Petrogenic Hydrocarbons in the Peace-Athabasca Delta and their Potential for Microbial Degradation

Michelle-Claire Roy

This thesis is submitted to the Faculty of Graduate and Postdoctoral Studies in partial fulfillment of the requirements for

Master of Science

In

Biology

Specializing in Chemical and Environmental Toxicology

Department of Biology
Faculty of Science
University of Ottawa

Supervisor:

Dr. Jules Blais (University of Ottawa)

Thesis Advisory Committee:

Dr. Alexandre Poulain (University of Ottawa)
Dr. Frances Pick (University of Ottawa)
Dr. Iain Lambert (Carleton University)

University of Ottawa

Ottawa, Canada

ABSTRACT

Microbial biodegradation is the primary mechanism by which petrogenic hydrocarbons (PHCs) are removed from the environment. Though hydrocarbon biodegradation is widely studied in marine systems, knowledge of how it occurs in freshwater systems is still lacking. The Peace-Athabasca Delta (PAD), located in northeastern Alberta, is an ideal location to study microbial hydrocarbon degradation since it has a long history of exposure to PHCs. What's more, these PHCs are predominately sourced from bituminous deposits and are therefore relevant to the Canadian Oil Sands Industry.

This thesis investigated the genetic potential for hydrocarbon degradation of PHCs via metagenomic reconstruction of microbial communities in lakes of the Peace and Athabasca Deltas, as well as reference lakes in the nearby boreal uplands. In order to properly evaluate the microbial community and its potential for hydrocarbon degradation, a comprehensive analysis of PHCs (including *n*-alkanes, polycyclic aromatic compounds (PACs), and petroleum biomarkers of terpanes, hopanes, and steranes) was performed. PHC analysis showed that *n*-alkanes in lake sediments from all three regions were highly similar and predominately biogenic, while PAC composition was significantly different in each region. Restricted-drainage lakes of the Athabasca Delta had the highest concentrations of PACs from petrogenic sources. Closed-drainage lakes in the Peace Delta had lower concentrations of PACs that likely originated from a mixture of pyrogenic and petrogenic sources. Closed-drainage lakes in the boreal upland region had the lowest concentrations of PACs likely sourced from pyrogenic wood combustion with traces of petrogenic PACs, possibly from atmospheric deposition of dust. Petroleum biomarkers of terpanes, hopanes, and steranes were successfully used to identify the long-range fluvial, and possibly atmospheric, transport of bituminous compounds more than one hundred kilometers

from their potential source. This validates the future use of these biomarkers in environmental forensics.

Microbial communities in all three regions under study were highly diverse, and their composition was significantly different in both sediment and water. Targeted gene analysis identified a total of 3885 genes involved in the degradation of *n*-alkanes and PACs in sediment and water. The results show that organic carbon, nitrogen, and sulfur content, as well as PAC and short-chain alkane concentrations were important chemical predictors of change in degradation gene composition. Furthermore, genes for anaerobic degradation of PHCs were identified in syntrophic bacteria, methanogens, nitrate and sulfate reducers, demonstrating the potential for syntrophic hydrocarbon degradation in PAD lakes. Though this thesis confirms the genetic potential for hydrocarbon degradation in PAD and boreal upland lakes, further research is necessary to determine whether these microbial communities can actively degrade the PHCs present in these lakes.

ACKNOWLEDGEMENTS AND DEDICATION

First and foremost, I would like to acknowledge that this research took place on Treaty 8 territory, the traditional land of the Mikisew Cree First Nation, Athabasca Chipewyan First Nation, and Métis Nation of Alberta Local #125. The Peace-Athabasca Delta is quite possibly one of the most beautiful places I have been in my life. I dedicate this thesis to the community of Fort Chipewyan. I hope that the research herein will contribute to the growing body of knowledge on the PAD and spark future research/regulation that will aid the community. I would like to express gratitude to Robert Grandjambe for being my guide in the field, as well as Liam McKay and the Community Based Monitoring team for field support.

None of this would be possible without my supervisor Jules M. Blais – thank you Jules for giving me this opportunity and the tools to make this thesis happen. Your curiosity and enthusiasm are what fueled me to keep going throughout my thesis. To my (unofficial) co-supervisor Charles W. Greer – thank you for welcoming me as part of the team in Montréal and guiding me through the journey of being an ecologist navigating through a multi-disciplinary project. I'd also like to acknowledge the many labs I was a part of for this thesis. Thank you to the Emergencies Science and Technology lab over at Environment and Climate Change Canada, Keval Shah for all of your help with extractions and Zeyu Yang for your guidance in GC interpretations. Thank you to the team at Energy, Mining and Environment of the National Research Council, Sylvie Sanschagrin and Julien Tremblay you were indispensable. Thank you for your help in the lab and metagenomics mastery. Thank you to the Blais Lab members. You have all played a crucial role in helping me along, encouraging me through thick and thin, and many much-needed beers after long day's work.

Most importantly, I must thank my partner Chloé Patry for her never-ending support and encouragement throughout this process. Thank you for your patience and tolerating my sleep-talking rants about metagenomics. You kept me sane and smiling these past 3 years. Thank you to my brother Martin Roy for your support. Even though we live on separate coasts I always knew I could count on you to pick up the phone. Lastly, to my parents Lorraine & Marc – thank you for believing in me and always supporting me in whatever I chose in life. You are truly the best parents a person could ask for. Thank you.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

a.s.l.	Above Sea Level
AHC	Agglomerative Hierarchical Clustering
AHS	Albian Heavy Synthetic oil
ANOVA	Analysis of Variance
AOS	Alberta Oil Sands
AOSR	Athabasca Oil Sands Region
ASMB	Alberta Sweet Mixed Blend
BASL	University of Alberta's Biogeochemical Analysis Service Laboratory
BCF	Bioconcentration Factor
BNT	Benzonaphthothiophene
C/N	Carbon to Nitrogen ratio
CNS	Carbon, Nitrogen, Sulfur
CPI	Carbon Preference Index
CPM	Count Per Million
DBT	Dibenzothiophene
DCM	Dichloromethane
DIC	Dissolved Inorganic Carbon
DOB	Diluted crude Oil sands Bitumen
DOC	Dissolved Organic Carbon
dszB	2'-hydroxybiphenyl-2-sulfinate desulfinase
dw	Dry Weight
ECCC	Environmental and Climate Change Canada
EDTA	Ethylenediaminetetraacetic Acid
FL	Fluoranthene
GAM	Gammacerane
GC-MS	Gas Chromatography coupled with Mass Spectrometry
IS	Internal Standard
KEGG	Kyoto Encyclopedia of Genes and Genomes
K_{ow}	Octanol-Water Partition Coefficient
LC_{50}	Median Lethal Concentration
NCBI	National Center for Biotechnology Information
NMDS	Non-metric Multidimensional Scaling
NRC	National Research Council of Canada
OC	Organic Carbon
PAC	Polycyclic Aromatic Compound
PAD	Peace-Athabasca Delta
P_{aq}	Proxy Ratio, or P-aqueous Ratio
PASH	Polycyclic Aromatic Sulfur Heterocycles

PCA	Principal Component Analysis
perMANOVA	Permutational Analysis of Variance
PHC	Petrogenic Hydrocarbon
PY	Pyrene
RDA	Redundancy Analysis
S/N	Signal to Noise
SIMPER	Similarity Percentages
TDN	Total Dissolved Nitrogen
TDP	Total Dissolved Phosphorus
TN	Total Nitrogen
TP	Total Phosphorus
UNESCO	United Nations Educational, Scientific and Cultural Organization
W.A.C.	William Andrew Cecil
WO_3	Tungsten Trioxide
χ^2	Chi-Squared

LIST OF TABLES

Chapter 1

Chapter 2

Chapter 3

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

PREFACE

This thesis is written as a collection of manuscripts according to the guidelines provided by the University of Ottawa's Faculty of Graduate and Postdoctoral Studies. Chapter 1 is an introduction to the thesis and provides the necessary background for the manuscripts that follow, as well as the primary aim, hypothesis, and predictions of the thesis. Chapter 2 and 3 each provide a brief introduction, methodology, results and discussion, and are formatted according to the American Chemical Society guidelines for authors provided by the journal of Environmental Science and Technology (ES&T). Chapter 4 summarizes each manuscript and concludes the thesis, elaborating on the primary aim and hypothesis introduced in Chapter 1, as well as concluding remarks and future directions for this research.

STATEMENT OF CONTRIBUTIONS

Chapter 2: Tracking Petrogenic Hydrocarbons in Lakes of the Peace-Athabasca Delta Using Petroleum Biomarkers

Authors: Michelle-Claire Roy[1], Joshua R. Thienpont[1], Zeyu Yang[4], Roland I. Hall[2], Brent B. Wolfe[3], Bruce P. Hollebone[4], Jules M. Blais[1]

Affiliations:

[1]University of Ottawa, Ottawa, Ontario, Canada

[2]Department of Biology, University of Waterloo, Waterloo, Ontario, Canada

[3]Department of Geography and Environmental Studies, Wilfrid Laurier University, Waterloo, Ontario, Canada

[4]Emergencies Science and Technology Section, Environment Canada, Ottawa, Ontario, Canada

Study Design ……	Michelle-Claire Roy
	Jules M. Blais
	Joshua R. Thienpont
	Roland I. Hall
	Brent B. Wolfe
Sample Collection ……	Michelle-Claire Roy
	Joshua R. Thienpont
	Roland I. Hall
	Robert Grandjambe
	Liam McKay
PHC Extraction ……..……..	Michelle-Claire Roy
	Bruce P. Hollebone
	Keval Shah
	Zeyu Yang
Data Analysis and Interpretation …….	Michelle-Claire Roy
	Zeyu Yang
	Jules M. Blais

Manuscript Preparation …… Michelle-Claire Roy

Peer Reviewers …… Jules M. Blais

Joshua R. Thienpont

Zeyu Yang

Roland I. Hall

Brent B. Wolfe

Bruce P. Hollebone

Chapter 3: Metagenomic Analysis of the Hydrocarbon Degradation Potential of Microbial Communities in Lakes of the Peace-Athabasca Delta, AB, Canada

Authors: Michelle-Claire Roy[1], Joshua R. Thienpont[1], Julien Tremblay[2], Jessica Wasserscheid[2], Charles W. Greer[2], and Jules M. Blais[1]

Affiliations:

[1]University of Ottawa, Ottawa, Ontario, Canada

[2]Energy, Mining and Environment, National Research Council Canada, Montréal, Québec, Canada

Study Design ……	Michelle-Claire Roy
	Jules M. Blais
	Charles W. Greer
	Joshua R. Thienpont
Sample Collection ……	Michelle-Claire
	Joshua R. Thienpont
	Roland I. Hall
	Robert Grandjambe
	Liam McKay
DNA Extraction ……	Michelle-Claire Roy
PHC Extraction and Analysis……..	Michelle-Claire Roy
	Bruce P. Hollebone
	Keval Shah
	Zeyu Yang
Metagenomic Library Prep and Sequencing ……	Michelle-Claire Roy
	Charles W. Greer
	Sylvie Sanschagrin
Genomics and Bioinformatics …….	Julien Tremblay
	Jessica Wasserscheid
Statistical Analyses ……..	Michelle-Claire Roy

	Julien Tremblay
	Mija Azdajic
Data Interpretation ……..	Michelle-Claire Roy
	Charles W. Greer
	Jules M. Blais
Manuscript Preparation ……	Michelle-Claire Roy
Peer Reviewers ……	Charles W. Greer
	Jules M. Blais
	Julien Tremblay
	Joshua R. Thienpont
	Jessica Wasserscheid

CHAPTER 1. GENERAL INTRODUCTION

1.1 The Peace-Athabasca Delta, AB, Canada

The Peace-Athabasca Delta (PAD), located in northeastern Alberta, is the world's largest inland freshwater boreal delta. This Ramsar Convention Wetland of International Importance spans over 5,000 km^2 of rivers, channels, lakes, grasslands, and marshes and shares much of its extent with Wood Buffalo National Park (est. 1922), a UNESCO World Heritage Site (Figure 1.1) (Timoney & Lee, 2009; Timoney, 2013c; Wolfe, Hall, Edwards, & Johnston, 2012). The PAD is a major nesting and staging area for waterfowl of all four North American migratory flyways and home to an identified 219 species of birds, 43 species of mammals, and 31 species of fish (Timoney, 2013b). The PAD also provides sustenance and spiritual well-being to First Nations and Métis people of the land, including the Mikisew Cree First Nation, Athabasca Chipewyan First Nation, and the Métis Nation of Alberta Local #125. Indigenous land users have relied on the PAD and its rich ecosystem for millennia and continue to do so today.

River deltas like the PAD result from the deposition of sediments carried by rivers. Two major rivers give rise to the PAD, the Peace and Athabasca rivers. The Peace River originates at the fork of the Finlay and Parsnip Rivers – what is now the Williston Reservoir - in northern British Columbia. The Athabasca River originates from the Columbia Icefield of the Alberta Rockies. A variety of land uses occurs in the catchment of these rivers, most notably the W.A.C. Bennett Dam that uses the Peace River for hydroelectric generation, and the Athabasca Oil Sands Region (AOSR) that surrounds the Athabasca River. These rivers provide floodwaters and discharge to the PAD, which recharge hundreds of water bodies, making them essential to the delta's health. Differences in their flow and hydrology give rise to distinct 'sectors' within the PAD with varying degrees of flooding activity and sensitivity to climate variability

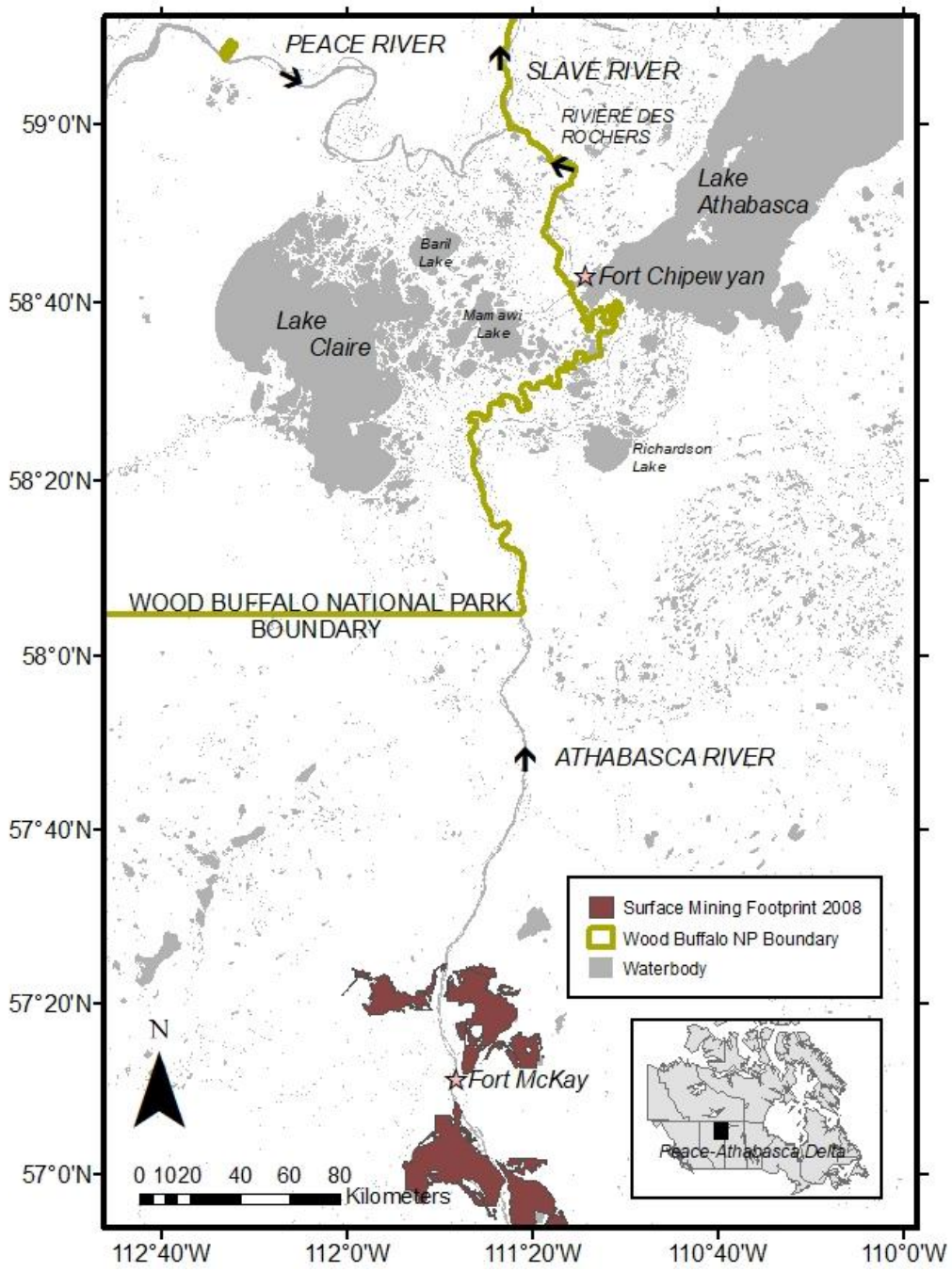

Figure 1.1. Map of the study region. Major rivers are shown with their dominant flow direction (black arrows) under normal, non-ice-jam conditions. Rivière des Rochers is particularly subject to flow reversal. The surface mining footprint of the AOS is shown in dark red (Global Forest Watch Canada, 2008). The outline of Wood Buffalo National Park is shown in green (Natural Resources Canada, 2008).

(Timoney, 2013c; Wolfe et al., 2012). The northern Peace Delta receives its water from Lake Athabasca and experiences episodic ice-jam floods of the Peace River that produces overland flooding due to ice accumulation during the spring thaw. River banks established from the deposition of sediments on the Peace River make it so that the Peace delta relies on ice jam flooding for periodic recharge of the landscape and perched lake basins. Water in the Peace Delta flows northward from Lake Athabasca through Rivière des Rochers and Rivière des Quatres Fourches to join the Peace River in its conflux with the Slave River, ultimately reaching the Slave Delta in the Northwest Territories. The southern active Athabasca Delta receives discharge from the Athabasca River, flowing northeast to Lake Athabasca. This portion of the delta receives floodwaters from its active rivers and channels as well as ice jam flooding. Major lakes in the delta include Athabasca, Mamawi, Claire, Richardson, and Baril.

Deltas are naturally prone to changes in water flow and sediment deposition over time. This dynamism also makes them sensitive to climate change and river management. The PAD is no exception to this, being subject to various stressors such as climate change and upstream industrial development including mining, forestry, agriculture, pulp mills and major resource extraction (Wolfe et al., 2012). Concerns that upstream industrial developments are contributing to contamination in the PAD have prompted research assessing the current state of the delta. Activities in the AOSR have previously been shown to emit polycyclic aromatic compounds (PACs), some of which are known to be toxic, carcinogenic, and mutagenic (Evans et al., 2016). It is unclear, however, if such compounds are widely distributed in the PAD, or if upstream activities are a significant source of these contaminants. To compound this issue, the lack of baseline data makes it difficult to assess any potential increase in PACs from industrial activities beyond what is naturally present in the watershed. Previous work has shown that the PAD is in

fact enriched in petrogenic hydrocarbons (PHCs) derived from bituminous sands, namely PACs, which originate from natural deposits and extraction activities of the oil sands industry (Hall et al., 2012; Jautzy, Ahad, Hall, et al., 2015; Timoney & Lee, 2011).

