolibacterales* ord. nov. and *Immundisolibacteraceae* fam. nov. Int. J. Syst. Evol. Microbiol. 67, 925–931.
- Cwiertny, D.M., Snyder, S.A., Schlenk, D., Kolodziej, E.P., 2014. Environmental Designer Drugs: When Transformation May Not Eliminate Risk. Environ. Sci. Technol. 48, 11737–11745.
- Davies, J.I., Evans, W.C., 1964. Oxidative metabolism of naphthalene by soil pseudomonads. The ring-fission mechanism. Biochem. J. 91, 251–261.

- de Voogt, P., Laane, R.W.P.M., 2009. Assessment of azaarenes and azaarones (oxidized azaarene derivatives) in the Dutch coastal zone of the North Sea. Chemosphere 76, 1067–1074.
- Dettmer, K., Aronov, P.A., Hammock, B.D., 2007. Mass spectrometry-based metabolomics. Mass Spectrom. Rev. 26, 51–78.
- Du, B., M. Lofton, J., T. Peter, K., D. Gipe, A., Andrew James, C., K. McIntyre, J., L. Scholz, N., E. Baker, J., P. Kolodziej, E., 2017. Development of suspect and non-target screening methods for detection of organic contaminants in highway runoff and fish tissue with high-resolution time-of-flight mass spectrometry. Environ. Sci. Process. Impacts 19, 1185–1196.
- Durant, J.L., Lafleur, A.L., Plummer, E.F., Taghizadeh, K., Busby, W.F., Thilly, W.G., 1998. Human Lymphoblast Mutagens in Urban Airborne Particles. Environ. Sci. Technol. 32, 1894–1906.
- Evans, W.C., Fernley, H.N., Griffiths, E., 1965. Oxidative metabolism of phenanthrene and anthracene by soil pseudomonads. The ring-fission mechanism. Biochem. J. 95, 819–831.
- Fan, T.W.-M., Lorkiewicz, P.K., Sellers, K., Moseley, H.N.B., Higashi, R.M., Lane, A.N., 2012. Stable isotope-resolved metabolomics and applications for drug development. Pharmacol. Ther. 133, 366–391.
- Fang, M., Ivanisevic, J., Benton, H.P., Johnson, C.H., Patti, G.J., Hoang, L.T., Uritboonthai, W., Kurczy, M.E., Siuzdak, G., 2015a. Thermal Degradation of Small Molecules: A Global Metabolomic Investigation. Anal. Chem. 87, 10935–10941.
- Fang, M., Webster, T.F., Stapleton, H.M., 2015b. Effect-Directed Analysis of Human Peroxisome Proliferator-Activated Nuclear Receptors (PPAR γ 1) Ligands in Indoor Dust. Environ. Sci. Technol. 49, 10065–10073.
- Fenner, K., Canonica, S., Wackett, L.P., Elsner, M., 2013. Evaluating Pesticide Degradation in the Environment: Blind Spots and Emerging Opportunities. Science 341, 752–758.
- Fernie, A.R., Trethewey, R.N., Krotzky, A.J., Willmitzer, L., 2004. Metabolite profiling: from diagnostics to systems biology. Nat. Rev. Mol. Cell Biol. 5, 763–769.
- Fetzner, S., 1998. Bacterial degradation of pyridine, indole, quinoline, and their derivatives under different redox conditions. Appl. Microbiol. Biotechnol. 49, 237–250.
- Fiehn, O., 2002. Metabolomics – the link between genotypes and phenotypes. Plant Mol. Biol. 48, 155–171.
- Fischer, A., Manefield, M., Bombach, P., 2016. Application of stable isotope tools for evaluating natural and stimulated biodegradation of organic pollutants in field studies. Curr. Opin. Biotechnol., Analytical biotechnology 41, 99–107.

- Flowers-Geary, L., Bleczinski, W., Harvey, R.G., Penning, T.M., 1996. Cytotoxicity and mutagenicity of polycyclic aromatic hydrocarbon o-quinones produced by dihydrodiol dehydrogenase. Chem. Biol. Interact. 99, 55–72.
- Fox, M.A., Olive, S., 1979. Photooxidation of Anthracene on Atmospheric Particulate Matter. Science 205, 582–583.
- Freund, D.M., Hegeman, A.D., 2017. Recent advances in stable isotope-enabled mass spectrometry-based plant metabolomics. Curr. Opin. Biotechnol., Analytical biotechnology 43, 41–48.
- Gallampos, C.M.J., Schymanski, E.L., Krauss, M., Ulrich, N., Bataineh, M., Brack, W., 2015. Multicriteria Approach To Select Polyaromatic River Mutagen Candidates. Environ. Sci. Technol. 49, 2959–2968.
- Giavalisco, P., Hummel, J., Lisec, J., Inostroza, A.C., Catchpole, G., Willmitzer, L., 2008. High-resolution direct infusion-based mass spectrometry in combination with whole ¹³C metabolome isotope labeling allows unambiguous assignment of chemical sum formulas. Anal. Chem. 80, 9417–9425.
- Gibson, D.T., Mahadevan, V., Jerina, D.M., Yogi, H., Yeh, H.J., 1975. Oxidation of the carcinogens benzo [a] pyrene and benzo [a] anthracene to dihydrodiols by a bacterium. Science 189, 295–297.
- Gillespie, A.M., Wang, S., McDonald, T., Garcia, S.G., Cosgriff, D., He, L.-Y., Huebner, H., Donnelly, K.C., 2007. Genotoxicity Assessment of Wood-Preserving Waste–Contaminated Soil Undergoing Bioremediation. Bioremediation J. 11, 171–182.
- Gillespie, I.M.M., Philp, J.C., 2013. Bioremediation, an environmental remediation technology for the bioeconomy. Trends Biotechnol. 31, 329–332.
- Gillis, R.G., Porter, Q.N., 1989. 5-Methoxyphenanthrene-4-carboxylic Acid. Aust. J. Chem. 42, 1007–1010.
- Grifoll, M., Selifonov, S.A., Gatlin, C.V., Chapman, P.J., 1995. Actions of a versatile fluorene-degrading bacterial isolate on polycyclic aromatic compounds. Appl. Environ. Microbiol. 61, 3711–3723.
- Grosse, S., Letzel, T., 2007. Liquid chromatography/atmospheric pressure ionization mass spectrometry with post-column liquid mixing for the efficient determination of partially oxidized polycyclic aromatic hydrocarbons. J. Chromatogr. A 1139, 75–83.
- Guo, K., Li, L., 2009. Differential ¹²C-/¹³C-isotope dansylation labeling and fast liquid chromatography/mass spectrometry for absolute and relative quantification of the metabolome. Anal. Chem. 81, 3919–3932.

