POLYCYCLIC AROMATIC HYDROCARBONS IN FLORIDA URBAN SOILS: FROM SOURCES TO EXPOSURES

By

PENG GAO

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2019
To My Parents, Teachers, and Mentors
ACKNOWLEDGMENTS

I would like to express my most sincere gratitude and appreciation to my advisor, Dr. Lena Q. Ma, for accepting me to her group and for her invaluable guidance, support, trust, patience and encouragement throughout my Ph.D. program. I also would like to thank my co-chair, Dr. P. Christopher Wilson, for his professional assistance and guidance in my research and career development. I would also like to thank my committee members Dr. Timothy G. Townsend, Dr. Stephen M. Roberts, and Dr. Nancy D. Denslow, for their valuable assistance, guidance and suggestions and the time they spent to ensure the smooth progress of my research. I would also like to thank my colleagues: Ky Gress, Letuzia de Oliveira, Evandro da Silva, Min Xu, Jing Su, Ying Guo, Yalan Liu, and Jingwei Fu for their kind supports during my research here.

In addition, I would like to thank my parents for their wise suggestions and emotional support as well as offering me all the opportunities to pursue my academic dreams. Moreover, a huge thank for Yeyuan as you are always there for me. Finally, there are many friends, who are of great supports in deliberating over our problems and findings, as well as providing a happy distraction to rest my mind outside of my research.
# TABLE OF CONTENTS

page

ACKNOWLEDGMENTS ................................................................................................................................. 4

LIST OF TABLES ........................................................................................................................................... 8

LIST OF FIGURES ......................................................................................................................................... 9

LIST OF ABBREVIATIONS .......................................................................................................................... 11

ABSTRACT .................................................................................................................................................. 13

CHAPTER

1 INTRODUCTION ....................................................................................................................................... 15

PAHs in Soils .................................................................................................................................................. 15
Emerging and Bioaccessible PAHs .................................................................................................................. 16
Sources of PAHs ........................................................................................................................................... 17
PAHs Exposures and Metabolomics ............................................................................................................... 19
Research Objectives ................................................................................................................................... 22

2 PAHS IN URBAN SOILS OF TWO LARGE CITIES: BACKGROUND CONCENTRATIONS, DISTRIBUTION, AND SOURCES .................................................................................................................. 25

Materials and Methods .................................................................................................................................. 25
Soil Sampling and Preparation ....................................................................................................................... 25
Chemicals and Sample Analysis ..................................................................................................................... 26
Data Analyses Quality Assurance ................................................................................................................ 27
Results and Discussion ..................................................................................................................................... 28
Concentrations and Characteristics of PAHs in Soils ...................................................................................... 28
Benzo[a]pyrene-equivalent Concentration and Distribution ...................................................................... 29
Source Identification and Apportionment ...................................................................................................... 30

3 EMERGING PAHS IN URBAN SOILS: CONCENTRATIONS, BIOACCESSIBILITY, AND SPATIAL DISTRIBUTION .......................................................................................................................... 39

Materials and Methods .................................................................................................................................. 39
Soil Sampling and Preparation ....................................................................................................................... 39
Chemicals and Sample Analysis ..................................................................................................................... 39
Bioaccessibility Analysis .................................................................................................................................. 40
Benzo[a]pyrene-equivalent Carcinogenicity ................................................................................................. 41
Data Analyses and Quality Assurance ......................................................................................................... 41
Results and Discussion ..................................................................................................................................... 42
Concentrations of Six Emerging PAHs .......................................................................................................... 42
Concentrations of Three Emerging cPAHs based on BaP-EQ ...................................................................... 43
## Distributions of Three Emerging cPAHs based on BaP-EQ

Distributions of Three Emerging cPAHs based on BaP-EQ ............................................ 44

## Bioaccessible Concentrations of Three Emerging cPAHs based on BaP-EQ

Bioaccessible Concentrations of Three Emerging cPAHs based on BaP-EQ .................................. 45

## 4 EMERGING AND LEGACY PAHS IN URBAN SOILS OF FOUR SMALL CITIES: CONCENTRATIONS, DISTRIBUTION, AND SOURCES

Materials and Methods .................................................................................................................. 51
Soil Sampling and Preparation ..................................................................................................... 51
Chemicals and Sample Analysis ................................................................................................... 52
Benzo[a]pyrene-equivalent Carcinogenicity ................................................................................. 53
Data Analyses and Quality Assurance ......................................................................................... 53

## Results and Discussions

Concentrations of Sixteen Legacy PAHs and Six Emerging PAHs ............................................. 55
BaP-EQ of Seven Legacy cPAHs and Six Emerging cPAHs ......................................................... 56
Source Identification and Apportionment ...................................................................................... 59

## 5 CONCLUSIONS

.................................................................................................................................................. 68

## APPENDIX

### A SOURCE IDENTIFICATION OF PAHS IN SOILS BASED ON STABLE CARBON ISOTOPIC SIGNATURES

Identification of PAH Sources ..................................................................................................... 70
Molecular Diagnostic Ratios ......................................................................................................... 70
Receptor Models ............................................................................................................................. 72
Stable Carbon Isotopic Signatures ............................................................................................... 73
Stable Carbon Isotopic Signatures of PAHs ............................................................................... 75
Biogenic PAHs ............................................................................................................................... 77
Petrogenic PAHs ............................................................................................................................. 79
Pyrogenic PAHs ............................................................................................................................. 80
Stable Carbon Isotopic Signatures of Soil PAHs ........................................................................ 83
PAHs in Rural and Urban Soils ..................................................................................................... 83
Formation and Degradation of PAHs ......................................................................................... 84
PAHs in Contaminated Soil ......................................................................................................... 87
Other Isotopic Signatures of Soil PAHs ....................................................................................... 88
Radioactive C Isotopic Ratios ....................................................................................................... 88
Stable Hydrogen Isotopic Ratios ................................................................................................. 89
Conclusions and Prospective Research ....................................................................................... 90

### B HUMAN EXPOSURE TO POLYCYCLIC AROMATIC HYDROCARBONS: METABOLIC PERSPECTIVE

Carcinogenesis and Transformation of PAHs in Humans and Animals ....................................... 100
Metabolomics Studies on Human Exposure to PAHs ................................................................ 104
Biomarkers .................................................................................................................................. 104
Analytical Approaches .................................................................................................................. 107
Data Analysis ................................................................................................................................. 113
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>Information of the six emerging PAHs in this study</td>
<td>24</td>
</tr>
<tr>
<td>2-1</td>
<td>Descriptive statistics for PAH concentrations in urban soils of Orlando and Tampa</td>
<td>33</td>
</tr>
<tr>
<td>2-2</td>
<td>Six diagnostic ratios of PAHs in urban soils of Orlando and Tampa</td>
<td>34</td>
</tr>
<tr>
<td>3-1</td>
<td>Total and bioaccessible concentrations of 3 emerging carcinogenic PAHs and USEPA 7 priority carcinogenic PAHs in urban soils of Orlando and Tampa based on BaP-EQ</td>
<td>47</td>
</tr>
<tr>
<td>4-1</td>
<td>Descriptive statistics for PAH concentrations in urban soils of four small cities</td>
<td>62</td>
</tr>
<tr>
<td>4-2</td>
<td>Six diagnostic ratios of PAHs in urban soils of four small cities</td>
<td>63</td>
</tr>
<tr>
<td>A-1</td>
<td>Source identification and apportionment of PAHs in soils based on molecular diagnostic ratios</td>
<td>92</td>
</tr>
<tr>
<td>A-2</td>
<td>Compilation of reference values of δ13C for biogenic bulk C, petrogenic PAHs, natural pyrogenic PAHs, and anthropogenic pyrogenic PAHs</td>
<td>94</td>
</tr>
<tr>
<td>B-1</td>
<td>Metabolomics studies on PAHs exposure</td>
<td>125</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1</td>
<td>Concentrations and compositions of 16 PAHs in urban soils of Orlando and Tampa.</td>
<td>35</td>
</tr>
<tr>
<td>2-2</td>
<td>$\Sigma 16$-PAHs and $\Sigma 7c$PAHs concentrations and spatial distributions in urban soils of Orlando and Tampa based on site-specific concentrations and IDW interpolation concentrations.</td>
<td>36</td>
</tr>
<tr>
<td>2-3</td>
<td>$\Sigma 7c$PAHs concentrations in Orlando and Tampa soils and comparisons with Florida residential and industrial soil cleanup target levels based on BaP-EQ.</td>
<td>37</td>
</tr>
<tr>
<td>2-4</td>
<td>The average mass contributions of each source to 16 PAHs in Tampa and Orlando urban soils based on USEPA PMF modeling.</td>
<td>38</td>
</tr>
<tr>
<td>3-1</td>
<td>Concentrations of 6 emerging PAHs in urban soils of Orlando and Tampa.</td>
<td>48</td>
</tr>
<tr>
<td>3-2</td>
<td>Total and bioaccessible $7c$PAHs and $3c$PAHs concentrations in urban soils of Orlando and Tampa based on BaP-EQ.</td>
<td>49</td>
</tr>
<tr>
<td>3-3</td>
<td>Spatial distributions of $\Sigma 3c$PAHs concentrations based on BaP-EQ in urban soils of Orlando and Tampa. Light blue circles represent the central business districts.</td>
<td>50</td>
</tr>
<tr>
<td>4-1</td>
<td>Concentrations of 22 PAHs in urban soils of four small cities.</td>
<td>64</td>
</tr>
<tr>
<td>4-2</td>
<td>$\Sigma 7c$PAHs and $\Sigma 3c$PAHs concentrations in urban soils of four small cities.</td>
<td>65</td>
</tr>
<tr>
<td>4-3</td>
<td>Comparisons with Florida residential and industrial soil cleanup target levels based on $\Sigma 7c$PAHs and the average contributions of each source to 22 PAHs in urban soils based on USEPA PMF model.</td>
<td>66</td>
</tr>
<tr>
<td>4-4</td>
<td>$\Sigma 7c$PAHs concentrations and spatial distributions in urban soils of four small cities based on site-specific concentrations and IDW interpolation concentrations.</td>
<td>67</td>
</tr>
<tr>
<td>A-1</td>
<td>A schematic of C isotopic fractionation of PAHs from CO2 to coal utilization including changes in $\delta^{13}C$ values in each step.</td>
<td>95</td>
</tr>
<tr>
<td>A-2</td>
<td>A schematic of PAHs formation processes: biogenic, petrogenic, and pyrogenic.</td>
<td>96</td>
</tr>
<tr>
<td>A-3</td>
<td>A biogeochemical cycle of stable C isotopic signatures in soil PAHs.</td>
<td>97</td>
</tr>
<tr>
<td>A-4</td>
<td>Compilation of reference values of $\delta^{13}C$ for biogenic bulk C, petrogenic PAHs, natural pyrogenic PAHs, and anthropogenic pyrogenic PAHs.</td>
<td>98</td>
</tr>
</tbody>
</table>
The proposed evaluation system of soil PAHs based on stable C isotopic........ 99

Biological active “bay” and “fjord” regions in PAH conformation using
dibenzo[a,l]pyrene as an example.......................................................... 126

Three metabolic activation pathways of benzo[a]pyrene to form DNA adduct
metabolites to cause mutation and detoxification where 7,8-oxide BaP and
7,8-dihydrodiol BaP are proximate carcinogens, 7,8-diol-9,10-epoxide BaP is
a strong mutagen and ultimate carcinogen.................................................. 127

Source-exposure-dose-effect outcome continuum: exposure biomarker is the
key to exposure reconstruction...................................................................... 128

Schemes for PAHs metabolomics experiment based on LC-MS/MS and/or
NMR approach ........................................................................................................ 129

Common multivariate statistical analysis in metabolomics studies............... 130

Metabolome alterations in human and animals by comparing control and
exposure groups. Red, blue, and purple box represent up regulations, down
regulations, and insignificant changes/dependent regulations in exposure
groups ................................................................................................................... 131
LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3cPAHs</td>
<td>3 Emerging Carcinogenic PAHs</td>
</tr>
<tr>
<td>7cPAHs</td>
<td>7 USEPA Carcinogenic PAHs</td>
</tr>
<tr>
<td>ACE</td>
<td>Acenaphthene</td>
</tr>
<tr>
<td>ACY</td>
<td>Acenaphthylene</td>
</tr>
<tr>
<td>AN</td>
<td>Anthanthrene</td>
</tr>
<tr>
<td>ANC</td>
<td>Anthracene</td>
</tr>
<tr>
<td>BaA</td>
<td>Benz[a]anthracene</td>
</tr>
<tr>
<td>BaP</td>
<td>Benzo[a]pyrene</td>
</tr>
<tr>
<td>BaP-EQ</td>
<td>Benzo[a]pyrene Equivalent</td>
</tr>
<tr>
<td>BbF</td>
<td>Benzo[b]fluoranthene</td>
</tr>
<tr>
<td>BcF</td>
<td>7H-benzo[c]fluorene</td>
</tr>
<tr>
<td>BkF</td>
<td>Benzo[k]fluoranthene</td>
</tr>
<tr>
<td>BP</td>
<td>Benzo[g,h,i]perylene</td>
</tr>
<tr>
<td>C</td>
<td>Carbon</td>
</tr>
<tr>
<td>CH</td>
<td>Chrysene</td>
</tr>
<tr>
<td>DeP</td>
<td>Dibenzo[a,e]pyrene</td>
</tr>
<tr>
<td>DhA</td>
<td>Dibenz[a,h]anthracene</td>
</tr>
<tr>
<td>DhP</td>
<td>Dibenzo[a,h]pyrene</td>
</tr>
<tr>
<td>DiP</td>
<td>Dibenzo[a,i]pyrene</td>
</tr>
<tr>
<td>DIP</td>
<td>Dibenzo[a,l]pyrene</td>
</tr>
<tr>
<td>FL</td>
<td>Fluorene</td>
</tr>
<tr>
<td>FLA</td>
<td>Fluoranthene</td>
</tr>
<tr>
<td>FSCTLs</td>
<td>Florida Soil Cleanup Target Levels</td>
</tr>
<tr>
<td>HMW</td>
<td>High Molecular Weight</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>IP</td>
<td>Indeno[1,2,3-cd]pyrene</td>
</tr>
<tr>
<td>LMW</td>
<td>Low Molecular Weight</td>
</tr>
<tr>
<td>NA</td>
<td>Naphthalene</td>
</tr>
<tr>
<td>PAHs</td>
<td>Polycyclic aromatic hydrocarbons</td>
</tr>
<tr>
<td>PH</td>
<td>Phenanthrene</td>
</tr>
<tr>
<td>PY</td>
<td>Pyrene</td>
</tr>
<tr>
<td>RPF</td>
<td>Relative potency factor</td>
</tr>
</tbody>
</table>
Polycyclic aromatic hydrocarbons (PAHs) are ubiquitous organic contaminants, with soil being the accumulative sink. This study determined the concentrations, distributions, and sources of 16 USEPA priority PAHs and 6 emerging PAHs in urban soils of two large cities: Orlando and Tampa and four small cities: Clay county, Ocala, Pensacola, and West Palm Beach in Florida, USA. Emerging PAHs include 3 carcinogenic (anthanthrene, 7H-benzo[c]fluorene, and dibenzo[a,l]pyrene; 3cPAHs) and 3 non-carcinogenic (dibenzo[a,e]pyrene, dibenzo[a,i]pyrene and dibenzo[a,h]pyrene) PAHs.

A total of 114 and 100 soil samples were collected from large cities and small cities, respectively. High molecular weight PAHs were dominated in all soil samples, indicating that anthropogenic sources contributed the most. The average $\sum$16-PAHs in urban soils of Clay county, Ocala, Pensacola, West Palm Beach, Orlando, and Tampa were 1821, 2748, 3115, 4055, 3227 and 4562 $\mu$g kg$^{-1}$, respectively. Based on benzo[a]pyrene equivalent (BaP-EQ), the 7 USEPA carcinogenic PAHs (7cPAHs) and 3 emerging carcinogenic PAHs (3cPAHs) in urban soils in Clay County averaged 223 and
3703, Ocala 319 and 4521, Pensacola 302 and 5423, West Palm Beach 449 and 5916, Orlando 452 and 7387, and Tampa 802 and 4943 µg kg\(^{-1}\), respectively. Although \(\sum \text{cPAHs}\) in 75-89% of samples were lower than the Florida Soil Cleanup Target Levels (FSCTLs) for industrial sites at 700 µg kg\(^{-1}\), \(\sum \text{3cPAHs}\) were 6-18 times greater than \(\sum \text{cPAHs}\). However, the oral bioaccessibility of PAHs based on n-butanol mild extraction in Orlando and Tampa urban soils were low, averaging 3.4–7.4%.

Based on molecular diagnostic ratios and PMF model, major sources of soil PAHs in all cities were similar, mainly from pyrogenic sources including vehicle emissions, and biomass, coal and coke combustion. Based on ArcGIS mapping, PAHs concentrations in soils near commercial sites, business districts, and high traffic roads were higher. By combining diagnostic ratios, PMF model, PAHs compositional profiles and PAHs spatial distributions, the main sources of PAHs in Florida urban soils mainly came from pyrogenic sources.

In short, it is important to consider the background concentrations of both legacy and emerging PAHs in urban soils as well as their bioaccessibility for soil remediation and human health risk assessment.
CHAPTER 1
INTRODUCTION

PAHs in Soils

Polycyclic aromatic hydrocarbons (PAHs) are persistent organic contaminants, with some being toxic and carcinogenic. Among the hundreds of PAHs, USEPA has designated 16 as priority contaminants. Due to their volatility, PAHs may be transported far from their original sources, accumulating in the environment, with soil being their primary reservoir and posing health risk to humans (Gan et al., 2009; Gao et al., 2018b).

Background concentrations of PAHs in soils are mainly from non-point sources (Azah et al., 2017; Teaf, 2008). PAHs in soils mainly result from incomplete combustion of organic C, far exceeding natural sources (Buczyńska et al., 2013; Kim et al., 2008). Urban soils are a major sink for PAHs but it is challenging to distinguish PAHs from site-related activities and non-point sources (Vane et al., 2014; Duan et al., 2015). It is important to determine the background concentrations of PAHs in urban soils since cleanup standards for PAHs in contaminated soils can be based on background concentrations. As such, studies have investigated PAH concentrations in large cities


around the world including New York, London, and Beijing (Vane et al., 2014; Azzolina et al., 2016; Li et al., 2016).

The Florida Soil Cleanup Target Level (FSCTL) for PAHs are 100 µg kg\(^{-1}\) and 700 µg kg\(^{-1}\) for residential and industrial soils based on benzo[a]pyrene-equivalent concentrations (BaP-EQ). This is because carcinogenic PAHs share a common toxicity mechanism but display different toxic potency (Teaf, 2008). Teaf et al. (2008) reported that BaP-EQ concentrations in two Florida soils were 1-5 mg kg\(^{-1}\) during site investigation, exceeding FSCTLs. One site is located in an old urban neighborhood near a highway, and the other site is near an asphalt parking lot, indicating that anthropogenic activities have contributed to the elevated PAHs. Banger et al. (2010) reported that the BaP-EQ concentrations in Miami urban soils were 175-281 µg kg\(^{-1}\), exceeding the FSCTL for residential soils. The results indicated that PAH concentrations in Florida varied greatly in the two studies, so a systematic study on PAH concentrations and their sources in Florida urban soils is needed.

**Emerging and Bioaccessible PAHs**

Among the hundreds of PAHs, USEPA designates 16 as priority contaminants since the 1970s. Therefore, current regulations mainly focus on the 16 PAHs. Recently, Andersson and Achten (2015) reported the risk based on additional PAHs. Although USEPA prepared a review draft to develop the relative potency factor (RPF) for 26 PAHs in 2010, little is known about the occurrence, bioavailability, and toxicity of those non-USEPA priority PAHs in soils (Richter-Brockmann and Achten, 2018).

Emerging contaminants are those that have not been studied compared with legacy contaminants, but they might be present in the environment and cause adverse effects on humans. In terms of PAHs, non-USEPA priority PAHs have not been
investigated in the past, so they can be considered emerging PAHs. This study include 6 emerging PAHs, 3 carcinogenic (anthanthrene-AN, 7H-benzo[c]fluorene-BcF, and dibenzo[a,l]pyrene-DlP: 3cPAHs) and 3 non-carcinogenic (dibenzo[a,e]pyrene-DeP, dibenzo[a,i]pyrene-DiP, and dibenzo[a,h]pyrene-DhP) (Table 1-1). To date, only one study determined the concentrations of the 6 emerging PAHs in 5 soil samples (Richter-Brockmann and Achten, 2018).

Besides total concentrations of PAHs in soils, it is also important to determine the bioavailable fractions of PAHs. This is because mounting evidence shows that the assessment of human exposure to ingested contaminants should be based on bioavailable not total concentrations. In vivo bioassays using animal models are effective to assess contaminant bioavailability in soils. However, due to their high cost and ethical concerns, they are unsuitable for large-scale testing. Consequently, simple and inexpensive in vitro assays have been developed as substitutes to predict contaminant bioavailability. These assays determine the portion of contaminants that are solubilized in the simulated human gastrointestinal tract, which is potentially available for absorption and is referred as bioaccessible (Zhang et al., 2017). As such, bioaccessibility is a measure of the physiological solubility of PAHs in simulated gastrointestinal tract of humans. In this study, the bioaccessible PAHs in soils were assessed by an in vitro n-butanol mild extraction (Duan et al., 2014; Ruby et al., 2016).

Sources of PAHs

Most PAHs in the environment result from incomplete combustion and pyrolysis processes of organic C (OC), including biomass, petroleum, and coals (Gan et al., 2009; Li et al., 2014a). Based on their origins, PAHs can come from natural and anthropogenic sources (Buczyńska et al., 2013; Wang et al., 2015a, 2015b; Cai et al.,
Natural sources include oil seeps from crude oil deposits, forest fires, volcanoes and erosion of ancient sediment. For example, some PAHs such as perylene are produced naturally from the biochemical transformation of OC (Abdel-shafy and Mansour, 2016). Anthropogenic PAHs are formed either by thermal alteration of OC or its incomplete combustion (Rothwell and Cooke, 2015; Abdel-shafy and Mansour, 2016). Today, the major sources of PAHs in the environment are from the human utilization of petroleum products and incomplete combustion of fossil fuels, biofuels or other forms of OC, far exceeding natural sources (Kim et al., 2008; Cai et al., 2017). As a result, PAH concentrations are elevated in soils, especially near urban and industrial areas that often have multiple sources of release (Duan et al., 2015).