1.2 Petrogenic Hydrocarbons in the Athabasca Oil Sands Region and Surrounding Watershed

1.2.1 Chemical Characterisation

In the environment, PHCs can originate directly from, or by combustion of, petroleum products such as crude oil, bitumen (a term synonymous to oil sands), diesel, as well as other refined oils. This class of hydrocarbons encompasses a wide variety of saturated and aromatic compounds including alkanes, PACs, asphaltenes, and resins (Yang et al., 2011). The contribution of hydrocarbons within each class of compounds will depend on the petroleum product of origin such that viscous oils like bitumen will contain a greater proportion of heavy PACs, asphaltenes and resins than lighter crude oils. Typically, research on chemical characterisation of petroleum, as well as source determination (environmental forensics) for PHCs, targets compounds that are most likely to be transported and remain in the environment, as well as compounds that can help distinguish different petroleum sources. These include a distribution of *n*-alkane compounds ranging from carbon chain-lengths *n*-C_9 to *n*-C_{40}, alkylated and unsubstituted PACs ranging from 2 to 6 rings, and petroleum biomarkers including select hopanes, terpanes, and steranes (Evans et al., 2016; Hall et al., 2012; Wang et al., 2014; Yang et al., 2011).

Petrogenic alkanes show distinct signatures depending on the petroleum product and its origin. Typically, more refined oil will contain a larger fraction of small-chain alkanes ranging from *n*-C_9 to *n*-C_{36}, whereas higher viscous oils such as bitumen will contain less overall *n*-

alkanes and range from *n*-C_{16} to *n*-C_{36} (Yang et al., 2011). When it comes to environmental samples, the carbon preference index (CPI) of the *n*-alkane distribution can help differentiate between petrogenic and biogenic sources of alkanes. This index has been widely applied in forensic science and can be calculated with the ratio of odd to even numbered *n*-alkanes. Petrogenic hydrocarbon sources will commonly have a CPI index near 1, whereas biogenic hydrocarbon sources will show higher CPI values from approximately 2 to 12 (Wang et al., 2009).

PACs are ubiquitous in the environment and derive from petrogenic, pyrogenic, and biogenic sources. Their composition provides information on their source of origin (Wang et al., 2014). As their name suggests, they are composed of multiple fused aromatic benzene rings and range from 2 or more rings with varying degrees of alkyl-substitution (Poster, Schantz, Sander, & Wise, 2006). Polycyclic aromatic sulfur heterocycles (PASHs) are also included under the umbrella term of PACs (*e.g.* dibenzothiophenes, benzonaphthothiophenes), many of which are associated with oil sands emissions (Jautzy, Ahad, Gobeil, et al., 2015; Lam et al., 2012; Mössner & Wise, 1999; Yang et al., 2011). PACs are formed naturally and anthropogenically by pyrogenic, petrogenic, or biogenic processes. Pyrogenic PACs are formed by pyrolysis, the high temperature burning of organic matter in low to no oxygen (350 °C to over 1200 °C). Sources of pyrogenic PACs include bitumen coke, coal tar, thermal cracking of petroleum residues, and incomplete combustion of wood, motor and fuel oils. Petrogenic PACs are formed at low temperatures and high pressures over millions of years (100 °C-150 °C). Typical sources include oil products, oceanic and freshwater spills, oil storage leaks, and small releases of fuel. Biogenic PACs include all PACs formed by plants, bacteria, or by the degradation of such vegetation (Abdel-Shafy & Mansour, 2015). Petrogenic processes operate slowly, and at low temperatures,

which favor the production of alkylated PACs to a much greater extent than pyrogenic and biogenic processes (Abdel-Shafy & Mansour, 2015; Wang et al., 2014). The distribution of alkylated PAC homologs based on the relative concentrations of C_0- to C_4-substituted PACs is often used for source determination. The patterns of alkylated PACs emitted from these sources are strikingly different however, with pyrogenic sources having a more characteristic slope distribution ($C_0 > C_1 > C_2 > C_3 > C_4$) and petrogenic sources typically showing bell-shaped or skewed ($C_0 < C_1 < C_2 < C_3$) patterns depending on their degree of biodegradation (Achten & Andersson, 2015). The main groups of alkylated PAC homologs targeted for oil forensics are those of naphthalene, phenanthrene, dibenzothiophene, fluorene, chrysene, and, more recently, benzonaphthothiophene (Kang, Lee, & Kwon, 2016).

Petroleum biomarkers can be emitted alongside other PHCs, aiding in the identification of source rocks and oil types. These geochemical fossils consist of hopanes, polycyclic terpanes, and steranes derived from previously living organisms and were synthesized during diagenesis of petroleum (Venosa, Suidan, King, & Wrenn, 1997; Wang, Stout, & Fingas, 2006). These compounds are highly resistant to degradation, which makes them a valuable tool to assess genesis, maturation, and biodegradation of petroleum products (Yang et al., 2011). In this same way, petroleum biomarkers can aid in identifying potential source rocks or petroleum sources with a similar distribution, or 'signature', of select hopanes, terpanes, and steranes (Table 1.1). This targeted set of compounds is unique from *n*-alkanes and PACs in that they were synthesized in petroleum reservoirs and are chemically distinct from their biological precursor molecules of hopanoids and steranes. Therefore, petroleum biomarkers offer higher specificity for petrogenic source identification than *n*-alkanes and PACs, which have multiple confounding sources in the environment. By targeting a combination of environmentally-relevant PHCs such as *n*-alkanes

and PACs in addition to petroleum biomarkers, a greater understanding of the nature and potential sources of PHCs in the PAD can be achieved.

Table 1.1. Petroleum biomarker compounds commonly used for oil forensics.

Name	Code	Target ion
Terpanes		
C_{21} terpane	C_{21} T	191
C_{22} terpane	C_{22} T	191
C_{23} terpane	C_{23} T	191
C_{24} terpane	C_{24} T	191
Hopanes		
C_{27} 18A-hopane II (Ts)	C_{27} Ts	191
C_{27} 17A-hopane (Tm)	C_{27} Tm	191
C_{29} αβ hopane	C_{29} αβ H	191
C_{30} αβ hopane	C_{30} αβ H	191
C_{31}(S) hopane	C_{31}(S) H	191
C_{31}(R) hopane	C_{31}(R) H	191
C_{32}(S) hopane	C_{32}(S) H	191
C_{32}(R) hopane	C_{32}(R) H	191
C_{33}(S) hopane	C_{33}(S) H	191
C_{33}(R) hopane	C_{33}(R) H	191
C_{34}(S) hopane	C_{34}(S) H	191
C_{34}(R) hopane	C_{34}(R) H	191
C_{35}(S) hopane	C_{35}(S) H	191
C_{35}(R) hopane	C_{35}(R) H	191
Gammacerane	GAM	191
Steranes		
C_{27}αββ steranes	C_{27}αββ S	218
C_{28}αββ steranes	C_{28}αββ S	218
C_{29}αββ steranes	C_{29}αββ S	218

1.2.2 Bitumen Formation and Deposits

The province of Alberta is underlain by rich deposits of bituminous sands and coal. Three major deposits of bituminous sands located in northern Alberta include the Athabasca, Cold Lake, and Peace River. These deposits account for approximately 293 125 million m^3 initial in-place reserves of crude bitumen (Alberta Energy Regulator, 2015). Bitumen formation in Alberta occurred over a period of hundreds of millions of years as a result of decaying plant and animal material from the late Devonian period. The oil formed during the Devonian period migrated northeast during the Cretaceous period due to Rocky Mountain uplift followed by significant drop in sea-level causing erosion of the foreland (Zhou, Huang, & Liu, 2008). During migration, oil deposits encountered sandstone and shale and were sealed by overlying Clearwater shales (Timoney, 2013a; Zhou et al., 2008). The deposits were effectively capped by overlying sandstone and shales, leading to the formation and maturation of the modern bituminous sand deposits (Conly, Crosley, & Headley, 2002). Bituminous coal deposits are prevalent in the Alberta foothills and account for approximately 94 billion tonnes of initial in-place reserves, the largest in Canada (Alberta Energy Regulator, 2015). These deposits were formed on a similar timescale as the bituminous sands, yet underwent coalification after the Jurassic period (ending ~150 million years ago) (Resource Delevopment and Geoscience Branch, 2004). Approximately 244 coal deposits exist in Alberta, some of which come in direct contact with the Peace River (Alberta Energy Regulator, 2015).

The Peace and Athabasca Rivers both incise natural Cretaceous bituminous sands and coal formations. Typically, these mixtures are trapped well below the surface. However, erosion of the landscape by rivers and channels can result in their exposure to the aquatic environment. This is especially the case in the Athabasca Valley where overburden, the overlying sediment of

the McMurray Formation, is thinnest (approx. 25-50 feet)(Conly et al., 2002; Hein, Cotterill, & Berhane, 2000). Consequently, the McMurray formation (bituminous sands) is eroded in some areas by the Athabasca River and its tributaries, namely at Boiler Rapids, the conflux of the MacKay River, and alongside the banks of the Clearwater River (Conly et al., 2002). The Gates and Gething lower Cretaceous formation (bituminous coal) is reportedly exposed at the Peace River canyon near Hudson's Hope, B.C. and along Smoky River, a tributary of the Peace River (Jautzy, Ahad, Hall, et al., 2015; Langenberg, 1983).

PACs emitted from these bituminous sands are of particular concern as they are potentially toxic and many have been found to be carcinogenic and mutagenic (Dupuis & Ucan-Marin, 2015). These compounds are highly lipophilic and tend to bind to organic matter and sediment (Timoney & Lee, 2011). In this way, the Peace and Athabasca rivers can carry dissolved and sediment-bound PACs downstream, and deposit them in the PAD. The relative proportion of PACs originating from erosion of these natural formations into the PAD is ambiguous (Conly et al., 2002; Jautzy, Ahad, Hall, et al., 2015). Few data exist on background concentrations of PHCs in the PAD before industrial development began. Paleolimnological studies examining historic levels of contaminants in lakes using sediment cores have identified the PAD as a repository for bitumen-associated PACs deposited by the Athabasca River, likely for millennia (Hall et al., 2012; Jautzy, Ahad, Hall, et al., 2015; Timoney & Lee, 2011)

1.2.3 Upstream Industrial Development

The Alberta oil sands industry generates the most concern for the PAD and surrounding environment. The largest oil sands operation, the AOSR, in Canada is primarily located near Fort McMurray less than 200 km south of the PAD on the headwaters of the Athabasca River. In 2013, these operations accounted for 220 km^2 of tailings ponds, managing 975.6 million m^3 of

fine liquid tailings (Government of Alberta, 2016). In addition to oil wells and *in situ* mining operations, several bitumen upgrading facilities are present in Alberta's north, heavily concentrated in and around Fort McMurray (Alberta Energy, 2017; Lands and Minerals Sector; National Energy Board, 2018). Toxic, carcinogenic, and mutagenic contaminants emitted to the surrounding area from such mining activities include PACs and heavy metals (Sb, As, Be, Cd, Cr, Cu, Pb, Hg, Ni, Se, Ag, Tl, and Zn) (Kelly et al., 2009, 2010). Their effects on water and air quality have only recently (~past 10 years) become an area of focus (Evans et al., 2016; Kelly et al., 2009).

Several studies have examined the effects of anthropogenic activity on the distribution of contaminants in the Athabasca Oil Sands Region (AOSR). Kelly et al. (2009) examined water and snowpack samples from the Athabasca River and its tributaries in 2008. Atmospheric deposition of PACs in snow ranged from 19 g/m^2 near upgrading facilities to 0.35 g/m^2 at sites outside the 50 km radius. Alkylated PACs associated with bitumen (*i.e.* petrogenic PACs including dibenzothiophenes, phenanthrenes/anthracenes, and fluorenes) dominated the total PAC mixtures in Athabasca River tributary water. Kurek et al. (2013) expanded on this study, analyzing alkylated PACs in lake sediments nearby and up to 90 km away from development. Their data show a range of 2.5 to 23-fold increase in total sediment PAC flux (measured in nanograms per square centimeter per year) since 1960, with a shift from predominately unsubstituted parent PACs (pyrogenic) to alkylated PAC species (petrogenic). The Athabasca River therefore crosses an increasingly large area of development, carrying a significant load of water and sediment and potentially delivering these PHC contaminants to the PAD.

1.2.4 Petrogenic Hydrocarbons in the Peace-Athabasca Delta

Research in the PAD has focussed primarily on characterising PACs from petrogenic sources and linking these to river discharge. Several techniques have been employed to resolve the sources of these PACs, including direct comparison to oil sands samples via similarity percentage (SIMPER) analysis, diagnostic ratios, as well as highly specialized isotopic techniques (Evans et al., 2016; Hall et al., 2012; Jautzy, Ahad, Gobeil, et al., 2015). Hall et al. (2012) compared the distributions of PACs in a 2007 Athabasca flood deposit to that of an oil sands sample and identified PACs associated with bitumen and river transportation (including C2-C4 dibenzothiphenes, C2-C4 fluoranthenes/pyrenes and C2 benz[*a*]anthracene/chrysenes). Further study of delta lake sediments by Jautzy, Ahad, Gobeil, et al. (2015) confirmed the presence of petrogenic PACs from mixed origins using compound-specific isotope analysis. Petroleum coke, a particulate dust emitted from bitumen upgrading processes, was identified as a potential source of PACs in Athabasca delta lakes in addition to forest fires and AOS bitumen. This alludes to the possibility that particulate dust may undergo long-range transport and delivery into lakes in the PAD.

River discharge and organic carbon content have been shown to correlate with PAC concentrations in the PAD (Evans et al., 2016; Hall et al., 2012; Timoney & Lee, 2011). Sediment cores collected by Hall et al. (2012) from two flood-prone lakes in the Athabasca sector ("PAD 31" and "PAD 23") point to the long-term presence (>200 years) of alkylated PACs, predating industrial development. They also show that PAC composition in sediment can differ significantly between flooding intervals. Seven species of PACs (C2-C4 dibenzothiophenes, C2-C4 fluoranthene/pyrenes, and C2 benz[*a*]anthrancenes/chrysenes) were shown to be more abundant when lakes experienced flood-prone periods (Hall et al., 2012). The

latter highlights the role that the Athabasca River plays in depositing sediment-bound PACs to delta lakes. Organic carbon content has been shown to be negatively correlated with concentration of river-associated PACs (listed above) (Hall et al., 2012; Jautzy, Ahad, Hall, et al., 2015). This negative correlation was attributed to river discharge, in that the influx of PACs into delta lakes from rivers is accompanied by low organic river sediments thereby decreasing the organic matter content of the lake sediments while increasing the PAC concentration (Hall et al., 2012).

1.3 Biodegradation of Petrogenic Hydrocarbons and Microbial Communities of the Athabasca Oil Sands Region

Many microorganisms have the ability to harvest carbon and energy from PHCs, allowing them to be used as raw materials for metabolism and growth. To do so, bacteria, and select archaea and fungi, have enzymes specific to particular degradation reactions encoded in their genetics (Das & Chandran, 2011; Grimes et al., 2011). Through the combined action of microbial communities, hydrocarbons such as branched and unbranched alkanes, and PACs can be biodegraded (Das & Chandran, 2011). The enzymes that facilitate these chemical reactions are generally categorized as assimilatory, respiratory, and dissimilatory. Assimilatory enzymes make use of the hydrocarbon source and reducing equivalents (*e.g.* nicotinamide adenine dinucleotide phosphate, NAD(P)H) to generate materials for their cells. Respiratory and dissimilatory enzymes make use of electron donors and hydrocarbon sources to either generate adenosine triphosphate or dissipate energy, respectively (Moreno-Vivián & Ferguson, 1998). These reactions occur in specific conditions depending on the nature of the hydrocarbons, nutrient availability, oxygen availability, temperature, pH, salinity, and the presence of other microorganisms (Grimes et al., 2011). This in turn influences the genetic makeup of microbial

communities, and the rate at which degradation of PHCs occurs. Additionally, previous exposure to hydrocarbons can lead to adaptation, increasing the hydrocarbon-degradation potential of the microbial community. This can occur via induction and/or depression of specific enzymes, genetic changes resulting in new metabolic capabilities, and selective enrichment of specific hydrocarbon-degrading organisms (Leahy & Colwell, 1990). Indigenous hydrocarbon degrading microorganisms have in fact been found in many hydrocarbon-laden environments, for example: deep sea sediments, freshwater ecosystems, tailings ponds, Arctic ecosystems, and permafrost (An et al., 2013; Bell et al., 2013; Kostka et al., 2011; Yakimov, Timmis, & Golyshin, 2007; Yang, Wen, Zhao, Shi, & Jin, 2014; Yergeau et al., 2012)

The degradation of PHCs can occur in both aerobic and anaerobic conditions through multiple pathways. When exposed to a mixture of PHCs, microbial communities tend to degrade the *n*-alkane fraction at a faster rate than polycyclic aromatics (Leahy & Colwell, 1990; Yang et al., 2011). Both classes of compounds are metabolized in stages. Activation of the hydrocarbon compound is first required to initialize metabolism which can be achieved aerobically through oxidation or hydroxylation, and anaerobically through fumarate addition, methylation, hydroxylation, or carboxylation of the hydrocarbon (Sierra-Garcia & de Oliveira, 2013). Following activation, the hydrocarbon can then be metabolized via multiple pathways. In the presence of oxygen, aliphatic compounds are typically oxidized to alcohols followed by oxidation to aldehydes that are converted to fatty acids. Aromatic intermediates that are hydroxylated can be catalyzed by either intra- or extra-dioxygenases and enter a variety of central pathways. Alternatively, aromatic compounds enter the CoA-thioester pathway and are further reduced (Vilchez-Vargas, Junca, & Pieper, 2010). In the absence of oxygen, fumarate addition to aliphatic compounds yields alkylsuccinate that undergoes further decarboxylation to

fatty acid thioesters (Widdel & Grundmann, 2010). Alternatively, carboxylation of aliphatics leads to the conversion of alkanes to fatty acids (Sierra-Garcia & de Oliveira, 2013). Fumarate addition and ring hydroxylation of aromatic compounds proceeds through pathways similar to aliphatics, yielding benzoyl-CoA whose aromatic ring is dearomatized followed by hydrolytic ring cleavage (Boll, Fuchs, & Heider, 2002). Finally, these pathways conclude with the formation of smaller molecules (*e.g.* acetyl-CoA) and incorporation into cells (*e.g.* via the tricarboxylic acid cycle) (Abbasian & Lockington, 2016).

Many enzymes and encoding genes for these pathways have been characterized for the degradation of aliphatic (*e.g. n*-alkanes) and aromatic (*e.g.* PACs) hydrocarbons. Monooxygenase enzymes are known to be involved in the initial activation of *n*-alkanes under aerobic conditions, and rubredoxin enzymes, hydroxylases, and dehydrogenases in their further metabolism (Sierra-Garcia & de Oliveira, 2013). Anaerobically, alkylsuccinate synthase is the most studied enzyme for the addition of fumarate to *n*-alkanes (Abbasian & Lockington, 2016). Enzymes involved in PAC degradation include oxygenase, dehydrogenase, and lignolitic enzymes, many of which operate at mesophilic temperatures (Haritash & Kaushik, 2009). Many enzymes are involved in the aerobic catabolism of alkylated PACs. Some of the reactions involve the oxidation of the methyl group to either an alcohol, aldehyde, or carboxylic acid, decarboxylation, demethylation, and deoxygenation (Seo, Keum, & Li, 2009). Seo et al. (2009) highlight that aerobic catabolism of this kind leads to the production of alkylsalicylate, or alkylphthalate. Anaerobic catabolism of alkylated PACs has been assessed for specific homologs including napthalenes (Annweiler et al., 2000) and alkyl-benzenes (Spormann & Widdel, 2000).