- Haeseler, F., Blanchet, D., Druelle, V., Werner, P., Vandecasteele, J.-P., 1999. Analytical Characterization of Contaminated Soils from Former Manufactured Gas Plants. Environ. Sci. Technol. 33, 825–830.
- Hegeman, A.D., Schulte, C.F., Cui, Q., Lewis, I.A., Huttlin, E.L., Eghbalnia, H., Harms, A.C., Ulrich, E.L., Markley, J.L., Sussman, M.R., 2007. Stable isotope assisted assignment of elemental compositions for metabolomics. Anal. Chem. 79, 6912–6921.
- Heitkamp, M.A., Cerniglia, C.E., 1989. Polycyclic aromatic hydrocarbon degradation by a *Mycobacterium* sp. in microcosms containing sediment and water from a pristine ecosystem. Appl. Environ. Microbiol. 55, 1968–1973.
- Heitkamp, M.A., Freeman, J.P., Miller, D.W., Cerniglia, C.E., 1988. Pyrene degradation by a *Mycobacterium* sp.: identification of ring oxidation and ring fission products. Appl. Environ. Microbiol. 54, 2556–2565.
- Hennessee, C.T., Li, Q.X., 2016. Effects of polycyclic aromatic hydrocarbon mixtures on degradation, gene expression, and metabolite production in four *mycobacterium* species. Appl. Environ. Microbiol. 82, 3357–3369.
- Henry, T.R., Wallace, K.B., 1996. Differential mechanisms of cell killing by redox cycling and arylating quinones. Arch. Toxicol. 70, 482–489.
- Herwijnen, R. van, Springael, D., Slot, P., Govers, H.A.J., Parsons, J.R., 2003. Degradation of Anthracene by *Mycobacterium* sp. Strain LB501T Proceeds via a Novel Pathway, through o-Phthalic Acid. Appl. Environ. Microbiol. 69, 186–190.
- Hiller, K., Metallo, C.M., Kelleher, J.K., Stephanopoulos, G., 2010. Nontargeted elucidation of metabolic pathways using stable-isotope tracers and mass spectrometry. Anal. Chem. 82, 6621–6628.
- Hollender, J., Schymanski, E.L., Singer, H.P., Ferguson, P.L., 2017. Nontarget Screening with High Resolution Mass Spectrometry in the Environment: Ready to Go? Environ. Sci. Technol. 51, 11505–11512.
- Hollender, J., Singer, H., Hernando, D., Kosjek, T., Heath, E., 2010. The Challenge of the Identification and Quantification of Transformation Products in the Aquatic Environment Using High Resolution Mass Spectrometry, in: Fatta-Kassinos, D., Bester, K., Kümmerer, K. (Eds.), *Xenobiotics in the Urban Water Cycle, Environmental Pollution*. Springer Netherlands, pp. 195–211.
- Hu, J., Adrion, A.C., Nakamura, J., Shea, D., Aitken, M.D., 2014. Bioavailability of (Geno)toxic Contaminants in Polycyclic Aromatic Hydrocarbon–Contaminated Soil Before and After Biological Treatment. Environ. Eng. Sci. 31, 176–182.
- Hu, J., Nakamura, J., Richardson, S.D., Aitken, M.D., 2012. Evaluating the effects of bioremediation on genotoxicity of polycyclic aromatic hydrocarbon-contaminated soil

- using genetically engineered, higher eukaryotic cell lines. Environ. Sci. Technol. 46, 4607–4613.
- Huang, X., Chen, Y.-J., Cho, K., Nikolskiy, I., Crawford, P.A., Patti, G.J., 2014. X13CMS: Global tracking of isotopic labels in untargeted metabolomics. Anal. Chem. 86, 1632–1639.
- Huba, A.K., Huba, K., Gardinali, P.R., 2016. Understanding the atmospheric pressure ionization of petroleum components: The effects of size, structure, and presence of heteroatoms. Sci. Total Environ.
- Hughes, T.J., Claxton, L.D., Brooks, L., Warren, S., Brenner, R., Kremer, F., 1998. Genotoxicity of bioremediated soils from the Reilly Tar site, St. Louis Park, Minnesota. Environ. Health Perspect. 106, 1427–1433.
- Hughey, C.A., Hendrickson, C.L., Rodgers, R.P., Marshall, A.G., Qian, K., 2001. Kendrick Mass Defect Spectrum: A Compact Visual Analysis for Ultrahigh-Resolution Broadband Mass Spectra. Anal. Chem. 73, 4676–4681.
- Huntscha, S., Hofstetter, T.B., Schymanski, E.L., Spahr, S., Hollender, J., 2014. Biotransformation of Benzotriazoles: Insights from Transformation Product Identification and Compound-Specific Isotope Analysis. Environ. Sci. Technol. 48, 4435–4443.
- Jakoncic, J., Jouanneau, Y., Meyer, C., Stojanoff, V., 2007. The catalytic pocket of the ring-hydroxylating dioxygenase from *Sphingomonas* CHY-1. Biochem. Biophys. Res. Commun. 352, 861–866.
- Jarabak, R., Harvey, R.G., Jarabak, J., 1998. Redox cycling of polycyclic aromatic hydrocarbon o-quinones: metal ion-catalyzed oxidation of catechols bypasses inhibition by superoxide dismutase. Chem. Biol. Interact. 115, 201–213.
- Jerina, D.M., Van Bladeren, P.J., Yagi, H., Gibson, D.T., Mahadevan, V., Neese, A.S., Koreeda, M., Sharma, N.D., Boyd, D.R., 1984. Synthesis and absolute configuration of the bacterial *cis*-1,2-, *cis*-8,9-, and *cis*-10,11-dihydrodiol metabolites of benz[*a*]anthracene formed by strain of *Beijerinckia*. J. Org. Chem. 49, 3621–3628.
- Ji, K., Kogame, T., Choi, K., Wang, X., Lee, J., Taniguchi, Y., Takeda, S., 2009. A novel approach using DNA-repair-deficient chicken DT40 cell lines for screening and characterizing the genotoxicity of environmental contaminants. Environ. Health Perspect. 117, 1737–1744.
- Jiang, H., Shen, Y.-M., Quinn, A.M., Penning, T.M., 2005. Competing Roles of Cytochrome P450 1A1/1B1 and Aldo-Keto Reductase 1A1 in the Metabolic Activation of (±)-7,8-Dihydroxy-7,8-dihydro-benzo[*a*]pyrene in Human Bronchoalveolar Cell Extracts. Chem. Res. Toxicol. 18, 365–374.
- Johnsen, A.R., Wick, L.Y., Harms, H., 2005. Principles of microbial PAH-degradation in soil. Environ. Pollut. 133, 71–84.

- Jones, M.D., Crandell, D.W., Singleton, D.R., Aitken, M.D., 2011. Stable-isotope probing of the polycyclic aromatic hydrocarbon-degrading bacterial guild in a contaminated soil. Environ. Microbiol. 13, 2623–2632.
- Jones, M.D., Singleton, D.R., Carstensen, D.P., Powell, S.N., Swanson, J.S., Pfaender, F.K., Aitken, M.D., 2008. Effect of incubation conditions on the enrichment of pyrene-degrading bacteria identified by stable-isotope probing in an aged, PAH-contaminated soil. Microb. Ecol. 56, 341–349.
- Jonsson, P., Johansson, A.I., Gullberg, J., Trygg, J., A, J., Grung, B., Marklund, S., Sjöström, M., Antti, H., Moritz, T., 2005. High-Throughput Data Analysis for Detecting and Identifying Differences between Samples in GC/MS-Based Metabolomic Analyses. Anal. Chem. 77, 5635–5642.
- Jouanneau, Y., Meyer, C., Duraffourg, N., 2016. Dihydroxylation of four- and five-ring aromatic hydrocarbons by the naphthalene dioxygenase from *Sphingomonas* CHY-1. Appl. Microbiol. Biotechnol. 100, 1253–1263.
- Jouanneau, Y., Micoud, J., Meyer, C., 2007. Purification and characterization of a three-component salicylate 1-hydroxylase from *Sphingomonas* sp. strain CHY-1. Appl. Environ. Microbiol. 73, 7515–7521.
- Kanaly, R.A., Bartha, R., 1999. Cometabolic mineralization of benzo[a]pyrene caused by hydrocarbon additions to soil. Environ. Toxicol. Chem. 18, 2186–2190.
- Kanaly, R.A., Harayama, S., 2000. Biodegradation of High-Molecular-Weight Polycyclic Aromatic Hydrocarbons by Bacteria. J. Bacteriol. 182, 2059–2067.
- Kazunga, C., Aitken, M.D., 2000. Products from the incomplete metabolism of pyrene by polycyclic aromatic hydrocarbon-degrading bacteria. Appl. Environ. Microbiol. 66, 1917–1922.
- Kazunga, C., Aitken, M.D., Gold, A., Sangaiah, R., 2001. Fluoranthene-2,3- and -1,5-diones are novel products from the bacterial transformation of fluoranthene. Environ. Sci. Technol. 35, 917–922.
- Keith, L.H., 2015. The Source of U.S. EPA's Sixteen PAH Priority Pollutants. Polycycl. Aromat. Compd. 35, 147–160.
- Kellogg, J.J., Todd, D.A., Egan, J.M., Raja, H.A., Oberlies, N.H., Kvalheim, O.M., Cech, N.B., 2016. Biochemometrics for Natural Products Research: Comparison of Data Analysis Approaches and Application to Identification of Bioactive Compounds. J. Nat. Prod. 79, 376–386.
- Kern, S., Fenner, K., Singer, H.P., Schwarzenbach, R.P., Hollender, J., 2009. Identification of Transformation Products of Organic Contaminants in Natural Waters by Computer-Aided Prediction and High-Resolution Mass Spectrometry. Environ. Sci. Technol. 43, 7039–7046.