Based on their formation process, PAHs from both natural and anthropogenic sources can be classified into three groups: pyrogenic, petrogenic, and biogenic (Buczyńska et al., 2013). Pyrogenic PAHs result from incomplete combustion of fossil fuels and biomass under high temperatures. They are released in the form of exhaust and solid residues, thereby ubiquitous in soils. Petrogenic PAHs originate from petroleum products such as crude oil, coal, and gasoline and are formed under relatively low temperatures during fossil fuel formation processes. Direct spillage from petroleum is a common petrogenic source of PAHs. In most cases, pyrogenic PAHs dominate over petrogenic PAHs due to human activities (Gan et al., 2009; Duan et al., 2015). Petrogenic PAHs are introduced into soils through accidental oil spills, release from tanker operations, and municipal runoff. Biogenic PAHs are produced during degradation of vegetative organic substances by plants, algae, and microorganisms. In
addition, they are produced during the slow transformation of OC in soils by plants and microorganisms (Abdel-shafy and Mansour, 2016).

Source characterization of PAHs can provide a clear link between contaminants and their corresponding sources and help to determine the potentially responsible party causing environmental pollution. Source apportionment determines the amount of contamination contributed by each party quantitatively to help regulators make more accurate decisions. For successful identification and apportionment of PAH sources in the environment, it is important to investigate the sources, chemistry, and fate of PAHs in soils (Stogiannidis and Laane, 2015).

Various models have been used to trace the sources of PAHs. Often natural sources of PAHs are dominated by low molecular weight PAHs (LMW-PAHs; 2-3 rings) whereas anthropogenic sources are dominated by high molecular weight PAHs (HMW-PAHs; >4 rings). As PAHs are often emitted as a group, the ratios of different PAHs can be used to trace their sources (Tobiszewski and Namie, 2012). Based on the diagnostic ratios of PAHs, atmospheric deposition is the main source of PAHs in soils. The positive matrix factorization (PMF) method, developed for atmospheric aerosol components, uses experimental uncertainties in data matrix to constrain solutions with positive values. As such, PMF handles missing values and explains data accuracy well regardless of PAH sources (Wang et al., 2015; Cai et al., 2017). For accurate source apportionment, two or three methods have been applied (Suman et al., 2016).

**PAHs Exposures and Metabolomics**

PAHs are carcinogenic, so they are of major risks to human health. According to USEPA, seven PAH compounds are probable human carcinogens, including benz[a]anthracene, benzo[a]pyrene (BaP), benzo[b]fluoranthene, benzo[k]fluoranthene,
chrysene, dibenz[a,h]anthracene, and indeno[1,2,3-cd]pyrene (Banger et al., 2010). Humans are exposed to PAHs through three main pathways: ingestion, inhalation, and dermal contact (Ma and Harrad, 2015; Ruby et al., 2016). Occupational exposure may occur to workers by breathing exhaust fumes or smoke (Fernando et al., 2016; Navarro et al., 2017).

Exposure to PAHs is linked with various adverse health effects including oxidative stress (Wang et al., 2015), diabetes (Yang et al., 2017), inflammation (Ferguson et al., 2017), infertility (Xia et al., 2009), cardiovascular disease (Jomova et al., 2012) and poor fetal development (Sexton et al., 2011). Short-term health effects include eye and skin irritation, nausea and vomiting, and inflammation while long-term include various cancers, DNA and proteins damage, and gene mutation (Abdel-Shafy and Mansour, 2015; Kim et al., 2013). Although unmetabolized PAHs are also toxic, a major concern is the ability of their reactive metabolites, such as epoxides and dihydrodiols, to bind to cellular proteins and DNA. The resulting biochemical disruption and cell damage lead to mutations, developmental malformations, tumors, and cancer (Ewa and Danuta, 2016; Kim et al., 2013; Moorthy et al., 2015). Since PAHs often exist as a group, understanding their composition differences helps to accurately assess their toxicity (Abdel-Shafy and Mansour, 2015). In addition, uncovering the dynamics of PAHs metabolisms and possible toxic effects help to better understand their impacts and control their metabolic pathways in humans.

Metabolomics studies the metabolites present in an organism, cell, or tissue by investigating the unique chemical fingerprints of a specific cellular process leaves behind, i.e., their small-molecule metabolite profiles (Rochfort, 2005). As such, it can
focus on the biochemical processes involving metabolites and biotransformation products of various chemicals. It is a quantitative measure of the multi-parametric metabolic responses of living systems to pathophysiological stimuli or genetic modification (Worley and Powers, 2015). The metabolome collects metabolites in a biological organism, which are the end-products of cellular processes. Gene expression data and proteomic analyses reveal the gene products produced in a cell, representing one aspect of cellular function. In addition, metabolic profiling provides a snapshot of cell physiology. Metabolic profiling of biofluids, cells, and tissues is a routine tool for biomarker discovery. Owing to the innovative developments in informatics and analytical technologies, and the integration of orthogonal biological approaches, it is possible to expand metabolomics to study the impacts of PAHs exposure on humans. Its inherent sensitivity and subtle alterations in biological pathways can be detected to provide insights into the mechanisms underlying various physiological conditions and aberrant processes (Johnson et al., 2016).

Environmental metabolomics is a relatively new technique to assess the biological consequences of chemical exposure (Lankadurai et al., 2013). This approach has the advantages to study organism-environment interactions and assess organism functions and health at the molecular level. As such, there are many applications for metabolomics in environmental sciences, ranging from understanding organismal responses to abiotic pressures to investigate the responses of organisms to environmental stressors. Environmental metabolomics often serves to detect the impacts of sub-acute toxicity in organisms prior to overt phenotype changes in addition to revealing the underlying mechanisms of the actions or synergistic effects of chemical
exposures, which often involve a complex mixture of compounds. Therefore, metabolite models can be used to characterize the endpoint of PAH-induced toxic reactions. Metabolomics analysis can be performed, with or without known metabolites being quantified, by using a comprehensive analysis of all metabolites. Non-targeted methods can detect the differences in pictorial description patterns, so it often provides information about the mechanisms, pathways, and biomarkers after chemical exposure (Bundy et al., 2009; Elie et al., 2015).

To avoid distortion and miss-discovery during metabolomics studies, various validations are required. While analytical validation is devoted to measurement examination and chemometric validation to test the reliability of statistical results obtained, biological validation consists of an evaluation of discovered biological knowledge. Biological validation constitutes the most relevant confirmation of the results. Follow-up analytical and biological validation studies to verify both the identification and biological reproducibility of the results are essential to confirm initial findings from metabolomics studies (Godzien et al., 2013).

**Research Objectives**

This study determined the background concentrations of 22 PAHs (16 legacy PAHs and 6 emerging PAHs) in urban soils of two large cities and four small cities in Florida, USA: Orlando and Tampa; Clay county, Ocala, Pensacola, and West Palm Beach. In addition, \( \sum_{7c} \text{PAH} \) concentrations based on BaP-EQ were compared with FSCTLs. The specific objectives were: 1) to determine the background and BaP-EQ concentrations of 22 PAHs in Florida urban soils; 2) to evaluate their oral bioaccessibility; 3) to investigate PAHs spatial distribution using site-specific
concentrations as well as the inverse distance weighted interpolation, and 4) to identify PAHs sources by using both diagnostic ratios and PMF model.
Table 1-1. Information of the six emerging PAHs in this study (Richter-Brockmann and Achten, 2018).

<table>
<thead>
<tr>
<th>PAHs</th>
<th>CASRN</th>
<th>Structure</th>
<th>Molecular weight (g/mol)</th>
<th>Average RPFs</th>
<th>Relative confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anthanthrene</td>
<td>191-26-4</td>
<td><img src="image1" alt="Structure" /></td>
<td>276.34</td>
<td>0.4</td>
<td>Medium</td>
</tr>
<tr>
<td>7H-benzo[c]fluorene</td>
<td>205-12-9</td>
<td><img src="image2" alt="Structure" /></td>
<td>216.28</td>
<td>20</td>
<td>Medium</td>
</tr>
<tr>
<td>Dibenzo[a,l]pyrene</td>
<td>191-30-0</td>
<td><img src="image3" alt="Structure" /></td>
<td>302.38</td>
<td>30</td>
<td>Medium</td>
</tr>
<tr>
<td>Dibenzo[a,i]pyrene</td>
<td>189-55-9</td>
<td><img src="image4" alt="Structure" /></td>
<td>302.38</td>
<td>0.6</td>
<td>Low</td>
</tr>
<tr>
<td>Dibenzo[a,e]pyrene</td>
<td>192-65-4</td>
<td><img src="image5" alt="Structure" /></td>
<td>302.38</td>
<td>0.4</td>
<td>Low</td>
</tr>
<tr>
<td>Dibenzo[a,h]pyrene</td>
<td>53-70-3</td>
<td><img src="image6" alt="Structure" /></td>
<td>302.38</td>
<td>0.9</td>
<td>Low</td>
</tr>
</tbody>
</table>
CHAPTER 2
PAHS IN URBAN SOILS OF TWO LARGE CITIES: BACKGROUND CONCENTRATIONS, DISTRIBUTION, AND SOURCES

Materials and Methods

Soil Sampling and Preparation

To determine the concentrations of PAHs in urban soils, we selected two representative large cities: Orlando and Tampa. While Orlando’s population is 277,173 with 111 mile² area, Tampa’s is 377,165 with 171 mile² area. To minimize point source contamination, we developed exclusion criteria. Samples excluded were areas of 1) residential yards; 2) low-lying areas with flooding or run-off; 3) <3 m of paved roads, parking lots, painted structures and visually-stained soil; 4) <10 m of gas stations, railroads, storm water collection ponds, and electrical substations; 5) <50 m of sites of incineration, wastewater, or solidwaste treatment facilities; and 6) <100 m of industrial properties or identified hazardous waste sites (Banger et al., 2010; Chirenje et al., 2002; Chung et al., 2014; Clement et al., 2015).

From March to May 2016, based on a randomized sampling plan, 50 soil samples (25 commercial and 25 public sites) in Orlando and 64 soil samples (32 commercial and 32 public sites) in Tampa were collected with similar sampling density. The commercial sites included restaurants, shopping malls, supermarkets, cinemas, and banks, and public sites include government facilities, courts, libraries, museums, and schools. We sampled the top 15-cm soil using composite samples. At each site, a 5 m × 5 m grid was created and 5 subsamples from four corners and the center were

taken to make 1 composite sample. The soil samples were taken using stainless steel shovels and placed in aluminum foil wraps, which were air-dried at room temperature for 3 days and sieved with a 2-mm sieve before analysis (Wang et al., 2015).

**Chemicals and Sample Analysis**

All chemicals were from Fisher Scientific (Hampton, NH), including 16 PAHs and their corresponding deuterated PAHs, 4 surrogates (decafluorobiphenyl, o-terphenyl, 6-methylnaphthalene, and perylene-d12), florisil, sodium sulfate anhydrous, and silica gel sorbent for pretreatment. All labware was made of glass or Teflon and washed with deionized water, acetone, and n-hexane before use. All solvents and chemicals were of HPLC grade or better.

Pretreatment and analysis of PAHs followed modified USEPA Method 3550C and 8270D. Five grams of soil sample was spiked with surrogate PAHs in a centrifuge tube, and then 20 mL 1:1 (v/v) acetone/n-hexane was added as extractant, with blank without soil samples. Each sample was vortexed for 1 minute and the mixture was subjected to ultrasonic treatment for 15 minutes. The sample was centrifuged at 2000 revolutions per minute at room temperature for 10 min. After centrifugation, the organic layer was siphoned out with a pipette, being extracted using 20 mL 1:1 (v/v) acetone/n-hexane for three times. The combined extractants were evaporated to 1 mL using a rotatory evaporator and cleaned up using a chromatographic column filled with anhydrous sodium sulfate, silica gel, and florisil. The extract was then eluted with 100 mL 1:1 (v/v) acetone/n-hexane through the column.

The collected sample was vacuum-evaporated and solvent-exchanged to n-hexane and then concentrated to 1 mL before analysis using a Trace Gas Chromatography (GC) Ultra Gas Chromatograph and DSQ2 Mass Spectrometer (MS;
Thermo Scientific). The GC column consisted of 30 m × 0.25 mm i.d. in 0.25 μm film thickness. A 1 μL aliquot of sample extract was injected in splitless mode. The oven temperature was set at 60˚C for 1 min, then ramped to 120˚C at 25˚C/min, further ramped to 160˚C at 10˚C/min, and finally ramped to 330˚C at 5˚C /min with a 1-min final hold time. The inlet temperature was at 320˚C, the ion source was at 250˚C, and the transfer line temperature was at 300˚C. The helium carrier gas was held at a constant rate of 0.3 mL/min. Ionization was carried out using the electron ionization mode and data were acquired using the selective ion mode. Identification of PAHs was based on selected ions and relative retention time between samples and standard solution containing individual PAHs.

**Data Analyses Quality Assurance**

The inverse distance weighted (IDW) interpolation was used to determine the distribution of soil PAHs based on ArcGIS 10.3. Besides IDW, diagnostic ratios and USEPA PMF 5.0 modeling were also used to trace PAH sources. The explanation of replicates and statistical methods can be found in literature (Vane et al., 2014; Cai et al., 2017).

The procedural, spike, and matrix blanks, and sample duplicates were analyzed routinely with samples, with no interference being detected. One laboratory blank and one duplicate were run with every six samples, with samples having difference >±15% being reanalyzed. The PAHs were quantified via external standard calibrations. Four surrogate standards were added to soil samples to monitor the procedures of sample extraction, cleanup, and analysis. The surrogate recoveries of perylene-d12 (5 rings), 6-methylchrysene (4 rings), o-terphenyl (3 rings), and decafluorobiphenyl (2 rings) were 61-110%, with standard deviations of 7-15%, and averaging 83, 80, 79 and 73%,
respectively. The limit of detection (0.9–3.6 µg kg\(^{-1}\)) was based on three times the signal-to-noise level of the chromatogram for blank sample with a sample size of 5 g and a final volume of 1 mL. All statistical analyses were performed using JMP®Pro 13.0 statistical software (SAS Institute, Inc., Cary, NC, USA).

**Results and Discussion**

**Concentrations and Characteristics of PAHs in Soils**

The concentrations of the 16 USEPA priority PAHs in Orlando and Tampa soils are shown in Figure 2-1. The results showed that PAHs in Tampa soils had higher concentrations than those in Orlando, especially for 3-5 ring PAHs. Most soils had PAH concentrations at hundreds of µg kg\(^{-1}\), but some were higher, especially in Tampa. Descriptive statistics of all PAHs are presented in Table 2-1. The average and highest \(\sum_{16}\)-PAHs in Orlando soils were 3227 and 30428 µg kg\(^{-1}\) while those in Tampa were 4562 and 58640 µg kg\(^{-1}\). Overall, the \(\sum_{16}\)-PAHs in 95% of soils samples from both cities were < 30000 µg kg\(^{-1}\).

In addition, the commercial sites had higher PAH concentrations than public sites in both cities (Figure 2-1A). A similar trend was observed in Miami soils, as PAH concentrations in commercial sites were higher than public sites (2364 vs. 1595 µg kg\(^{-1}\)) (Banger et al., 2010). This may be due to the larger traffic volumes on commercial sites than public sites (Azzolina et al., 2016). The average \(\sum_{16}\)-PAHs in Orlando and Tampa were similar to those in Nanjing, China (3330 µg kg\(^{-1}\), Wang et al., 2015), but higher than those in Miami, Florida (1869 µg kg\(^{-1}\), Banger et al., 2010) and lower than those in Bergen, Norway (6780 µg kg\(^{-1}\), Haugland et al., 2008). The fact that Miami generates more power via nuclear power plant while Orlando and Tampa use coal-fired power plant may have contributed to higher PAH concentrations in Orlando and Tampa. Based
on the compositional profiles of PAHs (Figure 2-1B), both cities were dominated by HMW-PAHs, especially 4-ring PAHs, similar to Miami soils (Banger et al., 2010). These results indicate that pyrogenic sources are dominant in both cities (Chen et al., 2016).

PAH concentrations along with their spatial distribution are shown in Figure 2-2. The three concentration levels were based on Jenks Natural Breaks algorithm in ArcGIS. Roads with annual average daily traffic (aadt) >50000 are considered high traffic. The highest PAH concentrations are in central business areas in both cities, so an increasing PAH concentration was evident from rural to urban areas (Figure 2-2). In addition, sites with high PAH concentrations were near high traffic roads in both cities, suggesting contributions from vehicle emissions. This is consistent with the higher average aadt in Tampa than Orlando (24144 vs. 21197).

**Benzo[a]pyrene-equivalent Concentration and Distribution**

Soil PAH concentrations are often converted to BaP-EQ before being compared to soil cleanup standards. This approach has been used to evaluate carcinogenic risks of PAHs in various studies (Banger et al., 2010; Bi et al., 2016). BaP-EQ concentrations of Orlando and Tampa soils are shown in Figure 2-3A. The BaP-EQ concentrations for most soils in both cities were similar, being <700 µg kg⁻¹. However, soils from Tampa generally had higher BaP-EQ concentrations than Orlando. The BaP-EQ concentrations were 4 to 3742 µg kg⁻¹ in Orlando, averaging 452 µg kg⁻¹. For Tampa, the BaP-EQ concentrations were non-detected to 9406 µg kg⁻¹, averaging 802 µg kg⁻¹. For both cities, BaP-EQ concentrations in commercials sites were higher than public sites. However, in Miami, Banger et al. (2010) found the average BaP-EQ concentrations in commercial soils were lower than public soils (175 vs. 281 µg kg⁻¹).
In addition, BaP-EQ concentrations in Orlando and Tampa were compared with FSCTLs (Figure 2-3B). About 60-62% of urban soils from both cities had greater BaP-EQ concentrations than residential FSCTL at 100 µg kg⁻¹. In addition, 20-25% of urban soils had greater BaP-EQ concentrations than the commercial FSCTL at 700 µg kg⁻¹. The data indicated that FSCTL based on BaP-EQ are close to background concentrations. Therefore, both FSCTLs and background concentrations should be considered during soil remediation. Strong correlations between BaP-EQ and ∑16-PAHs in both cities (R² > 0.97) were found, similar to Richter-Brockmann and Achten (2018).

The spatial distributions of BaP-EQ concentrations in both cities are shown in Figure 2-2. Based on FSCTLs, three concentrations were used: <100, 100 – 700, and >700 µg kg⁻¹. Similar to ∑16-PAHs, high BaP-EQ concentrations were found in high PAHs sites, near downtown and highways as highlighted by dark blue areas in both cities. The data suggested that pyrogenic sources were dominant in both cities, especially vehicle emissions (Morillo et al., 2007).

Source Identification and Apportionment

Diagnostic ratios provide source identification information whereas PMF modeling helps with both source identification and apportionment of soil PAHs. We have applied both methods to obtain more accurate results as each method has its own shortcomings.

We selected six different diagnostic ratios based on literature (Table 2-2; Tobiszewski and Namie, 2012). The average ratios in both cities were similar with pyrogenic sources being dominant including vehicle emission, and biomass and coal combustion. In addition, benzo[a]anthracene/(chrysene + benzo[a]anthracene) ratios in
Orlando soils and benzo[a]pyrene/benzo[g,h,i]perylene ratios in both cities were away from the critical values, indicating that vehicle emissions and combustion sources were dominant sources (Vane et al., 2014). Moreover, the ArcGIS maps in both cities showed high PAH concentrations were near high-traffic roads, consistent with our results.

Molecular diagnostic ratios are useful to explain the sources of PAHs in soils. However, the method alone is insufficient due to its uncertainty, so PMF modeling was used. The PMF analysis was run in the default robust mode to decrease the influence of extreme values. Each run was initialized with different starting points. The 100 random starting points were selected by random seed mode to examine 3–10 factors. Four factors were determined as the most fitted circumstances among 3-7 factors based on the molecular markers in different sources (Li et al., 2016). The major sources were petrogenic and pyrogenic including combustion of coal and coke, and vehicle emissions in both cities. Specifically, petrogenic source was dominated by NA, ACY, ACE, FL, PH, and CH; biomass combustion was dominated by FL, PH, ANC, FLA, and PY; coal and coke combustion was dominated by BaA, CH, BkF, BaP, BP, and DhA; and vehicle emissions was dominated by FLA, PY, BaA, CH, BbF, BkF, BaP, IP, and BP (Figure 2-4). The results showed that the major sources of PAHs included vehicle emissions, and coal and coke combustion, accounting for ~70% of PAHs. Though with similar contributions from traffic in both cities, Tampa had more coal and coke combustion than Orlando, likely due to more coal and coke combustion plants.