Hydrocarbon-degrading microorganisms have been studied in freshwater environments including freshwater lagoons, lakes, and rivers and their tributaries (Hadwin et al., 2006;

Jurelevicius et al., 2013; Yergeau et al., 2012). Much research has focussed on microcosms as a way to determine how indigenous microbial communities in sediment or water respond to hydrocarbon exposure from crude oil, naphthalene, and oil sands process water (Hadwin et al., 2006; Jurelevicius et al., 2013; Yergeau et al., 2013). Moving away from a microcosm approach, Yergeau et al. (2012) evaluated the impacts of oil sands extraction on microbial communities in tailings, the Athabasca River, and its tributaries, by analysing 16S rRNA sequences in sediment from 17 sites. Their results show that communities within a 10 km radius of extraction operations were more similar to each other than those beyond. In fact, microbial community composition shifted according to the level of hydrocarbon content in lake sediments, effectively lowering bacterial diversity, with some genera dominating hydrocarbon-contaminated lake sediments. Among them, the genus *Schumannella* had the strongest positive correlation with total straight-chain hydrocarbons, and *Sorangium* the strongest negative correlation with total aromatic hydrocarbons. The sensitivity of these microbial genera to specific hydrocarbons makes them potential bioindicators for contamination. Yergeau et al. (2012) point out that *Azonexus, Achromobacter,* and *Methanobrevibacter* could all serve as potential bioindicator taxa related to oil sands activity based on their strong positive correlation with PACs, naphthenic acids, and a mixture of hydrocarbons (PACs, naphthenic acids, straight-chained and aromatic hydrocarbons), respectively.

1.4 Thesis Objectives and Hypothesis

Biodegradation plays a key role in the cycling of PHCs in the environment. The PAD, having a long history of exposure to PHCs, particularly PACs, is an excellent system in which to study biodegradation. Studies performed in the AOSR have advanced the understanding of how PHCs from bituminous sands affect microbial communities, and vice-versa, in the watershed. However,

the hydrocarbon-degrading potential of the PAD has yet to be explored. Analyzing the taxonomy and genetic composition of microbial communities in lakes of the PAD is critical for understanding the potential for biodegradation of PHCs in freshwater systems.

The primary aim of this thesis is to evaluate the hydrocarbon-degrading potential of microbial communities in the Peace and Athabasca Delta by way of comparison to nearby lakes in the elevated boreal uplands. Lakes in the boreal uplands are not directly influenced by the Peace and Athabasca Rivers, and essentially serve as a reference. The comparison of select lakes in these three regions (Athabasca Delta, Peace Delta, and boreal uplands) helps contrast microbial communities and the potential influence of PHCs from different sources (mainly river-transported bitumen-associated hydrocarbons, a mixture of fossil-fuels and wood combustion, and atmospherically transported hydrocarbons, respectively). In order to guide this preliminary assessment of microbial communities in the PAD, analysis of environmentally-relevant PHCs and their potential sources was required. PHCs of interest consist of the dominant fractions prevailing from bituminous sands, and most commonly measured: *n*-alkanes and PACs. Petroleum biomarkers of terpanes, hopanes, and steranes were also analyzed for the first time in PAD lakes to determine their potential use as indicators of petrogenic source in a complex system such as the PAD.

The second chapter of this thesis provides a comprehensive analysis of PHCs (including *n*-alkanes, PACs, and the petroleum biomarkers of steranes, hopanes, and terpanes) in lake sediments of the Athabasca Delta, Peace Delta, and boreal uplands, and investigates the use of petroleum biomarkers as a tool to accurately identify the input and origin of PHCs in the PAD. The third chapter of this thesis analyzes the microbial community structure and hydrocarbon degradation potential of microbial communities in these same lakes. Analysis of microbial

communities and genetic composition was achieved via metagenomics, a genomics tool that enables us to recover the assemblage of genes and reconstruct taxonomic profiles from environmental samples. This novel approach allowed us to gain valuable information on the microbial communities in the PAD and how these may differ from boreal lakes of the region.

I hypothesize that lakes in the PAD that have been exposed to PHCs have significantly different microbial communities and gene assemblages from lakes in the boreal uplands. Accordingly, I predict that: (1) lakes hydrologically connected to the AOSR had a higher concentration and distribution of PHCs than lakes disconnected from the AOSR in the elevated boreal uplands region; and (2) this difference in PHC concentration and composition among lakes is accompanied by variations in the diversity and composition of hydrocarbon degradation genes.

1.5 References

Abbasian, F., & Lockington, R. (2016). A Review on the Genetics of Aliphatic and Aromatic Hydrocarbon Degradation, 224–250. http://doi.org/10.1007/s12010-015-1881-y

Abdel-Shafy, H. I., & Mansour, M. S. M. (2015). A review on polycyclic aromatic hydrocarbons: Source, environmental impact, effect on human health and remediation. Egyptian Journal of Petroleum, 25(1), 107–123. http://doi.org/10.1016/j.ejpe.2015.03.011

Achten, C., & Andersson, J. T. (2015). Overview of Polycyclic Aromatic Compounds (PAC). Polycyclic Aromatic Compounds, 35(2–4), 177–186. http://doi.org/10.1080/10406638.2014.994071

Alberta Energy. (2017). Upgraders and Refineries: Facts and Stats. Government of Alberta. Edmonton. Retrieved from https://open.alberta.ca/dataset/98c15cad-c5d9-4d96-b39c-423210a3050c/resource/b03c9445-267f-48e3-9395-ff192ec59cfd/download/fsrefiningupgrading.pdf

Alberta Energy Regulator. (2015). ST98-2015: Alberta's Energy Reserves 2014 and Supply/Demand Outlook 2015-2024. Calgary, AB. Retrieved from http://www.aer.ca/documents/sts/ST98/ST98-2015.pdf

An, D., Brown, D., Chatterjee, I., Dong, X., Ramos-Padron, E., Wilson, S., … Voordouw, G. (2013). Microbial community and potential functional gene diversity involved in anaerobic hydrocarbon degradation and methanogenesis in an oil sands tailings pond. Genome, 56(10), 612–618. http://doi.org/10.1139/gen-2013-0083

Annweiler, E., Materna, A., Safinowski, M., Kappler, A., Richnow, H. H., Michaelis, W., & Meckenstock, R. U. (2000). Anaerobic Degradation of 2-Methylnaphthalene by a Sulfate-Reducing Enrichment Culture. Applied and Environmental Microbiology, 66(12), 5329–5333. http://doi.org/10.1128/AEM.66.12.5329-5333.2000.Updated

Bell, T. H., Yergeau, E., Maynard, C., Juck, D., Whyte, L. G., & Greer, C. W. (2013). Predictable bacterial composition and hydrocarbon degradation in Arctic soils following diesel and nutrient disturbance. The ISME Journal, 7(6), 1200–1210. http://doi.org/10.1038/ismej.2013.1

Boll, M., Fuchs, G., & Heider, J. (2002). Anaerobic oxidation of aromatic compounds and hydrocarbons Boll, Fuchs and Heider 605. *Current Opinion on Chemical Biology*, *6*(5), 604–611.

Conly, F. M., Crosley, R. W., & Headley, J. (2002). Characterizing sediment sources and natural hydrocarbon inputs in the lower Athabasca River, Canada. Journal of Environmental Engineering and Science, 1, 187–199. http://doi.org/10.1139/s02-013

Das, N., & Chandran, P. (2011). Microbial degradation of petroleum hydrocarbon contaminants: an overview. Biotechnology Research International, 2011, 941810. http://doi.org/10.4061/2011/941810

Dupuis, A., & Ucan-Marin, F. (2015). A literature review on the aquatic toxicology of petroleum oil: an overview of oil properties and effects to aquatic biota. DFO Canadian Science Advisory Secretariat Research Document, 2015/007(vi), 52.

Evans, M., Davies, M., Janzen, K., Muir, D., Hazewinkel, R., Kirk, J., & De Boer, D. (2016). PAH distributions in sediments in the oil sands monitoring area and western Lake Athabasca: Concentration, composition and diagnostic ratios. Environmental Pollution, 213, 671–687. http://doi.org/10.1016/j.envpol.2016.03.014

Global Forest Watch Canada. (2008). Oil sands surface mining activity in Alberta, Canada up to 2008. Data Basin. Retrieved from https://databasin.org/datasets/ca774c8d651a46a1949bdbdb408a2561

Government of Alberta. (2016). Facts and Statistics. Retrieved from https://www.energy.alberta.ca/OS/AOS/Pages/FAS.aspx

Grimes, D. J., Hazen, T. C., King, G. M., Prince, R. C., Carolina, N., Ward, C. H. H., … Maloy, S. (2011). Microbes & Oil Spills. Environmental Protection, 1–13. http://doi.org/10.1016/S0378-777X(77)80110-8

Hadwin, A. K. M., Del Rio, L. F., Pinto, L. J., Painter, M., Routledge, R., & Moore, M. M. (2006). Microbial communities in wetlands of the Athabasca oil sands: Genetic and metabolic characterization. FEMS Microbiology Ecology, 55(1), 68–78. http://doi.org/10.1111/j.1574-6941.2005.00009.x

Hall, R. I., Wolfe, B. B., Wiklund, J. A., Edwards, T. W. D., Farwell, A. J., & Dixon, D. G. (2012). Has Alberta Oil Sands Development Altered Delivery of Polycyclic Aromatic Compounds to the Peace-Athabasca Delta? PLoS ONE, 7(9), e46089. http://doi.org/10.1371/journal.pone.0046089

Haritash, A. K., & Kaushik, C. P. (2009). Biodegradation aspects of Polycyclic Aromatic Hydrocarbons (PAHs): A review. Journal of Hazardous Materials, 169(1–3), 1–15. http://doi.org/10.1016/j.jhazmat.2009.03.137

Hein, F. J., Cotterill, D. K., & Berhane, H. (2000). An atlas of lithofacies of the McMurray Formation Athabasca Oil Sands deposit, northeastern Alberta; surface and subsurface. Earth Sciences Report 2000-07, 216. Retrieved from http://login.ezproxy.library.ualberta.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=geh&AN=2001-045977&site=ehost-live&scope=site

Jautzy, J. J., Ahad, J. M. E., Gobeil, C., Smirnoff, A., Barst, B. D., & Savard, M. M. (2015). Isotopic Evidence for Oil Sands Petroleum Coke in the Peace-Athabasca Delta.

Environmental Science and Technology, 49(20), 12062–12070. http://doi.org/10.1021/acs.est.5b03232

Jautzy, J. J., Ahad, J. M. E., Hall, R. I., Wiklund, J. A., Wolfe, B. B., Gobeil, C., & Savard, M. M. (2015). Source Apportionment of Background PAHs in the Peace-Athabasca Delta (Alberta, Canada) Using Molecular Level Radiocarbon Analysis. Environmental Science and Technology, 49(15), 9056–9063. http://doi.org/10.1021/acs.est.5b01490

Jurelevicius, D., Alvarez, V. M., Marques, J. M., Lima, L. R. F. de S., Dias, F. de A., & Seldin, L. (2013). Bacterial community response to petroleum hydrocarbon amendments in freshwater, marine, and hypersaline water-containing microcosms. Applied and Environmental Microbiology, 79(19), 5927–5935. http://doi.org/10.1128/AEM.02251-13

Kang, H. J., Lee, S. Y., & Kwon, J. H. (2016). Physico-chemical properties and toxicity of alkylated polycyclic aromatic hydrocarbons. Journal of Hazardous Materials, 312, 200–207. http://doi.org/10.1016/j.jhazmat.2016.03.051

Kelly, E. N., Schindler, D. W., Hodson, P. V, Short, J. W., Radmanovich, R., & Nielsen, C. C. (2010). Oil sands development contributes elements toxic at low concentrations to the Athabasca River and its tributaries. Proceedings of the National Academy of Sciences of the United States of America, 107(37), 16178–16183. http://doi.org/10.1073/pnas.1008754107

Kelly, E. N., Short, J. W., Schindler, D. W., Hodson, P. V, Ma, M., Kwan, A. K., & Fortin, B. L. (2009). Oil sands development contributes polycyclic aromatic compounds to the Athabasca River and its tributaries. Proceedings of the National Academy of Sciences of the United States of America, 106(52), 22346–51. http://doi.org/10.1073/pnas.0912050106

Kostka, J. E., Prakash, O., Overholt, W. A., Green, S. J., Freyer, G., Canion, A., … Huettel, M. (2011). Hydrocarbon-degrading bacteria and the bacterial community response in Gulf of Mexico beach sands impacted by the deepwater horizon oil spill. Applied and Environmental Microbiology, 77(22), 7962–7974. http://doi.org/10.1128/AEM.05402-11

Kurek, J., Kirk, J. L., Muir, D. C. G., Wang, X., Evans, M. S., & Smol, J. P. (2013). Legacy of a half century of Athabasca oil sands development recorded by lake ecosystems. Proceedings of the National Academy of Sciences of the United States of America, 110(5), 1761–6. http://doi.org/10.1073/pnas.1217675110

Lam, V., Li, G., Song, C., Chen, J., Fairbridge, C., Hui, R., & Zhang, J. (2012). A review of electrochemical desulfurization technologies for fossil fuels. Fuel Processing Technology, 98, 30–38. http://doi.org/10.1016/j.fuproc.2012.01.022

Lands and Minerals Sector; National Energy Board. (2018). Principal mineral areas, producing mines, and oil and gas fields in Canada [map]. Natural Resources Canada, “A” Series Map 900A, (ed. 67, 2017), 2 sheets. http://doi.org/https://doi.org/10.4095/306464

Langenberg, C. W. (1983). Structural and sedimentological framework of Lower Cretaceous coal-bearing rocks in the Grande Cache area, Alberta. Edmonton, AB: Alberta Geological Survey.

Leahy, J. G., & Colwell, R. R. (1990). Microbial degradation of hydrocarbons in the environment. Microbiological Reviews, 54(3), 305–315.

Moreno-Vivián, C., & Ferguson, S. J. (1998). Definition and distinction between assimilatory, dissimilatory and respiratory pathways. Molecular Microbiology, 29(2), 664–666. http://doi.org/10.1046/j.1365-2958.1998.00946.x

Mössner, S. G., & Wise, S. A. (1999). Determination of polycyclic aromatic sulfur heterocycles in fossil fuel-related samples. Analytical Chemistry, 71(1), 58–69. http://doi.org/10.1021/ac980664f

Natural Resources Canada. (2008). National Parks of Canada. Data Basin. Retrieved from https://databasin.org/datasets/a833b8954e5846189ec3005f54c0cf1c

Poster, D. L., Schantz, M. M., Sander, L. C., & Wise, S. A. (2006). Analysis of polycyclic aromatic hydrocarbons (PAHs) in environmental samples: A critical review of gas chromatographic (GC) methods. Analytical and Bioanalytical Chemistry, 386(4), 859–881. http://doi.org/10.1007/s00216-006-0771-0

Resource Delevopment and Geoscience Branch. (2004). Coalbed Gas Potential in British Columbia. Petroleum Geology, 1, 3–10. Retrieved from http://www.empr.gov.bc.ca/Mining/Geoscience/Coal/Documents/CoalbedGas/CBGPotential_2004-1.pdf

Seo, J. S., Keum, Y. S., & Li, Q. X. (2009). Bacterial degradation of aromatic compounds. International Journal of Environmental Research and Public Health, 6(1), 278–309. http://doi.org/10.3390/ijerph6010278

Sierra-Garcia, I. N., & de Oliveira, V. M. (2013). Microbial Hydrocarbon Degradation : Efforts to Understand Biodegradation in Petroleum Reservoirs. In Biodegradation - Engineering and Technology (pp. 47–72). IntechOpen. http://doi.org/10.5772/55920

Spormann, A. M., & Widdel, F. (2000). Metabolism of slkylbenzenes, alkanes, and other hydrocarbons in anaerobic bacteria. Biodegradation, 11, 85–105. http://doi.org/10.1023/A:1011122631799

Timoney, K. P. (2013a). Human History After 1900. In The Peace-Athabasca Delta: Portrait of a Dynamic Ecosystem (pp. 315–360). Edmonton: the University of Alberta Press.

Timoney, K. P. (2013b). Natural History: Fauna and Wildlife. In The Peace-Athabasca Delta: Portrait of a Dynamic Ecosystem (pp. 119–149). Edmonton: the University of Alberta Press.

Timoney, K. P. (2013c). The Delta's Physical Environment and Landforms. In The Peace-Athabasca Delta: Portrait of a Dynamic Ecosystem (pp. 15–56). Edmonton: University of Alberta Press.

Timoney, K. P., & Lee, P. (2009). Does the Alberta Tar Sands Industry Pollute? The Scientific Evidence. The Open Conservation Biology Journal, 3, 65–81. http://doi.org/10.2174/1874839200903010065

Timoney, K. P., & Lee, P. (2011). Polycyclic aromatic hydrocarbons increase in athabasca river delta sediment: Temporal trends and environmental correlates. Environmental Science and Technology, 45(10), 4278–4284. http://doi.org/10.1021/es104375d

Venosa, a D., Suidan, M. T., King, D., & Wrenn, B. a. (1997). Use of hopane as a conservative biomarker for monitoring the bioremediation effectiveness of crude oil contaminating a sandy beach. Journal of Industrial Microbiology and Biotechnology, 18(2–3), 131–139. http://doi.org/10.1038/sj.jim.2900304

Vilchez-Vargas, R., Junca, H., & Pieper, D. H. (2010). Metabolic networks, microbial ecology and "omics" technologies: Towards understanding in situ biodegradation processes. *Environmental Microbiology*, *12*(12), 3089–3104. http://doi.org/10.1111/j.1462-2920.2010.02340.x

Wang, Z., Stout, S. A., & Fingas, M. (2006). Forensic fingerprinting of biomarkers for oil spill characterization and source identification. Environmental Forensics, 7(2), 105–146. http://doi.org/10.1080/15275920600667104

Wang, Z., Yang, C., Kelly-Hooper, F., Hollebone, B. P., Peng, X., Brown, C. E., … Yang, Z. (2009). Forensic differentiation of biogenic organic compounds from petroleum hydrocarbons in biogenic and petrogenic compounds cross-contaminated soils and sediments. Journal of Chromatography A, 1216(7), 1174–1191. http://doi.org/10.1016/j.chroma.2008.12.036

Wang, Z., Yang, C., Parrott, J. L., Frank, R. A., Yang, Z., Brown, C. E., … Hewitt, L. M. (2014). Forensic source differentiation of petrogenic, pyrogenic, and biogenic hydrocarbons in Canadian oil sands environmental samples. Journal of Hazardous Materials, 271, 166–177. http://doi.org/10.1016/j.jhazmat.2014.02.021

Widdel, H. of H. and L. M. F., & Grundmann, O. (2010). Biochemistry of the Anaerobic Degradation of Non-Methane Alkanes. In K. N. Timmis (Ed.), *Handbook of Hydrocarbon and Lipid Microbiology*. Berlin: Springer.