- Kim, S.-J., Kweon, O., Jones, R.C., Freeman, J.P., Edmondson, R.D., Cerniglia, C.E., 2007. Complete and Integrated Pyrene Degradation Pathway in *Mycobacterium vanbaalenii* PYR-1 Based on Systems Biology. J. Bacteriol. 189, 464–472.
- Kim, Y.-H., Engesser, K.-H., Cerniglia, C.E., 2003. Two polycyclic aromatic hydrocarbon o-quinone reductases from a pyrene-degrading *Mycobacterium*. Arch. Biochem. Biophys. 416, 209–217.
- Kim, Y.-H., Freeman, J.P., Moody, J.D., Engesser, K.-H., Cerniglia, C.E., 2005. Effects of pH on the degradation of phenanthrene and pyrene by *Mycobacterium vanbaalenii* PYR-1. Appl. Microbiol. Biotechnol. 67, 275–285.
- Kluger, B., Bueschl, C., Neumann, N., Stückler, R., Doppler, M., Chassy, A.W., Waterhouse, A.L., Rechthaler, J., Kamplleitner, N., Thallinger, G.G., Adam, G., Krska, R., Schuhmacher, R., 2014. Untargeted profiling of tracer-derived metabolites using stable isotopic labeling and fast polarity-switching LC–ESI-HRMS. Anal. Chem. 86, 11533–11537.
- Knights, C.D., Peters, C.A., 2006. Multisubstrate biodegradation kinetics for binary and complex mixtures of polycyclic aromatic hydrocarbons. Environ. Toxicol. Chem. 25, 1746–1756.
- Kochany, J., Maguire, R.J., 1994. Abiotic transformations of polynuclear aromatic hydrocarbons and polynuclear aromatic nitrogen heterocycles in aquatic environments. Sci. Total Environ. 144, 17–31.
- Kolkman, A., Martijn, B.J., Vughs, D., Baken, K.A., van Wezel, A.P., 2015. Tracing nitrogenous disinfection byproducts after medium pressure UV water treatment by stable isotope labeling and high resolution mass spectrometry. Environ. Sci. Technol. 49, 4458–4465.
- Koolen, H.H.F., Swarthout, R.F., Nelson, R.K., Chen, H., Krajewski, L.C., Aeppli, C., McKenna, A.M., Rodgers, R.P., Reddy, C.M., 2015. Unprecedented Insights into the Chemical Complexity of Coal Tar from Comprehensive Two-Dimensional Gas Chromatography Mass Spectrometry and Direct Infusion Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. Energy Fuels 29, 641–648.
- Kraak, M.H.S., Ainscough, C., Fernández, A., van Vlaardingen, P.L.A., Voogt, P. de, Admiraal, W.A., 1997. Short-term and chronic exposure of the zebra mussel (*Dreissena polymorpha*) to acridine: effects and metabolism. Aquat. Toxicol. 37, 9–20.
- Krauss, M., Singer, H., Hollender, J., 2010. LC–high resolution MS in environmental analysis: from target screening to the identification of unknowns. Anal. Bioanal. Chem. 397, 943–951.
- Kunihiro, M., Ozeki, Y., Nogi, Y., Hamamura, N., Kanaly, R.A., 2013. Benz[*a*]anthracene biotransformation and production of ring fission products by *sphingobium* sp. strain KK22. Appl. Environ. Microbiol. 79, 4410–4420.

- Kuppusamy, S., Thavamani, P., Venkateswarlu, K., Lee, Y.B., Naidu, R., Megharaj, M., 2017. Remediation approaches for polycyclic aromatic hydrocarbons (PAHs) contaminated soils: Technological constraints, emerging trends and future directions. Chemosphere 168, 944–968.
- Kweon, O., Kim, S.-J., Holland, R.D., Chen, H., Kim, D.-W., Gao, Y., Yu, L.-R., Baek, S., Baek, D.-H., Ahn, H., Cerniglia, C.E., 2011. Polycyclic aromatic hydrocarbon metabolic network in *Mycobacterium vanbaalenii* PYR-1. J. Bacteriol. 193, 4326–4337.
- Kweon, O., Kim, S.-J., Jones, R.C., Freeman, J.P., Adjei, M.D., Edmondson, R.D., Cerniglia, C.E., 2007. A polyomic approach to elucidate the fluoranthene-degradative pathway in *Mycobacterium vanbaalenii* PYR-1. J. Bacteriol. 189, 4635–4647.
- Kweon, O., Kim, S.-J., Kim, D.-W., Kim, J.M., Kim, H., Ahn, Y., Sutherland, J.B., Cerniglia, C.E., 2014. Pleiotropic and Epistatic Behavior of a Ring-Hydroxylating Oxygenase System in the Polycyclic Aromatic Hydrocarbon Metabolic Network from *Mycobacterium vanbaalenii* PYR-1. J. Bacteriol. 196, 3503–3515.
- Lamberts, R.F., Christensen, J.H., Mayer, P., Andersen, O., Johnsen, A.R., 2008. Isomer-Specific Biodegradation of Methylphenanthrenes by Soil Bacteria. Environ. Sci. Technol. 42, 4790–4796.
- Lampi, M.A., Gurska, J., McDonald, K.I.C., Xie, F., Huang, X.-D., Dixon, D.G., Greenberg, B.M., 2006. Photoinduced toxicity of polycyclic aromatic hydrocarbons to *Daphnia magna*: Ultraviolet-mediated effects and the toxicity of polycyclic aromatic hydrocarbon photoproducts. Environ. Toxicol. Chem. 25, 1079–1087.
- Latino, D.A.R.S., Wicker, J., Gütlein, M., Schmid, E., Kramer, S., Fenner, K., 2017. Eawag-Soil in enviPath: a new resource for exploring regulatory pesticide soil biodegradation pathways and half-life data. Environ. Sci. Process. Impacts 19, 449–464.
- Leeming, M.G., Isaac, A.P., Pope, B.J., Cranswick, N., Wright, C.E., Ziogas, J., O’Hair, R.A.J., Donald, W.A., 2015. High-resolution twin-ion metabolite extraction (HiTIME) mass spectrometry: Nontargeted detection of unknown drug metabolites by isotope labeling, liquid chromatography mass spectrometry, and automated high-performance computing. Anal. Chem. 87, 4104–4109.
- LeFevre, G.H., Müller, C.E., Li, R.J., Luthy, R.G., Sattely, E.S., 2015. Rapid Phytotransformation of Benzotriazole Generates Synthetic Tryptophan and Auxin Analogs in Arabidopsis. Environ. Sci. Technol. 49, 10959–10968.
- Legler, J., van Velzen, M., Cenijn, P.H., Houtman, C.J., Lamoree, M.H., Wegener, J.W., 2011. Effect-Directed Analysis of Municipal Landfill Soil Reveals Novel Developmental Toxicants in the Zebrafish *Danio rerio*. Environ. Sci. Technol. 45, 8552–8558.
- Lemieux, C.L., Lambert, I.B., Lundstedt, S., Tysklind, M., White, P.A., 2008. Mutagenic hazards of complex polycyclic aromatic hydrocarbon mixtures in contaminated soil. Environ. Toxicol. Chem. 27, 978–990.