In short, the results showed that vehicle emissions as well as coal and coke combustion, are the major sources of PAHs in urban soils, similar to other large cities (Li et al., 2016). By coupling diagnostic ratios with PMF modeling, based on PAH
composition profiles and PAH spatial distribution, the major sources of soil PAHs in both cities were from pyrogenic sources, similar to that in Miami soils (Banger et al., 2010).
Table 2-1. Descriptive statistics for PAH concentrations in urban soils of Orlando and Tampa

<table>
<thead>
<tr>
<th>Total PAHs</th>
<th>Orlando (µg kg(^{-1}))</th>
<th>Tampa (µg kg(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Range</td>
<td>Median</td>
</tr>
<tr>
<td>∑16-PAHs</td>
<td>43-30428</td>
<td>1640</td>
</tr>
<tr>
<td>∑7-CPAHs</td>
<td>41-20509</td>
<td>442</td>
</tr>
<tr>
<td>∑LMW-PAHs</td>
<td>224-7444</td>
<td>590</td>
</tr>
<tr>
<td>∑HMW-PAHs</td>
<td>182-33107</td>
<td>1735</td>
</tr>
<tr>
<td>∑7cPAHs</td>
<td>4-3742</td>
<td>141</td>
</tr>
<tr>
<td>Diagnostic ratio</td>
<td>Tampa mean</td>
<td>Orlando mean</td>
</tr>
<tr>
<td>-------------------------------------------------------</td>
<td>------------</td>
<td>--------------</td>
</tr>
<tr>
<td>∑ LMW-PAHs/∑ HMW-PAHs</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>Anthracene/(Phenanthrene + Anthracene)</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Fluorene/(Pyrene + Fluorene)</td>
<td>0.2</td>
<td>0.4</td>
</tr>
<tr>
<td>Fluoranthene/(Pyrene + Fluoranthene)</td>
<td>0.5</td>
<td>0.7</td>
</tr>
<tr>
<td>Benzo[a]anthracene/(Chrysene + Benzo[a]anthracene)</td>
<td>0.4</td>
<td>0.8</td>
</tr>
<tr>
<td>Benzo[a]pyrene/Benzo[g,h,i]perylene</td>
<td>1.9</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2-2. Six diagnostic ratios of PAHs in urban soils of Orlando and Tampa
Figure 2-1. Concentrations (A) and compositions (B) of 16 PAHs in urban soils of Orlando and Tampa. Carcinogenic PAHs are marked as red. Plots show the median, 75th, and 90th percentiles as vertical boxes with error bars.
Figure 2-2. $\sum_{16}$-PAHs (A) and $\sum_{7c}$PAHs (B) concentrations and spatial distributions in urban soils of Orlando and Tampa based on site-specific concentrations and IDW interpolation concentrations.
Figure 2-3. $\sum 7$cPAHs concentrations in Orlando and Tampa soils (A) and comparisons with Florida residential and industrial soil cleanup target levels based on BaP-EQ (B). Plots show the median, 75th, and 90th percentiles as vertical boxes with error bars.
Figure 2-4. The average mass contributions of each source to 16 PAHs in Tampa (a-outer circle) and Orlando (b-inner circle) urban soils based on USEPA PMF modeling.
CHAPTER 3
EMERGING PAHS IN URBAN SOILS: CONCENTRATIONS, BIOACCESSIBILITY, AND SPATIAL DISTRIBUTION

Materials and Methods

Soil Sampling and Preparation

Soil sampling and preparation methods followed Liu et al. (2019). Specifically, soils in the top 15-cm in two large cities (Orlando and Tampa, Florida, USA) were collected from March to May 2016. Half of the samples were from commercial and the other half were from public sites, totaling 50 in Orlando and 64 in Tampa. The commercial sites included supermarkets, shopping malls, restaurants, cinemas, and banks, and public sites include libraries, museums, government facilities, courts, and schools. The soils were sampled using stainless steel shovels, which were placed in aluminum foil wraps, air-dried at room temperature for 3 d, and sieved with a 2-mm sieve (Liu et al., 2019).

Chemicals and Sample Analysis

All chemicals were from Fisher Scientific (Hampton, NH), including 6 emerging PAHs (BcF, AN, DeP, DiP, DiP, and DhP) and their corresponding deuterated PAHs, 4 surrogates (o-terphenyl, 6-methylchrysene, decafluorobiphenyl, and perylene-d12), florisil, sodium sulfate anhydrous, and silica gel sorbent for pretreatment (Liu et al., 2019). All solvents and chemicals were of HPLC grade or better. All labware was made of glass or Teflon, which were washed with DI water, acetone, and n-hexane before use.

Pretreatment and analysis of 6 emerging PAHs followed the USEPA Method 3550C and 8270D. Specifically, soil sample of 5 g was spiked with surrogate PAHs in 20 mL 1:1 (v/v) acetone/n-hexane. After vortexing for 1 minute and ultrasonic treatment for 15 minutes, samples were centrifuged at 2000 g for 10 min. The soil samples were extracted three times and the combined extractant was cleaned up and solvent-exchanged using n-hexane to 1 mL (Liu et al., 2019).

A Trace Gas Chromatography (GC) Ultra Gas Chromatograph and DSQ2 Mass Spectrometer (MS; Thermo Scientific) was used. Identification of PAHs was based on selected ions and relative retention time between samples and standard solutions containing individual PAHs. Total organic carbon of the soils were determined by dry combustion at 900°C with a Shimadzu total organic carbon analyzer (TOC-VCPH) connected to a solid sample introduction module (SSM-5000, Shimadzu Corp. Kyoto, Japan). For detailed information, please see Liu et al. (2019).

Bioaccessibility Analysis

The bioaccessibility of 6 emerging PAHs were determined using the n-butanol mild solvent extraction method, which was correlated with the swine model (Duan et al., 2014). Specifically, 10 g soil sample was mixed with 15 mL of butanol by vortexing for 50 s. The sample was centrifuged at 2000 g for 10 min and 1 mL of the liquid phase was then passed through a 0.45-µm Teflon filter. The extract was transferred to an autosampler vial for GC/MS analysis (Gomez-Eyles et al., 2012). PAHs bioaccessibility was calculated using the following equation:
Bioaccessibility (%) = bioaccessible PAH / total PAH * 100% \hspace{1cm} (3-1)

where bioaccessible PAH is the amount of PAH (µg) extracted by n-butanol solvent, and total PAH (µg) is the portion of PAHs extracted by acetone/n-hexane extraction.

**Benzo[a]pyrene-equivalent Carcinogenicity**

Since carcinogenic PAHs share a common mechanism but display different carcinogenicity, their RPFs are applied to calculate their toxicity based on benzo[a]pyrene carcinogenicity (BaP-EQ). This approach has been used to convert all cPAHs in soils into a single concentration, i.e., BaP-EQ. In addition, this approach has been used to assess risks from direct exposure (Liu et al., 2019; Wang et al., 2015). The cancer risk for a PAH mixture of concern is determined by multiplying the concentration of each PAH by its RPF based on BaP-EQ.

In this research, 7 USEPA priority carcinogenic PAHs (7cPAHs) and 3 emerging carcinogenic PAHs (3cPAHs) were determined based on BaP-EQ (Table 1) (Richter-Brockmann and Achten, 2018; Wang et al., 2015):

$$
\sum 7\text{cPAHs} = 0.001[\text{CH}] + 0.01[\text{BkF}] + 0.1[\text{BaA}] + 0.1[\text{BbF}] + 0.1[\text{IP}] + [\text{DhA}] + [\text{BaP}] \\
\sum 3\text{cPAHs} = 0.4[\text{AN}] + 20[\text{BcF}] + 30[\text{DIP}] \hspace{1cm} (3-2)
$$

**Data Analyses and Quality Assurance**

The procedural and matrix blanks, spike, and sample duplicates were analyzed with samples, with no interference being detected. One duplicate and one laboratory blank were run for every six samples, with samples having difference >±15% being reanalyzed. The PAHs were quantified via external standard calibrations. Four surrogate standards were added to soil samples to monitor the procedures of sample
extraction, cleanup, and analysis. The recoveries of 4 surrogate PAHs including perylene-d12, 6-methylchrysene, o-terphenyl, and decafluorobiphenyl were 61-110%, with standard deviations being 7-15%, averaging 83, 80, 79 and 73%, respectively. The detection limit (0.9–3.6 µg kg\(^{-1}\)) was based on three times the signal-to-noise level for the blank sample with a sample size of 5 g and a final volume of 1 mL.

All statistical analyses were performed using JMP®Pro 13.0 statistical software (SAS Institute, Inc., Cary, NC, USA). ArcGIS 10.3 was applied to determine the spatial distribution of PAHs in soils (Liu et al., 2019).

**Results and Discussion**

In this study, the concentrations, bioaccessibility, and spatial distribution of 6 emerging PAHs (AN, BcF, DlP, DeP, DiP, and DhP) in 114 urban soils of two large cities were determined. In total, 50 samples were from Orlando and 64 from Tampa. Among 6 emerging PAHs, 3 are carcinogenic (3cPAHs), with their RPF being 0.4 for AN, 20 for BcF, and 30 for DlP (Table 1-1), so the concentrations of 3cPAHs were also calculated based on BaP-EQ (Table 3-1).

**Concentrations of Six Emerging PAHs**

The concentrations of 6 emerging PAHs in Orlando and Tampa soils are shown in Figure 3-1. The results showed that 6 emerging PAHs were found in 90% of the soil samples in both cities. The concentrations of the 3 cPAHs (BcF, AN, and DlP) were 0-2185, 0-5438, and 0-807 µg kg\(^{-1}\), averaging 293, 284, and 134 µg kg\(^{-1}\) in Orlando soils, respectively. Those for Tampa soils were 0-1026, 0-7562, and 0-2118 µg kg\(^{-1}\), averaging 204, 393, and 207 µg kg\(^{-1}\), respectively. So the 3 emerging cPAHs had similar ranges in both cities as most samples were < 1000 µg kg\(^{-1}\). Tampa soils had higher AN and DIP but lower BcF concentrations compared with Orlando soils. Although
AN’s concentrations were high in both cities, it is less carcinogenic than BcF and DIP based on their RPF, i.e., 0.4 vs. 20 and 30 (Table 1-1). Though most soils had PAHs concentrations at hundreds of µg kg\(^{-1}\), some were higher than thousands of µg kg\(^{-1}\) in both cities. In addition, the concentrations of AN, BcF, and DIP had strong correlations among them in both cities (\(R^2 > 0.8\)), indicating similar sources (data not shown).

In our study, a total of 114 samples were collected from urban soils in two large cities. In comparison, only one study determined the concentrations of AN, BcF, and DIP in 5 soils (Richter-Brockmann and Achten, 2018). They reported AN (0, 6550, and 31390 µg kg\(^{-1}\)), BcF (250, 8200, and 20900 µg kg\(^{-1}\)), and DIP (200, 6890, and 27120 µg kg\(^{-1}\)) concentrations in 3 contaminated soils. In addition, the concentrations of AN, BcF and DIP were 0-210, 150-440, and 250-550 µg kg\(^{-1}\) in 2 urban soils. The concentrations of AN, BcF, and DIP in 2 urban soils were similar to ours, but those in 3 contaminated soils were much higher (Figure 3-1).

**Concentrations of Three Emerging cPAHs based on BaP-EQ**

In addition to total concentrations of PAHs, we also determined cPAHs concentrations based on BaP-EQ (Table 3-1). The concentrations of \(\Sigma 3\)cPAHs for most soils (95%) in both cities were similar, being <20 mg kg\(^{-1}\). However, those of USEPA \(\Sigma 7\)cPAHs for most soils (95%) in both cities were <3 mg kg\(^{-1}\) (Figure 3-2). Specifically, the concentrations of \(\Sigma 3\)cPAHs and \(\Sigma 7\)cPAHs were 0-134 and 0.004-3.74 mg kg\(^{-1}\) in Orlando, averaging 7.39 and 0.45 mg kg\(^{-1}\). They were 0-87.1 and 0-9.41 mg kg\(^{-1}\), averaging 4.94 and 0.80 mg kg\(^{-1}\) for Tampa. So, the concentrations of \(\Sigma 3\)cPAHs were 6-16 times greater than \(\Sigma 7\)cPAHs. Soils from Orlando had lower \(\Sigma 7\)cPAHs but higher \(\Sigma 3\)cPAHs concentrations than Tampa. Based on Richter-Brockmann and Achten
(2018), the concentrations of $\sum$3cPAHs were 11-1244 mg kg$^{-1}$ in 3 contaminated soils and 10.6-25.3 mg kg$^{-1}$ in 2 urban soils. So the $\sum$3cPAHs concentrations in 3 contaminated soils were higher than those in this study, whereas those in 2 urban soils were in similar ranges (Figure 3-2; Table 3-1).

Florida Soil Cleanup Target Level for soils contaminated with PAHs is set at 700 $\mu$g kg$^{-1}$ for commercial sites, which is based on BaP-EQ concentration of USEPA $\sum$7cPAHs (Liu et al., 2019). Among 114 samples of urban soils, based on $\sum$7cPAHs, 20-25% samples exceeded 700 $\mu$g kg$^{-1}$, but based on $\sum$3cPAHs, 61-76% soil samples exceeded 700 $\mu$g kg$^{-1}$. When the two were added together, i.e., $\sum$10cPAHs, the sample numbers were even higher (Table 3-1). However, the carcinogenicity of 3cPAHs are still in debate and further research is needed before these data can be used in risk assessment of contaminated soils.

**Distributions of Three Emerging cPAHs based on BaP-EQ**

Like the data based on $\sum$7cPAHs, the commercial sites in both cities had higher concentrations of $\sum$3cPAHs than public sites (Figure 3-2). Similar trends were observed by Banger et al. (2010) and Liu et al. (2019). This may be due to the higher traffic volumes in commercial sites than in public sites (Azzolina et al., 2016). However, the much greater concentrations of $\sum$3cPAHs than $\sum$7cPAHs in both cities (Figure 3-2) showed that $\sum$3cPAHs had significant impacts on risk assessment due to their high RPFs (Table 1-1). The results indicated that risk assessment based on $\sum$7cPAHs concentrations may underestimate the risk of carcinogenic PAHs in soils. However, there are still uncertainties associated with the carcinogenicity of 3cPAHs (Richter-Brockmann and Achten, 2018).
In addition to their concentrations, the spatial distributions of $\Sigma3cPAHs$ concentrations in both cities are shown in Figure 3-3. The concentrations of $\Sigma3cPAHs$ were higher than those of $\Sigma7cPAHs$ in 75-86% soils in both cities, especially at sites near downtown areas. The maps showed that high $\Sigma3cPAHs$ concentrations were near downtown and busy highways in both cities. In addition, strong correlations were found between $\Sigma7cPAHs$ and $\Sigma3cPAHs$ in both cities ($R^2 > 0.8$), again indicating similar sources. Pyrogenic sources were the primary sources of 7cPAHs in both cities, the data suggested that pyrogenic sources were probably the dominant source for 3cPAHs as well (Liu et al., 2019).

**Bioaccessible Concentrations of Three Emerging cPAHs based on BaP-EQ**

Though high concentrations of total PAHs have been detected in many soil samples, their presences in soils do not always mean toxicity due to their low bioavailability. Animal models have been used to determine bioavailable PAHs in soils, but they are costly and time-consuming. In this study, we used n-butanol extraction to determine bioaccessible PAHs, which has been correlated with animal models and can be used to predict bioavailable PAHs (Duan et al., 2014; Gomez-Eyles et al., 2012).

In general, bioaccessibility of PAHs in all soils was low, averaging 3.4 to 11% (Table 3-1). The bioaccessibility of 3cPAHs and 7cPAHs in Orlando soils were 0.2-33% and 0.1-75%, averaging 3.4 and 11%, respectively (Table 3-1). Those for Tampa soils were 0.3-64% and 0.2-70%, averaging 7.4 and 11%, respectively. The results showed that, though the concentrations of $\Sigma3cPAHs$ were much greater than those of $\Sigma7cPAHs$, their bioaccessibility was lower than that of 7cPAHs (Table 2). Overall, the average bioaccessibility was $<11\%$, similar to other research (Li et al., 2017; Zhang et
al., 2017). In addition, we also determined the total OC in Orlando and Tampa soils, which were at 0.7-8.1% and 0.8-6.9%. Since total OC and the bioaccessibility of 7cPAHs were similar in two cities while that of 3cPAHs differed more, it may suggest that the bioaccessibility of 7cPAHs were more related to the total OC than 3cPAHs.

Bioaccessible $\Sigma 3cPAHs$ and $\Sigma 7cPAHs$ in Orlando soils were 0-1252 and 0-361 $\mu g$ kg$^{-1}$, averaging 47 and 20 $\mu g$ kg$^{-1}$, respectively. Those for Tampa soils were 0-272 and 0-90 $\mu g$ kg$^{-1}$, averaging 85 and 20 $\mu g$ kg$^{-1}$, respectively (Table 3-1 & Figure 3-2). Unlike total PAH concentrations, bioaccessible $\Sigma 3cPAHs$ and $\Sigma 7cPAHs$ in 95% soils were similar, being <200 $\mu g$ kg$^{-1}$. However, bioaccessible $\Sigma 3cPAHs$ were still 2-4 times higher than those of $\Sigma 7cPAHs$ in both cities. Like total concentrations, commercial sites had higher bioaccessible PAHs concentrations than public sites (Figure 3-2). However, no correlation was found between total and bioaccessible concentrations based on BaP-EQ in both cities.

Among 114 soils, 20-25% samples exceeded 700 $\mu g$ kg$^{-1}$ based on Bap-EQ of $\Sigma 7cPAHs$ concentrations, but 61-76% samples exceeded 700 $\mu g$ kg$^{-1}$ based on Bap-EQ of $\Sigma 3cPAHs$ concentrations. However, when bioaccessible $\Sigma 3cPAHs$ concentrations were used, only 2 samples in Orlando exceeded 700 $\mu g$ kg$^{-1}$ (Table 3-1). The data indicated that it is important to consider not only the 3 emerging cPAHs but also their bioaccessibility in assessing their risk in contaminated soils.
Table 3-1. Total and bioaccessible concentrations of 3 emerging carcinogenic PAHs (3cPAHs) and USEPA 7 priority carcinogenic PAHs (7cPAHs) in urban soils of Orlando and Tampa based on BaP-EQ

<table>
<thead>
<tr>
<th>Total PAHs</th>
<th>Orlando (µg kg⁻¹ BaP-EQ)</th>
<th>Tampa (µg kg⁻¹ BaP-EQ)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Range</td>
<td>Median</td>
</tr>
<tr>
<td>∑3cPAHs</td>
<td>0-133859</td>
<td>4612</td>
</tr>
<tr>
<td>∑7cPAHs</td>
<td>4-3742</td>
<td>141</td>
</tr>
<tr>
<td>∑10cPAHs</td>
<td>4-137601</td>
<td>4655</td>
</tr>
<tr>
<td>Bioaccessible PAHs</td>
<td>Range (%)</td>
<td>Orlando (µg kg⁻¹ BaP-EQ)</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>Median</td>
</tr>
<tr>
<td>∑3cPAHs</td>
<td>0-1252 (0.2-33)</td>
<td>23</td>
</tr>
<tr>
<td>∑7cPAHs</td>
<td>0-361 (0.1-75)</td>
<td>11</td>
</tr>
<tr>
<td>∑10cPAHs</td>
<td>0-1613</td>
<td>35</td>
</tr>
</tbody>
</table>
Figure 3-1. Concentrations of 6 emerging PAHs in urban soils of Orlando (A) and Tampa (B). The first three are carcinogenic PAHs (red). Plots show the median, 75th, and 90th percentiles as vertical boxes with error bars.
Figure 3-2. Total and bioaccessible 7cPAHs and 3cPAHs concentrations in urban soils of Orlando (AB) and Tampa (CD) based on BaP-EQ. Plots show the median, 75th, and 90th percentiles as vertical boxes with error bars.
Figure 3-3. Spatial distributions of $\Sigma 3$PAHs concentrations based on BaP-EQ in urban soils of Orlando and Tampa. Light blue circles represent the central business districts.
CHAPTER 4
EMERGING AND LEGACY PAHS IN URBAN SOILS OF FOUR SMALL CITIES:
CONCENTRATIONS, DISTRIBUTION, AND SOURCES

Materials and Methods

Soil Sampling and Preparation

To determine the concentrations of PAHs in urban soils, we selected four representative small cities in Florida: Clay County, Ocala, Pensacola, and West Palm Beach. Clay County has a population of 212,230 and an area of 644 square miles; Ocala has a population of 59,110 and an area of 46 square miles; Pensacola has a population of 52,590 and an area of 41 square miles; and West Palm Beach has a population of 110,222 and an area of 58 square miles. In contrast, large cities have a larger population and larger areas, such as Orlando has a population of 277,173 and an area of 111 square miles and Tampa has a population of 377,165 and an area of 171 square miles (Liu et al., 2019). To minimize point source pollution, we have developed exclusion criteria. We excluded: 1) residential areas; 2) floods or runoff in low-lying areas; 3) <3 m of paved roads, parking lots, painted structures and visually contaminated soil; 4) <10 m of gas station, railway, stormwater collection ponds, and substations; 5) <50 m incineration, wastewater or solid waste treatment facilities; 6) <100 m industrial characteristics or identified hazardous waste sites (Gao et al., 2019a).

From May to July 2018, 26 soil samples (13 commercial and 13 public places) in Clay County and West Palm Beach, and 24 soil samples (12 Commercial and 12 public spaces) from Pensacola and Ocala were collected according to a random sampling

plan. Commercial sites include restaurants, shopping malls, supermarkets, cinemas and
banks, and public places include government facilities, courts, libraries, museums, and
schools. We used a composite method to sample the top 15 cm of soil. At each site, we
created a 5 m x 5 m grid to make 1 composite sample from the 4 corners and 1 center
subsamples. Soil samples were taken using stainless steel shovels and placed in
aluminum foil wraps, which were air dried for 3 days at room temperature and passed
through a 2-mm sieve prior to analysis.

**Chemicals and Sample Analysis**

All chemicals were from Fisher Scientific (Hampton, NH, USA) including 22 PAHs
and their corresponding deuterated PAHs, 4 surrogates (o-terphenyl,
decafluorobiphenyl, 6-methylchrysene, and perylene-d12), florisil, sodium sulfate
anhydrous, and silica gel sorbent for pretreatment. All solvents and chemicals were
HPLC grade or higher. All labware was made of glass or Teflon, which was washed by
deionized water, acetone, and n-hexane prior to use. Pretreatment and analyses of
PAHs followed the improved USEPA Methods 3550C and 8270D. Specifically, ~5 g of
soil sample was added to the glass centrifuge tube with surrogate PAHs, and then 20
mL 1:1 (v/v) acetone/n-hexane was added as the extractant, and the blank was free of
soils. Each sample was vortexed for 1 minute and sonicated for 15 minutes. The sample
was centrifuged at 2000 rpm for 10 minutes at room temperature. The soil samples
were extracted three times, the combined extractants were cleaned up and solvent
exchanged by n-hexane to 1 mL.