Wolfe, B. B., Hall, R. I., Edwards, T. W. D., & Johnston, J. W. (2012). Developing temporal hydroecological perspectives to inform stewardship of a northern floodplain landscape subject to multiple stressors: paleolimnological investigations of the Peace–Athabasca Delta. Environmental Reviews, 20, 191–210. http://doi.org/10.1139/a2012-008

Yakimov, M. M., Timmis, K. N., & Golyshin, P. N. (2007). Obligate oil-degrading marine bacteria. Current Opinion in Biotechnology, 18(3), 257–266. http://doi.org/10.1016/j.copbio.2007.04.006

Yang, C., Wang, Z., Yang, Z., Hollebone, B., Brown, C. E., Landriault, M., & Fieldhouse, B. (2011). Chemical Fingerprints of Alberta Oil Sands and Related Petroleum Products. Environmental Forensics, 12(2), 173–188. http://doi.org/10.1080/15275922.2011.574312

Yang, S., Wen, X., Zhao, L., Shi, Y., & Jin, H. (2014). Crude oil treatment leads to shift of bacterial communities in soils from the deep active layer and upper permafrost along the China-Russia Crude Oil Pipeline route. PLoS ONE, 9(5), 12–14. http://doi.org/10.1371/journal.pone.0096552

Yergeau, E., Lawrence, J. R., Sanschagrin, S., Roy, J. L., Swerhone, G. D. W., Korber, D. R., & Greera, C. W. (2013). Aerobic biofilms grown from athabasca watershed sediments are inhibited by increasing concentrations of bituminous compounds. Applied and Environmental Microbiology, 79(23), 7398–7412. http://doi.org/10.1128/AEM.02216-13

Yergeau, E., Lawrence, J. R., Sanschagrin, S., Waiser, M. J., Korber, D. R., & Greer, C. W. (2012). Next-generation sequencing of microbial communities in the athabasca river and its tributaries in relation to oil sands mining activities. Applied and Environmental Microbiology, 78(21), 7626–7637. http://doi.org/10.1128/AEM.02036-12

Zhou, S., Huang, H., & Liu, Y. (2008). Biodegradation and origin of oil sands in the western Canada sedimentary basin. Petroleum Science, 5(2), 87–94. http://doi.org/10.1007/s12182-008-0015-3

CHAPTER 2. TRACKING PETROGENIC HYDROCARBONS IN LAKES OF THE PEACE-ATHABASCA DELTA, AB, CANADA USING PETROLEUM BIOMARKERS

Michelle-Claire Roy[1], Joshua R. Thienpont[1], Zeyu Yang[4], Roland I. Hall[2], Brent B. Wolfe[3], Bruce P. Hollebone[4], Jules M. Blais[1]

[1]University of Ottawa, Ottawa, Ontario, Canada

[2]Department of Biology, University of Waterloo, Waterloo, Ontario, Canada

[3]Department of Geography and Environmental Studies, Wilfrid Laurier University, Waterloo, Ontario, Canada

[4]Emergencies Science and Technology Section, Environment Canada, Ottawa, Ontario, Canada

2.1 Abstract

The Peace-Athabasca Delta (PAD) receives a mixture of hydrocarbons from biogenic, pyrogenic, and petrogenic processes. Source apportionment in the PAD has focussed on polycyclic aromatic compounds (PACs), which are ubiquitous in the environment and susceptible to weathering. In contrast, petroleum biomarkers of terpanes, hopanes, and steranes are degradation-resistant organic compounds found uniquely in petroleum products that can identify the input and origin of petrogenic hydrocarbons (PHCs). We provide in-depth analysis of environmentally-relevant PHCs (including *n*-alkanes, PACs, and petroleum biomarkers) in surficial sediments of strategically selected lakes in the Athabasca and Peace deltas and adjacent boreal uplands. Alkanes were found to be predominately biogenic in all lakes. PAC sources were identified as wood combustion in boreal upland lakes, a mixture of petrogenic and pyrogenic combustion in Peace Delta lakes, and predominately petrogenic in Athabasca Delta lakes. Using multivariate analyses, raw Alberta oil sands was identified as a potential source of PHCs in the Athabasca Delta. Biomarkers of terpanes and hopanes were identified in closed-drainage lakes of the Peace Delta and boreal uplands, possibly transported and deposited atmospherically via dust particles. These findings validate the use of petroleum biomarkers as tracers for bituminous sands in PAD surficial sediments and their potential use in paleolimnological investigations.

2.2 Introduction

The Peace-Athabasca Delta (PAD) is the world's largest freshwater boreal delta formed at the confluence of two major rivers, the Peace and Athabasca. The Peace River originates from the headwaters regulated by the Williston Reservoir (location of the W.A.C. Bennett Dam) in northern British Columbia. The Athabasca River originates from Alberta's Columbia Icefield and crosses through the Athabasca Oil Sands region (AOSR) before terminating at the PAD where it enters Lake Athabasca (Timoney, 2013c). Concerns have been mounting in recent decades about effects of potential stressors on the PAD, including climate-driven hydrological change, river regulation, and upstream mining-related activities (Evans et al., 2016; Kelly et al., 2009; Timoney, 2013c). The PAD is situated on the traditional territory of the Mikisew Cree First Nations, the Athabasca Chipewyan First Nations, and Métis Local #125, all of whom practice traditional land use. There are concerns for the health of aboriginal peoples because the Peace and Athabasca rivers are exposed to petrogenic hydrocarbons (PHCs) naturally (via channel incision and erosion of bituminous sands and coal deposits) (Conly et al., 2002; Jautzy, Ahad, Hall, et al., 2015) and potentially via anthropogenic activities (via mining-related activities in the AOSR) (Jautzy, Ahad, Gobeil, et al., 2015; Kelly et al., 2009; Kurek et al., 2013).

Previous studies in the PAD investigating the source of PHCs and temporal patterns of deposition have focussed on polycyclic aromatic compounds (PACs), as many of these persistent organic compounds are known to be capable of bioaccumulating in aquatic organisms(Dupuis & Ucan-Marin, 2015). Paleolimnological studies have identified lakes connected to the Athabasca River as repositories for bitumen-associated river-transported PACs (Hall et al., 2012; Timoney & Lee, 2011), as well as petroleum coke-derived PACs (Jautzy, Ahad, Gobeil, et al., 2015; Manzano et al., 2017). Due to the ubiquitous nature of PACs in the environment, these studies

have applied various techniques for source apportionment in the PAD including diagnostic ratios and compounds (Evans et al., 2016; Hall et al., 2012), radiocarbon analysis (Jautzy, Ahad, Hall, et al., 2015), and compound-specific isotope analysis (Jautzy, Ahad, Gobeil, et al., 2015). However, each PAC congener is influenced by different weathering processes including evaporative loss, photooxidation, and biodegradation, further complicating source apportionment of PACs in areas with extensive weathering, such as the PAD (Wang et al., 2007).

Petroleum biomarkers may provide a powerful tool to accurately and definitively identify the source of PHCs (including PACs) in the environment. Terpanes, hopanes, and steranes are organic biomarker compounds formed and preserved exclusively in petroleum reservoirs, and are commonly used in oil forensics (Zumberge, 1987). These compounds have been traditionally applied to oil spill forensics to track source and differentiate weathered oils (Wang et al., 2007). Due to their resistance to biodegradation, petroleum biomarkers remain unaltered when exposed to weathering processes (Wang, Fingas, Blenkinsopp, Sergy, Landriault, Sigouin, Foght, et al., 1998), in contrast to PACs (Wang et al., 2007). For example, a study on the persistence of oil 25 years after the Nipisi pipeline spill in Lesser Slave Lake, Alberta, determined that biomarker composition was nearly unaffected when comparing surface (0-4 cm) to subsurface (30-40 cm) sediments, in contrast to advanced weathering of PACs and *n*-alkanes (Wang, Fingas, Blenkinsopp, Sergy, Landriault, Sigouin, & Lambert, 1998). Thus, applying techniques that are known source identifiers in oil spill forensics can provide unique insights and elicit greater confidence in identifying the geochemical origin of hydrocarbon contaminants in the PAD compared to employing diagnostic ratios to determine the petrogenic origin of PHCs.

This study reports the analysis of an expanded set of environmentally-relevant PHCs (including *n*-alkanes, PACs, and the petroleum biomarkers of steranes, hopanes, and terpanes) in

surface sediments of strategically-selected lakes in the PAD. These include two flood-prone restricted-drainage lakes in the Athabasca Delta, two closed-drainage lakes in the Peace Delta, and three elevated lakes in the adjacent boreal uplands that are not influenced by the Peace or Athabasca rivers (Figure 2.1). This design allows us to analyse the composition of PHCs in three regions suspected of receiving hydrocarbons from different sources (mainly river-transported bitumen-associated hydrocarbons, a mixture of fossil-fuels and wood combustion, and atmospherically transported hydrocarbons, respectively). The main objective of this research is to evaluate the potential use of petroleum biomarkers as a tool to accurately identify the input and origin of PHCs in complex, broadly-sourced regions, such as the PAD. Using these techniques, we show that raw Alberta oil sands is an important source of PHCs to river-influenced lakes in the Athabasca Delta. We also show that petroleum biomarkers and associated PHCs may be entering closed-drainage lakes in the Peace-Delta and boreal uplands via atmospheric deposition.

2.3 Methodology

2.3.1 Study Sites and Sample Collection

The study was conducted in the Peace-Athabasca Delta and in an adjacent elevated boreal region northeast of Fort Chipewyan on the Canadian Shield, unofficially named “Uplands”. Study sites were selected in the Athabasca Delta (PAD 30 and 31), the Peace Delta (PAD 1 and 3), and in the Uplands region (UP 3, 7, and 9). Lakes PAD 30 and 31 (210 m above sea level (a.s.l.) and 213 m a.s.l., respectively) adjacent to Mamawi Creek downstream of the Athabasca River, and are flood-prone lakes which receive floodwaters from the Athabasca River on an approximately annual basis (Remmer, Klemt, Wolfe, & Hall, 2018). Sediment cores from lake PAD 31 were previously taken by Hall et al. (2012) where the depositional history of PACs was

studied in relation to the paleohydrological record. Lakes PAD 1 and 3 (212 m a.s.l. and 213 m a.s.l., respectively) are adjacent to Rivière des Rochers and are typically closed-drainage lakes that receive episodic inputs of Peace River floodwaters when ice-jams develop on the Peace and Slave rivers along the northern margin of the PAD. Communication with a local knowledge holder indicated lake PAD 1 last flooded in 1996, with nearby lake PAD 3 likely reflecting a similar flood record. Uplands lakes 3, 7, and 9 are in an elevated region (256 m a.s.l, 254 m a.s.l, and 273 m a.s.l, respectively) with no direct connection to the Peace or Athabasca rivers.

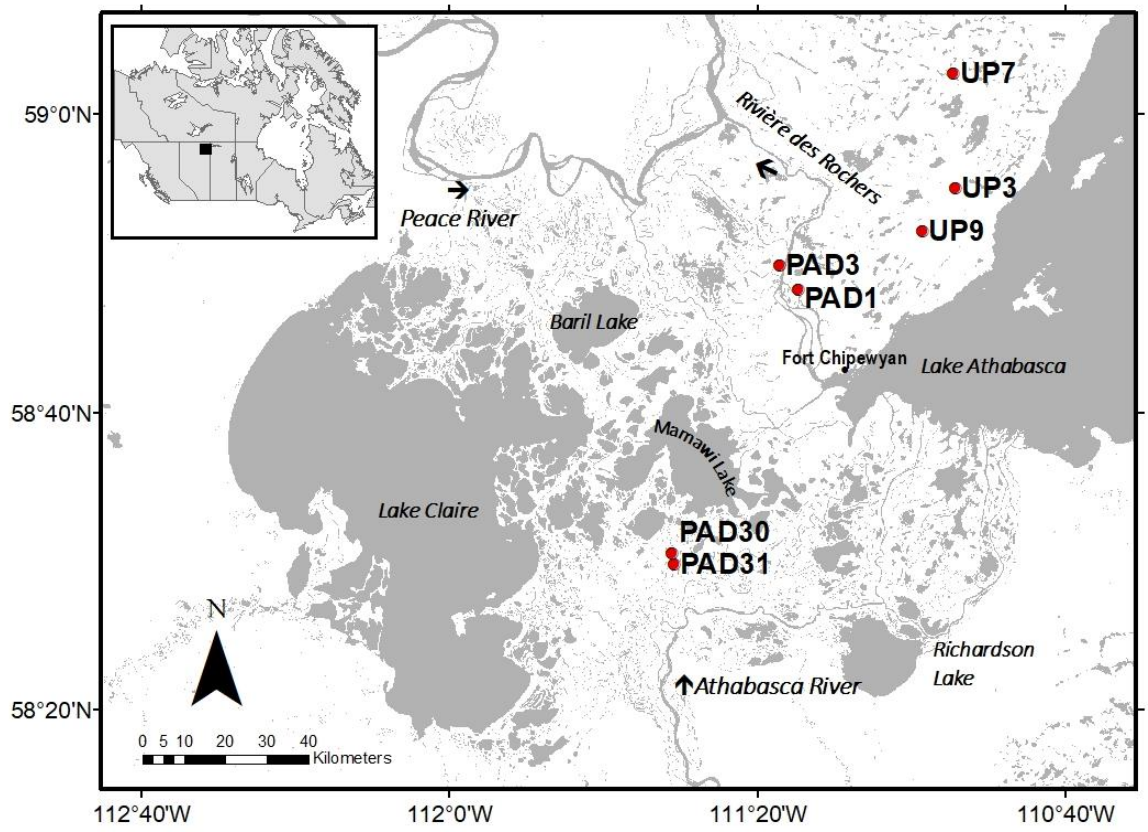

Figure 2.1. Map of the Study Region. Map identifying locations of sampling sites in the Peace-Athabasca Delta: PAD 1 (58°48'N, 111°14'W), PAD 3 (58°49'N, 111°17'W), PAD 30 (58°30'N, 111°31'W), PAD 31 (58°29'N, 111°30'W), and sampling sites outside of the Delta: Uplands 3 (58°55'N, 110°54'W), Uplands 7 (59°02'N, 110°54'W), and Uplands 9 (58°52'N, 110°58'W). Also shown are the Athabasca and Peace rivers and their respective flow directions, under normal, non-flood conditions.

UP 3 and 7 are closed-drainage lakes while UP 9 is a headwater lake connected by a small creek to a larger lake three metres below. These three lakes are situated on the Canadian Shield with bedrock parent material, including numerous exposed granite outcrops. Based on the flood histories of these three subsets of lakes, and their differing degrees of hydrological connectivity to bituminous sands, we predicted that the PHC-burden would be higher and significantly different in lake sediments with higher flood frequency in the Athabasca Delta, when compared to closed-drainage lakes in the Peace Delta which have not flooded since 1996. Since lakes in the elevated boreal upland region are outside of the Peace-Athabasca flood zone, we predicted that these lakes would have the lowest PHC-burden. These boreal upland lakes also allowed us to characterize PHCs coming from predominately atmospheric sources outside of the delta, which further informed source apportionment in such a broadly-sourced region.

Triplicate surface sediment samples (~0-5 cm) were collected from each lake July 11th – 13th 2017 using a mini-Glew gravity corer. Each 5 cm interval was transferred into separate Whirl-Pak® bags. Sediment samples were kept at 4 °C in the field until they could be thoroughly homogenized and subsampled into 50 mL centrifuge tubes, and frozen at -20 °C. Sediments remained frozen and were kept in the dark until elemental analysis and extraction of PHCs. Surface water samples were collected for water chemistry in 4 L carboys at approximately 10 cm below the surface. Water samples were kept at 4 °C in the field and processed within 24 h.

2.3.2 Sediment Elemental Analysis (%CNS)

A subsample of surface sediment was sent to G.G. Hatch Stable Isotope Laboratory (University of Ottawa) for elemental analysis of organic carbon, nitrogen and sulfur content (%CNS). In preparation, a small subsample of sediment was freeze-dried, transferred to an acid desiccator (H_2O:HCl) for 72 h, rinsed three times with distilled water, and freeze-dried to remove

inorganic carbon from the subsample. Upon arrival at the Hatch Laboratory, 5-10 mg of dry sediments were combined with ~20 mg powdered tungsten trioxide (WO_3) in a tin capsule and analyzed using a VarioEL cube elemental analyzer (Elementar Americas Inc., New York).

2.3.3 Extraction and Analysis of PHCs

Extraction and analysis of PHCs in surface sediments was performed in collaboration with Environment and Climate Change Canada (River Road, Ottawa, ON). Target compounds for PHC analysis included *n*C9-*n*C40 alkanes, 15 unsubstituted PACs, 7 homologous series of alkyl-substititued PACs (C0-C4 naphthalenes, C0-C4 phenanthrenes, C0-C3 dibenzothiophenes, C0-C3 fluorenes, C0-C4 fluoranthenes, C0-C4 benzonaphthothiophenes, and C0-C3 chrysenes) and biomarkers of terpanes, hopanes, and steranes (Table S2.1). Prior to extraction, water content was determined by drying a subsample of sediment (2-5 g wet) at 60 °C overnight (~17 h) and calculating weight differential. For hydrocarbon extraction, sediment (~5 g wet) was transferred to a pre-cleaned thimble and mixed with approximately 10 g anhydrous sodium sulphate. Samples were spiked with 100 µL recovery surrogate (200 µg/mL ortho-terphenyl, 200 µg/mL d_{50}-tetracosane ($C_{24}D_{50}$), and 10 µg/mL of each of the five deuterated PACs (d_8-naphthalene, d_{10}-acenaphthene, d_{10}-phenanthrene, d_{12}-benz[*a*]anthracene and d_{12}-perylene)), loaded onto Soxhlet with 200 mL 9:1 dichloromethane(DCM):acetone, and extracted for approximately 17 hours.

Sample extracts were exchanged into hexane, concentrated to 2 mL, and transferred to silica gel chromatographic columns (3 g) for cleanup as per Yang et al. (2011). Saturates (F1 fraction) were eluted with 20 mL hexane. Aromatics (F2 fraction) were eluted with 15 mL DCM:hexane. F1 and F2 fractions were evaporated to 0.9 mL and 0.8 mL (respectively) under a gentle stream of nitrogen. The final product was spiked with 100 µL of each internal standard (F1: 200 µg/mL 5-α-androstane. F2: 200 µg/mL of 5-α–androstane and 10 µg/mL d_{14}-terphenyl).

Both fractions were analyzed by gas chromatography coupled with mass spectrometry (GC-MS) with a final pre-injection volume of 1 mL.

Mean surrogate recoveries for all samples were 85.0 ± 8.0% for $C_{24}D_{50,}$ 56.3 ± 19.2% for d_8-naphthalene, 70.4 ± 11.0% for d_{10}-acenaphthene, 78.9 ± 11.0% for d_{10}-phenanthrene, 78.4 ± 11.7% for d_{12}-benz[*a*]anthracene, and 84.2 ± 33.5% for d_{12}-perylene. Calibrations for instrument stability were executed every 10 injections with standard solutions, and daily calibrations were < 20% for all target analyses. Limit of quantitation (LOQ) for this study was 0.08-1.3 ng/g dry weight (dw) for *n*-alkanes, 0.02–0.2 ng/g dw for PACs, and 0.3-0.7 ng/g dw for petroleum biomarkers (defined as the ratio of signal to noise (S/N) >10).

2.3.4 Statistical Analyses

Statistical analyses for this study were performed using the R statistical computing environment (v3.3.2). Wherever the assumptions of a t-test could not be met (normality, sample size, and/or variance equality), the Kruskal-Wallis rank sum test was applied using the "kruskal.test" function of the 'stats' package. Biomarker and PAC data were transformed using the "decostand" function in the 'vegan' package prior to multivariate statistics with the Hellinger method because it is commonly recommended for clustering and ordination, giving low weight to rare values (*i.e.*, zeroes in the data) (Legendre & Gallagher, 2001). Agglomerative hierarchical clustering (AHC) and *K*-means partitioning were used to reveal underlying patterns in the data that describe the statistical similarity between sediment samples and potential sources of oil relevant to the PAD watershed from a previous study by Yang et al. (2011). Samples from that study included raw Alberta Oil Sands (AOS1, AOS2, AOS3), diluted crude oil sands bitumen (DOB), Albian Heavy Synthetic (AHS), and Alberta Sweet Mixed Blend (ASMB) (Yang et al., 2011). AHC analysis was performed on Hellinger-transformed biomarker and PAC data using

the “agnes” function of the ‘cluster’ package, with the “ward” method for minimum-variance clustering. Following AHC, *K*-means partitioning was conducted on these same data to determine the number of clusters to be formed that minimized the total error sum-of-squares in each dataset. The within-group sum-of-squares was first plotted for each partition in the data using the ‘kmeans’ method (Everitt & Hothorn, 2010), and the optimal number of groups was identified by an inflection point in the resulting curve. Once the optimal number of groups was determined, the final cluster solution was calculated using the “kmeans” function of the ‘cclust’ package. Finally, these clusters were applied to the data, and fitted into a principal component analysis (PCA) using the “prcomp” function in the ‘vegan’ package. Clusters were represented on the AHC cluster dendrograms using the “cutree” function of the ‘stats’ package.