- Lemieux, C.L., Long, A.S., Lambert, I.B., Lundstedt, S., Tysklind, M., White, P.A., 2015. In vitro mammalian mutagenicity of complex polycyclic aromatic hydrocarbon mixtures in contaminated soils. Environ. Sci. Technol. 49, 1787–1796.
- Lemieux, C.L., Lynes, K.D., White, P.A., Lundstedt, S., Öberg, L., Lambert, I.B., 2009. Mutagenicity of an aged gasworks soil during bioslurry treatment. Environ. Mol. Mutagen. 50, 404–412.
- Lintelmann, J., França, M.H., Hübner, E., Matuschek, G., 2010. A liquid chromatography–atmospheric pressure photoionization tandem mass spectrometric method for the determination of azaarenes in atmospheric particulate matter. J. Chromatogr. A 1217, 1636–1646.
- Liu, J., Wang, Y., Jiang, B., Wang, L., Chen, J., Guo, H., Ji, R., 2013. Degradation, metabolism, and bound-residue formation and release of tetrabromobisphenol A in soil during sequential anoxic–oxic incubation. Environ. Sci. Technol. 47, 8348–8354.
- Lu, Z., Gan, J., 2014. Isomer-Specific Biodegradation of Nonylphenol in River Sediments and Structure-Biodegradability Relationship. Environ. Sci. Technol. 48, 1008–1014.
- Luan, T., Fang, S., Zhong, Y., Lin, L., Chan, S.M.N., Lan, C., Tam, N.F.Y., 2007. Determination of hydroxy metabolites of polycyclic aromatic hydrocarbons by fully automated solid-phase microextraction derivatization and gas chromatography–mass spectrometry. J. Chromatogr. A 1173, 37–43.
- Lundstedt, S., Bandowe, B.A.M., Wilcke, W., Boll, E., Christensen, J.H., Vila, J., Grifoll, M., Faure, P., Biache, C., Lorgeoux, C., Larsson, M., Frech Irgum, K., Ivarsson, P., Ricci, M., 2014. First intercomparison study on the analysis of oxygenated polycyclic aromatic hydrocarbons (oxy-PAHs) and nitrogen heterocyclic polycyclic aromatic compounds (N-PACs) in contaminated soil. TrAC Trends Anal. Chem. 57, 83–92.
- Lundstedt, S., Haglund, P., Öberg, L., 2003. Degradation and formation of polycyclic aromatic compounds during bioslurry treatment of an aged gasworks soil. Environ. Toxicol. Chem. 22, 1413–1420.
- Lundstedt, S., White, P.A., Lemieux, C.L., Lynes, K.D., Lambert, I.B., Öberg, L., Haglund, P., Tysklind, M., 2007. Sources, Fate, and Toxic Hazards of Oxygenated Polycyclic Aromatic Hydrocarbons (PAHs) at PAH- contaminated Sites. AMBIO J. Hum. Environ. 36, 475–485.
- Luo, S., Chen, B., Lin, L., Wang, X., Tam, N.F.-Y., Luan, T., 2014. Pyrene degradation accelerated by constructed consortium of bacterium and microalga: effects of degradation products on the microalgal growth. Environ. Sci. Technol. 48, 13917–13924.
- Machala, M., Ciganek, M., Bláha, L., Minksová, K., Vondráček, J., 2001. Aryl hydrocarbon receptor-mediated and estrogenic activities of oxygenated polycyclic aromatic hydrocarbons and azaarenes originally identified in extracts of river sediments. Environ. Toxicol. Chem. 20, 2736–2743.

- Mackay, D., Shiu, W.-Y., Ma, K.-C., Lee, S.C., 2006. Handbook of Physical-Chemical Properties and Environmental Fate for Organic Chemicals, Second Edition. CRC Press.
- Mahaffey, W.R., Gibson, D.T., Cerniglia, C.E., 1988. Bacterial oxidation of chemical carcinogens: formation of polycyclic aromatic acids from benz[*a*]anthracene. Appl. Environ. Microbiol. 54, 2415–2423.
- Mahler, B.J., Ingersoll, C.G., Van Metre, P.C., Kunz, J.L., Little, E.E., 2015. Acute Toxicity of Runoff from Sealcoated Pavement to *Ceriodaphnia dubia* and *Pimephales promelas*. Environ. Sci. Technol. 49, 5060–5069.
- Manzano, C.A., Marvin, C., Muir, D., Harner, T., Martin, J., Zhang, Y., 2017. Heterocyclic Aromatics in Petroleum Coke, Snow, Lake Sediments, and Air Samples from the Athabasca Oil Sands Region. Environ. Sci. Technol.
- Markley, J.L., Brüschweiler, R., Edison, A.S., Eghbalnia, H.R., Powers, R., Raftery, D., Wishart, D.S., 2017. The future of NMR-based metabolomics. Curr. Opin. Biotechnol., Analytical biotechnology 43, 34–40.
- Marshall, A.G., Rodgers, R.P., 2008. Petroleomics: Chemistry of the underworld. Proc. Natl. Acad. Sci. 105, 18090–18095.
- McConkey, B.J., Duxbury, C.L., Dixon, D.G., Greenberg, B.M., 1997. Toxicity of a pah photooxidation product to the bacteria *Photobacterium phosphoreum* and the duckweed *Lemna gibba*: Effects of phenanthrene and its primary photoproduct, phenanthrenequinone. Environ. Toxicol. Chem. 16, 892–899.
- McCoull, K.D., Rindgen, D., Blair, I.A., Penning, T.M., 1999. Synthesis and Characterization of Polycyclic Aromatic Hydrocarbon o-Quinone Depurinating N7-Guanine Adducts. Chem. Res. Toxicol. 12, 237–246.
- Meador, J.P., Stein, J.E., Reichert, W.L., Varanasi, U., 1995. Bioaccumulation of Polycyclic Aromatic Hydrocarbons by Marine Organisms, in: Ware, G.W. (Ed.), *Reviews of Environmental Contamination and Toxicology, Reviews of Environmental Contamination and Toxicology*. Springer New York, pp. 79–165.
- Meehan, T., Straub, K., Calvin, M., 1977. Benzo[*a*]pyrene diol epoxide covalently binds to deoxyguanosine and deoxyadenosine in DNA. Nature 269, 725–727.
- Men, Y., Han, P., Helbling, D.E., Jehmlich, N., Herbold, C., Gulde, R., Onnis-Hayden, A., Gu, A.Z., Johnson, D.R., Wagner, M., Fenner, K., 2016. Biotransformation of Two Pharmaceuticals by the Ammonia-Oxidizing Archaeon *Nitrososphaera gargensis*. Environ. Sci. Technol.
- Meyer, S., Cartellieri, S., Steinhart, H., 1999. Simultaneous Determination of PAHs, Hetero-PAHs (N, S, O), and Their Degradation Products in Creosote-Contaminated Soils. Method Development, Validation, and Application to Hazardous Waste Sites. Anal. Chem. 71, 4023–4029.