Instrumental analysis was performed using a Trace Gas Chromatography (GC)
Ultra Gas Chromatograph and DSQ2 Mass Spectrometer (MS; Thermo Scientific). The
GC column consisted of 30 m × 0.25 mm i.d. in 0.25 μm film thickness. A 1 μL aliquot of
the sample extract was injected in a splitless mode. The oven temperature was set at 60°C for 1 minute, then the temperature was raised to 120°C at 25°C/min, 160°C at 10°C/min, and 330°C at 5°C/min with 1-min final hold time. The inlet temperature was 320°C, the ion source temperature was 250°C, the transfer line temperature was 300°C, and the helium carrier gas was maintained at 0.3 mL/min. Ionization was performed using the electron ionization mode and fragments were obtained using a selective ion mode. The identifications of PAHs were based on characteristic ions and relative retention time between samples and the standards.

**Benzo[a]pyrene-equivalent Carcinogenicity**

Since carcinogenic PAHs share a common mechanism but have different carcinogenicity, their RPFs (relative potency factor) are used to determine their carcinogenicity based on BaP-EQ. This method has been used to convert all cPAH in a soil to a single concentration, namely BaP-EQ. In addition, this method has been used to assess the risk of direct exposure from PAHs in soils. The cancer risk of the PAH mixture of interest was determined by multiplying the concentration of each PAH by its RPF based on BaP-EQ (Teaf, 2008). In this study, the BaP-EQs of seven USEPA legacy carcinogenic PAHs (7cPAHs) and three emerging carcinogenic PAHs (3cPAHs) were calculated (Gao et al., 2019a).

**Data Analyses and Quality Assurance**

The procedural, spike, and matrix blanks, and sample duplicates were analyzed routinely with no interference being detected. Since the background values were below detection limits, no subtraction was needed. One laboratory blank and one replicate were run with every six samples, with samples having >±15% difference being reanalyzed. The PAHs were quantified by external standard calibrations with standard
PAHs as well as internal standards with deuterated PAHs. Four surrogate standards were added to the soil samples to monitor the procedures for sample extraction, cleanup, and analyses. The surrogate recoveries of perylene-d12 (5 rings), 6-methylchrysene (4 rings), o-terphenyl (3 rings), and decafluorobiphenyl (2 rings) were 61-110%, and the standard deviations were 7-15%, with an average of 83%, 80%, 79%, and 73%, respectively. The detection limits (0.9–3.6 µg kg\(^{-1}\)) were based on three times the signal-to-noise levels of the chromatogram for the blank sample with 5 g sample size and 1 mL final volume.

All statistical analyses were performed using JMP\textsuperscript{®} Pro13.0 and SigmaPlot 14 statistical software. Correlation analysis and t-test for multiple comparisons were applied. The inverse distance weighting (IDW) interpolation was used to estimate the distribution of soil PAHs based on ArcGIS 10.3. In addition to IDW, diagnostic ratios and USEPA PMF 5.0 model were also used to trace PAH sources. The explanation of replicates and statistical methods can be found in Wang et al. (2015), Yu et al. (2015) and Vane et al. (2014).

**Results and Discussions**

So far, two studies have investigated the concentrations of 16 USEPA legacy PAHs in large cities in Florida, USA (Orlando, Tampa, and Miami), however, their concentrations in small cities are unknown (Banger et al., 2010; Liu et al., 2019). In addition, two studies focused on 6 emerging PAHs in urban soils. While Gao et al. (2019) investigated the concentrations of 6 emerging PAHs in 114 urban soils of large cities (Orlando and Tampa), Richter-Brockmann and Achten (2018) determined AN, BcF, and DIP concentrations in 5 soil samples. However, their concentrations in small cities are unknown.
Concentrations of Sixteen Legacy PAHs and Six Emerging PAHs

The concentrations of 16 legacy PAHs in urban soils of four small cities varied greatly (Figure 4-1). Among 4 cities, PAHs in West Palm Beach had the highest concentrations, especially the 3-5 ring PAHs. Most soils had PAH concentrations at several hundred µg kg⁻¹, but some of them were high, especially those in West Palm Beach. Compared to Pensacola and Ocala, West Palm Beach has higher annual average daily traffic (AADT) and population (AADT: 12200 vs. 8200 & 5500; population: 110222 vs. 52590 & 59110). The data are consistent with the highest concentrations of PAHs in urban soils of West Palm Beach.

The descriptive statistics of 22 PAHs concentrations are presented in Table 4-1. The average concentrations of ∑16-PAHs in urban soils in Clay county, Ocala, Pensacola, and West Palm Beach were 1821, 2748, 3115, and 4055 µg kg⁻¹, which were lower than those in Orlando and Tampa. This was probably due to more human activities in large cities, with anthropogenic sources being the major sources of PAHs in urban soils (Liu et al., 2019). In addition, the highest ∑16-PAHs in urban soils in Clay county, Ocala, Pensacola, and West Palm Beach were 7909, 11451, 17698 and 30691 µg kg⁻¹, respectively. Overall, the ∑16-PAHs in 95% of soils samples from 4 small cities were < 10000 µg kg⁻¹ while those in 95% of soils samples from Orlando and Tampa were < 30000 µg kg⁻¹ (Liu et al., 2019). As a result, the ∑16-PAHs concentrations in small cities were lower than those in large cities in Florida. Based on the PAHs compositional profiles (Figure 4-1), HMW-PAHs were dominants in all samples in small cities, especially 4-ring PAHs, similar to Orlando, Tampa, and Miami soils (Banger et al., 2010; Liu et al., 2019). These results indicate that PAHs in all cities in Florida were dominated by pyrogenic sources.
The concentrations of 6 emerging PAHs are shown in Figure 4-1. The results showed that all 6 emerging PAHs were detected in samples, similar to those in Orlando and Tampa (Gao et al., 2019a). The concentrations of BcF, AN, and DlP (3 emerging cPAHs) were 32-555, 52-561, and 52-248 µg kg\(^{-1}\) in Clay county. Those for Ocala were 34-219, 57-794, and 53-350 µg kg\(^{-1}\); Pensacola were 34-555, 56-886, and 54-465 µg kg\(^{-1}\) and West Palm Beach were 35-552, 54-1585, and 54-628 µg kg\(^{-1}\). In comparison, their concentrations were 1.8–2185, 1.8–5438, and 1.8–807 µg kg\(^{-1}\) in Orlando soils and 1.8–1026, 1.8–7562, and 1.8–2118 µg kg\(^{-1}\) in Tampa soils (Gao et al., 2019a).

Similar to 7cPAHs, urban soils in small cities had lower concentrations of 3 cPAHs than those in large cities in Florida. Richter-Brockmann and Achten (2018) determined the concentrations of AN, BcF and DlP in two urban soils at 0-210, 150-440, and 250-550 µg kg\(^{-1}\), which were similar to our results.

Similar ranges of 3cPAHs were found in urban soils of four small cities as most samples were < 300 µg kg\(^{-1}\), but most samples in large cities (Tampa and Orlando) were higher at < 1000 µg kg\(^{-1}\) (Gao et al., 2019a). Though PAHs concentrations were of several hundred µg kg\(^{-1}\) in most soils, few were high, especially in West Palm Beach. Like legacy PAHs, urban soils in West Palm Beach had the highest 3cPAHs concentrations among 4 cities. However, they were 2-3 times lower than those in Orlando and Tampa. Although AN concentrations were relatively high in all soils, it is less carcinogenic than BcF and DlP based on their RPFs, namely, 0.4 vs. 20 and 30, similar to Orlando and Tampa soils (Gao et al., 2019a).

**BaP-EQ of Seven Legacy cPAHs and Six Emerging cPAHs**

PAHs concentrations in soils are often converted to benzo[a]pyrene-equivalent (BaP-EQ) when comparing with soil cleanup standards. This approach has been widely
used to determine PAHs carcinogenic risks in various studies (Banger et al., 2010; Wang et al., 2015). BaP-EQ concentrations in urban soils of four small cities are shown in Figure 4-2. The $\sum$7cPAHs concentrations were similar, being <700 µg kg$^{-1}$ in most soils. Nevertheless, soils from West Palm Beach again had the highest $\sum$7cPAHs concentrations among the 4 small cities (Table 4-1; Figure 4-2). The $\sum$7cPAHs concentrations were 83-950, 125-1572, 102-1537 and 116-3200 µg kg$^{-1}$ in urban soils in Clay county, Ocala, Pensacola and West Palm Beach (Table 4-1).

The $\sum$7cPAHs concentrations in public sites were significantly lower than commercials sites in Clay County (p<0.05) possibly due to their higher traffics (Azzolina et al., 2016) (Figure 4-2). In addition, FSCTLs were compared with $\sum$7cPAHs concentrations (Figure 4-3A). Their concentrations in 70-89% soil samples from small cities exceeded the residential FSCTL at 100 µg kg$^{-1}$ BaP-EQ. Besides, their concentrations in 11-13% soil samples exceeded the industrial FSCTL at 700 µg kg$^{-1}$ BaP-EQ. The results indicated that FSCTLs are similar to the background concentrations of PAHs, like large cities Orlando and Tampa. However, fewer samples were higher than industrial FSCTL, indicating small cities had lower $\sum$7cPAHs concentrations than large cities in Florida.

The $\sum$7cPAHs concentrations and spatial distributions in urban soils of four small cities based on site-specific concentrations and IDW interpolation concentrations are shown in Figure 4-4. The sites were divided into three categories (<100, 100–700, and >700 µg kg$^{-1}$) based on FSCTLs. Highways with AADT >50000 are identified as high traffic roads (Liu et al., 2019). Like Orlando and Tampa, the highest $\sum$7cPAHs concentrations were in downtown areas in small cities and the IDW results showed
increasing $\sum_7$PAHs concentrations from rural to urban areas (Figure 4-4). Most sites having high $\sum_7$PAHs were close to high traffic roads, suggesting vehicle emission was one of the dominant sources in small cities like large cities Orlando and Tampa.

Among 6 emerging PAHs, 3 are carcinogenic (AN, BcF, and DlP RPFs: 0.4, 20, and 30) so their concentrations were also determined based on BaP-EQ as $\sum_3$PAHs (Table 4-1; Figure 4-2). The $\sum_3$PAHs concentrations in most soils (95%) in four small cities were similar, being $\leq 2000 \mu g kg^{-1}$ and $\sim 10$ times lower than those in Orlando and Tampa soils (Gao et al., 2019a). In contrast, the $\sum_7$PAHs concentrations in most soils (95%) in four cities were $< 1000 \mu g kg^{-1}$, which is half of $\sum_3$PAHs (Figure 4-2). As such, the differences between $\sum_3$PAHs and $\sum_7$PAHs in four small cities were less than those in Orlando and Tampa. Specifically, the concentrations of $\sum_3$PAHs and $\sum_7$PAHs averaged 223 and 3703, 319 and 4521, 302 and 5423 and 449 and 5916 $\mu g$ kg$^{-1}$, respectively, for urban soils in Clay county, Ocala, Pensacola, and West Palm Beach. The concentrations of $\sum_3$PAHs were 13-18 times greater than $\sum_7$PAHs due to their similar concentrations but higher RPFs of 3 emerging cPAHs, with the differences being greater than those in Orlando and Tampa soils (6-16 times). In addition, soils from West Palm Beach again had both higher $\sum_7$PAHs and $\sum_3$PAHs concentrations than others (Table 4-1; Figure 4-2). In comparison, the concentrations of $\sum_3$PAHs and $\sum_7$PAHs averaged 7387 and 452 $\mu g kg^{-1}$ for Orlando and 4943 and 802 $\mu g kg^{-1}$ for Tampa (Gao et al., 2019a). As a result, small cities had both lower $\sum_7$PAHs and $\sum_3$PAHs concentrations than large cities in Florida. In addition, $\sum_3$PAHs concentrations were 1060-2530 $\mu g kg^{-1}$ in two urban soils based on Richter-Brockmann and Achten (2018), similar to those in Orlando and Tampa soils (Table 4-1; Figure 4-2).
Based on the BaP-EQ concentration of USEPA $\Sigma$7cPAHs, FSCTL for industrial sites in contaminated soils is at 700 µg kg$^{-1}$. Among 100 soil samples from four small cities, 11-13% samples exceeded 700 µg kg$^{-1}$ based on $\Sigma$7cPAHs, but all exceeded 700 µg kg$^{-1}$ based on $\Sigma$3cPAHs, which were worse than Orlando and Tampa soils (61-76%; Gao et al., 2019a). However, the $\Sigma$3cPAHs in small cities were much lower than those in large cities in Florida. Further research is needed before the results can be used for risk assessment in contaminated soils.

Combing the data of both large and small cities, the concentrations of $\Sigma$3cPAHs was much higher than that of $\Sigma$7cPAHs in all cities, indicating that $\Sigma$3cPAHs had a significant impact on risk assessment. However, there are still uncertainties about the carcinogenicity and bioavailability of emerging cPAHs. Though emerging PAHs had higher concentrations than legacy PAHs in urban soils, their oral bioaccessibility was relatively low, averaging 3.4–7.4% in Orlando and Tampa (Gao et al., 2019a).

Source Identification and Apportionment

Diagnostic ratios provide source identification information and average ratios can represent the major sources, while PMF model facilitates source identification and apportionment of PAHs in soils. We applied both methods to obtain more accurate results since each method has its own pros and cons (Gao et al., 2018a).

We chose six different diagnostic ratios according to literature (Table 4-2; Tobiszewski and Namie, 2012). The results of average diagnostic ratios showed that the major sources in four small cities were similar, being dominated by pyrogenic sources including petroleum emission, vehicle emission, and coal and biomass combustions. In addition, the ratios were around the critical values, indicating that more
mixed anthropogenic sources were dominant than soils in large city like Orlando and Tampa (Vane et al., 2014; Su et al., 2019). The ArcGIS maps also showed some high PAHs concentrations were not near high-traffic roads, being consistent with these results (Figure 4-4).

Molecular diagnostic ratios can be applied to interpret the sources of PAHs in soils. Nevertheless, due to its uncertainty, one approach alone is insufficient, so the PMF model was applied. The model ran in the default robust mode to reduce the effects of extreme values. One hundred random starting points were chosen by random seed mode. Based on molecular markers from different sources, four of the 3-7 factors are the most suitable environment (Yu et al., 2015; Wang et al., 2015). The results indicated that the main sources of PAHs in all small cities were pyrogenic sources including coal, coke, and biomass combustions, and vehicle emissions (Figure 4-3B). Specifically, PAHs from petrogenic sources included NA, ACY, ACE, FL, PH, and CH. In addition, FL, PH, ANC, FLA, and PY were mainly from biomass combustions, BaA, CH, BkF, BaP, BP and DhA coal and coke combustions, and FLA, PY, BaA, CH, BbF, BkF, BaP, IP and BP vehicle emissions.

The results indicated that anthropogenic sources of PAHs included biomass, coal and coke combustion, and vehicle emissions, contributing 68-89% of PAHs (Figure 4-3B). Specifically, PAHs sources in urban soils of Pensacola were dominated by vehicle emissions, those in Ocala and West Palm Beach were dominated by biomass combustion, and those in Clay county were dominated by petrogenic sources, which is consistent with its lowest PAHs concentrations among four cities (Table 4-1; Figure 4-1). Although the contributions from traffic emissions in all cities were similar, coal and
coke combustions contributed more in Pensacola, Orlando, and Tampa than in other cities, mainly due to the presence of more coal and coke burning plants. In addition, large cities had more contributions from biomass combustion because they have more biomass-to-energy facilities. The fact that West Palm Beach generates more electricity from nuclear power plants while coal-fired power plants in other small cities such as Pensacola and Clay County may contribute more coal and coke sources (Banger et al., 2010; Liu et al., 2019).

Furthermore, in all small cities, $\sum 7cPAHs$ and $\sum 3cPAHs$ were correlated ($R^2 > 0.7$) and $3cPAHs$ also correlated with each other ($R^2 > 0.7$), indicating similar sources. By combining diagnostic ratios with the PMF model, based on the PAH compositional profiles and PAHs spatial distributions, the main sources of PAHs in urban soils of four small cities came from pyrogenic sources. In addition, the results indicated that biomass, coal and coke combustions, as well as vehicle emissions, were the dominant sources of PAHs in small cities soils, similar to large cities (Banger et al., 2010; Liu et al., 2019).
Table 4-1. Descriptive statistics for PAH concentrations in urban soils of four small cities

<table>
<thead>
<tr>
<th>Total PAHs</th>
<th>Range</th>
<th>Clay County (µg kg(^{-1}))</th>
<th>West Palm Beach (µg kg(^{-1}))</th>
<th>Ocala (µg kg(^{-1}))</th>
<th>Pensacola (µg kg(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Range</td>
<td>Median</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Σ16-PAHs</td>
<td>797-7909</td>
<td>1085</td>
<td>1821</td>
<td>5873</td>
<td>4055</td>
</tr>
<tr>
<td>Σ7cPAHs</td>
<td>83-950</td>
<td>118</td>
<td>223</td>
<td>617</td>
<td>449</td>
</tr>
<tr>
<td>Σ3cPAHs</td>
<td>2267-16161</td>
<td>2687</td>
<td>3703</td>
<td>5916</td>
<td>617</td>
</tr>
<tr>
<td>Total PAHs</td>
<td>1133-30691</td>
<td>2262</td>
<td>4055</td>
<td>3155</td>
<td>1444</td>
</tr>
<tr>
<td>Σ16-PAHs</td>
<td>950-11451</td>
<td>1444</td>
<td>2748</td>
<td>5873</td>
<td>201</td>
</tr>
<tr>
<td>Σ7cPAHs</td>
<td>125-1572</td>
<td>201</td>
<td>319</td>
<td>2808</td>
<td>319</td>
</tr>
<tr>
<td>Σ3cPAHs</td>
<td>2334-13511</td>
<td>3212</td>
<td>4521</td>
<td>3155</td>
<td>3212</td>
</tr>
<tr>
<td>Total PAHs</td>
<td>922-17698</td>
<td>1326</td>
<td>3115</td>
<td>4111</td>
<td>137</td>
</tr>
<tr>
<td>Σ16-PAHs</td>
<td>102-1537</td>
<td>137</td>
<td>302</td>
<td>354</td>
<td>137</td>
</tr>
<tr>
<td>Σ7cPAHs</td>
<td>2349-21261</td>
<td>3325</td>
<td>5423</td>
<td>4948</td>
<td>3325</td>
</tr>
</tbody>
</table>

62
Table 4-2. Six diagnostic ratios of PAHs in urban soils of four small cities (Tobiszewski and Namie, 2012; Liu et al., 2019)

<table>
<thead>
<tr>
<th>Diagnostic ratio</th>
<th>Clay County mean</th>
<th>West Palm Beach mean</th>
<th>Ocala mean</th>
<th>Pensacola mean</th>
<th>Critical range</th>
</tr>
</thead>
<tbody>
<tr>
<td>∑ LMW-PAHs/∑ HMW-PAHs</td>
<td>0.6</td>
<td>0.4</td>
<td>0.5</td>
<td>0.4</td>
<td>&lt;1 pyrogenic &gt;1 petrogenic</td>
</tr>
<tr>
<td>Anthracene/(Phenanthrene + Anthracene)</td>
<td>0.5</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>&lt;0.1 petrogenic &gt;0.1 pyrogenic</td>
</tr>
<tr>
<td>Fluorene/(Pyrene + Fluorene)</td>
<td>0.5</td>
<td>0.3</td>
<td>0.4</td>
<td>0.4</td>
<td>&lt;0.5 petroleum emission &gt;0.5 diesel emission</td>
</tr>
<tr>
<td>Fluoranthene/(Pyrene + Fluoranthene)</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>&lt;0.4 petrogenic &gt;0.5 vehicle emission</td>
</tr>
<tr>
<td>Benzo[a]anthracene/(Chrysene + Benzo[a]anthracene)</td>
<td>0.4</td>
<td>0.2</td>
<td>0.3</td>
<td>0.3</td>
<td>&lt;0.2 petrogenic &lt;0.4 biomass/coal combustion &gt;0.4 vehicle emission</td>
</tr>
<tr>
<td>Benzo[a]pyrene/Benzo[g,h,i]perylene</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.7</td>
<td>&lt;0.6 non-vehicle emission &gt;0.6 vehicle emission</td>
</tr>
</tbody>
</table>
Figure 4-1. Concentrations of 22 PAHs in urban soils of four small cities. 7 legacy cPAHs are marked as red and 3 emerging cPAHs are marked as pink. Plots show the median, 75th, and 90th percentiles as vertical boxes with error bars.
Figure 4-2. Σ7cPAHs and Σ3cPAHs concentrations in urban soils of four small cities. Plots show the median, 75th, and 90th percentiles as vertical boxes with error bars.
Figure 4-3. Comparisons with Florida residential and industrial soil cleanup target levels based on $\Sigma$7cPAHs (A) and the average contributions of each source to 22 PAHs in urban soils based on USEPA PMF model (B) (Liu et al., 2019).
Figure 4-4. $\sum c$PAHs concentrations and spatial distributions in urban soils of four small cities based on site-specific concentrations and IDW interpolation concentrations.
BACKGROUND

Background concentrations, compositional profiles, benzo[a]pyrene-equivalent concentrations, spatial distributions, and source identification and apportionment of 16 legacy PAHs in urban soils were investigated. Total PAH concentrations in two Florida large cities (Tampa and Orlando) were similar to other large cities (Miami). BaP-EQ concentrations in most soils were lower than FSCTLs. Diagnostic ratios and PMF modeling indicated that the major sources of soil PAHs were pyrogenic, including vehicle emissions, and coal and coke combustion. PAH compositional profiles and ArcGIS maps also showed traffic impacted PAH concentrations and distributions in urban soils. The data suggested that traffic control as well as reduction in coal and coke combustions can help to reduce PAH accumulation in urban soils.