2.4 Results and Discussion

2.4.1 Sediment Elemental Analysis (%CNS)

A summary of results for chemistry of the surficial lake sediment samples is provided in the supplemental information (Table S2.2, Figure S2.1). Briefly, sediment organic carbon varied across samples, with lowest percentage in lakes PAD 30 and 31 (5.7 and 3.1%, respectively), and the highest value in lake UP 7 (35.1%). Total nitrogen content follows a similar trend, with lake PAD 31 having the lowest nitrogen content (0.21%) and UP 7 the highest (3.14%). Sulfur content was highest in lake sediments from PAD 1 and 3 (0.84 and 1.01%, respectively), and lowest in uplands lakes UP 3 and UP 9 (0.09 and 0.07%, respectively). The carbon to nitrogen ratio (C/N) in lake sediments ranges from 9.89 to 14.94, which lies between values for algal (4-10) and terrestrial (>20) organic matter indicating a dominance of autochthonous (in-lake) production with some influence from allochthonous (external) organic matter (Meyers & Teranes, 2002). Lower organic carbon and nitrogen content in Athabasca Delta lake sediments

suggests greater influence of inorganic sediment-laden river floodwaters, whereas higher organic carbon and nitrogen in Peace Delta and boreal upland sediments suggests the absence of flooding (Remmer et al., 2018). This conclusion is consistent with the flood histories and hydrological regimes described above.

2.4.2 n-Alkanes

n-Alkane concentrations vary among lakes and were primarily derived from biogenic (non-petrogenic) sources. Total *n*-alkane concentration ranges from 11.60 to 141.53 μg/g OC in boreal upland lake sediment samples, 111.72 to 115.51 μg/g OC in Peace Delta lake sediments, and 109.40 to 118.90 μg/g OC in Athabasca Delta lake samples (Table S2.3). No significant difference in the total concentrations of $\sum nC_9$-nC_{40} alkanes was found among sediments from the three regions. However, short chain length ($\sum nC_9$-nC_{16}) alkane concentrations are significantly different when comparing samples from the Peace Delta (mean 3.50 ± 1.53 μg/g OC), Athabasca Delta (mean 7.61 ± 4.63 μg/g OC), and boreal upland (mean 0.63 ± 0.83 μg/g OC) (Kruskal-Wallis test: $\chi^2 = 13.5$, $p < 0.01$) (Figure 2.2a). The carbon preference index (CPI), the ratio of odd to even carbon-numbered alkanes (Wang et al., 2009), was above three in all lakes (3.1-10.0), indicating a biogenic source of *n*-alkanes in all samples, despite the higher concentration of nC$_9$-nC$_{16}$ alkanes measured in PAD samples. The proxy ratio (P_{aq}), also known as the P-aqueous ratio, can help distinguish between terrestrial (< 0.1), emergent macrophyte (0.1-0.4) and submerged/floating macrophyte (0.4-1.0) sources of *n*-alkanes(Ficken, Li, Swain, & Eglinton, 2000). When measured in sediments, this ratio holds less value in distinguishing between these three sources since sediments often contain a mixture of two or more of these *n*-alkane sources. Based on the P_{aq}, *n*-alkanes in sediments from lakes UP 3, 7, 9, and lake PAD 30 were indicative

of a mixture of emergent and submerged/floating macrophytes, while those from lakes PAD 1, 3, and 31 indicated a higher content of submerged/floating macrophytes (Table S2.3).

Distribution patterns of the entire range of nC_9-nC_{40} alkanes further confirm input from biogenic sources (Figure S2.2). The relative proportions of *n*-alkanes were highly similar between samples, with a carbon preference for odd carbon-numbered alkanes apparent. The patterns in these samples bear resemblance to those taken farther upstream on the Steepbank River (Wang et al., 2014). The relatively high concentration of nC_{17} observed in all sediment samples is likely attributed to algae as this odd carbon-numbered alkane has been identified as an algal biomarker in lake sediments (Meyers & Ishiwatari, 1993). Dominance of nC_{23}, nC_{25}, nC_{27}, and nC_{29} alkanes in all sediment samples also indicates the presence of macrophytes and higher plants (Fang et al., 2014; Meyers & Ishiwatari, 1993). The nC_{23}/nC_{29} ratio described by Ficken et al. (2000) follows a similar conclusion as the P_{aq} ratio, where lakes PAD 1, 3, 31, and also UP 7 had a higher nC_{23} content, indicating a greater influence of submerged aquatic macrophytes. Lakes UP 3, 9, and PAD 30 had higher concentrations of nC_{29} alkanes, indicating a more emergent or terrestrial source of *n*-alkanes. The evidence for biogenic sources of alkanes in the present study is further supported by field and documented observations that lakes in the PAD support abundant macrophytes, with willows, grasses, and sedges present along the shorlines (K. Timoney, 2008). Overall, the CPI and P_{aq} ratios, distribution patterns, and dominance of odd carbon-numbered alkanes in the nC_{23}-nC_{29} range indicate a predominant source of alkanes in sediments is likely a mixture of submerged and emergent macrophytes with some input from terrestrial organic matter in lakes for all three regions.

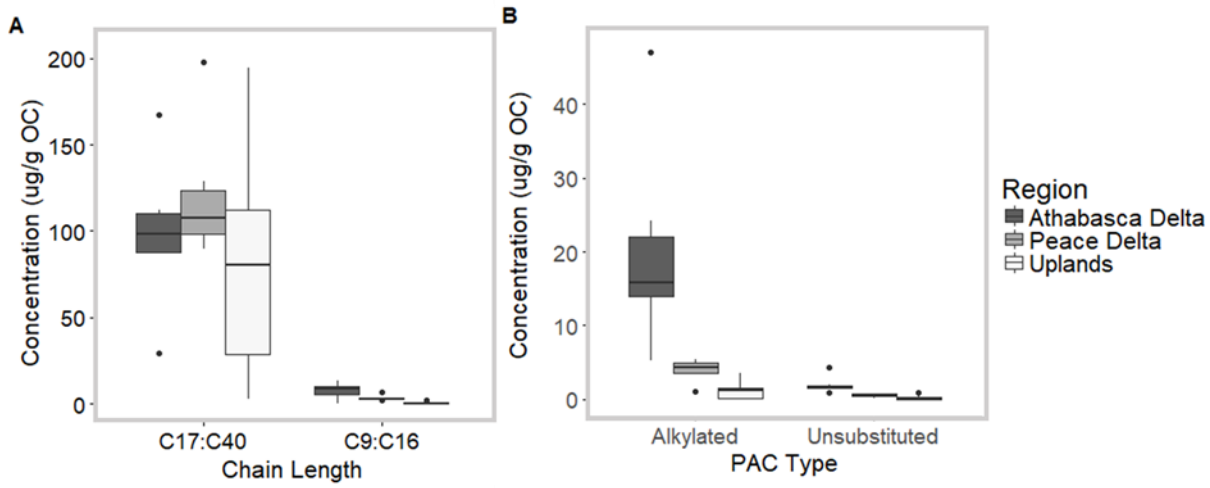

Figure 2.2. Boxplots showing A) the concentrations (µg/g OC) of nC$_9$-nC$_{16}$ alkanes and nC$_{17}$-nC$_{40}$ alkanes, and B) the concentrations (µg/g OC) of the sum of 15 unsubstituted PACs alkylated PACs (7 homologous series) separated by region for the Athabasca Delta (dark grey, n=6), Peace Delta (medium grey, n=6) and uplands (light grey, n=9) sediment samples.

2.4.3 Polycyclic Aromatic Compounds

Sediment PAC concentrations were significantly different among the three regions studied (Table S2.4). Total PAC concentration was lowest in boreal upland sediments (0.08-2.69 µg/g OC), slightly higher in Peace Delta sediments (3.77-5.06 µg/g OC), and highest in Athabasca Delta sediments (12.89-31.61 µg/g OC). The concentrations of PACs in sediment samples were significantly different between the Athabasca Delta (PAD 30 and 31), Peace Delta (PAD 1 and 3), and boreal uplands samples (UP 3, 7, and 9) for the sums of 15 unsubstituted PACs (Kruskal-Wallis test: $\chi^2 = 14.23$, $p < 0.001$) and alkylated PACs (Kruskal-Wallis test: $\chi^2 = 15.1$, $p < 0.001$) (Figure 2.2b). Upon further analysis of the homologous groups of alkylated PACs, the concentrations of each of the seven homologous series were significantly different among Athabasca Delta (highest), Peace Delta, and boreal upland (lowest) samples (Figure S2.3).

Geographical separation among sampling sites (*i.e.*, Athabasca Delta, Peace Delta, and boreal upland) was also apparent when analyzing the composition of individual PACs in sediments and their potential sources (Figure S2.3 and S2.4). Alkyl PAC composition (Figure S2.4) showed a mixture of distribution patterns in the PAD that can provide information on their sources and degree of weathering. Namely, upwards slope distribution of dibenzothiophenes (DBTs) and fluorenes in PAD 30 and 31 indicated advanced weathering stages (Michel & Hayes, 1999), and downward slope distributions of fluorenes in PAD 1 and 3 indicated pyrogenic sources of PACs (Wang et al., 2014). Fluoranthene (FL) and pyrene (PY) are two compounds that account for some of the separation among the three groups, and can be used to distinguish between pyrogenic and petrogenic sources of PACs by way of traditional diagnostic ratios (*e.g.* FL/(FL+PY)) (Tobiszewski & Namieśnik, 2012). The higher abundance of fluoranthene versus pyrene in boreal upland samples was indicative of wood combustion sources (mean FL/(FL+PY) = 0.69). Equal parts fluoranthene and pyrene in Peace Delta sediments indicated a mixture of petrogenic and pyrogenic combustion sources (mean FL/(FL+PY) = 0.48), and a higher abundance of pyrene versus fluoranthene in Athabasca Delta sediments indicated a petrogenic source of PACs (mean FL/(FL+PY) = 0.36)(Evans et al., 2016), which is consistent with paleolimnological records of PAD 31 (Hall et al., 2012). In terms of alkylated homologs, sediments sampled from lakes PAD 30 and 31 in the Athabasca Delta were the only samples to contain the entire homologous series of benzonaphthothiophene (BNT), arranged in a bell-shaped pattern (Figure S2.4). C_0-C_4 BNTs were the largest proportion of total PAC concentration in lakes PAD 30 and 31, accounting for 25 and 38% of the total, respectively. These polycyclic aromatic sulfur heterocycles (PASHs) are a major component of the organic sulfur content in crude oils (Mössner & Wise, 1999) and bitumen (Lam et al., 2012), and are associated with

petcoke dust (Manzano et al., 2017). Microbial formation of BNTs via the desulfurization of benzo[b]thiophenes has also been reported (Kropp, Gonçalves, Andersson, & Fedorak, 1994).

The high concentration of C_0-C_4 BNTs (mean 3.25 and 12.13 μg/g OC for PAD 30 and 31, respectively) along with C_0-C_3 DBTs (mean 1.07 and 2.86 μg/g OC for PAD 30 and 31, respectively), sets these two lakes in the Athabasca Delta apart from the rest of the dataset (Table S2.4). Alkylated dibenzothiophenes (C_2-C_4) have been previously identified as indicators of river-transported bitumen-associated PACs(Hall et al., 2012). However, this is the first study conducted in the Peace-Athabasca Delta's watershed to include and identify BNTs in the analysis of PACs. Far-field delivery of C_0 and C_1 BNTs via atmospheric deposition of petcoke dust has been demonstrated in surface sediments of the Athabasca River basin (Manzano et al., 2017). However, the presence of the entire series of C_0-C_4 BNTs in lakes PAD 30 and 31 may indicate the importance of the Athabasca River watershed in delivering these compounds. The higher concentration of BNTs relative to DBTs can be expected, as BNTs are more resistant to biodegradation than the latter and tend to increase in relative abundance, and to dominate the PAC fraction in late weathering stages (Michel & Hayes, 1999). The toxicity and mutagenicity of unsubstituted and methyl-substituted BNTs have been assessed several times (Eastmond, Booth, & Lee, 1984; Mcfall et al., 1984) and were reviewed elsewhere(Kropp & Fedorak, 1998) . Methyl-substituted BNTs having mutagenic potential, as determined by a Salmonella/microsome mutagenicity test (rat-liver homogenate metabolic activator), include 1- and 3-methylbenzo[*b*]naphtho[1,2-*d*]thiophene, 1- and 6-methylbenzo[*b*]naphtho[2,1-*d*]thiophene, and 4-methylbenzo[*b*]naphtho[2,3-*d*]thiophene(Mcfall et al., 1984). Benzo[*b*]naphtho[1,2-*d*]thiophene in particular was found to exhibit acute toxicity in *Daphnia magna* (LC_{50} = 0.220 mg/L) and had a maximum bioconcentration factor (BCF) 1.6 times that of chrysenes, in the

upper ranges of the values reported by the study (Eastmond et al., 1984). It is also important to note that the octanol-water coefficients (log K_{ow}) for unsubstituted BNTs range from 5.19-5.34 (Achten & Andersson, 2015). According to the Canadian Sediment Quality Guidelines, these log K_{ow} values are within the optimal range for bioaccumulation (Canadian Council of Ministers of the Environment, 1999). This group of alkylated PACs, previously undocumented in past studies in the PAD, may be important to take into consideration because their concentration in sediments is approximately three to four times that of DBTs (Table S2.5). Further investigation of the occurrence of BNTs and their distribution in the Peace-Athabasca Delta watershed is therefore warranted.

2.4.4 Petroleum Biomarkers

Petroleum biomarkers of terpanes, hopanes, and steranes further confirmed the regional separation of sediment samples among lakes in the Athabasca Delta, the Peace Delta, and the boreal uplands regions (Figure 2.3). Sediments from lakes in the boreal upland region had a mean biomarker concentration of 5.68 ng/g dw in UP 3, 369.22 ng/g dw in UP 7, and 48.37 ng/g dw in UP 9 (Table S2.5). Biomarkers identified include C_{22}-C_{24} terpanes, C_{27}-C_{31} hopanes, and gammacerane in UP 7 and UP 9. No steranes were identified in boreal upland sediments. Biomarkers including C_{22}-C_{24} terpanes, C_{27}-C_{31} hopanes, gammacerane, and $C_{27\text{-}29}$ αββ steranes were identified in Peace Delta lake sediments (mean concentrations of 137.27 and 76.27 ng/g dw for PAD 1 and 3, respectively). No steranes were detected in PAD 3. However, low concentrations of steranes were identified in PAD 1 (mean 8.82 ng/g dw). The presence of steranes at such low concentrations accompanied by a lack of C_{32}-C_{35} does not support a dominant petrogenic source of hydrocarbons in lake PAD 1 sediments. The presence of these petroleum biomarkers in surface sediments of closed-drainage lakes, especially in the boreal

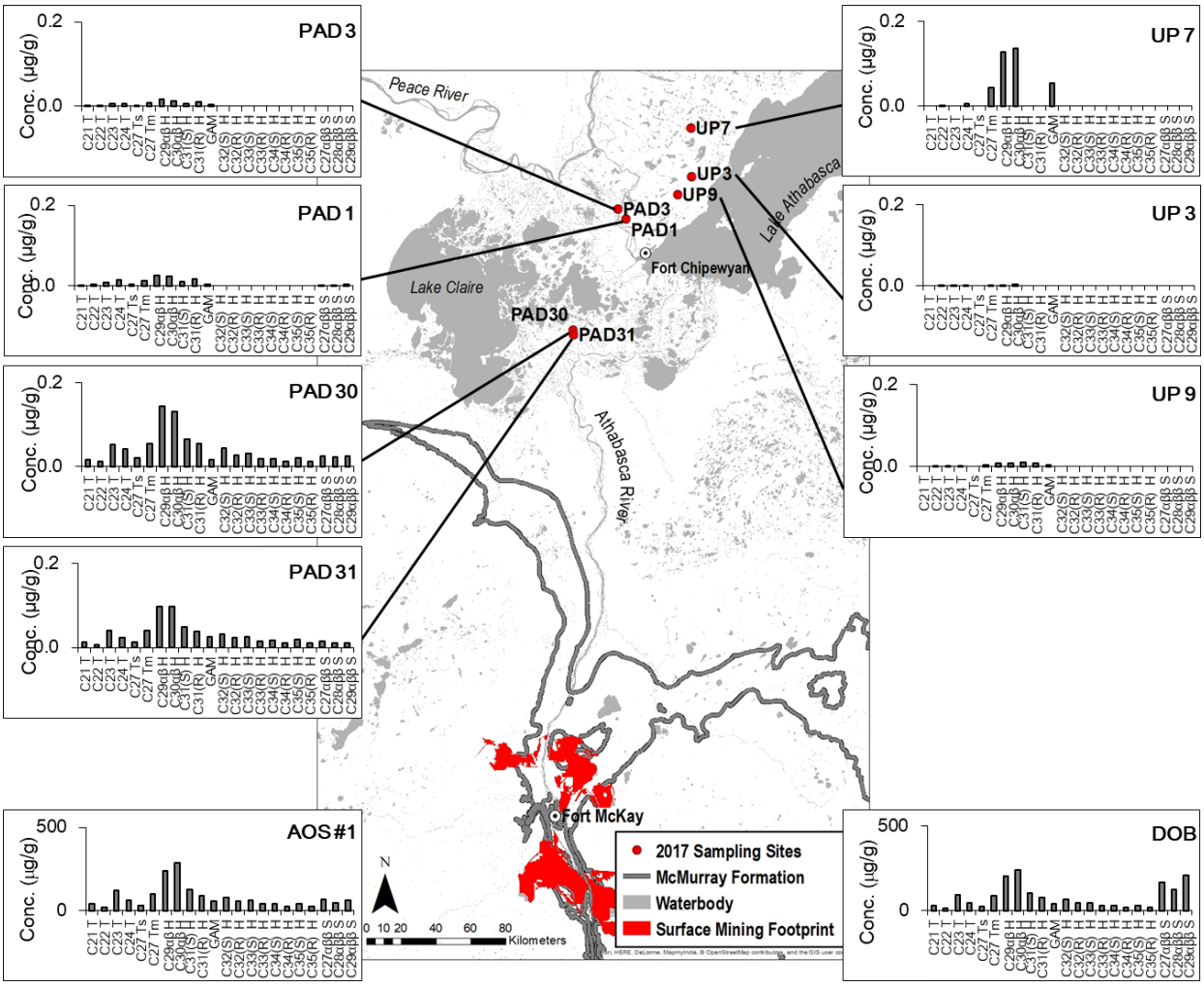

Figure 2.3. Map showing the location of study sites (red dots), the extent of the McMurray Formation (dark grey line with black outline) from Okulich and Fallas (Okulitch & Fallas, 2007), and the surface mining footprint of the AOS (shown in red) (Global Forest Watch Canada, 2008). Graphs in the margins show the biomarker distributions for terpanes (T), hopanes (H), gammacerane (GAM), and steranes (S) in surficial sediments of all study sites (mean µg/g dw, n=3 samples per lake) and two samples (raw Alberta oil sands (AOS #1) and diluted bitumen (DOB)) from the Yang et al. (2011) study (µg/g dw) for reference.

upland region, suggests their delivery via atmospheric deposition. Transport of petrogenic hopanoids and steranes has been shown via aerosols of vehicular emissions (*e.g.* lubricants) (Rogge, Hildemann, Mazurek, Cass, & Simoneit, 1993) and wood biomass burning (Omar, Abas, & Tahir, 2001). Other potential methods of transport for these compounds that are relevant

to the region include local emissions from energy generation (*e.g.* fuel oil) and long-range transport of secondary aerosols (Liggio et al., 2016) as well as petcoke dust (Jautzy, Ahad, Gobeil, et al., 2015).