- Meyer, S., Steinhart, H., 2001. Fate of PAHs and Hetero-PAHs during Biodegradation in a Model Soil/Compost-System: Formation of Extractable Metabolites. Water, Air, Soil Pollut. 132, 215–231.
- Mihelcic, J.R., Luthy, R.G., 1988. Degradation of polycyclic aromatic hydrocarbon compounds under various redox conditions in soil-water systems. Appl. Environ. Microbiol. 54, 1182–1187.
- Mohler, R.E., O'Reilly, K.T., Zemo, D.A., Tiwary, A.K., Magaw, R.I., Synowiec, K.A., 2013. Non-Targeted Analysis of Petroleum Metabolites in Groundwater Using GC×GC–TOFMS. Environ. Sci. Technol. 47, 10471–10476.
- Montes, R., Aguirre, J., Vidal, X., Rodil, R., Cela, R., Quintana, J.B., 2017. Screening for Polar Chemicals in Water by Trifunctional Mixed-Mode Liquid Chromatography–High Resolution Mass Spectrometry. Environ. Sci. Technol. 51, 6250–6259.
- Moody, J.D., Freeman, J.P., Doerge, D.R., Cerniglia, C.E., 2001. Degradation of Phenanthrene and Anthracene by Cell Suspensions of Mycobacterium sp. Strain PYR-1. Appl. Environ. Microbiol. 67, 1476–1483.
- Morasch, B., Hunkeler, D., Zopfi, J., Temime, B., Höhener, P., 2011. Intrinsic biodegradation potential of aromatic hydrocarbons in an alluvial aquifer – potentials and limits of signature metabolite analysis and two stable isotope-based techniques. Water Res. 45, 4459–4469.
- Moschet, C., Lew, B.M., Hasenbein, S., Anumol, T., Young, T.M., 2017. LC- and GC-QTOF-MS as Complementary Tools for a Comprehensive Micropollutant Analysis in Aquatic Systems. Environ. Sci. Technol. 51, 1553–1561.
- Mueller, J.G., Chapman, P.J., Pritchard, P.H., 1989. Creosote-contaminated sites. Their potential for bioremediation. Environ. Sci. Technol. 23, 1197–1201.
- Mueller, J.G., Middaugh, D.P., Lantz, S.E., Chapman, P.J., 1991. Biodegradation of creosote and pentachlorophenol in contaminated groundwater: chemical and biological assessment. Appl. Environ. Microbiol. 57, 1277–1285.
- Müller, H., Andersson, J.T., Schrader, W., 2005. Characterization of High-Molecular-Weight Sulfur-Containing Aromatics in Vacuum Residues Using Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. Anal. Chem. 77, 2536–2543.
- Myers, A.L., Watson-Leung, T., Jobst, K.J., Shen, L., Besevic, S., Organtini, K., Dorman, F.L., Mabury, S.A., Reiner, E.J., 2014. Complementary Nontargeted and Targeted Mass Spectrometry Techniques to Determine Bioaccumulation of Halogenated Contaminants in Freshwater Species. Environ. Sci. Technol. 48, 13844–13854.
- Nakata, H., Sakai, Y., Miyawaki, T., Takemura, A., 2003. Bioaccumulation and Toxic Potencies of Polychlorinated Biphenyls and Polycyclic Aromatic Hydrocarbons in Tidal Flat and Coastal Ecosystems of the Ariake Sea, Japan. Environ. Sci. Technol. 37, 3513–3521.

- Neilson, A.N., 1995. PAHs and Related Compounds: Chemistry. Springer.
- Neumann, N.K.N., Lehner, S.M., Kluger, B., Bueschl, C., Sedelmaier, K., Lemmens, M., Krska, R., Schuhmacher, R., 2014. Automated LC-HRMS(/MS) approach for the annotation of fragment ions derived from stable isotope labeling-assisted untargeted metabolomics. Anal. Chem. 86, 7320–7327.
- Newton, S., McMahan, R., Stoeckel, J.A., Chislock, M., Lindstrom, A., Strynar, M., 2017. Novel Polyfluorinated Compounds Identified Using High Resolution Mass Spectrometry Downstream of Manufacturing Facilities near Decatur, Alabama. Environ. Sci. Technol. 51, 1544–1552.
- Nikolaou, K., Masclet, P., Mouvier, G., 1984. Sources and chemical reactivity of polynuclear aromatic hydrocarbons in the atmosphere — A critical review. Sci. Total Environ. 32, 103–132.
- O’Connell, S.G., Haigh, T., Wilson, G., Anderson, K.A., 2013. An analytical investigation of 24 oxygenated-PAHs (OPAHs) using liquid and gas chromatography–mass spectrometry. Anal. Bioanal. Chem. 405, 8885–8896.
- Ouyang, X., Weiss, J.M., de Boer, J., Lamoree, M.H., Leonards, P.E.G., 2017. Non-target analysis of household dust and laundry dryer lint using comprehensive two-dimensional liquid chromatography coupled with time-of-flight mass spectrometry. Chemosphere 166, 431–437.
- Park, J., Ball, L.M., Richardson, S.D., Zhu, H.-B., Aitken, M.D., 2008. Oxidative mutagenicity of polar fractions from polycyclic aromatic hydrocarbon–contaminated soils. Environ. Toxicol. Chem. 27, 2207–2215.
- Park, J.-H., Gopishetty, S., Szewczuk, L.M., Penning, T.M., 2004. Formation of 8-oxo-7,8-dihydro-2’-deoxyguanosine (8-oxo-dGuo) by PAH *o*-quinones: Involvement of reactive oxygen species and copper (II)/copper (I) redox cycling. Cancer Res. 64, 739–740.
- Parshikov, I.A., Netrusov, A.I., Sutherland, J.B., 2012. Microbial transformation of azaarenes and potential uses in pharmaceutical synthesis. Appl. Microbiol. Biotechnol. 95, 871–889.
- Patti, G.J., Tautenhahn, R., Siuzdak, G., 2012. Meta-analysis of untargeted metabolomic data from multiple profiling experiments. Nat. Protoc. 7, 508–516.
- Pearlman, R.S., Yalkowsky, S.H., Banerjee, S., 1984. Water Solubilities of Polynuclear Aromatic and Heteroaromatic Compounds. J. Phys. Chem. Ref. Data 13, 555–562.
- Penning, T.M., Burczynski, M.E., Hung, C.-F., McCoull, K.D., Palackal, N.T., Tsuruda, L.S., 1999. Dihydrodiol Dehydrogenases and Polycyclic Aromatic Hydrocarbon Activation: Generation of Reactive and Redox Active *o*-Quinones. Chem. Res. Toxicol. 12, 1–18.

- Penning, T.M., Ohnishi, S.T., Ohnishi, T., Harvey, R.G., 1996. Generation of Reactive Oxygen Species during the Enzymatic Oxidation of Polycyclic Aromatic Hydrocarbon trans-Dihydrodiols Catalyzed by Dihydrodiol Dehydrogenase. Chem. Res. Toxicol. 9, 84–92.
- Pereira, W.E., Rostad, C.E., Updegraff, D.M., Bennett, J.L., 1988. Microbial Transformations of Azaarenes in Creosote-Contaminated Soil and Ground Water: Laboratory and Field Studies. Water Sci. Technol. 20, 17–23.
- Pereira, W.E., Rostad, C.E., Updegraff, D.M., Bennett, J.L., 1987. Fate and movement of azaarenes and their anaerobic biotransformation products in an aquifer contaminated by wood-treatment chemicals. Environ. Toxicol. Chem. 6, 163–176.
- Pfenning, C., Esch, H.L., Fliege, R., Lehmann, L., 2016. The mycotoxin patulin reacts with DNA bases with and without previous conjugation to GSH: implication for related α,β -unsaturated carbonyl compounds? Arch. Toxicol. 90, 433–448.
- Pitts, J.N., Lokensgard, D.M., Harger, W., Fisher, T.S., Mejia, V., Schuler, J.J., Scorziell, G.M., Katzenstein, Y.A., 1982. Mutagens in diesel exhaust particulate identification and direct activities of 6-nitrobenzo[a]pyrene, 9-nitroanthracene, 1-nitropyrene and 5H-phenanthro[4,5-bcd]pyran-5-one. Mutat. Res. Lett. 103, 241–249.
- Pothuluri, J.V., Evans, F.E., Heinze, T.M., Cerniglia, C.E., 1996. Formation of sulfate and glucoside conjugates of benzo[e]pyrene by *Cunninghamella elegans*. Appl. Microbiol. Biotechnol. 45, 677–683.
- Qu, S., Kolodziej, E.P., Long, S.A., Gloer, J.B., Patterson, E.V., Baltrusaitis, J., Jones, G.D., Benchetler, P.V., Cole, E.A., Kimbrough, K.C., Tarnoff, M.D., Cwiertny, D.M., 2013. Product-to-Parent Reversion of Trenbolone: Unrecognized Risks for Endocrine Disruption. Science 342, 347–351.
- Rehmann, K., Hertkorn, N., Kettrup, A.A., 2001. Fluoranthene metabolism in *Mycobacterium* sp. strain KR20: identity of pathway intermediates during degradation and growth. Microbiology 147, 2783–2794.
- Reineke, A.-K., Göen, T., Preiss, A., Hollender, J., 2007. Quinoline and Derivatives at a Tar Oil Contaminated Site: Hydroxylated Products as Indicator for Natural Attenuation? Environ. Sci. Technol. 41, 5314–5322.
- Richardson, S.D., 2010. Effects of in situ bioremediation strategies on the biodegradation and bioavailability of polycyclic aromatic hydrocarbons in weathered manufactured gas plant soil (Ph.D.). The University of North Carolina at Chapel Hill, United States -- North Carolina.
- Richardson, S.D., Lebron, B.L., Miller, C.T., Aitken, M.D., 2011. Recovery of Phenanthrene-Degrading Bacteria after Simulated in Situ Persulfate Oxidation in Contaminated Soil. Environ. Sci. Technol. 45, 719–725.