Besides the 16 legacy PAHs, total and bioaccessible concentrations based on Bap-EQ, and spatial distributions of 6 emerging PAHs in urban soils of two large cities were also investigated (Tampa and Orlando). The 3 emerging carcinogenic PAHs (anthanthrene, 7H-benzo[c]fluorene, and dibenzo[a,l]pyrene) were common in Florida urban soils. In addition, ArcGIS maps showed that anthropogenic sources were the major sources of 3 carcinogenic PAHs. Although the concentrations of those emerging

---


cPAHs were high (up to 1344 mg kg\(^{-1}\)), their bioaccessibility were low (up to 11%). Therefore, on one hand, we may underestimate the risks of PAHs when only considering 16 USEPA priority PAHs as some cPAHs are not included. On the other hand, we may overestimate their risks when using their total concentrations instead of bioaccessible concentrations due to their low bioaccessibility. In short, further research should focus on emerging PAHs in urban soils as well as their bioavailability.

Beside large cities, total and BaP-EQ concentrations, spatial distributions, and source identification of 22 PAHs in urban soils of four small cities in Florida were determined (Clay County, Ocala, Pensacola, and West Palm Beach). The data showed that the 3 emerging cPAHs (BcF, AN, and DIP) were detected in urban soils of four small cities, with \(\sum 3\)cPAHs concentrations being higher than \(\sum 7\)cPAHs in all samples. Background PAH concentrations in small cities were lower than those of large cities (Tampa and Orlando). The \(\sum 7\)cPAHs concentrations in most small cities soils did not exceed FSCTLs. Diagnostic ratios and PMF model showed that the main sources of PAHs in those soils were pyrogenic, including coal, biomass and coke combustion and vehicle emissions. The compositional profile of the PAHs and the ArcGIS map also indicated that human activities affected the PAHs concentration and distribution in urban soils of small cities. The results implied that traffic control and reductions in coal, biomass and coke combustions help to reduce the accumulation of both legacy and emerging PAH in urban soils. However, further research on the occurrence, bioavailability, and carcinogenicity of emerging carcinogenic PAHs are needed.
Identification of PAH Sources

Among environmental matrices, soil is the major sink for accumulation and degradation of PAHs (Gan et al., 2009; Lau et al., 2010). As such, much research has focused on the identification and apportionment of PAHs sources in soils. Three methods have been used, including 1) molecular diagnostic ratios, 2) receptor models and 3) stable C isotopic signatures (Yu et al., 2015; Yuan et al., 2015; Bi et al., 2016; Devi et al., 2016; Dudhagara et al., 2016; Montuori et al., 2016; Suman et al., 2016). For accurate source identification and apportionment of PAHs, two or three methods have often been applied (Wang et al., 2015a, 2015b; Cai et al., 2017).

Molecular Diagnostic Ratios

Analysis of molecular diagnostic ratios is the most traditional and common technique to determine PAHs sources in the environment (Tobiszewski and Namie, 2012; Vane et al., 2014; Zheng et al., 2014; Clement et al., 2015; Devi et al., 2016). As PAHs are often emitted as a group of compounds, which vary with emission source, the relative concentration ratios of different PAHs can be used to trace their sources (Tobiszewski and Namie, 2012). During low-temperature (T) processes (e.g., wood burning), low molecular weight (LMW) PAHs (2-3 rings) are often formed whereas higher molecular weight (HMW) PAHs (>3 rings) are emitted during high-T processes (e.g., fuel combustion), resulting in different molecular diagnostic ratios of PAHs. At high
temperatures, organic C produces reactive radicals to form stable PAHs, which are less alkylated with more aromatic rings than petrogenic PAHs (Tobiszewski and Namie, 2012; Stogiannidis and Laane, 2015).

The molecular ratios of different PAHs have been used to trace the sources of PAHs in soils (Banger et al., 2010). For example, the ratios of LMW-PAHs containing 2–3 benzene rings to HMW-PAHs containing >3 benzene rings have been used in order to distinguish the dominances of petrogenic and pyrogenic sources (Table A-1). In addition, the ratios of individual PAHs can also be used to trace their sources. For example, the ratios of fluoranthene to fluoranthene + pyrene at <0.4 and >0.4 indicate petrogenic and pyrogenic sources (Table A-1). Similarly, phenanthrene to anthracene ratio of <10 is indicative of pyrogenic sources and >10 indicates petrogenic sources of PAHs (Tobiszewski and Namie, 2012). This fingerprinting technique can be based on parent PAHs or their derivatives. The PAHs used to calculate the diagnostic ratios are stable with similar molecular weight, therefore having similar properties. However, some studies showed that the ratios of PAHs may change during biogeochemical processes in soils such as chemical transformation and biodegradation (Tobiszewski and Namie, 2012; Stogiannidis and Laane, 2015).

Although this technique has been used to trace soil PAHs, it may only be applicable to stable PAHs, which don’t change during weathering and degradation processes (Tobiszewski and Namie, 2012). Since the main source of soil PAHs is from atmospheric deposition, the ratios may also depend on the altitude of sampling sites. Nevertheless, fingerprinting technique based on molecular diagnostic ratios of PAHs is important for source identification of PAHs. However, its accuracy has been criticized as
ratios often overlap with each other (Galarneau, 2008; Stogiannidis and Laane, 2015). In addition, those who use the diagnostic ratios fail to consider the impacts of PAHs transformation and biodegradation on the ratios.

**Receptor Models**

Receptor models, which are based on mathematical or statistical tools, have been used to identify and quantify PAH sources, especially for source apportionment due to the statistical abilities. They use the physical and chemical properties of the environmental matrix at the source and receptor to determine the presence of all sources and quantify their percentage of the receptor concentrations. Receptor models have been widely used for source determination of PAHs, which include Positive Matrix Factorization (PMF) (Wang et al., 2013, 2015b; Lang et al., 2015; Teixeira et al., 2015; Bi et al., 2016), Chemical Mass Balance (CMB) (Teixeira et al., 2015), principal component analysis (PCA) (Sharma et al., 2007; Devi et al., 2016), UNMIX (Lang et al., 2015), and Multilinear Engine 2 (Xu et al., 2016). However, these models were developed for air quality, not soil quality. They may work well when contaminants are mainly from atmospheric deposition, the main source of soil PAHs. But when other sources (e.g., water) and environmental factors (e.g., biodegradation) are involved, these models lack validation and may be limited due to the matrix effects of soil (Wang et al., 2013, 2015b; Lang et al., 2015; Teixeira et al., 2015; Bi et al., 2016).

Among receptor models for source identification and apportionment of POPs, USEPA developed several including CMB, UNMIX, and PMF models. The CMB model allocates receptor concentrations to corresponding sources based on a source profile database, but UNMIX and PMF produce source profiles and compare them with reference profiles. These models were developed for atmospheric receptors but have
been used to assess soil quality. For example, statistical analysis based on PCA has been used for PAH source apportionment in soils (Devi et al., 2016). As such, they do not always fit well with receptors in soil matrix, necessitating optimization and modification for their application to soils.

**Stable Carbon Isotopic Signatures**

Stable C isotopic ratios are used to corroborate with molecular diagnostic ratios and/or receptor models for source identification of PAHs in soils. However, development of fingerprinting techniques based on isotopic ratios for PAHs source identification in the environment has been slow. Three isotopic ratios can be used to trace sources of PAHs including δ\textsubscript{13}C, Δ\textsuperscript{14}C, and δ\textsuperscript{2}H, which are present in all aromatic hydrocarbons including PAHs. Among these isotopic signatures, δ\textsuperscript{13}C and Δ\textsuperscript{14}C are calculated based on the ratio of isotopes \textsuperscript{13}C: \textsuperscript{12}C and \textsuperscript{14}C: \textsuperscript{12}C while δ\textsuperscript{2}H is calculated based on the ratio of stable isotopes of \textsuperscript{2}H: \textsuperscript{1}H. These ratios are expressed as parts per thousand (per mil, ‰) due to the small differences in isotope ratios.

Stable C isotopic ratios, i.e., δ\textsuperscript{13}C values, can be based on specific compounds such as PAHs or bulk C in a matrix such as soil or plant. Analyzing the specific δ\textsuperscript{13}C value of an organic compound from different matrices helps to identify its origin. In contrast, the bulk δ\textsuperscript{13}C value of a given sample represents an average of all C compounds, therefore an average of all sources. Allocation of contaminants to a specific source allows appropriate risk reduction and helps to identify responsible parties causing the contamination (Buczyńska et al., 2013).

To obtain isotopic signatures for specific compounds instead of bulk isotopic signatures, compound-specific isotope analysis is needed. Schmidt et al. (Schmidt et al., 2004) reviewed a method based on gas chromatography-combustion-isotope ratio
mass spectrometry (GC-C-IRMS) to determine the sources of organic chemicals.
Buczyńska et al. (Buczyńska et al., 2013) reviewed compound-specific isotope analysis based on stable C isotopes in PAHs to determine their source apportionment in the environment. Based on the characteristics of $\delta^{13}C$ values of different PAHs in different matrices, changes in isotopic signatures of PAHs during their formation, transportation, and degradation can be determined. Based on the differences in $\delta^{13}C$ values, PAHs in soils from natural sources can be separated from anthropogenic sources, thereby helping to determine if a site is contaminated or not. In addition, to improve the source identification of PAHs in soils, all three isotopic signatures ($\delta^{13}C$, $\Delta^{14}C$, and $\delta^2H$) can be used together (Bosch et al., 2015).

Though compound-specific isotope analysis for source identification and apportionment of PAHs was first proposed several decades ago, its development has been slow compared with other methods. This is partially due to its limitations including the requirement of expensive GC-C-IRMS equipment, a more complicated pretreatment process, and the lack of standard analytical methods and an isotopic signature database. However, the isotopic ratio fingerprinting technique is a useful yet underutilized method for source identification and apportionment of PAHs in the environment. More research are needed to better understand PAHs sources in the environment.

In short, all three methods have been employed to trace sources of PAHs in the environment. Each technique has advantages and limitations. A summary of the molecular diagnostic ratios of different PAHs is shown Table A-1, indicating that this technique has an abundant database compared with stable C isotopic signature
technique. Since the isotopic signatures of PAHs are more complicated compared with molecular diagnostic ratios, an expanded reference database should be developed to better use the fingerprinting technique based on stable C isotopic ratios.

**Stable Carbon Isotopic Signatures of PAHs**

Carbon has 15 known isotopes from $^8$C to $^{22}$C, with $^{12}$C and $^{13}$C being stable isotopes. The three naturally-occurring C isotopes are 12, 13 and 14. Based on their concentrations, stable isotopes $^{12}$C and $^{13}$C are present at a 99:1 ratio in the environment. Radioactive $^{14}$C is produced by cosmic radiation in the upper atmosphere and is transferred to the earth by wet and/or dry deposition, which can be absorbed by living biota (Lichtfouse, *et al.*, 1997). Compared with isotopes $^{12}$C and $^{13}$C, $^{14}$C concentrations are negligible in the environment. $^{14}$C has a long half-life of 5,700 years and its radioactivity is detectable. Because only live biota can absorb $^{14}$C, its concentrations have been used as an indicator of radiometric dating for biomaterials in archaeology. In environmental forensics, both $^{13}$C and $^{14}$C have been successfully applied to trace sources of contaminants (Lichtfouse, *et al.*, 1997; Kumata *et al.*, 2006; Bosch *et al.*, 2015).

The isotopic reference for stable C is based on Vienna Pee Dee Belemnite (VPDB), with the $\delta^{13}$C of VPDB set as 0‰. The most common way to express the delta value of stable C is shown below.

$$\delta^{13}C = \{[(^{13}C/^{12}C)_{\text{sample}} / (^{13}C/^{12}C)_{\text{VPDB}}] - 1\} \times 1000‰$$

(A-1)

VPDB is a hypothetical standard reference for Vienna Pee Dee Belemnite, which is a cretaceous belemnite (*Belemnitella Americana*), from the Pee Dee formation in South Carolina. It is commonly used as an international standard for stable C measurement.
(Buczyńska et al., 2013). However, according to the latest IUPAC guidelines, the term of multiplying by 1000‰ should not be included in the calculation of δ\textsuperscript{13}C to improve its uniformity (Coplen, 2011).

Since the stable C isotopic ratios of PAHs from the same origin are similar, δ\textsuperscript{13}C values can be used to trace the origin of PAHs in environmental matrices, which range from -15 to -37‰ (Mcraea et al., 1996; O’Malley et al., 1997; Mcrae et al., 1999; Sun et al., 2003a; Guillon et al., 2013). However, the impacts of natural weathering, biodegradation, and experimental pretreatment procedures on C signatures of PAHs are still unclear. It is unlikely that the C isotopic signatures of PAHs in tar samples are altered significantly during weathering or biodegradation (Sun et al., 2003b). In addition, no significant fractionation of isotopic C occurs during soil pretreatments including extraction and cleanup, suggesting that soils maintain their original δ\textsuperscript{13}C values during weathering and analyses (Kim et al., 2005). However, the δ\textsuperscript{13}C values may vary with extraction and cleanup methods, so a standardized method is necessary for using compound-specific δ\textsuperscript{13}C values to trace sources of soil PAHs (Graham et al., 2006). In addition, soil weathering processes such as vaporization, photolysis, and biodegradation also alter the δ\textsuperscript{13}C values of soil PAHs (O’Malley et al., 1994; Yanik et al., 2003). Since atmospheric deposition is a major pathway for soil PAHs, the alteration of δ\textsuperscript{13}C values of PAHs during their transportation from air particulate matters to soil are still unknown. Moreover, the influences of organic matter and clay minerals on C isotopic signatures of PAHs also remain unclear. Further research investigating changes in δ\textsuperscript{13}C values of PAHs in soils can help better understand the fractionation of stable C isotopes during their transportation and transformation in soils.
Fractionation of C isotopes occurs during the formation, transportation, and degradation of PAHs, resulting in different δ\(^{13}\)C values or isotopic signatures depending on the environmental conditions. The effects of environmental processes on fractionation of C isotopes in soils can be divided into two categories: kinetic and equilibrium. The kinetic effect is caused by different reaction rates between light and heavy isotopes, as lighter isotopes exhibit faster reaction rates and can be retained in incomplete physical and biological processes. Physical processes such as sorption, dissolution, evaporation, diffusion, and condensation while biological processes may be related to preferential utilization of a particular isotope by enzymes. Therefore, the influence of the kinetic effect depends on how complete a reaction is. When two processes have similar reaction rates, then equilibrium effect may come into play. This may occur during bonds changing chemical processes such as redox transformation and hydrolysis. In this case, heavier isotopes are enriched in compounds with stronger bonds including compounds with higher redox state or fewer coordinated positions. The equilibrium effect usually is less important at higher temperatures but is dominant at lower temperatures as the reaction moves towards the direction of lower enthalpy (Wiederhold, 2015).

**Biogenic PAHs**

Biogenic PAHs are produced during degradation of vegetative organic substances by plants, algae, and microorganisms (Abdel-shafy and Mansour, 2016). Based on photosynthetic pathways, plants are divided into three categories: C\(_3\), C\(_4\), and crassulacean acid metabolism (CAM) plants. In C\(_3\) plants, C fixation produces 3-phosphoglycerate from CO\(_2\), water, and ribulose bisphosphate, which occurs in all plants as the first step of Calvin-Benson cycle. Examples of C\(_3\) plants include wheat,
peanut, coffee, tea, and various vegetables and fruits. In C₄ woody plants, the C originates directly from sugar and other biological molecules such as malic acid instead of the air. Examples of C₄ plants include corn, sorghum, millet, sugarcane, and various grasses. CAM is a more specialized photosynthetic pathway that evolved in some plants as an adaptation to drought conditions. In CAM plants, leaf stomata remain closed to reduce evapotranspiration during the daytime, but open to collect CO₂ in the evening (O’Leary, 1988).

Carbon fractionation in biogenic PAHs occurs during C absorption by biota from CO₂ in the environment due to preferential uptake of C isotopes by enzymes. The duration of C absorption under environmental temperatures is relatively short, so the kinetic effect dominates. Plant enzymes react more readily with lighter $^{12}$CO₂ from the atmosphere, causing fractionation of C isotopes, i.e., $^{12}$C enrichment in plant biomass. In addition, C₄ plants absorb $^{13}$C faster than C₃ plants. As a result, C₄ plants accumulate more $^{13}$C relative to C₃ plants, making the $\delta^{13}$C values of biogenic PAHs in C₃ plants lower than those in C₄ plants (Hobbie and Werner, 2004). For example, O’Leary et al. (O’Leary, 1988) first reported the bulk $\delta^{13}$C values for C₃, CAM, and C₄ plants as -24 to -33‰, -10 to -20‰, and -10 to -16‰, respectively. Similarly, O’Malley et al. (O’Malley et al., 1997) reported the bulk $\delta^{13}$C values are -28.7 to -31.8‰ in C₃ plants, but they increase to -15.5 to -23.1‰ in C₄ plants. In addition, Skrzypek et al. (Skrzypek et al., 2008) reported the bulk $\delta^{13}$C values in different plants range from -20.4 to -28.9‰. Since we lack compounds specific $\delta^{13}$C values for PAHs, the $\delta^{13}$C values of mixed plants may represent those of biogenic PAHs, show a median value around -25‰ (Hobbie and Werner, 2004). However, little research is available regarding the bulk $\delta^{13}$C
values in the other biota such as microbes as well as the $\delta^{13}C$ values of PAHs in all plants.

**Petrogenic PAHs**

Petrogenic PAHs are formed during thermal maturation of ancient fossilized organic matter into fossil fuels such as crude oil and coal under relatively low temperatures. In addition, when petroleum products are produced from crude oil, more PAHs are produced (Saber *et al.*, 2005). Therefore, petrogenic PAHs are present in petroleum and coal as byproducts. Compared with the bulk $\delta^{13}C$ values in plants at -20.4 to -28.9‰ based on Skrzypek *et al.* (Skrzypek *et al.*, 2008), those in fossil fuels are relatively lower. For instance, they are -22.8 to -31.9‰ in kerogens, -23.2 to -27.1‰ in coals (Redding and Schoell, 1980), and -25.5 to -28.3‰ and -28.0 to -29.2‰ in different peats (Jedrysek and Skrzypek, 2005; Skrzypek *et al.*, 2008). However, since those data are based on bulk $\delta^{13}C$ values which are the ratios based on all C in fossil fuels, they are different from the specific $\delta^{13}C$ values of PAHs.

Compared to bulk $\delta^{13}C$ values, the specific $\delta^{13}C$ values of PAHs are slightly higher with narrower ranges. For example, the $\delta^{13}C$ values of PAHs in coals are -23.2 to -25.7‰, -23.0 to -25.4‰, and -23.2 to -25.6‰ for 2-3 rings, 4 rings, and 5-6 rings PAHs, respectively. After pyrolysis, they decrease to -24.5 to -28.6‰, -25.0 to -28.2‰, and -26.1 to -29.4‰ (McRae *et al.*, 1998). The $\delta^{13}C$ values of PAHs are -20 to -27‰ in creosote and -25 to -29‰ in manufactured gas plant (MGP) tar (Saber *et al.*, 2005). These data show that lower $\delta^{13}C$ values are observed in pyrogenic PAHs compared with biogenic bulk C in plants. The temperatures during petrogenic PAH formation may vary from 60 to 150°C. These temperatures are higher than those in biogenic processes (20-40°C) but lower than those in the pyrogenic processes (>200°C). Compared with
biogenic PAHs, the formation of petrogenic PAHs takes much longer (up to million years), so both kinetic and equilibrium effects may play a role during fractionation of C isotopes. In short, lower δ^{13}C values are expected for petrogenic PAHs compared with biogenic PAHs that are from the same origin, ranging from -23 to -29‰.

**Pyrogenic PAHs**

Pyrogenic PAHs result from incomplete combustion of OM and are temperature-dependent so their sources are complicated, resulting from various anthropogenic activities. Two factors affect the δ^{13}C values of PAHs during the pyrogenic processes. They are: 1) the characteristics of OM such as type, precursor substance and aging time, and 2) formation conditions such as temperature, pressure, and duration (Wang *et al.*, 2016). More PAHs are formed and more fractionation of C isotopes occurs under higher temperatures due to the kinetic effect. It induce less fractionation at higher temperature. PAHs formation mechanisms differ during pyrolysis and combustion reactions and can induce different bond formation pathways and potential different isotopic effects.

During the pyrogenic process, the kinetic effect is more important as the reaction often finishes quickly without reaching equilibrium. During this process, the formation of lighter $^{12}$C-$^{12}$C bonds are faster relative to heavier bonds that contain $^{13}$C (requires less energy based on thermodynamics), resulting in fewer $^{13}$C relative to $^{12}$C in the products (Guillon *et al.*, 2013; Bosch *et al.*, 2015; Wang *et al.*, 2016). In addition, during ring condensation of PAHs, $^{12}$C is preferentially incorporated into PAHs (Mcrae *et al.*, 1999). As evidence, a significant decrease in $^{13}$C has been observed in combusted materials compared to original materials (Sun *et al.*, 2003a; Bosch *et al.*, 2015; Wang *et al.*, 2016).
As expected, the $\delta^{13}$C values of pyrogenic PAHs vary with materials of different origins (Bosch et al., 2015). For instance, the mean $\delta^{13}$C values of PAHs from gasoline exhaust are -21.1 to -28.1‰ (Okuda et al., 2002; Sun et al., 2003a), probably due to different gasoline origins in addition to different combustion conditions. The mean $\delta^{13}$C values of PAHs are 28.7 ± 1.4‰ and 25.3 ± 1.6‰ from C₃ plant combustion and liquid fossil fuel combustion compared to 23.2 ± 1.1‰, and 26.8 ± 1.3‰ for coal pyrolysis at 650°C and 900°C (Bosch et al., 2015).