Biomarker distributions and their percent composition in PAD 30 and 31 strongly resembled raw Alberta oil sands (AOS) samples #1-3 reported by Yang et al. (2011) (Figure 2.3). Mean biomarker concentrations in Athabasca Delta surficial sediments were 837.28 ng/g dw in PAD 30 and 633.59 ng/g dw in lake PAD 31. Biomarker concentrations in PAD 30 and 31 were approximately four orders of magnitude lower than that of AOS#1 from Yang et al. (2011). Nonetheless, their percent composition and diagnostic ratios are highly similar (Figure 2.3). Ratios of petroleum biomarkers for C_{23} terpane /C_{30}αβ hopane and C_{24} terpane /C_{30}αβ hopane, C_{27}Ts/C_{27}Tm hopanes (though slightly higher), and C_{27}ααβ/ C_{29}ααβ steranes in PAD 30 and 31 (Table S2.5) closely matched those of AOS#1 (0.42, 0.22, 0.29, 1.16, respectively) (Yang et al., 2011). The slight "V" shaped arrangement of steranes, marked by a lower concentration of C_{28}ααβ steranes in PAD 30 and 31 (Figure 2.3), has been demonstrated in bitumen and oils sourced from Albian rocks in the lower Cretaceous series (Creaney & Allan, 1992). This indicates that the PHCs in these lakes may originate from these deposits. Some dissimilarities exist between these sediments and samples of raw AOS. For instance, ratios of C_{23}/C_{24} terpane were lower in PAD 30 and 31, whereas ratios of C_{29}αβ/C_{30}αβ hopane and C_{32}(S)/C_{32}(R) hopane are higher in PAD 30 and 31 than they are in raw AOS#1 (Table S2.5). As previously mentioned, the contribution of C_{27}-C_{30} hopanes via bacteria (Peter et al., 2015; Zumberge, 1987), and the extent of weathering (as made evident by PAC composition), may explain why these ratios do not match those of raw AOS samples.

Agglomerative hierachical clustering (AHC) and *k*-means partitioning of standardized biomarker data confirmed raw AOS as a probable source of PHCs in Athabasca Delta lake sediments. This can be illustrated via ordination by the high degree of similarity and tight clustering between biomarker composition in Athabasca Delta lakes (PAD 30 and 31) and those from raw AOS samples (Figure 2.4a). Compounds that explain the highest proportion of variance within the ordination space include C_{30}αβ hopane (most susceptible to change via bacteria), C_{31}(S) and (R) hopanes (least prone to degradation within the $C_{31\text{-}35}$ range of hopanes), and $C_{27\text{-}29}$αββ steranes. Even though other potential source oils (*i.e.*, DOB, AHS, and ASMB) appeared in the same cluster determined by *k*-means partitioning, AHC analysis demonstrates the within-group distance (or dissimilarity) that exists between these biomarker compositions, illustrated by a branch separation in the dendrogram (Figure 2.5a). Visualization of the clustered sites in an

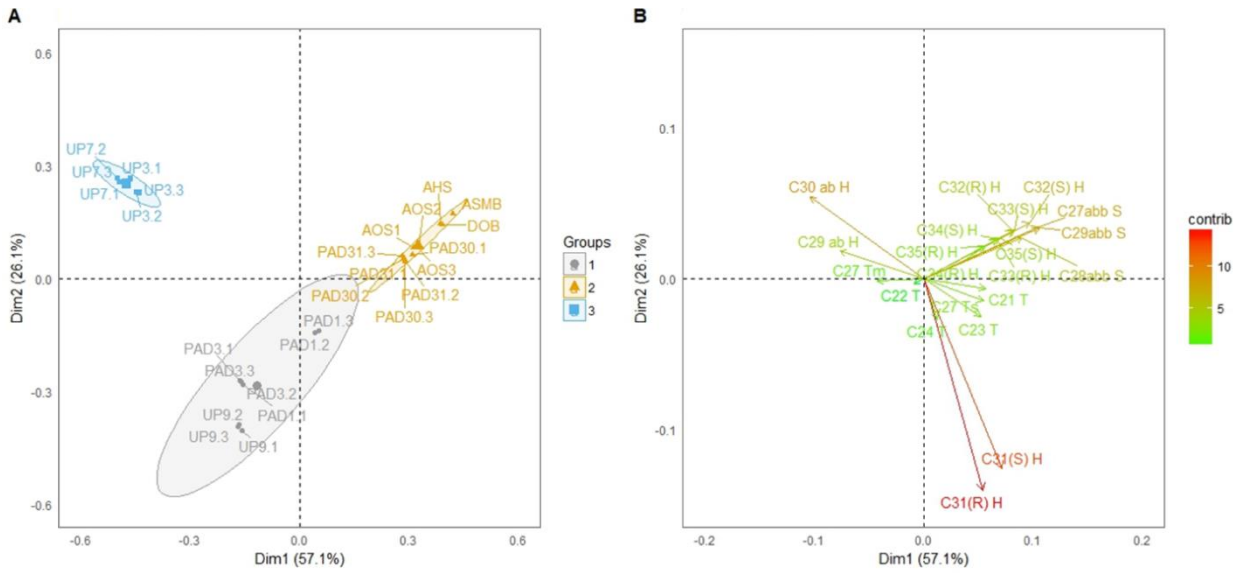

Figure 2.4. Ordination plots of principal components ('Dim') 1 and 2 (83.2% of total variance explained) of biomarker data in surficial lake sediments samples in 2017 and potential source oils (AOS1, AOS2, AOS3, AHS, ASMB, and DOB), n=27, showing A) sample scores, and B) biomarkers as explanatory variables. 'Groups' represents clusters obtained from *k*-means partitioning. Confidence ellipses are drawn for the 95% confidence interval. 'Contrib' represents the contribution to the total variance in the ordination for each vector as a percentage.

ordination of principal components 1 and 3 (Figure S2.5) confirms this separation as well. This distance is likely due to the higher contribution of steranes in crudes and diluted bitumen.

K-means partitioning of alkylated PAC composition forms five clusters and an inherently different AHC dendrogram from the biomarker data (Figure 2.5b). This illustrates the interference of alternate (non-petrogenic) sources of PACs (*e.g.* naphthalenes and phenanthrenes from forest fires) and weathering processes in the clustering of sites, which makes source identification more difficult. Relying solely on PAC compositions for this type of analysis would lead to the false assumption that PACs in Athabasca Delta sites may be coming from crude or diluted bitumen products when biomarker signatures do not support this. It is important to note that this AHC analysis was based on only five homologous series of alkylated PACs reported by Yang et al. (2011) (naphthalenes, phenanthrenes, dibenzothiophenes, fluorenes, and chrysenes),

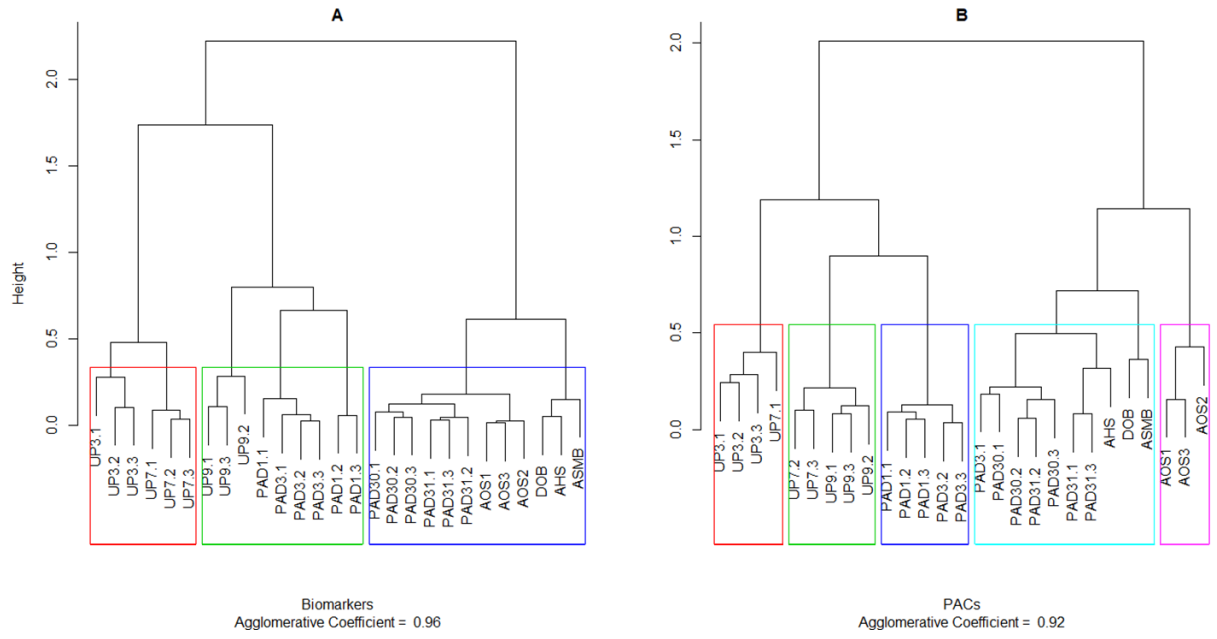

Figure 2.5. Cluster dendrograms of agglomerative hierarchical clustering using Ward's method for A) biomarker composition and B) PAC composition in 2017 lake sediment samples and potential source oils (AOS1, AOS2, AOS3, AHS, ASMB, and DOB) (Yang et al., 2011), n=27. Trees are cut into the appropriate number of clusters (as determined by *k*-means partitioning). Height is representative of the Euclidean distance.

excluding benzonaphthothiophenes and fluoranthene, further illustrating the importance of an expanded suite of PAC compounds.

Using a multivariate approach, we confirm that petroleum biomarkers can be used in the PAD to accurately identify the input and origin of petrogenic hydrocarbons from a set of potential source oils. We also provide further evidence that petrogenic PACs in these Athabasca Delta surficial lake sediments are likely of origin resembling raw AOS, and chemically distinct from sediments sampled in closed-drainage lakes of the Peace Delta and boreal uplands. In conjunction with hydrology, petroleum biomarkers indicate that PHCs in restricted-drainage lakes of the Athabasca Delta (PAD 30 and 31) may be delivered by the Athabasca River's watershed, whereas PHCs in closed-drainage lakes of the Peace Delta (PAD 1 and 3) and boreal uplands (UP 3, 7, and 9) may be delivered via atmospheric deposition. Obtaining and characterizing biomarkers from the Peace River oil sands deposit as well as oil sands fugitive dust (*e.g.* petcoke dust) and locally combusted fuel should become a priority for research into PHC distribution in the PAD. Since biomarkers are well conserved and resistant to biodegradation, a paleolimnological assessment of petroleum biomarkers could provide valuable insights into the origin and historical deposition of petrogenic materials in these ecologically important lake ecosystems. Analyzing petroleum biomarker content in a wider range of PAD lake sediments could elucidate the role that hydrological regimes (*e.g.* flooding) play in delivering PHCs throughout this complex ecosystem. In addition, extending the chronology (*i.e.*, downcore) passed the surficial sediments analysed in this study could help assess whether PHC deposition has changed through time since the onslaught of upstream industrial activities.

2.5 References

Achten, C., & Andersson, J. T. (2015). Overview of Polycyclic Aromatic Compounds (PAC). Polycyclic Aromatic Compounds, 35(2–4), 177–186. http://doi.org/10.1080/10406638.2014.994071

Canadian Council of Ministers of the Environment. (1999). Canadian Sediment Quality Guidelines for the Protection of Aquatic Life: Polycyclic Aromatic Hydrocarbons (PAHs). Canadian Environmental Quality Guidelines. Winnipeg.

Conly, F. M., Crosley, R. W., & Headley, J. (2002). Characterizing sediment sources and natural hydrocarbon inputs in the lower Athabasca River, Canada. Journal of Environmental Engineering and Science, 1, 187–199. http://doi.org/10.1139/s02-013

Creaney, S., & Allan, J. (1992). Petroleum systems in the foreland basin of Western Canada. In R. W. Macqueen & D. A. Leckie (Eds.), AAPG Memoir 55: Foreland basins and fold belts (pp. 279–308). American Association of Petroleum Geologists. http://doi.org/https://doi.org/10.1306/M55563

Dupuis, A., & Ucan-Marin, F. (2015). A literature review on the aquatic toxicology of petroleum oil: an overview of oil properties and effects to aquatic biota. DFO Canadian Science Advisory Secretariat Research Document, 2015/007(vi), 52.

Eastmond, D. A., Booth, G. M., & Lee, M. L. (1984). Toxicity, accumulation, and elimination of polycyclic aromatic sulfur heterocycles in Daphnia magna. Archives of Environmental Contamination and Toxicology, 13(1), 105–111. http://doi.org/10.1007/BF01055652

Evans, M., Davies, M., Janzen, K., Muir, D., Hazewinkel, R., Kirk, J., & De Boer, D. (2016). PAH distributions in sediments in the oil sands monitoring area and western Lake Athabasca: Concentration, composition and diagnostic ratios. Environmental Pollution, 213, 671–687. http://doi.org/10.1016/j.envpol.2016.03.014

Everitt, B. S., & Hothorn, T. (2010). Cluster Analysis: Classifying Romano-British Pottery and Exoplanets. In A Handbook of Statistical Analyses Using R (2nd ed., pp. 315–334). Taylor and Francis Group.

Fang, J., Wu, F., Xiong, Y., Li, F., Du, X., An, D., & Wang, L. (2014). Source characterization of sedimentary organic matter using molecular and stable carbon isotopic composition of n-alkanes and fatty acids in sediment core from Lake Dianchi, China. Science of the Total Environment, 473–474, 410–421. http://doi.org/10.1016/j.scitotenv.2013.10.066

Ficken, K. J., Li, B., Swain, D. L., & Eglinton, G. (2000). An n-alkane proxy for the sedimentary input of submerged/floating freshwater aquatic macrophytes. Organic Geochemistry, 31(7–8), 745–749. http://doi.org/10.1016/S0146-6380(00)00081-4

Hall, R. I., Wolfe, B. B., Wiklund, J. A., Edwards, T. W. D., Farwell, A. J., & Dixon, D. G. (2012). Has Alberta Oil Sands Development Altered Delivery of Polycyclic Aromatic Compounds to the Peace-Athabasca Delta? PLoS ONE, 7(9), e46089. http://doi.org/10.1371/journal.pone.0046089

Jautzy, J. J., Ahad, J. M. E., Gobeil, C., Smirnoff, A., Barst, B. D., & Savard, M. M. (2015). Isotopic Evidence for Oil Sands Petroleum Coke in the Peace-Athabasca Delta. Environmental Science and Technology, 49(20), 12062–12070. http://doi.org/10.1021/acs.est.5b03232

Jautzy, J. J., Ahad, J. M. E., Hall, R. I., Wiklund, J. A., Wolfe, B. B., Gobeil, C., & Savard, M. M. (2015). Source Apportionment of Background PAHs in the Peace-Athabasca Delta (Alberta, Canada) Using Molecular Level Radiocarbon Analysis. Environmental Science and Technology, 49(15), 9056–9063. http://doi.org/10.1021/acs.est.5b01490

Kelly, E. N., Short, J. W., Schindler, D. W., Hodson, P. V, Ma, M., Kwan, A. K., & Fortin, B. L. (2009). Oil sands development contributes polycyclic aromatic compounds to the Athabasca River and its tributaries. Proceedings of the National Academy of Sciences of the United States of America, 106(52), 22346–51. http://doi.org/10.1073/pnas.0912050106

Kropp, K. G., & Fedorak, P. M. (1998). A review of the occurrence, toxicity, and biodegradation of condensed thiophenes found in petroleum. Canadian Journal of Microbiology, 44(7), 605–22. http://doi.org/10.1139/w98-045

Kropp, K. G., Gonçalves, J. A., Andersson, J. T., & Fedorak, P. M. (1994). Microbially Mediated Formation of Benzonaphthothiophenes from Benzo[b]thiophenes. Applied and Environmental Microbiology, 60(10), 3624–3631.

Kurek, J., Kirk, J. L., Muir, D. C. G., Wang, X., Evans, M. S., & Smol, J. P. (2013). Legacy of a half century of Athabasca oil sands development recorded by lake ecosystems. Proceedings of the National Academy of Sciences of the United States of America, 110(5), 1761–6. http://doi.org/10.1073/pnas.1217675110

Lam, V., Li, G., Song, C., Chen, J., Fairbridge, C., Hui, R., & Zhang, J. (2012). A review of electrochemical desulfurization technologies for fossil fuels. Fuel Processing Technology, 98, 30–38. http://doi.org/10.1016/j.fuproc.2012.01.022

Legendre, P., & Gallagher, E. D. (2001). Ecologically meaningful transformations for ordination of species data. Oecologia, 129(2), 271–280. http://doi.org/10.1007/s004420100716

Liggio, J., Li, S.-M., Hayden, K., Taha, Y. M., Stroud, C., Darlington, A., … Gentner, D. R. (2016). Oil sands operations as a large source of secondary organic aerosols. Nature, 1–16. http://doi.org/10.1038/nature17646

Manzano, C. A., Marvin, C., Muir, D., Harner, T., Martin, J., & Zhang, Y. (2017). Heterocyclic Aromatics in Petroleum Coke, Snow, Lake Sediments, and Air Samples from the Athabasca Oil Sands Region. Environmental Science and Technology, 51(10), 5445–5453. http://doi.org/10.1021/acs.est.7b01345

Mcfall, T., Booth, G. M., Lee, M. L., Tominaga, Y., Pratap, R., Tedjamulia, M., & Castle, R. N. (1984). Mutagenic activity of methyl-substituted tri- and tetracyclic aromatic sulfur heterocycles heterocycles. Mutation Research, 135, 97–103.

Meyers, P. A., & Ishiwatari, R. (1993). Lacustrine organic geochemistry-an overview of indicators of organic matter sources and diagenesis in lake sediments. Organic Geochemistry, 20(7), 867–900. http://doi.org/10.1016/0146-6380(93)90100-P

Meyers, P. A., & Teranes, J. L. (2002). Sediment Organic Matter. In W. M. Last & J. P. Smol (Eds.), Tracking Environmental Change Using Lake Sediments (2nd ed., pp. 239–269). Springer. http://doi.org/DOI https://doi-org.proxy.bib.uottawa.ca/10.1007/0-306-47670-3_9

Michel, J., & Hayes, M. O. (1999). Weathering patterns of oil residueseight years after the Exxon Valdez oil spill. Marine Pollution Bulletin, 38(10), 855–863.

Mössner, S. G., & Wise, S. A. (1999). Determination of polycyclic aromatic sulfur heterocycles in fossil fuel-related samples. Analytical Chemistry, 71(1), 58–69. http://doi.org/10.1021/ac980664f

Okulitch, A. V, & Fallas, K. M. (2007). GIS dataset of Phanerozoic geological features, Slave River, Peace River, and Lake Athabasca map areas, Northwest Territories, Alberta, and Saskatchewan. Geological Survey of Canada. http://doi.org/https://doi.org/10.4095/223451

Omar, N. Y. M. J., Abas, M. R. Bin, & Tahir, N. M. (2001). Heavy molecular-weight organic compounds in the atmosphere: The hopanes. Malaysian Journal of Analytical Sciences, 7(1), 203–208.