- Richnow, H.H., Annweiler, E., Koning, M., Lüth, J.-C., Stegmann, R., Garms, C., Francke, W., Michaelis, W., 2000. Tracing the transformation of labelled [1-¹³C]phenanthrene in a soil bioreactor. Environ. Pollut. 108, 91–101.
- Richnow, H.H., Eschenbach, A., Mahro, B., Seifert, R., Wehrung, P., Albrecht, P., Michaelis, W., 1998. The use of ¹³C-labelled polycyclic aromatic hydrocarbons for the analysis of their transformation in soil. Chemosphere 36, 2211–2224.
- Riva, M., Tomaz, S., Cui, T., Lin, Y.-H., Perraudin, E., Gold, A., Stone, E.A., Villenave, E., Surratt, J.D., 2015. Evidence for an unrecognized secondary anthropogenic source of organosulfates and sulfonates: gas-phase oxidation of polycyclic aromatic hydrocarbons in the presence of sulfate aerosol. Environ. Sci. Technol. 49, 6654–6664.
- Rotander, A., Kärrman, A., Toms, L.-M.L., Kay, M., Mueller, J.F., Gómez Ramos, M.J., 2015. Novel Fluorinated Surfactants Tentatively Identified in Firefighters Using Liquid Chromatography Quadrupole Time-of-Flight Tandem Mass Spectrometry and a Case-Control Approach. Environ. Sci. Technol. 49, 2434–2442.
- Ruttkies, C., Schymanski, E.L., Wolf, S., Hollender, J., Neumann, S., 2016. MetFrag relaunched: incorporating strategies beyond in silico fragmentation. J. Cheminformatics 8, 3.
- Schaum, J., Schuda, L., Wu, C., Sears, R., Ferrario, J., Andrews, K., 2003. A national survey of persistent, bioaccumulative, and toxic (PBT) pollutants in the United States milk supply. J. Expo. Sci. Environ. Epidemiol. 13, 177–186.
- Schmidt, S.N., Christensen, J.H., Johnsen, A.R., 2010. Fungal PAH-metabolites resist mineralization by soil microorganisms. Environ. Sci. Technol. 44, 1677–1682.
- Schollée, J.E., Schymanski, E.L., Avak, S.E., Loos, M., Hollender, J., 2015. Prioritizing Unknown Transformation Products from Biologically-Treated Wastewater using High-Resolution Mass Spectrometry, Multivariate Statistics, and Metabolic Logic. Anal. Chem.
- Schweigert, N., Zehnder, A.J.B., Eggen, R.I.L., 2001. Chemical properties of catechols and their molecular modes of toxic action in cells, from microorganisms to mammals. Environ. Microbiol. 3, 81–91.
- Schymanski, E.L., Bataineh, M., Goss, K.-U., Brack, W., 2009. Integrated analytical and computer tools for structure elucidation in effect-directed analysis. TrAC Trends Anal. Chem., Applying combinations of chemical analysis and biological effects to environmental and food samples - I 28, 550–561.
- Schymanski, E.L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H.P., Hollender, J., 2014a. Identifying small molecules via high resolution mass spectrometry: communicating confidence. Environ. Sci. Technol. 48, 2097–2098.
- Schymanski, E.L., Singer, H.P., Longrée, P., Loos, M., Ruff, M., Stravs, M.A., Ripollés Vidal, C., Hollender, J., 2014b. Strategies to Characterize Polar Organic Contamination in

- Wastewater: Exploring the Capability of High Resolution Mass Spectrometry. Environ. Sci. Technol. 48, 1811–1818.
- Schymanski, E.L., Singer, H.P., Slobodnik, J., Ipolyi, I.M., Oswald, P., Krauss, M., Schulze, T., Haglund, P., Letzel, T., Grosse, S., Thomaidis, N.S., Bletsou, A., Zwiener, C., Ibáñez, M., Portolés, T., Boer, R. de, Reid, M.J., Onghena, M., Kunkel, U., Schulz, W., Guillon, A., Noyon, N., Leroy, G., Bados, P., Bogialli, S., Stipaničev, D., Rostkowski, P., Hollender, J., 2015. Non-target screening with high-resolution mass spectrometry: critical review using a collaborative trial on water analysis. Anal. Bioanal. Chem. 407, 6237–6255.
- Seixas, G.M., Andon, B.M., Hollingshead, P.G., Thilly, W.G., 1982. The aza-arenes as mutagens for *Salmonella typhimurium*. Mutat. Res. Toxicol. 102, 201–212.
- Sharma, V., Collins, L.B., Chen, T.-H., Herr, N., Takeda, S., Sun, W., Swenberg, J.A., Nakamura, J., 2016. Oxidative stress at low levels can induce clustered DNA lesions leading to NHEJ mediated mutations. Oncotarget 7, 25377–25390.
- Shuttleworth, K.L., Cerniglia, E., 1995. Environmental aspects of PAH biodegradation. Appl. Biochem. Biotechnol. 54, 291–302.
- Sikkema, J., Bont, J.A. de, Poolman, B., 1995. Mechanisms of membrane toxicity of hydrocarbons. Microbiol. Rev. 59, 201–222.
- Singh, R.R., Lester, Y., Linden, K.G., Love, N.G., Atilla-Gokcumen, G.E., Aga, D.S., 2015. Application of Metabolite Profiling Tools and Time-of-Flight Mass Spectrometry in the Identification of Transformation Products of Iopromide and Iopamidol during Advanced Oxidation. Environ. Sci. Technol. 49, 2983–2990.
- Singleton, D.R., Dickey, A.N., Scholl, E.H., Wright, F.A., Aitken, M.D., 2016. Complete Genome Sequence of a Bacterium Representing a Deep Uncultivated Lineage within the Gammaproteobacteria Associated with the Degradation of Polycyclic Aromatic Hydrocarbons. Genome Announc. 4, e01086-16.
- Singleton, D.R., Hu, J., Aitken, M.D., 2012. Heterologous Expression of Polycyclic Aromatic Hydrocarbon Ring-Hydroxylating Dioxygenase Genes from a Novel Pyrene-Degrading Betaproteobacterium. Appl. Environ. Microbiol. 78, 3552–3559.
- Singleton, D.R., Richardson, S.D., Aitken, M.D., 2011. Pyrosequence analysis of bacterial communities in aerobic bioreactors treating polycyclic aromatic hydrocarbon-contaminated soil. Biodegradation 22, 1061–1073.
- Singleton, D.R., Sangaiah, R., Gold, A., Ball, L.M., Aitken, M.D., 2006. Identification and quantification of uncultivated Proteobacteria associated with pyrene degradation in a bioreactor treating PAH-contaminated soil. Environ. Microbiol. 8, 1736–1745.
- Sleno, L., 2012. The use of mass defect in modern mass spectrometry. J. Mass Spectrom. 47, 226–236.