Although there are some overlaps among $\delta^{13}$C values for different materials, the mean $\delta^{13}$C values generally decrease in the order: C₄-plant combustion (-15 to -19‰) > low-T coal pyrolysis at 650°C (-22 to -24‰) > liquid fossil fuel combustion (-21 to -29‰) > high-T coal pyrolysis at 900°C (-25 to -29‰) > C₃-plant combustion (-27 to -35‰) (Mcrae et al., 1999; Bosch et al., 2015). Based on this order, clearly, temperature plays an important role in controlling the $\delta^{13}$C values of PAHs during the pyrogenic processes. As temperature controls the extent of aromatic ring condensation, it impacts the $\delta^{13}$C values of PAHs (Mcrae et al., 1999). For example, the $\delta^{13}$C values of PAHs in coal at ~-25‰ decrease with increasing temperature from -24.0 to -25.8‰ after low-T carbonization at ~640°C to -25.1 to -26.5‰ after high-T carbonization at ~900°C, to -27.5 to -29.5‰ after gasification at 700-1600°C and to -29.8 to -30.9‰ after combustion at ~2,200°C. The $\delta^{13}$C values of PAHs from low-T combustion are similar to those in parent coal materials, but with increasing temperature from 640°C to 900°C, lower $\delta^{13}$C values are observed. In other words, the higher the temperature, the lower $\delta^{13}$C values of resulted PAHs.
A schematic on fractionation of C isotopes of PAHs from CO$_2$ to coal utilization is shown in Figure A-1 (Mcrae et al., 1999). Considering low-T carbonization at ~500°C is not a typical anthropogenic pyrogenic process, the median δ$^{13}$C value of anthropogenic pyrogenic PAHs is set as -29‰ for high-T pyrolysis at ~900°C to combustion at ~2,200°C. However, natural pyrogenic sources such as forest fires and volcanic eruption may be considered as low-T carbonization or pyrolysis since the temperature of forest fire is <800°C and hot lava is <1,200°C. In addition, only limited materials are combusted at high-T during these processes, with the rest going through low-T pyrolysis through diffused heat or embers. Since the combusted plants are the mixture of C$_4$ grasses and C$_3$ woody species during forest fires, the resulting δ$^{13}$C values in PAHs from forest fires are usually higher than biomass (C$_3$ woody species) combustion. Additionally, O`Malley et al. found that the δ$^{13}$C values in burned biomass from forest fires are similar to those of parent materials, which are usually >-30‰ (O`Malley et al., 1997). However, the δ$^{13}$C values of PAHs from anthropogenic biomass combustion are < -30‰ since the majority of the anthropogenic activities utilize temperatures >650°C. This supports the hypothesis that forest fires can be considered as low-T biomass pyrolysis, which produces δ$^{13}$C values similar to parent materials. As further evidence, the δ$^{13}$C values of PAHs in combusted C$_3$ plants are lower at -27 to -29‰ compared to C$_4$ plants at -15 to -17‰ (Saber et al., 2005). Likewise, Kawashima and Haneishi reported the δ$^{13}$C values for PAHs at -28.0 to -34.7‰ for C$_3$ plant combustion, decreasing to -16.1 to -19.3‰ for C$_4$ plant combustion. Therefore, a value of -26‰ may represent the δ$^{13}$C values of natural pyrogenic PAHs since C$_3$ plants are dominant on the planet (Kawashima and Haneishi, 2012).
In short, both the origin of source materials and formation process play important roles in controlling C isotopic fractionation of PAHs during biogenic, petrogenic, and pyrogenic processes. In addition, the $\delta^{13}C$ values of natural pyrogenic PAHs at -15 to -29‰ are higher than those of anthropogenic pyrogenic PAHs at -16 to -37‰. Therefore, natural PAHs (including biogenic, petrogenic and natural pyrogenic) tend to have higher $\delta^{13}C$ values while anthropogenic PAHs (including petrogenic and anthropogenic pyrogenic) tend to have lower $\delta^{13}C$ values, especially anthropogenic pyrogenic PAHs formed under relatively higher temperatures.

**Stable Carbon Isotopic Signatures of Soil PAHs**

Compared to stable C isotopic ratio measurements of PAH sources, less research has focused on stable C isotopic ratio measurement in soils. However, the $\delta^{13}C$ values of PAHs in soils are useful for tracing sources of PAHs (Lichtfouse, *et al.*, 1997; Mcrae *et al.*, 1999; Wilcke *et al.*, 2002; Sun *et al.*, 2003b; Saber *et al.*, 2005; Graham *et al.*, 2006; Bosch *et al.*, 2015).

**PAHs in Rural and Urban Soils**

Lichtfouse *et al.* (1997) measured the $\delta^{13}C$ values of OM and PAHs in two rural soils (Lichtfouse, *et al.*, 1997). The bulk $\delta^{13}C$ values in OM and specific $\delta^{13}C$ values of PAHs in soil A are -25.0 to -26.3‰ and -25.9 to -26.0‰, respectively. The corresponding values are -20.3 to -25.9‰ and -27.0 to -27.2‰ in soil B. Bosch *et al.* determined the $\delta^{13}C$ values of PAHs in forest soils, which are -23.0 to -25.3‰ with a mean of 24.0 ± 0.1‰ (Bosch *et al.*, 2015). Since the values are in the range of -23.0 to -29.4‰ for combusted coal based on references, coal combustion was the main source of PAHs in this case, with low-T pyrolysis being dominant.
Unlike rural soils, more research has focused on urban soils where more human activities are present. Compared to those in rural soils, the mean δ^{13}C values for PAHs in urban soils are lower. For example, Wilcke et al. (Wilcke et al., 2002) reported values of -24.8 to -26.2‰ and -26.8 to -35.1‰ for naphthalene and perylene in gas works and urban soils, respectively. Sun et al. (2003) reported values of -25.2 to -29.7‰ for soils contaminated with coal tar and fuels, compared to -22.1 to -30.5‰ for soils contaminated with gas works tar. Similarly, values of -28 to -32‰ and -25 to -29‰ were reported for MGP-contaminated soils (Sun et al., 2003b; Saber et al., 2005; Graham et al., 2006).

In short, urban soils and contaminated soils tend to have lower δ^{13}C values of PAHs (-22 to -35‰), while remote and rural soils tend to have higher δ^{13}C values of PAHs (-20 to -27‰) due to reduced influence of anthropogenic activities.

**Formation and Degradation of PAHs**

Stable C isotopic fractionations occur during formation and degradation of PAHs in the environment. The δ^{13}C values are impacted by various biogeochemical processes. In addition, the uncertainty in environment-dependent C isotopic fractionation warrants more research.

As mentioned earlier, stable C fractionations happen during PAHs formation process including biogenic, petrogenic, and pyrogenic. During biogenic formation processes, more ^{12}C than ^{13}C are absorbed by biota to form organic C such as carbohydrate, protein, and fat. As such, the C coming from plant biomass, which are the sources of PAHs, have more ^{12}C than CO\textsubscript{2}. As a result, δ^{13}C depletion occurs during biogenic PAHs formation processes, i.e., from CO\textsubscript{2} to LMW-PAHs. During petrogenic processes, more PAHs may be produced similar to biogenic PAHs, with more benzene
rings being added to biogenic PAHs. Since all C in fossils come from biota, they are also more enriched with $^{12}$C than CO$_2$. Furthermore, C fractionations continue during the petrogenic formation process, with more $^{13}$C being depleted from PAHs during the process. Regarding PAHs, pyrogenic formation processes are of most environmental concern mainly due to their anthropogenic nature. Since pyrogenic PAHs come from biomass or fossils, which are $^{12}$C enriched, along with the kinetic isotope effect, more $^{12}$C are enriched in pyrogenic PAHs during incomplete combustions or high-T burning (Richter and Howard, 2000). In short, more and more $^{12}$C are enriched from CO$_2$ to pyrogenic PAHs, with $^{13}$C being depleted during the processes (Figure A-2). Therefore, $^{12}$C in PAHs decreases from biogenic, to petrogenic and to pyrogenic PAHs, so are the $\delta^{13}$C values.

The C in PAHs originates from atmospheric CO$_2$, which may end up as CO$_2$ after its complete degradation in a full cycle (Figures 5-2 & 5-3). The $\delta^{13}$C values of atmospheric CO$_2$ vary with sampling area but are often >-9‰. In general, the $\delta^{13}$C values decrease with increasing CO$_2$ from fossil fuels this is because $\delta^{13}$C depletion/$\delta^{12}$C enrichment occurs during pyrogenic processes (Gleason and Kyser, 1984; Sturm et al., 2005). The kinetic effects on $\delta^{13}$C values of biogenic PAHs are discussed earlier. However, the effect of C transfer from plants to animals through food chains on C isotopic fractionation is still unclear. How PAHs C isotopic fractionation is related to the biological uptake processes is also unknown. During transformation by animals and microbes, some PAHs are metabolized to CO$_2$ and released back into the atmosphere. In this case, C isotopic fractionation is impacted by the equilibrium effect due to the long-term degradation.
While biota including plants release biogenic PAHs into the soil (Hobbie and Werner, 2004), biomass burning produces pyrogenic PAHs (Guillon et al., 2013). Petrogenic PAHs can be produced during the conversion of plant biomass to soil OM, which depends on the environment. In addition, petrogenic PAHs can form during formation of fossil fuels. By natural seepage or anthropogenic leaking, petrogenic PAHs may be released into the soil as non-aqueous phase liquid or coal particles. Pyrogenic PAHs are produced during natural disasters such as forest fires and volcanic eruptions and also through human activities such as coal combustion and traffic emissions.

For pyrogenic PAHs, their major pathway to enter soils is by atmospheric deposition. PAHs may volatilize and condense in soils or travel with the fine particulate matter in the air and accumulate in soils. Based on their properties, LMW-PAHs tend to travel further compared with HMW-PAHs (Yuan et al., 2015). Once in soils, PAHs can be absorbed by soil OM and minerals, reducing availability and slowing down degradation (Kuppusamy et al., 2016). Lastly, under the influence of long-term equilibrium effect, PAHs may be converted to CO₂ and released back to the atmosphere. In short, as stable ¹³C depletes during the transportation and transformation from CO₂ to pyrogenic PAHs, the δ¹³C values are significantly decreased. However, during PAH degradation, δ¹³C values may increase again if PAHs are degraded to CO₂.

It is understandable that the δ¹³C values for PAHs originating from plants and fossil fuels have a large range due to different plant species and fossil fuels. As such, the δ¹³C values of biogenic, petrogenic and pyrogenic PAHs may overlap and may not always alter as expected. Nevertheless, the general trend of the δ¹³C values of PAHs...
during their biogeochemical cycle should generally decrease from a specific given source, i.e., CO$_2$ → biogenic PAHs → petrogenic PAHs → pyrogenic PAHs. In other words, it is a process of $\delta^{13}$C depletion or $\delta^{12}$C enrichment (Figure A-3).

**PAHs in Contaminated Soil**

The sources of soil PAHs include biogenic, petrogenic, and pyrogenic. Anthropogenic pyrogenic processes such as high-T incomplete combustion of biomass and fossil fuels significantly decrease the $\delta^{13}$C values of soil PAHs. For example, the $\delta^{13}$C values of PAHs are -20 to -27‰ in rural soils, and they decrease to -22 to -35‰ in contaminated soils. Furthermore, the $\delta^{13}$C values of PAHs are significantly lower at -31 to -62‰ in contaminated sediments (McRae *et al.*, 2000). The data support the hypothesis that anthropogenic PAHs in contaminated soils have significantly lower $\delta^{13}$C values.

Based on the literature, the mean $\delta^{13}$C values of -25‰, -27‰, and -29‰ can be used to roughly represent biogenic, petrogenic, and anthropogenic pyrogenic PAHs, with those of natural pyrogenic PAHs having -26‰ (Table A-2 & Figure A-4). These reference values can be used to separate natural from anthropogenic PAHs (Figure A-5). The boundary limits of $\delta^{13}$C values for natural biogenic and anthropogenic pyrogenic PAHs are -25‰ and -29‰, which represent the lowest values for natural sources and highest values for anthropogenic sources. In this case, if the $\delta^{13}$C values of soil PAHs are >-25‰, they are mostly from natural sources. On the other hand, if they are <-29‰, they are primarily from anthropogenic sources. If they are between -25 to -29‰, they are from the combination of natural and anthropogenic sources.
Other Isotopic Signatures of Soil PAHs

Due to the limitations of stable C isotope analysis, additional isotope systems such as radiocarbon analysis and stable H isotopic ratios are being used to assess sources and transformation processes of PAHs in soils. By comparing the $\Delta^{14}$C and $\delta^{2}$H values of PAHs between sources and environmental matrices via statistical approaches, source identification and apportionment can be improved. However, due to the limited research on $\Delta^{14}$C and $\delta^{2}$H values of PAHs in soil, they have even more scant database than stable C signatures of PAHs (Bosch et al., 2015).

Radioactive C Isotopic Ratios

Based on the literature, there is still a lack of $\Delta^{14}$C values for PAHs, and not all available studies are based on soils. Compared to stable C isotopic ratios, the ranges for $\Delta^{14}$C values are more variable. For example, the $\Delta^{14}$C values of PAHs are 100 to 150‰ (biomass), -400 to -600‰ (peat), and -950 to -1000‰ (coal and liquid fuel) for various OM (Bosch et al., 2015). Based on the $\Delta^{14}$C values of PAHs range from -550 to -934‰ in sediments, combustion of fossil fuels is probably the dominant source (Mandalakis et al., 2004). The $\Delta^{14}$C signatures range from -381 to -388‰ for atmospheric PAHs in Sweden and range from -888 to -914‰ for atmospheric PAHs in Greece and Croatia (Manolis Mandalakis, Örjan Gustafsson, Tomas Alsberg, Anna-Lena Egebäck, Christopher M. Reddy, Li Xu, Jana Klanova, Ivan Holoubek, 2005). The $\Delta^{14}$C values of PAHs in airborne particulate matter range from -514 to -787‰, indicating that fossil fuel combustion is the dominant source (Kumata et al., 2006). The $\Delta^{14}$C values of atmospheric PAHs are -288 to -568‰, showing a major contribution of non-fossil material combustion (Zencak et al., 2007); while Sheesley et al. (Sheesley et al., 2009) reported $\Delta^{14}$C values of -58.0 to -138‰, indicating biomass combustion as the
dominant source. The $\Delta^{14}C$ values for PAHs in the coke plant particulate matter are -911 to -990‰, indicating fossil fuel combustion dominated source (Xu et al., 2012). The $\Delta^{14}C$ signature of alkylated PAHs in sediment samples are -849 to -962‰, indicating a petrogenic dominated source (Jautzy et al., 2015).

**Stable Hydrogen Isotopic Ratios**

Similar to the $\delta^{13}C$ values of PAHs, the reference values for $\delta^2H$ of PAHs have been investigated (Redding and Schoell, 1980; Mastalerz and Schimmelmann, 2002; Sun et al., 2003a; Jedrysek and Skrzypek, 2005; Skrzypek et al., 2008; Zheng et al., 2011; Vitzthum et al., 2012; Bosch et al., 2015). The $\delta^2H$ values of PAHs from gasworks coal tar combustion, high-T coal combustion, jet fuel combustion, and gasoline combustion are -32.2 to -49.3‰, -65.3 to -81.1‰, -60.2 to -74.3‰, and -47.0 to -61.5‰, respectively (Sun et al., 2003a). Based on the results, $\delta^2H$ values of PAHs have opposite trend compared with $\delta^{13}C$ values of PAHs: $\delta^{13}C$ values of PAHs tend to decrease with increasing temperature, while $\delta^2H$ values of PAHs tend to increase. The mean $\delta^2H$ values of PAHs are -62‰ (liquid fossil fuel combustion), -73‰ (high-T coal pyrolysis), -94‰ (C$_3$ plant combustion), -129‰ (coal combustion), and -159‰ (peat combustion) based on Bosch et al. (Bosch et al., 2015). In addition, $^2H$ enrichment and $^{13}C$ depletion occur simultaneously during fossil fuel combustion, resulting from dehydration during PAHs formation (Sun et al., 2003a). Since C-H bonds are weaker than C-$^2H$ bonds, C-H bonds form at a faster rate compared to C-$^2H$ bonds (Sun et al., 2003a). Similar to $\Delta^{14}C$ values of PAHs, limited research is available regarding $\delta^2H$ values of PAHs, resulting in a scant database.

For accurate source identification and apportionment of PAHs, often two different approaches have been used including molecular diagnostic ratios, receptor models, and
isotopic ratios. Similarly, dual or triple isotopic ratios have been used for source identification and apportionment of PAHs. For example, Bosch et al. (Bosch et al., 2015) applied triple isotopic ratios ($\delta^{13}$C, $\Delta^{14}$C, and $\delta^{2}$H) of PAHs for source identification and apportionment in soils. The $\delta^{13}$C values of PAHs are -23.0 to -25.3‰ with a mean of 24.0 ± 0.1‰, the $\Delta^{14}$C values are -768 to -960‰ with a mean of 892 ± 37‰, and $\delta^{2}$H values are -53 to -263‰ with a mean of 129 ± 44‰. They confirmed that soil PAHs are mainly from fossil fuels based on the high depletion of $^{14}$C. In contrast to C signature results, the H signature showed a higher variability in soil samples. However, the authors did not find correlations between the $\delta^{13}$C and $\delta^{2}$H values of PAHs (Bosch et al., 2015).

Conclusions and Prospective Research

Three major approaches for source identification and apportionment of PAHs in soils have been used, which include molecular diagnostic ratios, receptor models, and stable C isotopic signatures. The pros and cons of these methods are discussed, and the specific $\delta^{13}$C values of PAHs have advantages compared with other two approaches for being more accurate and applicable to soils. Based on the $\delta^{13}$C values of PAHs, a biogeochemical cycle of PAHs based on stable C isotopic signature in soils and a systematic approach based on stable C isotopic signature to determine PAHs sources are proposed. The differences among three major sources of PAHs in soils are illustrated, i.e., biogenic, petrogenic and pyrogenic. In addition, $\Delta^{14}$C values and $\delta^{2}$H values have also been briefly introduced.

Based on the data, stable C isotopic signatures based specific $\delta^{13}$C values of PAHs is a useful but highly undervalued technique for source identification and appointment of soil PAHs. However, its limitation includes technological limitations such
as the instrument requirement of GC-C-IRMS and a lack of standard pretreatment methodology, which may further hinder development. Further research should build an abundant database based on different isotopic signatures of PAHs to better understand the biogeochemical transformation of PAHs and to trace PAHs sources in soils.
Table A-1. Source identification and apportionment of PAHs in soils based on molecular diagnostic ratios (Oliveira et al., 2011; Tobiszewski and Namie, 2012; Biache et al., 2014).

<table>
<thead>
<tr>
<th>Diagnostic ratios</th>
<th>Pyrogenic</th>
<th>Petrogenic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total low (2-3 rings) / high (&gt;3 rings) molecular weight PAHs</td>
<td>&lt;1</td>
<td>&gt;1</td>
</tr>
<tr>
<td>Anthracene/(Phenanthrene + Anthracene)</td>
<td>&gt;0.1</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>Fluorene/(Pyrene + Fluorene)</td>
<td>N/A</td>
<td>&lt;0.5: petrol &gt;0.5: diesel 0.4-0.5: vehicle</td>
</tr>
<tr>
<td>Fluoranthene/(Pyrene + Fluoranthene)</td>
<td>0.4-0.5: fossil fuel &gt;0.5: grass, wood &amp; coal</td>
<td>&lt;0.4</td>
</tr>
<tr>
<td>Benzo[a]anthracene/(Chrysene + Benzo[a]anthracene)</td>
<td>0.2-0.35: coal &gt;0.35: vehicle</td>
<td>&lt;0.2</td>
</tr>
<tr>
<td>Benzo[a]pyrene/Benzo[g,h,i]perylene</td>
<td>&lt;0.6: non-traffic &gt;0.6: traffic</td>
<td></td>
</tr>
<tr>
<td>Indeno[1,2,3-c,d]pyrene</td>
<td>0.2-0.5: petrol &gt;0.5: biomass</td>
<td>&lt;0.2</td>
</tr>
<tr>
<td>/(Benzo[g,h,i]perylene + Indeno[1,2,3-c,d]pyrene)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenanthrene/Anthracene</td>
<td>&lt;10</td>
<td>&gt;10</td>
</tr>
<tr>
<td>Fluoranthene/Pyrene</td>
<td>&gt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Diagnostic ratios</td>
<td>Pyrogenic</td>
<td>Petrogenic</td>
</tr>
<tr>
<td>------------------------------------------</td>
<td>----------------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Benzo[e]pyrene/Benzo[a]pyrene particles</td>
<td>1: fresh air</td>
<td>N/A</td>
</tr>
<tr>
<td>Fluoranthene/Benzo[e]pyrene</td>
<td>3.5: vehicle</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>0.22-0.55:</td>
<td>&gt;0.5: diesel</td>
</tr>
<tr>
<td>Benzo[b]fluoranthene/Benzo[k]fluoranthene</td>
<td>gasoline</td>
<td>N/A</td>
</tr>
</tbody>
</table>
| Pyrene/Benzo[a]pyrene                    | <1: gasoline                           | 10: diesel | N/A
Table A-2. Compilation of reference values of δ13C (‰) for biogenic bulk C, petrogenic PAHs, natural pyrogenic PAHs, and anthropogenic pyrogenic PAHs (O’Malley et al., 1994, 1997; McRae et al., 1998; Mcrae et al., 1999; Saber et al., 2005; Skrzypek et al., 2008; Bosch et al., 2015).