Peter, H., Ingermann, H., Bryld, M., Fyhn, W., Henrik, L., Hovikoski, J., & Abatzis, I. (2015). Organic Geochemistry Novel saturated hexacyclic C 34 and C 35 hopanes in lacustrine oils and source rocks. Organic Geochemistry, 87, 107–118. http://doi.org/10.1016/j.orggeochem.2015.07.013

Remmer, C. R., Klemt, W. H., Wolfe, B. B., & Hall, R. I. (2018). Inconsequential effects of flooding in 2014 on lakes in the Peace-Athabasca Delta (Canada) due to long-term drying. Limnology and Oceanography, 63(4), 1502–1518. http://doi.org/10.1002/lno.10787

Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., & Simoneit, B. R. T. (1993). Sources of Fine Organic Aerosol. 2. Noncatalyst and Catalyst-Equipped Automobiles and Heavy-Duty Diesel Trucks. Environmental Science and Technology, 27(4), 636–651. http://doi.org/10.1021/es00041a007

Timoney, K. (2008). Factors influencing wetland plant communities during a flood-drawdown cycle in the Peace-Athabasca Delta, Northern Alberta, Canada. Wetlands, 28(2), 450–463. http://doi.org/10.1672/07-45.1

Timoney, K. P. (2013). The Delta's Physical Environment and Landforms. In The Peace-Athabasca Delta: Portrait of a Dynamic Ecosystem (pp. 15–56). Edmonton: University of Alberta Press.

Timoney, K. P., & Lee, P. (2011). Polycyclic aromatic hydrocarbons increase in athabasca river delta sediment: Temporal trends and environmental correlates. Environmental Science and Technology, 45(10), 4278–4284. http://doi.org/10.1021/es104375d

Tobiszewski, M., & Namieśnik, J. (2012). PAH diagnostic ratios for the identification of pollution emission sources. Environmental Pollution, 162, 110–119. http://doi.org/10.1016/j.envpol.2011.10.025

Wang, Z., Fingas, M., Blenkinsopp, S., Sergy, G., Landriault, M., Sigouin, L., & Lambert, P. (1998). Study of the 25-year-old Nipisi oil spill: Persistence of oil residues and comparisons between surface and subsurface sediments. Environmental Science and Technology, 32(15), 2222–2232. http://doi.org/10.1021/es971070h

Wang, Z., Fingas, M., Blenkinsopp, S., Sergy, G., Landriault, M., Sigouin, L., … Westlake, D. W. S. (1998). Comparison of oil composition changes due to biodegradation and physical weathering in different oils. Journal of Chromatography A, 809(1–2), 89–107. http://doi.org/10.1016/S0021-9673(98)00166-6

Wang, Z., Yang, C., Fingas, M., Hollebone, B., Hyuk Yim, U., & Ryoung Oh, J. (2007). Petroleum biomarker fingerprinting for oil spill characterization and source identification. In Oil Spill Environmental Forensics (pp. 73–146). Elsevier. http://doi.org/10.1016/B978-012369523-9.50007-0

Wang, Z., Yang, C., Kelly-Hooper, F., Hollebone, B. P., Peng, X., Brown, C. E., … Yang, Z. (2009). Forensic differentiation of biogenic organic compounds from petroleum hydrocarbons in biogenic and petrogenic compounds cross-contaminated soils and sediments. Journal of Chromatography A, 1216(7), 1174–1191. http://doi.org/10.1016/j.chroma.2008.12.036

Wang, Z., Yang, C., Parrott, J. L., Frank, R. A., Yang, Z., Brown, C. E., … Hewitt, L. M. (2014). Forensic source differentiation of petrogenic, pyrogenic, and biogenic hydrocarbons in Canadian oil sands environmental samples. Journal of Hazardous Materials, 271, 166–177. http://doi.org/10.1016/j.jhazmat.2014.02.021

Yang, C., Wang, Z., Yang, Z., Hollebone, B., Brown, C. E., Landriault, M., & Fieldhouse, B. (2011). Chemical Fingerprints of Alberta Oil Sands and Related Petroleum Products. Environmental Forensics, 12(2), 173–188. http://doi.org/10.1080/15275922.2011.574312

Zumberge, J. E. (1987). Prediction of source rock characteristics based on terpane biomarkers in crude oils : A multivariate statistical approach. Geochimica et Cosmochimica Acta, 51(6), 1625–1637.

2.6 Supplemental Information

Table S2.1. Biomarkers of terpanes, steranes, and hopanes targeted in 2017 sediment samples.

Name	Code	Target ion
Terpanes		
C21 terpane	C21 T	191
C22 terpane	C22 T	191
C23 terpane	C23 T	191
C24 terpane	C24 T	191
Hopanes		
C27 18A-hopane II (Ts)	C27 Ts	191
C27 17A-hopane (Tm)	C27 Tm	191
C29 αβ hopane	C29 αβ H	191
C30 αβ hopane	C30 αβ H	191
C31(S) hopane	C31(S) H	191
C31(R) hopane	C31(R) H	191
C32(S) hopane	C32(S) H	191
C32(R) hopane	C32(R) H	191
C33(S) hopane	C33(S) H	191
C33(R) hopane	C33(R) H	191
C34(S) hopane	C34(S) H	191
C34(R) hopane	C34(R) H	191
C35(S) hopane	C35(S) H	191
C35(R) hopane	C35(R) H	191
Gammacerane	GAM	191
Steranes		
C27αββ steranes	C27αββ S	218
C28αββ steranes	C28αββ S	218
C29αββ steranes	C29αββ S	218

Table S2.2. Mean values percent organic matter (carbon, nitrogen, and sulfur) and carbon to nitrogen ration (C/N) in July 2017 surface sediment samples from each lake.

	PAD 1	PAD 3	PAD 30	PAD 31	UP 3	UP 7	UP 9
Carbon (%)	11.67	13.17	5.70	3.10	14.13	35.10	11.23
Nitrogen (%)	1.19	1.22	0.50	0.21	1.41	3.14	0.89
Sulfur (%)	0.84	1.01	0.16	0.10	0.09	0.18	0.13
C/N	9.81	10.80	11.40	14.76	10.02	11.18	12.62

Table S2.3. Mean concentrations of *n*-Alkanes from triplicate lake sediment samples in µg/g OC.

	PAD 1	PAD 3	PAD 30	PAD 31	UP 3	UP 7	UP 9
n-Alkanes (µg/g OC)							
n-C_9	0.25	0.25	0.82	0.58	ND	0.04	0.22
n-C_{10}	0.37	0.37	1.07	0.74	ND	0.04	0.19
n-C_{11}	0.66	0.50	1.78	1.10	ND	0.07	0.26
n-C_{12}	0.39	0.30	0.87	0.73	ND	0.03	0.12
n-C_{13}	0.42	0.37	0.71	0.67	0.01*	0.02	0.05
2,6,10-trimethyldodecane	0.09	0.08	0.17	0.09	$1.72x10^{-3}$*	0.02	0.02
n-C_{14}	0.29	0.22	0.57	0.65	$4.85x10^{-3}$*	0.02	0.07
n-C_{15}	0.47	0.39	0.72	1.19	0.05	0.18	0.16
n-C_{16}	0.31	0.29	0.54	0.65	0.03	0.12	0.10
2,6,10-trimethylpentadecane	0.10	0.08	0.23	0.08	$3.21x10^{-3}$*	0.02	0.02
n-C_{17}	2.61	3.32	4.18	3.23	1.27	6.17	1.42
Pristane	0.32	0.31	0.60	0.50	0.01*	0.09	0.12
n-C_{18}	0.60	0.60	0.86	0.71	0.08	0.51	0.21
Phytane	1.46	1.52	2.10	1.25	0.09	1.05	0.37
n-C_{19}	1.58	1.39	1.26	1.07	0.21	1.14	0.69
n-C_{20}	0.65	0.56	0.98	0.75	0.34	0.56	0.48
n-C_{21}	5.31	3.68	2.52	2.23	0.50	3.84	1.89
n-C_{22}	1.64	1.26	2.29	1.82	0.17	1.80	0.79
n-C_{23}	20.80	15.24	7.47	18.67	0.64	16.67	5.00
n-C_{24}	4.74	3.45	4.99	4.75	0.28	4.24	1.61
n-C_{25}	27.52	25.75	17.61	26.01	2.24	27.81	16.74
n-C_{26}	2.68	2.46	6.23	2.78	0.28	6.16	1.79
n-C_{27}	22.33	23.44	21.19	20.14	3.30	36.33	23.72
n-C_{28}	1.93	2.00	4.48	1.85	0.22	3.65	1.28
n-C_{29}	13.30	16.84	17.08	10.36	1.43	18.17	14.76
n-C_{30}	0.54	0.68	2.58	0.80	0.04	1.27	0.42
n-C_{31}	3.08	4.94	8.55	4.14	0.30	8.66	6.32
n-C_{32}	0.17	0.19	1.59	0.36	0.01*	0.48	0.29
n-C_{33}	0.69	1.09	2.96	1.20	0.07	2.09	1.90
n-C_{34}	0.05	0.04	0.60	0.11	$4.10x10^{-3}$*	0.07	0.02
n-C_{35}	0.07	0.09	0.56	0.19	0.01*	0.10	0.04
n-C_{36}	ND	ND	0.22	ND	ND	0.05	$3.87x10^{-3}$*
n-C_{37}	0.02	ND	0.15	ND	ND	0.03	$4.52x10^{-3}$*
n-C_{38}	ND	ND	0.09	ND	ND	ND	ND
n-C_{39}	0.07	0.02	0.23	ND	0.01*	0.03	0.03
n-C_{40}	ND	ND	0.05	ND	ND	ND	ND
Total n-alkanes (µg/g OC)	115.51	111.72	118.90	109.40	11.60	141.53	81.11
P_{aq}							
(C_{23}+C_{25}) / (C_{23}+C_{25}+C_{29}+C_{31})	0.75	0.65	0.49	0.75	0.62	0.62	0.51
CPI	6.9	7.8	3.1	5.4	6.8	6.4	10.0
C_{23}/C_{29}	1.56	0.90	0.44	1.80	0.45	0.92	0.34

* Below upper range of LOQ (1.3 ng/g dry)

ND = not detected

Table S2.4. Mean concentrations of 7 homologous series of alkylated PACs and 15 unsubstituted PACs from triplicate lake sediment samples in μg/g OC.

Compounds	Abbreviations	PAD 1	PAD 3	PAD 30	PAD 31	UP 3	UP 7	UP 9
Alkylated PACs (μg/g OC)								
Naphthalene								
C0-Naphthalene	C0-N	0.23	0.14	0.38	0.41	ND	0.01	0.05
C1-Naphthalene	C1-N	0.48	0.27	0.49	0.63	ND	0.03	0.08
C2-Naphthalene	C2-N	0.60	0.35	0.56	0.78	$1.48x10^{-3}$*	0.03	0.10
C3-Naphthalene	C3-N	0.39	0.24	0.43	0.63	$3.39x10^{-3}$	0.04	0.14
C4-Naphthalene	C4-N	0.19	0.10	0.25	0.45	$8.69x10^{-4}$*	0.04	0.10
Phenanthrene								
C0-Phenanthrene	C0-P	0.29	0.20	0.34	0.49	0.03	0.14	0.30
C1-Phenanthrene	C1-P	0.37	0.28	0.48	0.79	$8.70x10^{-3}$	0.14	0.27
C2-Phenanthrene	C2-P	0.24	0.19	0.49	0.93	$3.18x10^{-3}$	0.09	0.18
C3-Phenanthrene	C3-P	0.14	0.16	0.50	1.43	$1.56x10^{-3}$*	0.05	0.09
C4-Phenanthrene	C4-P	0.04	0.05	0.38	1.17	$3.07x10^{-4}$*	0.01	0.01
Dibenzothiophene								
C0-Dibenzothiophene	C0-D	0.05	0.04	0.05	0.09	$7.87x10^{-4}$*	0.01	0.02
C1-Dibenzothiophene	C1-D	0.06	0.05	0.13	0.25	$1.01x10^{-3}$*	0.02	0.04
C2-Dibenzothiophene	C2-D	0.06	0.06	0.31	0.71	ND	0.02	0.05
C3-Dibenzothiophene	C3-D	0.03	0.04	0.58	1.81	ND	0.01	0.02
Fluorene								
C0-Fluorene	C0-F	0.19	0.13	0.14	0.23	$4.78x10^{-3}$	0.03	0.05
C1-Fluorene	C1-F	0.15	0.11	0.24	0.31	$2.65x10^{-3}$	0.11	0.34
C2-Fluorene	C2-F	0.11	0.10	0.31	0.41	$3.94x10^{-3}$	0.05	0.09
C3-Fluorene	C3-F	0.08	0.07	0.31	0.70	$2.47x10^{-3}$*	0.03	0.06
Fluoranthene								
C0-Fluoranthene	C0-Fl	0.06	0.05	0.11	0.14	$5.19x10^{-3}$	0.05	0.09
C1-Fluoranthene	C1-Fl	0.11	0.10	0.24	0.54	$9.52x10^{-5}$*	0.03	0.04
C2-Fluoranthene	C2-Fl	0.10	0.09	0.32	0.85	$6.89x10^{-5}$*	0.01	0.02
C3-Fluoranthene	C3-Fl	0.04	0.06	0.26	0.75	ND	$2.66x10^{-3}$	$2.36x10^{-3}$*
C4-Fluoranthene	C4-Fl	0.02	0.03	0.14	0.39	ND	$1.43x10^{-3}$	ND

Table S2.4 (cont'd). Mean concentrations of 7 homologous series of alkylated PACs and 15 unsubstituted PACs from triplicate lake sediment samples in μg/g OC.

Compounds	Abbreviations	PAD 1	PAD 3	PAD 30	PAD 31	UP 3	UP 7	UP 9
Benzonaphthothiophene								
C0-Benzonaphthothiophene	C0-B	0.02	0.02	0.07	0.19	$5.24x10^{-5}$*	$3.47x10^{-3}$	$5.43x10^{-3}$
C1-Benzonaphthothiophene	C1-B	0.10	0.08	0.53	1.67	ND	ND	ND
C2-Benzonaphthothiophene	C2-B	0.05	0.08	0.81	3.34	ND	ND	ND
C3-Benzonaphthothiophene	C3-B	ND	ND	1.34	4.55	ND	ND	ND
C4-Benzonaphthothiophene	C4-B	ND	ND	0.50	2.39	ND	ND	ND
Chrysene								
C0-Chrysene	C0-C	0.06	0.05	0.13	0.25	$9.53x10^{-4}$*	0.01	0.02
C1-Chrysene	C1-C	0.10	0.08	0.25	0.59	$7.55x10^{-5}$*	$3.73x10^{-3}$	$8.08x10^{-3}$
C2-Chrysene	C2-C	0.06	0.06	0.32	0.84	ND	$4.38x10^{-3}$	ND
C3-Chrysene	C3-C	0.02	0.02	0.12	0.32	ND	ND	ND
Total alkylated PACs		4.44	3.30	11.51	29.03	0.07	0.98	2.18
15 Unsubstituted PACs (μg/g OC)								
Biphenyl	Bph	0.05	0.03	0.07	0.09	$3.49x10^{-4}$*	$5.77x10^{-7}$	0.02
Acenaphthylene	Acl	$6.54x10^{-3}$	$3.98x10^{-3}$	0.01	0.01	ND	$2.57x10^{-4}$*	$4.14x10^{-4}$*
Acenaphthene	Ace	0.01	0.01	0.02	0.03	ND	$2.18x10^{-3}$	$1.86x10^{-3}$*
Anthracene	An	0.01	0.01	0.03	0.05	ND	0.01	0.02
Fluoranthene	Fl	0.06	0.05	0.11	0.14	$5.19x10^{-3}$	0.05	0.09
Pyrene	Py	0.07	0.05	0.20	0.24	$1.31x10^{-3}$*	0.03	0.06
Benz[*a*]anthracene	BaA	0.01	0.01	0.02	0.06	ND	$5.52x10^{-4}$*	$9.90x10^{-4}$*
Benzo[*b*]fluoranthene	BbF	0.04	0.03	0.09	0.18	$8.12x10^{-5}$*	$2.50x10^{-3}$	$4.15x10^{-3}$
Benzo[*k*]fluoranthene	BkF	0.01	0.01	0.02	0.03	$2.29x10^{-5}$*	$1.09x10^{-3}$	$1.82x10^{-3}$*
Benzo[*e*]pyrene	BeP	0.04	0.03	0.07	0.14	$5.41x10^{-5}$*	$1.51x10^{-3}$	$3.04x10^{-3}$
Benzo[*a*]pyrene	BaP	0.02	0.01	0.03	0.09	$5.46x10^{-6}$*	$4.12x10^{-4}$*	$1.17x10^{-3}$*
Perylene	Pe	0.25	0.18	0.60	1.22	$3.49x10^{-4}$*	$6.61x10^{-3}$	0.31
Indeno[1,2,3-*cd*]pyrene	IP	$6.11x10^{-3}$	0.01	0.04	0.11	ND	$4.25x10^{-4}$*	ND
Dibenzo[*ah*]anthracene	DA	$3.21x10^{-3}$	$5.19x10^{-3}$	0.01	0.04	ND	ND	ND
Benzo[*ghi*]perylene	BgP	0.03	0.03	0.07	0.15	ND	$5.54x10^{-3}$	$1.45x10^{-3}$*
Total 15 Unsubstituted PACs (μg/g OC)		0.62	0.47	1.38	2.58	0.01	0.12	0.51
Total PACs (μg/g OC)		5.06	3.77	12.89	31.61	0.08	1.10	2.69

* Below upper range of LOQ (1.3 ng/g dry weight), ND = not detected

Table S2.5. Mean concentrations of petroleum biomarkers from triplicate lake samples in ng/g dw.

Compounds	Abbreviations	PAD 1	PAD 3	PAD 30	PAD 31	UP 3	UP 7	UP 9
Petroleum Biomarkers (ng/g dw)								
C21 terpane	C21 T	2.24	1.39	15.94	13.32	ND	ND	ND
C22 terpane	C22 T	3.87	1.52	11.80	6.56	0.21*	1.81	0.52*
C23 terpane	C23 T	9.07	5.47	51.83	40.58	0.26*	ND	1.11
C24 terpane	C24 T	14.15	5.40	42.93	23.58	0.21*	5.83	1.85
C27 Ts	C27 Ts	3.51	2.42	20.92	13.13	ND	ND	ND
C27 Tm	C27 Tm	12.87	8.00	55.91	40.76	0.58*	43.35	3.72
C29 ab Hopane	C29 ab H	26.49	17.65	143.78	98.19	1.52	127.62	7.86
C30 ab Hopane	C30 ab H	24.23	13.50	130.20	98.07	2.89	135.93	9.06
C31(S) hopane	C31(S) H	11.26	6.41	66.05	48.24	ND	ND	11.39
C31(R) hopane	C31(R) H	16.87	10.15	54.02	39.25	ND	ND	9.27
Gammacerane	GAM	3.89	4.38	16.96	25.55	ND	54.69	3.59
C32(S) hopane	C32(S) H	ND	ND	44.73	31.28	ND	ND	ND
C32(R) hopane	C32(R) H	ND	ND	27.52	22.45	ND	ND	ND
C33(S) hopane	C33(S) H	ND	ND	31.10	24.55	ND	ND	ND
C33(R) hopane	C33(R) H	ND	ND	18.98	15.04	ND	ND	ND
C34(S) hopane	C34(S) H	ND	ND	19.11	16.12	ND	ND	ND
C34(R) hopane	C34(R) H	ND	ND	13.24	10.49	ND	ND	ND
C35(S) hopane	C35(S) H	ND	ND	20.21	18.22	ND	ND	ND
C35(R) hopane	C35(R) H	ND	ND	13.41	11.11	ND	ND	ND
C27aßß steranes	C27aßß S	2.53	ND	26.08	14.53	ND	ND	ND
C28aßß steranes	C28aßß S	2.55	ND	23.86	10.97	ND	ND	ND
C29aßß steranes	C29aßß S	3.74	ND	24.68	11.61	ND	ND	ND
Sum of terpanes (ng/g dw)		128.45	76.27	798.66	596.49	5.68	369.22	48.37
Sum of steranes (ng/g dw)		8.82	ND	74.62	37.10	ND	ND	ND
Sum biomarkers (ng/g dw)		137.27	76.27	873.28	633.59	5.68	369.22	48.37
Diagnostic Ratios								
C_{23}/C_{24}		0.15	NA	0.33	0.84	0.81	1.21	1.60
C_{23}/C_{30}		0.09	NA	0.13	0.39	0.41	0.40	0.41
C_{24}/C_{30}		0.60	0.06	0.38	0.54	0.51	0.34	0.26
C_{29}/C_{30}		0.53	0.94	0.89	1.12	1.31	1.11	1.01
$C_{31}(S)/C_{31}(R)$		ND	ND	1.20	0.71	0.63	1.22	1.24
$C_{32}(S)/C_{32}(R)$		ND	ND	ND	ND	ND	1.64	1.40
Ts/Tm		NA	NA	NA	0.27	0.30	0.37	0.33
C_{27}abb/ C_{29}abb		ND	ND	ND	0.67	ND	1.09	1.17

* Below upper range of LOQ (0.7 ng/g dw)

ND = not detected, NA = not applicable

Table S2.6. Concentrations in µg/g of petroleum biomarkers from Yang et al. (2011) study samples including raw Alberta oil sands (AOS #1, #2, and #3), diluted crude oil sands bitumen (DOB), Albian Heavy Synthetic (AHS), and Alberta Sweet Mixed Blend (ASMB).