- Stringfellow, W.T., Aitken, M.D., 1995. Competitive metabolism of naphthalene, methylnaphthalenes, and fluorene by phenanthrene-degrading pseudomonads. Appl. Environ. Microbiol. 61, 357–362.
- Strynar, M., Dagnino, S., McMahan, R., Liang, S., Lindstrom, A., Andersen, E., McMillan, L., Thurman, M., Ferrer, I., Ball, C., 2015. Identification of Novel Perfluoroalkyl Ether Carboxylic Acids (PFECAs) and Sulfonic Acids (PFESAs) in Natural Waters Using Accurate Mass Time-of-Flight Mass Spectrometry (TOFMS). Environ. Sci. Technol. 49, 11622–11630.
- Sun, K., Liang, S., Kang, F., Gao, Y., Huang, Q., 2016. Transformation of 17 β -estradiol in humic acid solution by ϵ -MnO₂ nanorods as probed by high-resolution mass spectrometry combined with ¹³C labeling. Environ. Pollut. 214, 211–218.
- Sutherland, J.B., Heinze, T.M., Pearce, M.G., Deck, J., Williams, A.J., Freeman, J.P., 2009. Biotransformation of acridine by *Mycobacterium vanbaalenii*. Environ. Toxicol. Chem. 28, 61–64.
- Švábenský, R., Oravec, M., Šimek, Z., 2009. Determination of polycyclic aromatic nitrogen heterocycles in soil using liquid chromatography/tandem mass spectrometry. Int. J. Environ. Anal. Chem. 89, 167–181.
- Sverdrup, L.E., Nielsen, T., Krogh, P.H., 2002. Soil Ecotoxicity of Polycyclic Aromatic Hydrocarbons in Relation to Soil Sorption, Lipophilicity, and Water Solubility. Environ. Sci. Technol. 36, 2429–2435.
- Szeliga, J., Dipple, A., 1998. DNA Adduct Formation by Polycyclic Aromatic Hydrocarbon Dihydrodiol Epoxides. Chem. Res. Toxicol. 11, 1–11.
- Tao, S., Lu, X., Levac, N., Bateman, A.P., Nguyen, T.B., Bones, D.L., Nizkorodov, S.A., Laskin, J., Laskin, A., Yang, X., 2014. Molecular Characterization of Organosulfates in Organic Aerosols from Shanghai and Los Angeles Urban Areas by Nanospray-Desorption Electrospray Ionization High-Resolution Mass Spectrometry. Environ. Sci. Technol. 48, 10993–11001.
- Tautenhahn, R., Patti, G.J., Rinehart, D., Siuzdak, G., 2012. XCMS Online: A web-based platform to process untargeted metabolomic data. Anal. Chem. 84, 5035–5039.
- Tejeda-Agredano, M.C., Gallego, S., Vila, J., Grifoll, M., Ortega-Calvo, J.J., Cantos, M., 2013. Influence of the sunflower rhizosphere on the biodegradation of PAHs in soil. Soil Biol. Biochem. 57, 830–840.
- Tian, Z., Gold, A., Nakamura, J., Zhang, Z., Vila, J., Singleton, D.R., Collins, L.B., Aitken, M.D., 2017a. Nontarget analysis reveals a bacterial metabolite of pyrene implicated in the genotoxicity of contaminated soil after bioremediation. Environ. Sci. Technol. 51, 7091–7100.

- Tian, Z., Vila, J., Wang, H., Bodnar, W., Aitken, M.D., 2017b. Diversity and Abundance of High-Molecular-Weight Azaarenes in PAH-Contaminated Environmental Samples. Environ. Sci. Technol. 51, 14047–14054.
- Trost, B.M., Kinson, P.L., 1975. Perturbed [12]annulenes. Derivatives of dibenzo[cd,gh]pentalene. J. Am. Chem. Soc. 97, 2438–2449.
- Tullis, D.L., Banerjee, S., 1984. Covalent hydration: A possible mechanism for aza-arene carcinogenesis. Cancer Lett. 23, 241–244.
- US EPA, O., 2015a. Search for Superfund Sites Where You Live [WWW Document]. US EPA.
- US EPA, O., 2015b. Superfund Remedy Report [WWW Document]. US EPA.
- Van Herwijnen, R., De Graaf, C., Govers, H.A.J., Parsons, J.R., 2004. Estimation of kinetic parameter for the biotransformation of three-ring azaarenes by the phenanthrene-degrading strain *Sphingomonas* SP LH128. Environ. Toxicol. Chem. 23, 331–338.
- van Herwijnen, R., Wattiau, P., Bastiaens, L., Daal, L., Jonker, L., Springael, D., Govers, H.A.J., Parsons, J.R., 2003. Elucidation of the metabolic pathway of fluorene and cometabolic pathways of phenanthrene, fluoranthene, anthracene and dibenzothiophene by *Sphingomonas* sp. LB126. Res. Microbiol. 154, 199–206.
- Vila, J., Grifoll, M., 2009. Actions of *Mycobacterium* sp. Strain AP1 on the Saturated- and Aromatic-Hydrocarbon Fractions of Fuel Oil in a Marine Medium. Appl. Environ. Microbiol. 75, 6232–6239.
- Vila, J., López, Z., Sabaté, J., Minguillón, C., Solanas, A.M., Grifoll, M., 2001. Identification of a Novel Metabolite in the Degradation of Pyrene by *Mycobacterium* sp. Strain AP1: Actions of the Isolate on Two- and Three-Ring Polycyclic Aromatic Hydrocarbons. Appl. Environ. Microbiol. 67, 5497–5505.
- Vila, J., Tauler, M., Grifoll, M., 2015. Bacterial PAH degradation in marine and terrestrial habitats. Curr. Opin. Biotechnol., Environmental biotechnology • Energy biotechnology 33, 95–102.
- Vughes, D., Baken, K.A., Kolkman, A., Martijn, A.J., Voogt, P. de, 2016. Application of effect-directed analysis to identify mutagenic nitrogenous disinfection by-products of advanced oxidation drinking water treatment. Environ. Sci. Pollut. Res. 1–14.
- Wammer, K.H., Peters, C.A., 2006. A molecular modeling analysis of polycyclic aromatic hydrocarbon biodegradation by naphthalene dioxygenase. Environ. Toxicol. Chem. 25, 912–920.
- Wang, Y., Xu, J., Shan, J., Ma, Y., Ji, R., 2017. Fate of phenanthrene and mineralization of its non-extractable residues in an oxic soil. Environ. Pollut. 224, 377–383.