<table>
<thead>
<tr>
<th>Biogenic bulk C</th>
<th>Petrogenic PAHs</th>
<th>Natural Pyrogenic PAHs</th>
<th>Anthropogenic pyrogenic PAHs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C₄ and CAM plants</td>
<td>Mixed plants</td>
<td>C₃ plants</td>
</tr>
<tr>
<td>-10 to -23</td>
<td>-20 to -29</td>
<td>-24 to -33</td>
<td>-20 to -27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-25 (mostly C₃ plants)</td>
<td>-27</td>
<td>-26 (mostly C₃ plants)</td>
<td>-29</td>
</tr>
</tbody>
</table>
Figure A-1. A schematic of C isotopic fractionation of PAHs from CO2 to coal utilization including changes in δ13C values in each step.
Figure A-2. A schematic of PAHs formation processes: biogenic (a), petrogenic (b), and pyrogenic (c). $^{12}\text{C}\text{O}_2$ has larger arrow due to its predominant absorption and utilization by biota. X represents electron donor element. Red arrows represent dominant kinetic effects, and purple arrows represent mixtures of effects or uncertain environment-dependent isotope effects. Pure $^{12}\text{C}$ molecules are red, pure $^{13}\text{C}$ molecules are blue, $^{12}\text{C}$-enriched molecules are dark purple, more $^{12}\text{C}$-enriched molecules are red-purple, and most $^{12}\text{C}$-enriched molecules are dark red. Naphthalene represents biogenic PAHs, phenanthrene represents petrogenic PAHs, and chrysene represents pyrogenic PAHs since higher temperature produces higher molecular weight PAHs.
Figure A-3. A biogeochemical cycle of stable C isotopic signatures in soil PAHs. Red arrows represent dominant kinetic effects, blue arrows dominant equilibrium effects, purple arrows represent mixtures of effects or uncertain environment-dependent isotope effects. The large black arrow represents stable C depletion during formation of pyrogenic PAHs from CO$_2$ and the small black arrows represent $\delta^{13}$C values from high to low between C$_3$ and C$_4$ plants and between low- and high-temperature pyrolysis and incomplete combustion. Bold C$_3$ plants are dominant plants and bold incomplete combustion is the most common use of fossil fuels.
Figure A-4. Compilation of reference values of $\delta^{13}C$ (‰) for biogenic bulk C, petrogenic PAHs, natural pyrogenic PAHs, and anthropogenic pyrogenic PAHs. Green, black, orange, red color represent biogenic bulk C, petrogenic PAHs, natural pyrogenic PAHs, and anthropogenic pyrogenic PAHs, respectively. Biogenic bulk C and natural pyrogenic PAHs are considered as C$_3$ plants predominated (O’Malley et al., 1994, 1997; McRae et al., 1998; Mcrae et al., 1999; Saber et al., 2005; Skrzypek et al., 2008; Bosch et al., 2015).
Figure A-5. The proposed evaluation system of soil PAHs based on stable C isotopic
HUMAN EXPOSURE TO POLYCYCLIC AROMATIC HYDROCARBONS: METABOLOMICS PERSPECTIVE

Carcinogenesis and Transformation of PAHs in Humans and Animals

PAHs are ubiquitous chemicals in the environment. Some carcinogenic PAHs are genotoxic by inducing mutations to initiate cancer while others are not genotoxic but they enhance cancers (Abdel-Shafy and Mansour, 2015). PAHs that initiate cancer are often modified by enzymes into biotransformation products that react with DNA, leading to mutations. In other words, PAHs acquire carcinogenicity only after they have been activated by xenobiotic-metabolizing enzymes to reactive biotransformation products, which can attack cellular DNA. As such, the alteration of DNA sequence in genes regulating cell replication may increase the possibility of cancer and other diseases (Moorthy et al., 2015). The biological activity of PAHs is related to their structures, which are formed between angular condensed aromatic rings resulting from regional distortions with maximal impact, termed as “fjord” and “bay” regions (Figure B-1) (Ewa and Danuta, 2016). Though their reactivity depends on the density of electron charges, geometric distortions in molecules also influences charge distribution and indirectly its reactivity. While PAHs with “fjord” regions (e.g., dibenzo[a,l]pyrene) are non-planar, which bind preferentially to adenine nucleotides, PAHs with a “bay” regions (e.g., BaP) are planar, which bind to guanine nucleotides (Xue and Warshawsky, 2005). Furthermore, increase in their non-planarity lowers their ability to be biotransformed to reactive forms, which produce DNA-damaging adducts. Mutagenic biotransformation

products of PAHs include diol epoxides, quinones, and radical PAH cations, which can bind to DNA to form bulky complexes called DNA adducts. While stable adducts may lead to DNA replication errors, unstable adducts react with DNA strands by removing purine bases (either adenine or guanine). Such mutations, if not repaired, can transform genes encoding normal proteins into cancer-causing oncogenes. In addition, quinones can also repeatedly generate reactive oxygen species (ROS), which may independently damage DNA (Ewa and Danuta, 2016).

Among enzymes, those in the cytochrome P450 (CYP) family including CYP1A, CYP1B, CYP2C, and CYP2E may metabolize PAHs to diol epoxides. Exposure to PAHs can increase the production of cytochrome enzymes, which convert PAHs into mutagenic diol epoxides. In this pathway, PAH molecules bind to the aryl hydrocarbon receptor (AhR) to activate the transcription factor, thereby increasing the production of cytochrome enzymes. As a result, PAHs can form several OH-PAHs isomers, which are excreted and can serve as tracers for different CYP isoform expression and activity. On the other hand, the activity of these enzymes may at times protect against PAHs’ toxicity and also play an important role to detoxify PAHs (Xue and Warshawsky, 2005; Moorthy et al., 2015). Since CYP family controls the metabolic activation of PAHs, other dietary and chemical exposures can also impact PAHs metabolism, including pharmaceuticals and drugs that are inducers of CYP isoforms, or may function as inhibitors or competitors of CYP isoforms. Thus, lifestyle and dietary patterns play an important role in modulating the biotransformation, metabolism, and detoxification of PAHs in humans.
The most common mechanisms of metabolic activation of PAHs involve many metabolites due to diverse enzyme action at different phases, i.e., phase I (activation) and phase II (conjugation). As such, PAHs are biotransformed by both phase I enzymes and peroxidases, which produce DNA-reactive metabolites, and phase II enzymes, which form polar conjugates. In phase I oxidation, reactions are catalyzed by cytochrome P450 enzymes and hydroxylated by epoxide hydrolase (Ewa and Danuta, 2016). CYP1A1 and CYP1B1 are highly inducible by PAHs via activation of AhR. The AhR is present in the cytoplasm as a complex with other proteins such as heat shock protein 90 (Hsp90), p23, and AhR-interacting protein. Having formed a complex with PAHs, Hsp90 is released and an AhR-PAH complex is translocated to the nucleus. There, the AhR-PAH complex creates a heterodimer with an AhR nuclear translocator (ARNT), which binds to DNA at xenobiotic response elements situated in the promoter region of CYP1A and CYP1B genes (Moorthy et al., 2015). Therefore, AhR plays an important role in PAH-mediated tumorigenesis (Xue and Warshawsky, 2005; Tarantini et al., 2011). The obtained diol epoxides are hydrophilic, which can easily dissolve in water. For this reason, they can involve in phase II reactions, which couple the metabolites with endogenous compounds — sulfuric acid, glucuronic acid, and glutathione (Moorthy et al., 2015; Ewa and Danuta, 2016). Some PAHs metabolic intermediates show genotoxic and carcinogenic properties (Xue and Warshawsky, 2005). For example, proximate carcinogens 7,8-oxide BaP and 7,8-dihydrodiol BaP, and mutagen and ultimate carcinogen 7,8-diol-9,10-epoxide BaP have been identified after BaP exposure (Figure B-2) (Ewa and Danuta, 2016; Kim et al., 2013).
In addition, the reactive biotransformation products of PAHs may also induce the formation of protein adducts in cells, which may affect the normal activities of these proteins. PAHs biotransformation products may also elevate ROS, which can directly affect DNA, lipids, or proteins to initiate carcinogenesis (Kafferlein et al., 2010). CYPs also contain a ‘peroxidase-like’ activity, which catalyzes the 1-electron oxidation of PAHs at a specific carbon position to produce radical cations. These PAH-derived radical cations, though short-lived, can react with DNA and cause mutations (Hrycay and Bandiera, 2012). In the AhR pathway, dihydrodiol dehydrogenase, a member of the aldo-keto reductase superfamily, catalyzes dehydrogenation of PAH-diol. Further oxidation of catechols generates o-quinones. Redox cycling of quinones can form ROSs, which could also lead to carcinogenesis via oxidative DNA damage (Moorthy et al., 2015; Ewa and Danuta, 2016).

In short, metabolism of PAHs occurs in all tissues and involves several possible pathways. Metabolism of PAHs has been studied extensively both in vitro and in vivo. The metabolism products include epoxide intermediates, dihydrodiols, phenols, quinones, and various combinations. While phenols, quinones, and dihydrodiols can all be conjugated to form glucuronides and sulfate esters, quinones also form glutathione conjugates. The pathways for metabolic activation of PAHs can form three carcinogens: dihydrodiol epoxides requiring 2 CYP-catalyzed oxidations and epoxide hydrolase, phenols via radical cations by 1-electron oxidation, and o-quinones via catechols by involving aldo-keto reductases to form ROSs (Moorthy et al., 2015).
Metabolomics Studies on Human Exposure to PAHs

Biomarkers

It is important to understand the interplay between chemicals and humans to evaluate the effects of human activities on the environment, as well as to evaluate the effects of chemicals on human health. When investigating human health after exposure to chemicals, it is critical to establish the relationship between the magnitude of exposure and the incidence of adverse outcomes at various biological endpoints. A major challenge in studying the impacts of PAHs exposure on human health is to determine the relationship between exposure and the prevalent biological endpoints of adverse events such as cancer and irritation. This relationship can only be established if all elements on the source-exposure-dose-effect continuum are connected (Figure B-3).

For adverse health consequences of PAHs exposures, chemicals must be released from the source and transported through the media to reach the body's receptors or target sites. In addition, the chemicals should have sufficient accumulation in the target tissue in an organism to show an effect (Tan et al., 2012).

A biomarker is a substance that can be measured in a biological sample and is correlated to a metric of interest, such as disease- and drug-related biomarkers, and biomarkers of effect and susceptibility. Exposure biomarkers measure accessible biological media (e.g., blood and urine) to infer exposures to chemicals. An exposure biomarker can be the chemical itself, its metabolite, or changes of an endogenous species responding to the exposure. Since exposure biomarkers provide direct evidence of human exposure to a chemical, they have been used to reconstruct exposures in the workplace for decades. As analytical techniques to measure biomarker concentrations continue to advance, allowing detection of more chemicals at ever-lower
concentrations, they have been used to observe trends of exposures to chemicals over time and among different populations (Campo et al., 2010; Tan et al., 2012).

Different methods have been developed to assess their internal levels after human exposure to PAHs. In many studies, biotransformation products of pyrene, such as 1-hydroxy (1-OH) pyrene, have been widely used as urinary biomarkers for PAHs exposure (Hu et al., 2012; Kim et al., 2013; Keir et al., 2017). More importantly, pyrene is present in all PAHs mixtures at relatively high concentrations and its content in PAHs is fairly constant in certain environments (Sexton et al., 2011; Yasir et al., 2017). However, 1-OH pyrene cannot always be used to predict the extent of exposure to BaP or other carcinogenic PAHs (Kim et al., 2013). As such, as PAHs compositions vary among sources, urinary 1-OH pyrene alone cannot reflect the overall exposure to PAHs. In addition, while 1-OH pyrene can mimics the absorption of particle-bound PAHs through the lungs and skin, 1-OH naphthalene and 1-OH phenanthrene are biomarkers of inhaled PAHs. Thus, multiple metabolites and biotransformation products of PAHs need to be measured to more accurately assess PAHs exposure in humans (Abdel-Shafy and Mansour, 2015; Kim et al., 2013). Besides 1-OH pyrene, various hydroxylated-PAHs including mono-, poly-, and multi-hydroxy PAHs have been used as biomarkers in PAHs exposure studies, such as 2-OH naphthalene, 2-OH fluorine, 9-OH phenanthrene, and 2-OH biphenyl (Table B-1; Fernando et al., 2016; Keir et al., 2017; Lin et al., 2016; Navarro et al., 2017).

Besides metabolites, parent PAHs have also been used as biomarkers to estimate PAHs exposure. So both hydroxylated biotransformation products and unmetabolized PAHs in urine are useful biomarkers for PAHs exposure (Campo et al.,
2010; De Craemer et al., 2016). It is important to know that the concentrations of parent PAHs or their metabolites and biotransformation products in an organism depends on not only external exposure but also on their absorption, metabolism, bioconversion, detoxification, and excretion by the organism, which may differ among PAHs and varies significantly between subjects (Abdel-Shafy and Mansour, 2015; Kim et al., 2013). For instance, most higher molecular weight and lipophilic PAHs, such as BaP, are prone to bioaccumulation in adipose tissue, as well as being preferentially excreted in stool but not urine. Obviously, stool specimen collection is far less convenient to perform than random single-spot urine specimens.

In addition, BaP-DNA adducts in peripheral lymphocytes and proteins such as albumin have also been used to estimate their biotransformation products. As such, various PAH-DNA adducts have been used as biomarkers (Pratt et al., 2011). Since binding of biotransformation products of electrophilic PAHs to DNA is a key step to develop cancer, measurement of DNA adducts can estimate PAHs exposure and metabolite responses (Elie et al., 2015). Similar to DNA adducts, protein adducts have also been used as biomarkers in some studies (Castano-Vinyals et al., 2004). Besides, other chemicals related to PAHs exposure have also been used as biomarkers, which include but are not limited to inflammatory, angiogenic, and oxidative stress biomarkers (Ferguson et al., 2017; Gauthier et al., 2015; Wang et al., 2015).

To ensure the accuracy of selected biomarkers for PAHs exposure, biological validation is often used. It involves the adverse effects specifically related to the discovered biomarkers, which may be required before their acceptance. So, biological validation is becoming increasingly important. This is in part driven by increasing
number of false positive biomarkers. While it is reassuring when the altered metabolism can be directly related to the exposure, it is not always possible to understand the biological origin of the metabolic changes, given the somewhat dormant state of metabolism studies. In addition, though some biomarkers are well used, they are not well understood (Nagana Gowda et al., 2013).

In conclusion, various studies applied different biomarkers in different models in PAHs exposure studies, which are predominated by hydroxy-PAHs. Since each biomarker has its own pros and cons in each case, multiple biomarkers with biological validations and untargeted analysis are recommended in PAHs exposure studies to ensure more reliable results.

**Analytical Approaches**

Knowing the early molecular events in an exposure-effect continuum provides valuable information to intervene and change the outcome. Metabolomics is a promising molecular profiling technology to detect the biological consequences after PAHs exposure by identifying metabolic biomarkers that correlate with exposure and predict health endpoints, which is termed a meet-in-the-middle strategy. For high-throughput metabolomics technologies, they aim to reveal the metabolic characteristics of perturbations to biological systems by profiling small-molecules (<1 kDa) in biological samples. At the final “omic” level in a biological system, with clear functions of metabolites, the metabolome can provide the most “functional” information of the omics technologies (Johnson et al., 2016; Zhang et al., 2017). Such markers help to elucidate the mechanisms underlying the associated toxicity and disease etiology in humans, which could ultimately inform follow-up monitoring of populations exposed to pollutants. In addition, compared to targeted methods, untargeted global metabolomics does not
focus on a specific chemical. Rather, it provides a high-throughput and unbiased evaluation of the metabolic responses of organisms to the pollutants, making it useful when pollutant’s compositions are complicated and unknown. Metabolomics has been used to investigate the health effects of human exposure to organic pollutants. However, to our knowledge, little is known about the metabolic responses in humans after exposing to PAHs.

Mass spectrometry- (MS) and nuclear magnetic resonance- (NMR) based metabolomics approaches have been coupled with multivariate statistical data analyses to investigate PAHs perturbations in the systemic metabolism of humans (Figure B-4) (Lankadurai et al., 2013; Simpson and McKelvie, 2009; Wang et al., 2015). Data processing and statistical analysis, metabolite identification and quantification, and biological interpretation are three major processes for big data computational analysis of metabolomics research. After uploading data, mass spectra or NMR peaks are picked, realigned and annotated. The data are deconvoluted using computational tools to remove instrumental and chemical noise to provide biologically-relevant information. Initial putative metabolite identifications can be made based on the accurate mass-to-charge ratio (m/z) of ions. In addition, the analysis of biofluids, cells, and tissues reveals quantitative metabolite changes. Metabolites can be mapped and analyzed in metabolic pathways to relate metabolites to each other via interconnected biological pathways, thereby providing potential targets for further mechanistic studies. The combination of metabolomics, orthogonal biological analysis and isotope-assisted deciphering of pathways helps to ascertain the mechanisms of the aberrant phenotype. Network modeling and pathway-mapping tools can also help to understand the role metabolites
play in relation to each other and in biological aberrations. Thereafter, metabolites can be placed into context with upstream genes and proteins to lead mechanistic investigations (Johnson et al., 2016).

Mass spectrometry is widely used as a metabolomics analysis platform because of its high sensitivity, reproducibility, and versatility. It measures molecules and fragments to confirm their identity. It is obtained by measuring the m/z of ions formed by inducing charge loss or gain from a neutral species. Complex mixtures of metabolite-containing samples can be measured directly using separation methods such as liquid chromatography (LC) and gas chromatography (GC) followed by MS. Direct injection has been used in high-throughput metabolomics. However, due to hundreds of thousands of ions that may exist in metabolomics experiments, chromatographic separations should be used to minimize signal suppression and increase the sensitivity of MS. In addition, the use of retention times can further assist metabolite identification. In addition to m/z and retention time, ion recognition is facilitated by the fragmentation pattern that can be obtained by tandem MS (Johnson et al., 2016; Liu et al., 2017; Urbancova et al., 2017).

Different from targeted metabolomics, untargeted or global metabolomics analysis assesses the metabolites extracted from a sample, which can reveal novel and unanticipated perturbations. Untargeted analyses are most effective when implemented in a high-resolution MS to facilitate structural characterization of the metabolites. It offers an unbiased means to examine the relationship between interconnected metabolites from multiple pathways. However, it is impossible to obtain all metabolite classes simultaneously as many factors affect metabolite recovery, which depends on
the functional group of the metabolite. In addition, there are many unknown metabolites that remain unannotated in metabolite databases. Depending on pH, solvent, column chemistry and ionization technique used, untargeted metabolomics can provide a Therefore, by fine-tuning various parameters, untargeted metabolomics can focus on a subset of the metabolome. On the other hand, targeted metabolomics analyses measure the concentrations of a predefined set of metabolites. A standard curve for a concentration range of the metabolite of interest is prepared to gain accurate quantification. Such analysis helps to obtain the concentrations of metabolites identified by untargeted metabolomics. In addition, analytical validations, which are performed based on analytical procedures, are required to guarantee accurate and reliable concentrations (Johnson et al., 2016; Wang et al., 2015; Yang et al., 2016).

During analysis, analytes in a metabolomics sample are comprised of highly complex mixtures, which can be simplified prior to detection by analyte separation. Separation procedures achieve various goals, including separate the analytes that cannot be resolved by the detector, reduce ion suppression in MS analysis, and use retention time as identifying information. Mass spectrometry is used to identify and quantify metabolites after optional separation by GC, LC, high-performance LC (HPLC), and capillary electrophoresis (CE) (Cheng et al., 2013; Ling et al., 2014; Urbancova et al., 2017). GC-MS is the first hyphenated technique being developed. Identification leverages the distinct patterns of analyte fragments, which serves as a mass spectral fingerprint. Various libraries allow for identification of a metabolite according to its fragmentation pattern. In addition, LC-MS is a common instrument in metabolomics analysis. It has advantages including rapid analysis, quantification, and diversity, so LC-
MS provides a powerful platform to identify metabolites of biological perturbation. In addition, its function can be broadened when coupled with genomic approaches to link genetic variation to phenotypic characteristics (Elie et al., 2015). Highly selective and sensitive methods are often needed for OH-PAHs or PAHs analysis in biological specimens using "targeted" analysis often using GC or LC coupled with tandem MS, whereas "untargeted" metabolite analysis often relies on alternative mass analyzers that allow full-scan data acquisition with high mass resolution and accuracy, such as time-of-flight or orbitrap-based mass analyzers. In addition, tandem MS based on targeted analysis is not typically used for non-targeted metabolite profiling in contrast to the targeted analysis of PAHs or OH-PAHs in urine, which requires greater specificity and sensitivity. However, a major challenge in MS-based metabolomics is the identification of unknown metabolites of clinical or biological significance since a large fraction of the urine metabolome remains unknown without chemical standards or mass spectral database entries in public databases. Therefore, further research to overcome this obstacle is urgently needed.

Unlike MS, NMR spectroscopy does not rely on analyte separation, so samples can be recovered for further analyses. Various small molecule metabolites can be measured simultaneously. In this sense, NMR serves as a universal detector. Its main advantages are high analytical reproducibility and simplicity of sample preparation (Lin et al., 2015; Zhang et al., 2017). However, a major limitation of NMR is its low sensitivity, which ultimately limits the metabolome coverage as compared to high-resolution MS. Although NMR may not be as sensitive as MS methods, as one of the most-used methods in metabolomics research, $^1$H NMR-based metabolomics is an
attractive tool due to its simple sample preparation, high reproducibility and fast analysis (Katsiadaki et al., 2010; Zhang et al., 2017). This is likely because $^1$H NMR is highly reproducible, more versatile and less susceptible to sample biases. In addition, it is not as selective as some MS methods, which means it can analyze a wide variety of components in a reasonable time frame. Furthermore, some MS methods may require derivatization steps, which are more time-consuming than required for NMR spectroscopy. Overall, these NMR- and MS-based methods can both detect, identify, and quantify the metabolic changes in organisms (Ling et al., 2014).

For reliable analysis of low concentrations of OH-PAHs and their isomers in urine samples, sample pretreatment is important, including enzyme deconjugation for GC-MS, and solid-phase or liquid extraction for sample enrichment and chemical derivatization for GC-MS. Furthermore, urine samples are often preserved with boric acid or sodium azide to prevent bacterial activity if samples are not stored under -80°C. This is also the case for untargeted metabolite profiling since the lack of pre-analytical standardization can contribute to false discoveries and bias. Also, the need for quality control and quality assurance approaches in metabolomics are required, which include general analytical QA/QC, and biological and statistical validation processes.

Both MS and NMR have been successfully applied in metabolomics studies of PAHs exposure. For instance, Elie et al. (2015) utilized LC-MS/MS to conduct an untargeted metabolomics investigation into the in vivo metabolome profiles of developing zebrafish (Danio rerio) after being exposed to benz[a]anthracene (BAA) or benz[a]anthracene- 7,12-dione (BAQ) (Elie et al., 2015). In addition, Wang et al. (2015) determined 9 urine PAHs metabolites by a sensitive LC–MS/MS approach to evaluate
the PAHs exposure level of each individual. The metabolic profiles were characterized via an LC–MS-based metabolomics approach (Wang et al., 2015). Besides LC-MS/MS, GC-MS and GC-MS/MS have also been applied in PAHs study. For example, Fernando et al. (2016) evaluated firefighters’ exposure to PAHs from smoke during their training exercises at burning houses. The air and skin wipe samples were analyzed on a GC-MS while urine samples were analyzed on a GC-MS/MS (Fernando et al., 2016). In addition, Lin et al. (2015) established a naphthalene tolerance model to investigate the metabolic responses between tolerant and non-tolerant mice using NMR-based metabolomics. Similarly, Zhang et al. (2017) found that intratracheal instillation of PAH-enriched PM$_{2.5}$ induced apparent systemic metabolic changes in serum and urine samples of rats using NMR-based metabolomics approach. Moreover, Ling et al. (2014) deconvoluted the metabolites related to naphthalene intervention in various organs using both NMR and LC-MS/MS (Ling et al., 2014) (Table B-1).

**Data Analysis**

Because metabolomics approaches generate large data sets, computational tools for processing and interpretation are important. Since big data processing, statistical analysis, metabolite identification and biological explanation related issues are not trivial, there are now automated tools to accelerate computational workflows. They help to process chemical information and calculate metabolomics results effectively, which can support the uploading, processing, statistical analysis and metabolite identification of experimental data. When coupled with bioinformatics tools, they help to place metabolites in a biological context. Metabolomics analysis, especially untargeted metabolomics, can result in complicated data sets. They contain information on thousands of ions generated in MS for each sample, where ions represent the intact
metabolites or its fragments, adducts or isotopes. Therefore, applying computing tools to reduce the redundancy of these complex datasets and to identify the most relevant metabolites is important (Bundy et al., 2009; Johnson et al., 2016; Lankadurai et al., 2013).

Metabolomics aims to provide a global snapshot of all small-molecule metabolites in cells and biological fluids. Such a process can minimize observational biases inherent to more focused studies of metabolism. However, the staggeringly large amounts of information of such global analyses introduce a challenge of its own. Therefore, gaining biologically relevant conclusions from a given metabolomics dataset requires a specialized data analysis. Multivariate analysis involves observation and analysis of multiple statistical outcome variables at a time. In design and analysis, the technique is used to perform trade studies across multiple dimensions while taking into account the effects of all variables on the responses of interest (Schervish, 1987). This approach can be useful to find meaning in metabolomics datasets using methods such as principal component analysis (PCA) and partial least squares (PLS), where spectral features contributing most to variation or separation are identified for further analysis. Techniques such as PCA and PLS provide an essential platform to rapidly interpret information-rich spectral datasets to infer biological conclusions. Through proper application of preprocessing transformations, best choice of analysis algorithms, and rational application of validation metrics, multivariate analysis can provide a powerful means for biological understanding and exploration of complex, multi-parametric metabolic systems. Nevertheless, misunderstandings and misuse of multivariate analysis can result in misleading or erroneous biological inferences. In addition,
metabolomics still has various data challenges left to be solved, but machine learning chemometrics methods may have much to offer for the big data analyses of metabolomics (Worley and Powers, 2015).

As mentioned above, multivariate analysis techniques are widely used in big data analyses of PAHs metabolomics studies, which include PCA-discriminant analysis (PCA-DA) (Elie et al., 2015), PLS-DA (Fernando et al., 2016; Ling et al., 2014), and orthogonal projections to latent structures-DA (OPLS-DA) (Zhang et al., 2017). In addition, global metabolic pathways can show the disorders of metabolism in exposure groups. While heat maps show the alteration of various metabolomes among different groups, volcano plots can be useful to show fold change vs. significance as p-value (Figure B-5) (Elie et al., 2015; Wang et al., 2015; Fernando et al., 2016; Zhang et al., 2017).

Among all multivariate analysis techniques, PCA is arguably the most widely used for metabolic fingerprinting and, in fact, chemometrics in general. Its objective is to arrive at a linear transformation that preserves the variance of the original data in a lower dimensionality in the output data. While the unsupervised nature of the PCA algorithm provides a means to achieve unbiased dimension reduction, its application only reveals group structure when the within-group variation is less than between-group variation. Therefore, supervised forms of discriminant analysis such as PLS and PLS-DA that rely on class membership of each observation are also commonly applied in metabolic fingerprinting experiments. The utilization of class memberships in PLS-DA allows the algorithm to better expose separations between classes in score space. However, variations not directly correlated with class memberships coded matrix is still
present in the scores. This complicates the interpretation of PLS-DA scores and loading plots, especially as the number of classes increases. To addresses this interpretability problem, OPLS incorporates an orthogonal signal correction filter into a PLS model, thereby effectively separating predictive variation from uncorrelated variation (Worley and Powers, 2015).

Since both PLS and OPLS have an inherent tendency to over-fit models to data, they can identify excellent class separation in completely random variables. For both, validation is a critical step in ensuring model reliability. Truly honest model validation requires dividing the data into a training set to build a model and a validation set to assess model prediction ability, where the validation set cannot be used to generate the trained model (Broadhurst and Kell, 2006). As such, internal cross-validation is routinely employed, where the leave-one-out method is a common choice. However, it has been demonstrated that leave-one-out internal cross-validation should be abandoned in favor of the more consistent leave-n-out method. (Eriksson et al., 2000; Eliasson et al., 2011). In addition, multivariate statistical methods require the use of appropriate data pre-processing procedure including scaling, normalization, log transformation and quality control filtering of metabolomics datasets together with rigorous cross-validation or permutation testing of models especially without independent replication or hold-out sets for external validation.

For example, to reveal metabolic differences between PAH-enriched PM$_{2.5}$ exposure and control group, Zhang et al. (2017) constructed OPLS-DA models for investigation. In OPLS-DA models, clear clustering and separation of PM$_{2.5}$-H NMR and control group were observed in both serum and urine samples (Zhang et al., 2017).
Also, to investigate the metabolic mechanism change after PM$_{2.5}$ exposure, metabolic pathways of the significantly-changed metabolites were analyzed using the “pathway analysis” module. The results showed 18 metabolic pathways were disturbed in PM$_{2.5}$-exposed rats. In the study of Wang et al. (2015), to achieve the maximum separation, the metabolomics profiles of the exposed and control groups were compared using supervised multivariate OPLS-DA analysis. As a result, the OPLS-DA scatter plot showed that the control group could be clearly separated from the exposed group based on 1400 UHPLC-MS peaks. In addition, although significant between-subject and between-site variability in chemical exposure exist, based on supervised multivariate analysis using PLS-DA of log-transformed/autoscaled data, Ling et al. (2014) applied PLS-DA score plot to show the clustering of the LC-MS/MS spectra of ceramide- and phosphorylcholine-containing lipid extracted from various tissues of ICR mice 48 h after exposing to naphthalene.

To directly evaluate the induced metabolic alterations and determine the differences in metabolic profiles among the groups, Elie et al. (2015) applied both PLS-DA and PCA-DA techniques to analyze data sets. Based on the results represented as score plots, exposed and control groups displayed a metabolic profile clustering together in an area of the plot. Together, the two-dimensional score plots of the PLS-DA and PCA-DA models demonstrated that the benz[a]anthracene (BAA) and benz[a]anthracene- 7,12-dione (BAQ) groups were separated from each other and from the control group. The clear separation from the control group indicated that the metabolic perturbations were associated with BAA and BAQ exposure. Based on the global metabolic disorders, Elie et al. (2015) showed the most relevant pathways in
zebrafish influenced by BAA and BAQ exposure. PCA-DA was also applied by Lin et al. (2015) in their study of the lungs, liver, and kidneys in mice after naphthalene exposures. The PCA score plot showed the grouping of the NMR spectra of bronchial alveolar lavage fluid, and hydrophilic and hydrophobic metabolites extracted from mouse organs after naphthalene exposure.

**Animal Models**

Though many studies have investigated PAHs metabolism, few have used metabolomics approaches. As an emerging and useful method, more research is needed in this area. Several studies have been done using animal models including mouse (Lin et al., 2015; Ling et al., 2014), rat (Zhang et al., 2017) and fish (Elie et al., 2015; Katsiadaki et al., 2010) (Table B-1). For example, Ling et al. (2014) investigated the metabolites related to naphthalene exposure in mouse organs using NMR and LC-MS/MS. Male mice were intraperitoneally dosed with olive oil and a low and high dose of naphthalene. After 48 h, the lungs, liver, and kidneys were analyzed for metabolic responses. The metabolites were extracted and non-targeted profiles were obtained using NMR. Low NMR resolution limited the identification of hydrophobic metabolites. Therefore, LC-MS/MS-based metabolomics was applied to profile phosphatidylcholine-containing lipids and sphingolipids. Chemometric analysis revealed that succinate and lactate were increased in the lungs, suggesting that energy metabolism and antioxidation were increased following naphthalene exposure. In the liver, antioxidative stress-related metabolites were increased to overcome oxidative stress during naphthalene biotransformation and detoxification. However, the elevation of glutathione protected kidneys from reactive naphthalene-metabolite-induced injury. Significant alteration of hydrophobic metabolites revealed lung and liver were the target organs of
naphthalene exposure. In addition, MS data demonstrated that phosphatidylcholine and ceramide species were altered in the lungs and liver, whereas only phosphatidylcholine alterations were observed in the kidneys. Further, elevated numbers of unsaturated bonds and fatty acyl chains in both ceramides and phosphatidylcholines were determined to reduce cellular membrane rigidity and facilitate trafficking of recovery elements into the cell for rejuvenation. Hence, the complementary results of NMR- and MS-based metabolomics effectively characterized naphthalene-induced changes in various organs (Ling et al., 2014).

Besides mouse, Zhang et al. (2017) investigated the toxicity mechanisms of PAHs in PM$_{2.5}$ by exploring the endogenous metabolic changes and possible influence on metabolic pathways in rats after intratracheal instillation of PM$_{2.5}$ using an NMR-based metabolomics approach. Total PAHs concentration was 1042 μg/g PM$_{2.5}$ samples, including high levels of BaP (87.6 μg/g), indeno[1,2,3-cd]pyrene (107 μg/g), and benzo[g,h,i]peryene (90.6 μg/g) in PM$_{2.5}$ samples. Histopathology showed cellular edema in the liver and glomerulus atrophy in the kidneys of PM$_{2.5}$-exposed rats. They also analyzed the metabolite changes in the serum and urine using $^1$H NMR technique coupled with multivariate statistical analysis. Reduced levels of lactate, alanine, dimethylglycine, creatine, glycine and histidine in serum, together with increased levels of citrate, arginine, hippurate, and allantoin and decreased levels of threonine, lactate, alanine, acetate, succinate, trimethylamine and formate in urine were observed for rats exposed to PM$_{2.5}$. The PM$_{2.5}$-affected metabolic pathways included glycine, serine and threonine metabolism, glyoxylate and dicarboxylate metabolism, citrate cycle, and nitrogen and methane metabolism. (Zhang et al., 2017).
In addition to mouse and rat, Elie et al. (2015) used an untargeted metabolomics approach to examine the metabolomic profiles of developing zebrafish after exposing to 4 μM benz[a]anthracene (BAA) or benz[a]anthracene-7,12-dione (BAQ). By integrating multivariate, univariate and pathway analyses, 63 altered metabolites were detected after 5 d of exposure. The metabolomics data revealed that BAA and BAD exposures were associated with changes known to affect protein biosynthesis, mitochondrial dysfunction (oxidative stress), neural development, and disturbance in vascular development and cardiac development function (cardiac toxicity). Their transcriptomics and genomics data were incorporated to provide a more comprehensive view of the relationship between PAHs and oxy-PAHs exposures on vertebrate development. The utility of LC-MS-based metabolomics combined with the developmental zebrafish model provided mechanistic insights into the connections between chemical exposure and associated effects on organisms. Application of this approach to more structurally diverse PAHs and PAHs derivatives could expand the utility of metabolomics to predict the structure-activity relationships and hazard potential of the different and ubiquitous class of contaminants (Elie et al., 2015).

In addition to animals, Wang et al. (2018) developed a metabolomics strategy to assess the combined effects of PAHs and short-chain chlorinated paraffins (SCCPs) on the metabolism in human hepatoma HepG2 cells. Based on their results, a total of 21 metabolic pathways were significantly influenced after exposing to PAHs and SCCPs. For example, PAHs exposure perturbed the citric acid cycle and purine metabolism. The combined exposure significantly impacted several lipid metabolic pathways, including glycerophospholipid, linoleic acid, α-linolenic acid and arachidonic acid metabolisms.
Coupled with multivariate statistical analyses, Wang et al. (2015) applied an LC-MS-based metabolomics method to study the disturbances in human systemic metabolism after exposing to PAHs. This was achieved by analyzing urine samples of a large population living in an coking industry polluted area and also a control area. Metabolic changes responding to PAHs exposure were used to investigate potential metabolic biomarkers. Exposure of PAHs in the same population was assessed by determining nine urinary PAHs metabolites using LC-MS/MS. As a result, correlation analyses between individual PAHs exposure and its metabolic changes were performed to access the dose-effect relationship. The results showed that 18 identified metabolites associated with amino acid, purine, lipid, and glucuronic acid metabolism were altered in the exposed group. In addition, they concluded that 1-OH phenanthrene and dodecadienylcarnitine can be used as sensitive and reliable biomarkers for PAHs exposure and their metabolic consequences, respectively. These results affirmed that PAHs exposure causes oxidative stress-related effects in humans (Wang et al., 2015).

For biological validation of those metabolomics results, different models need to compare with each other and also with other PAHs metabolism and biotransformation studies. For example, lipid peroxidation, protein breakdown, mitochondrial dysfunction, antioxidants depletion, purine metabolism interruption, and citric acid cycle perturbation were observed in human PAHs exposures studies (Wang et al., 2015; Wang et al., 2018). Similarly, protein biosynthesis perturbation and mitochondrial dysfunction were found in zebrafish after PAHs exposure (Elie et al., 2015). In addition, both mouse and rat studies revealed that oxidative damage, amino acids metabolism disturbance, and energy metabolism interruption occur after PAHs exposure (Figure B-6). Although
observed lipid peroxidation was not mentioned in mouse, rat, or zebrafish studies, several studies showed that lipid peroxidation and lipid-related dysfunctions were induced by PAHs exposure in humans (Fu et al., 2012; Lin et al., 2016). Based on the current database, metabolism regulations were not fully studied and lack of overlap between different models makes it hard to validate human exposure to PAHs using animal model. Since different metabolic pathways have been studied based on different models, making some results confusing and conflicting, more comprehensive metabolomics studies on PAHs exposure using different models are encouraged.

Using various orthogonal techniques, targets identified with metabolomics can be further validated and investigated in more detail. For example, other ‘omics’ approaches, including (epi)genomics, transcriptomics, and proteomics, can reveal further mechanistic insights into metabolite-related phenotypic alterations. Various orthogonal methods also allow targeting of metabolic pathways and can be used to influence metabolite levels and to interfere with metabolic pathways. These methods can be targeted at the gene level and designed for silencing gene expression by applying techniques such as CRISPR–Cas-mediated knock outs or RNA interference. Alternatively, the use of antimetabolites can affect the metabolic pathways at the protein level. Manipulating sources of exposure to different chemicals can also influence the metabolome and provide further mechanistic insights. For example, using antibiotics or germ-free models with species-specific inoculation reveals the direct effect of the microbiome on metabolic pathways. Similarly, immunomodulators can be used to alter the efficacy of the host immune system to respond to both the resident microbiota and
pathogens, and their metabolites. This collectively opens up possibility to better understand and control metabolism eventually (Johnson et al., 2016).

In short, adverse effects have been found after PAHs exposure based on human and animal studies. More precisely, oxidative stress and damage, protein biosynthesis perturbations, organ dysfunction and metabolism disturbances were found in humans and animals. Moreover, some metabolites may not be impacted by PAHs exposure, so conflicting results were found for some metabolites based on different methods and different models. It is reasonable since different analytical methods, endpoints and models may have different metabolomics reactions even under similar PAHs exposure (Wang et al., 2015; Elie et al., 2015; Zhang et al., 2017). However, the impacts of low level and chronic PAHs exposures of broader metabolism was yet well understood.

**Conclusions**

PAHs and their derivatives are ubiquitous in the environment. Due to their toxic, carcinogenic, and mutagenic nature, it is important to study their biological effects and associated mechanisms in humans. As discussed in this review, metabolomics analysis is a powerful tool to achieve these goals. Results demonstrated the utility of LC-MS/MS and NMR based metabolomics in combination with various statistical models to provide better biological insights into the link between PAHs exposure and the associated impacts on humans. Applying these approaches to more PAHs and PAHs derivatives can extend the metabolomics to predict the structure-activity relationships and may provide helpful information for metabolism and biotransformation of similar organic pollutants in humans. Moreover, by applying big data multivariate analysis methods, we can better understand how PAHs and their derivatives alter biochemical pathways to cause adverse health effects in humans. In addition, complementary "omics"
technologies such as genomics, proteomics are also needed to better understand the significance of metabolic perturbations induced by PAHs exposures. Also, biochemical interpretation of metabolic alterations after PAHs exposures is challenging, especially when relying on single time point or cross-sectional studies. In short, these metabolomics studies provide valuable information for toxicology and molecular biology research, which help to understand, influence and control the metabolisms of PAHs.
<table>
<thead>
<tr>
<th>Exposure</th>
<th>Model</th>
<th>Endpoints</th>
<th>Biomarkers</th>
<th>Analysis approaches</th>
<th>Statistical analysis</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAHs in cigarettes</td>
<td>Human</td>
<td>Urine</td>
<td>1-OH naphthalene, 2-OH naphthalene, 2-OH fluorene, 1-OH phenanthrene, 2-OH phenanthrene, 3-OH phenanthrene, 4-OH phenanthrene, 1-OH pyrene, and 6-OH chrysene</td>
<td>UHPLC-MS</td>
<td>OPLS-DA</td>
<td>Wang et al., 2015</td>
</tr>
<tr>
<td>PAHs in wood smoke</td>
<td>Human</td>
<td>Urine</td>
<td>1-OH naphthalene, 2-OH naphthalene, 9, 3, and 2-OH fluorene, 4, 3, and 2-OH phenanthrene, 3-OH fluoranthene, and 1-OH pyrene</td>
<td>GC-MS &amp; GC-MS/MS</td>
<td>PLS-DA</td>
<td>Fernando et al., 2016</td>
</tr>
<tr>
<td>Naphthalene</td>
<td>Mouse</td>
<td>Urine, serum, lung, liver, and kidney</td>
<td>Untargeted</td>
<td>^H NMR &amp; LC-MS/MS</td>
<td>PLS-DA</td>
<td>Ling et al., 2014</td>
</tr>
<tr>
<td>PAHs in PM_{2.5}</td>
<td>Rat</td>
<td>Lung, liver, and kidney Serum and urine</td>
<td>Untargeted</td>
<td>^H NMR</td>
<td>PCA-DA</td>
<td>Lin et al., 2015</td>
</tr>
<tr>
<td>Benz[a]anthracene and benz[a]anthracene-7,12-dione</td>
<td>Zebrafish</td>
<td>Entire larvae</td>
<td>Untargeted</td>
<td>LC-MS/MS</td>
<td></td>
<td>Elie et al., 2015</td>
</tr>
<tr>
<td>1,2:5,6-Dibenzanthracene</td>
<td>Stickleback</td>
<td>Liver and kidney</td>
<td>Untargeted</td>
<td>^H NMR</td>
<td>PCA-DA</td>
<td>Williams et al., 2009</td>
</tr>
</tbody>
</table>
Figure B-1. Biological active “bay” (green) and “fjord” (red) regions in PAH conformation using dibenzo[a,l]pyrene as an example (Ewa and Danuta, 2016).
Figure B-2. Three metabolic activation pathways of benzo[a]pyrene to form DNA adduct metabolites to cause mutation and detoxification where 7,8-oxide BaP and 7,8-dihydrodiol BaP are proximate carcinogens, 7,8-diol-9,10-epoxide BaP is a strong mutagen and ultimate carcinogen (Ewa and Danuta, 2016 and Kim et al., 2013).
Figure B-3. Source-exposure-dose-effect outcome continuum: exposure biomarker is the key to exposure reconstruction (Tan et al., 2012).
Figure B-4. Schemes for PAHs metabolomics experiment based on LC-MS/MS and/or NMR approach (Lankadurai et al., 2013, Simpson and McKelvie, 2009 and Wang et al., 2015).
Figure B-5. Common multivariate statistical analysis in metabolomics studies (Wang et al., 2015, Elie et al., 2015, Fernando et al., 2016 and Zhang et al., 2017).
Figure B-6. Metabolome alterations in human and animals by comparing control and exposure groups. Red, blue, and purple box represent up regulations, down regulations, and insignificant changes/dependent regulations in exposure groups (Ling et al., 2014, Lin et al., 2015, Wang et al., 2015, Elie et al., 2015, and Zhang et al., 2017).
LIST OF REFERENCES


Peng Gao was born in Shenzhen, China in 1991 as the single child of his parents. He obtained his B.S. in Chemistry at the China Agricultural University in 2013. After that, he went to Case Western Reserve University and earned his M.S. in chemistry in 2015. After being recruited by Dr. Lena Q. Ma, he switched his major to environmental soil chemistry and joined the Soil and Water Sciences Department at the University of Florida to pursue his Ph.D. degree in August 2015. He worked on PAHs in Florida urban soils and expected to earn his Ph.D. from the University of Florida in August 2019.