- Wang, Z., Fingas, M., Blenkinsopp, S., Sergy, G., Landriault, M., Sigouin, L., Foght, J., Semple, K., Westlake, D.W.S., 1998. Comparison of oil composition changes due to biodegradation and physical weathering in different oils. J. Chromatogr. A 809, 89–107.
- Wei, C., Bandowe, B.A.M., Han, Y., Cao, J., Zhan, C., Wilcke, W., 2015. Polycyclic aromatic hydrocarbons (PAHs) and their derivatives (alkyl-PAHs, oxygenated-PAHs, nitrated-PAHs and azaarenes) in urban road dusts from Xi'an, Central China. Chemosphere 134, 512–520.
- Weiss, J.M., Hamers, T., Thomas, K.V., Linden, S. van der, Leonards, P.E.G., Lamoree, M.H., 2009. Masking effect of anti-androgens on androgenic activity in European river sediment unveiled by effect-directed analysis. Anal. Bioanal. Chem. 394, 1385–1397.
- Weiss, J.M., Simon, E., Stroomberg, G.J., Boer, R. de, Boer, J. de, Linden, S.C. van der, Leonards, P.E.G., Lamoree, M.H., 2011. Identification strategy for unknown pollutants using high-resolution mass spectrometry: Androgen-disrupting compounds identified through effect-directed analysis. Anal. Bioanal. Chem. 400, 3141–3149.
- Wicker, J., Fenner, K., Ellis, L., Wackett, L., Kramer, S., 2010. Predicting biodegradation products and pathways: a hybrid knowledge- and machine learning-based approach. Bioinformatics 26, 814–821.
- Wiegman, S., Termeer, J.A.G., Verheul, T., Kraak, M.H.S., de Voogt, P., Laane, R.W.P.M., Admiraal, W., 2002. UV Absorbance Dependent Toxicity of Acridine to the Marine Diatom *Phaeodactylum tricorutum*. Environ. Sci. Technol. 36, 908–913.
- Wiegman, S., van Vlaardingen, P.L.A., Bleeker, E.A.J., de Voogt, P., Kraak, M.H.S., 2001. Phototoxicity of azaarene isomers to the marine flagellate *Dunaliella tertiolecta*. Environ. Toxicol. Chem. 20, 1544–1550.
- Wilcke, W., 2000. SYNOPSIS Polycyclic Aromatic Hydrocarbons (PAHs) in Soil — a Review. J. Plant Nutr. Soil Sci. 163, 229–248.
- Wilhelm, M., Matuschek, G., Kettrup, A., 2000. Determination of basic nitrogen-containing polynuclear aromatic hydrocarbons formed during thermal degradation of polymers by high-performance liquid chromatography–fluorescence detection. J. Chromatogr. A 878, 171–181.
- Willumsen, P.A., Johansen, J.E., Karlson, U., Hansen, B.M., 2005. Isolation and taxonomic affiliation of N-heterocyclic aromatic hydrocarbon-transforming bacteria. Appl. Microbiol. Biotechnol. 67, 420–428.
- Wincent, E., Jönsson, M.E., Bottai, M., Lundstedt, S., Dreij, K., 2015. Aryl Hydrocarbon Receptor Activation and Developmental Toxicity in Zebrafish in Response to Soil Extracts Containing Unsubstituted and Oxygenated PAHs. Environ. Sci. Technol. 49, 3869–3877.

- Wischmann, H., Steinhart, H., 1997. The formation of PAH oxidation products in soils and soil/compost mixtures. Chemosphere 35, 1681–1698.
- Xia, T., Korge, P., Weiss, J.N., Li, N., Venkatesen, M.I., Sioutas, C., Nel, A., 2004. Quinones and Aromatic Chemical Compounds in Particulate Matter Induce Mitochondrial Dysfunction: Implications for Ultrafine Particle Toxicity. Environ. Health Perspect. 112, 1347–1358.
- Xiao, H., Krauss, M., Floehr, T., Yan, Y., Bahlmann, A., Eichbaum, K., Brinkmann, M., Zhang, X., Yuan, X., Brack, W., Hollert, H., 2016. Effect-Directed Analysis of Aryl Hydrocarbon Receptor Agonists in Sediments from the Three Gorges Reservoir, China. Environ. Sci. Technol. 50, 11319–11328.
- Xue, W., Siner, A., Rance, M., Jayasimhulu, K., Talaska, G., Warshawsky, D., 2002. A Metabolic Activation Mechanism of 7H-Dibenzo[c,g]carbazole via O-Quinone. Part 2: Covalent Adducts of 7H-Dibenzo[c,g]carbazole-3,4-dione with Nucleic Acid Bases and Nucleosides. Chem. Res. Toxicol. 15, 915–921.
- Xue, W., Warshawsky, D., 2005. Metabolic activation of polycyclic and heterocyclic aromatic hydrocarbons and DNA damage: A review. Toxicol. Appl. Pharmacol. 206, 73–93.
- Yamada, K., Suzuki, T., Hakura, A., Mizutani, T., Saeki, K., 2004. Metabolic activation of 10-aza-substituted benzo[a]pyrene by cytochrome P450 1A2 in human liver microsomes. Mutat. Res. Toxicol. Environ. Mutagen. 557, 159–165.
- Yang, Y., Woodward, L.A., Li, Q.X., Wang, J., 2014. Concentrations, Source and Risk Assessment of Polycyclic Aromatic Hydrocarbons in Soils from Midway Atoll, North Pacific Ocean. PLOS ONE 9, e86441.
- Young, L.Y., Cerniglia, C.E., 1995. Microbial Transformation and Degradation of Toxic Organic Chemicals. Wiley.
- Yu, D., Berlin, J.A., Penning, T.M., Field, J., 2002. Reactive Oxygen Species Generated by PAH o-Quinones Cause Change-In-Function Mutations in p53. Chem. Res. Toxicol. 15, 832–842.
- Yue, S., Ramsay, B.A., Brown, R.S., Wang, J., Ramsay, J.A., 2015. Identification of Estrogenic Compounds in Oil Sands Process Waters by Effect Directed Analysis. Environ. Sci. Technol. 49, 570–577.
- Zhang, Z., Sangaiah, R., Gold, A., M. Ball, L., 2011. Synthesis of uniformly ¹³C-labeled polycyclic aromatic hydrocarbons. Org. Biomol. Chem. 9, 5431–5435.
- Zhong, Y., Luan, T., Lin, L., Liu, H., Tam, N.F.Y., 2011. Production of metabolites in the biodegradation of phenanthrene, fluoranthene and pyrene by the mixed culture of *Mycobacterium* sp. and *Sphingomonas* sp. Bioresour. Technol. 102, 2965–2972.

- Zhong, Y., Luan, T., Wang, X., Lan, C., Tam, N.F.Y., 2007. Influence of growth medium on cometabolic degradation of polycyclic aromatic hydrocarbons by *Sphingomonas* sp. strain PheB4. Appl. Microbiol. Biotechnol. 75, 175–186.
- Zhong, Y., Luan, T., Zhou, H., Lan, C., Tam, N.F.Y., 2006. Metabolite production in degradation of pyrene alone or in a mixture with another polycyclic aromatic hydrocarbon by *Mycobacterium* sp. Environ. Toxicol. Chem. 25, 2853–2859.
- Zhong, Y., Zou, S., Lin, L., Luan, T., Qiu, R., Tam, N.F.Y., 2010. Effects of pyrene and fluoranthene on the degradation characteristics of phenanthrene in the cometabolism process by *Sphingomonas* sp. strain PheB4 isolated from mangrove sediments. Mar. Pollut. Bull. 60, 2043–2049.
- Zhou, R., Tseng, C.-L., Huan, T., Li, L., 2014. IsoMS: Automated processing of LC-MS data generated by a chemical isotope labeling metabolomics platform. Anal. Chem. 86, 4675–4679.
- Zhu, H., Li, Y.B., Trush, M.A., 1995. Characterization of Benzo[a]pyrene Quinone-Induced Toxicity to Primary Cultured Bone Marrow Stromal Cells from DBA/2 Mice: Potential Role of Mitochondrial Dysfunction. Toxicol. Appl. Pharmacol. 130, 108–120.
- Zielinska-Park, J., Nakamura, J., Swenberg, J.A., Aitken, M.D., 2004. Aldehydic DNA lesions in calf thymus DNA and HeLa S3 cells produced by bacterial quinone metabolites of fluoranthene and pyrene. Carcinogenesis 25, 1727–1733.
- Zonja, B., Delgado, A., Pérez, S., Barceló, D., 2015. LC-HRMS Suspect Screening for Detection-Based Prioritization of Iodinated Contrast Media Photodegradates in Surface Waters. Environ. Sci. Technol. 49, 3464–3472.