Compounds	Abbreviations	AOS #1	AOS #2	AOS #3	DOB	AHS	ASMB
Petroleum Biomarkers		µg/g TSEM	µg/g TSEM	µg/g TSEM	µg/g oil	µg/g oil	µg/g oil
C_{21} terpane	C_{21} T	39.5	37.5	39.7	30.9	15.9	16
C_{22} terpane	C_{22} T	18.7	18.2	18.6	14.1	6.78	6.29
C_{23} terpane	C_{23} T	120	113	120	92.3	46.9	41.3
C_{24} terpane	C_{24} T	62	58.9	62.8	47.8	24.2	21.2
C_{27} 18A-hopane II (Ts)	C_{27} Ts	29.3	27.7	29.3	25.4	13.1	31.1
C_{27} 17A-hopane (Tm)	C_{27} Tm	102	95	101	85.9	54.4	33.2
C_{29} αβ hopane	C_{29} αβ H	239	230	237	200	97.1	86.8
C_{30} αβ hopane	C_{30} αβ H	288	268	283	240	114	97.2
C_{31}(S) hopane	C_{31}(S) H	126	120	125	101	46.3	37.2
C_{31}(R) hopane	C_{31}(R) H	91.6	88.4	91.8	75.5	35.7	25.9
C_{32}(S) hopane	C_{32}(S) H	55.3	50	52	42.6	21.2	12.8
C_{32}(R) hopane	C_{32}(R) H	79.5	75.3	78.5	65.4	30.4	26.9
C_{33}(S) hopane	C_{33}(S) H	58.4	55.1	57.8	47.6	23.6	19.9
C_{33}(R) hopane	C_{33}(R) H	59.8	58	58.5	47.1	24.2	18.3
C_{34}(S) hopane	C_{34}(S) H	39.1	37.2	38	31.7	16.1	12.1
C_{34}(R) hopane	C_{34}(R) H	39.4	36.9	39.7	31.4	17.8	14.5
C_{35}(S) hopane	C_{35}(S) H	25.2	23.9	24.6	19.6	10.8	9.03
C_{35}(R) hopane	C_{35}(R) H	41.6	39.9	42.2	32	19	11.8
Gammacerane	GAM	27.4	26.3	28.2	21.3	13.6	8.11
C_{27}αββ steranes	C_{27}αββ S	69.80	80.20	75	166.00	76.5	106.0
C_{28}αββ steranes	C_{28}αββ S	48.70	54.60	52	124.00	56.5	62.0
C_{29}αββ steranes	C_{29}αββ S	60.0	70.20	68	206.0	85	122
Sum of terpanes		1541.8	1459.3	1527.70	1251.6	631.08	529.63
Sum of steranes		178.5	205	195.10	496	217.9	290
Sum biomarkers		1720.3	1664.3	1722.80	1747.6	848.98	819.63
Diagnostic Ratios							
C_{23}/C_{24}		1.94	1.92	1.91	1.93	1.94	1.95
C_{23}/C_{30}		0.42	0.42	0.42	0.38	0.41	0.42
C_{24}/C_{30}		0.22	0.22	0.22	0.20	0.21	0.22
C_{29}/C_{30}		0.83	0.86	0.84	0.83	0.85	0.89
C_{31}(S)/C_{31}(R)		1.38	1.36	1.36	1.34	1.30	1.44
C_{32}(S)/C_{32}(R)		1.36	1.37	1.36	1.37	1.29	1.35
Ts/Tm		0.29	0.29	0.29	0.30	0.24	0.94
C_{27}abb/ C_{29}abb		1.16	1.14	1.10	0.81	0.90	0.87

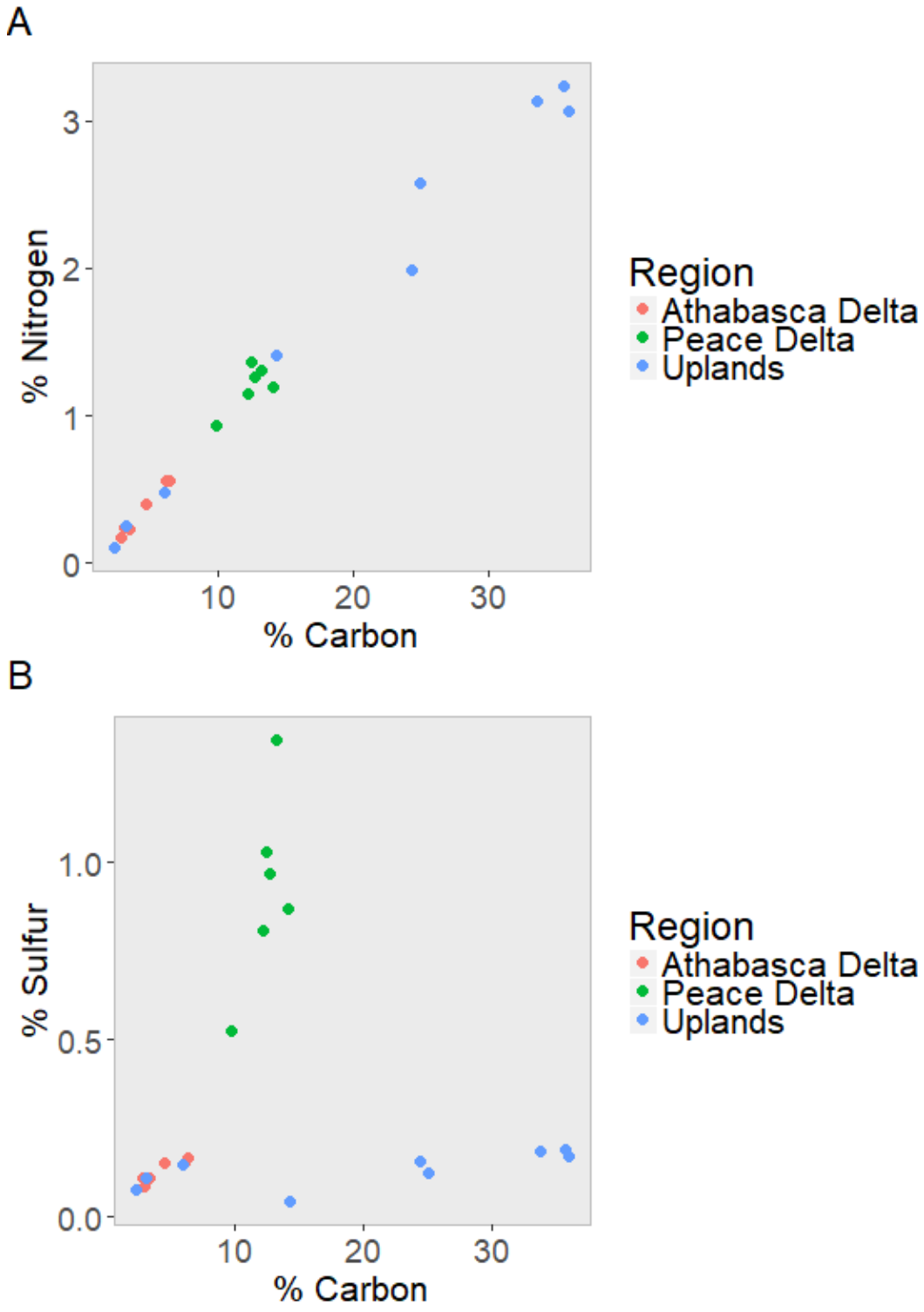

Figure S2.1. Scatterplots showing A) percent organic carbon versus nitrogen, and B) percent organic carbon versus sulfur, in July 2017 lake sediments from the Peace Delta (green, n=6), Athabasca Delta (red, n=6), and boreal uplands (blue, n=9).

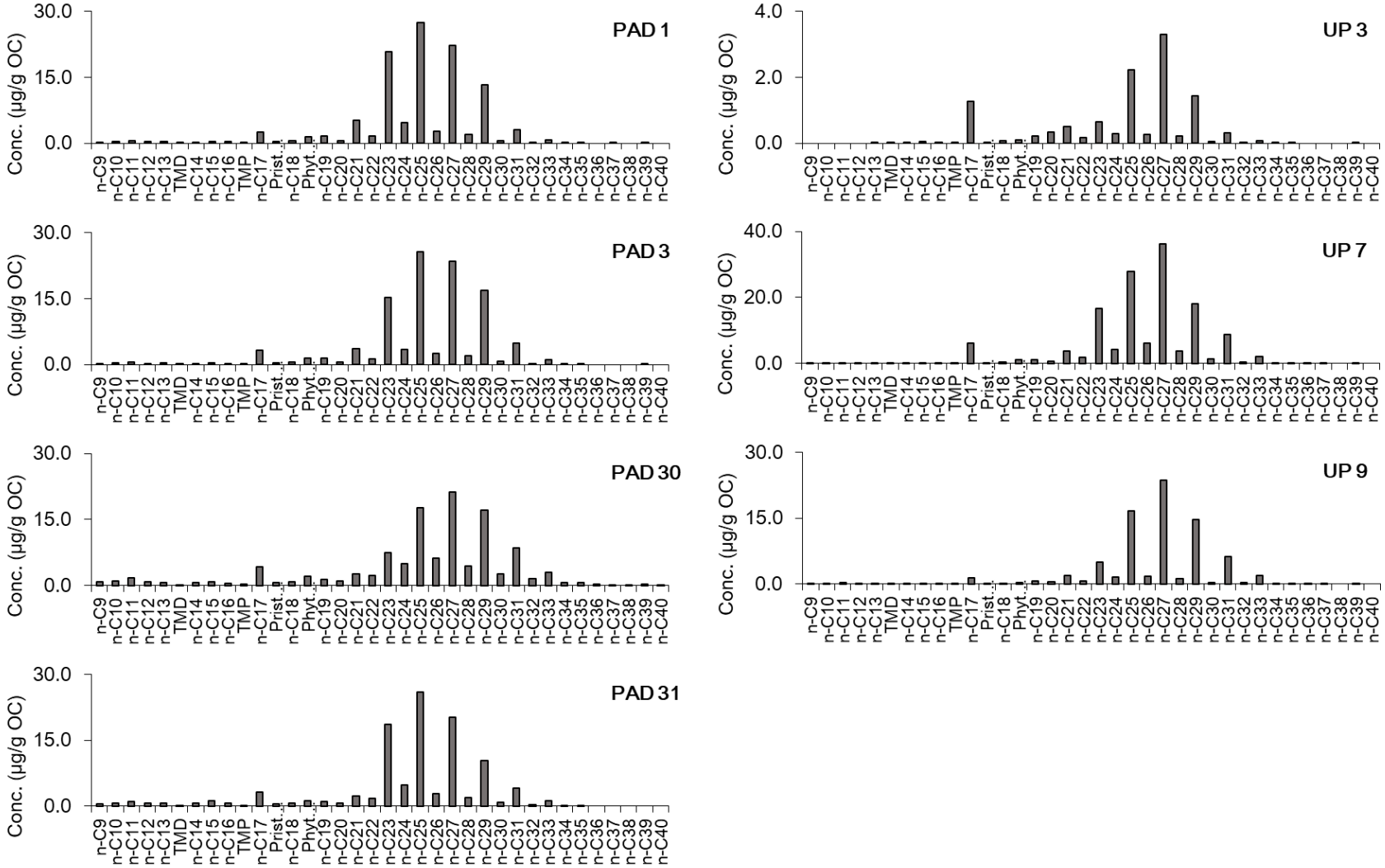

Figure S2.2. Distributions of *n*C_9-*n*C_{40} alkane concentrations including 2,6,10-trimethyldodecane (TMD), 2,6,10-trimethylpentadecane (TMP), pristane (prist), and phytane (phyt) in sediment samples. Concentrations represent the mean of triplicates from each lake (n=3 samples per lake).

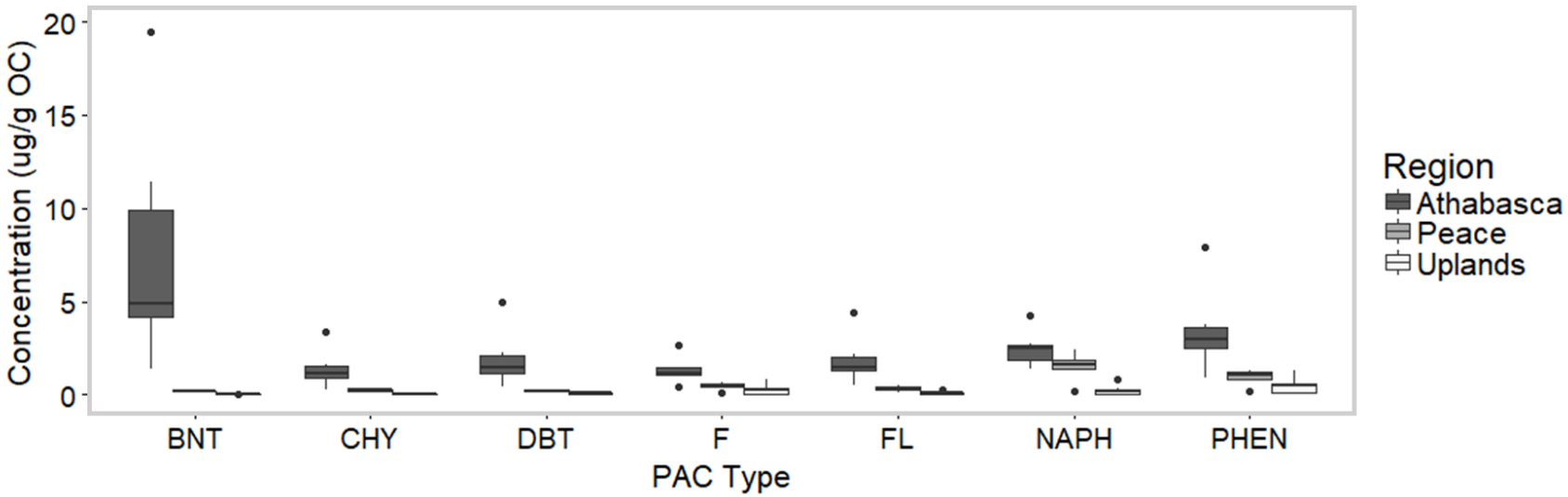

Figure S2.3. Boxplot showing the mean concentrations (in µg/g OC) of the 7 series of alkylated PAC homologs in Athabasca Delta samples (dark grey, n=6), Peace Delta samples (medium grey, n=6), and Uplands samples (light grey, n=9). PAC homologous series represented include C0-C4 benzonaphthothiophene (BNT), C0-C3 chrysene (CHY), C0-C3 dibenzothiophene (DBT), C0-C3 fluorene (F), C0-C4 fluoranthene (FL), C0-C4 naphthalene (NAPH), and C0-C4 phenanthrene (PHEN). Outliers are shown with black dots.

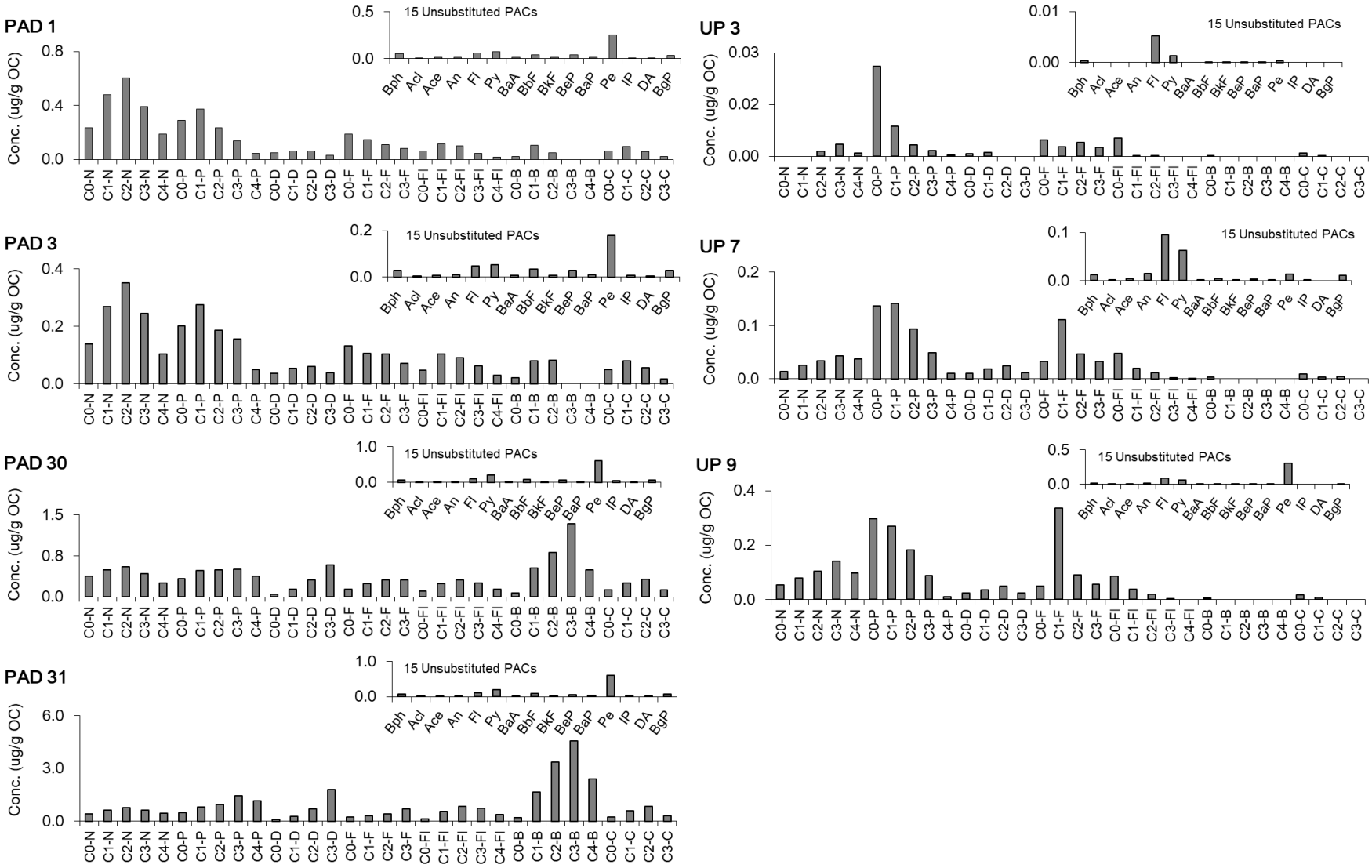

Figure S2.4. Distributions of PAC concentrations including 7 alkylated homologous series (large panels) and 15 unsubstituted PACs (upper right panels) in sediment samples. Concentrations represent the mean of triplicates from each lake in µg/g OC, n=3 samples per lake.

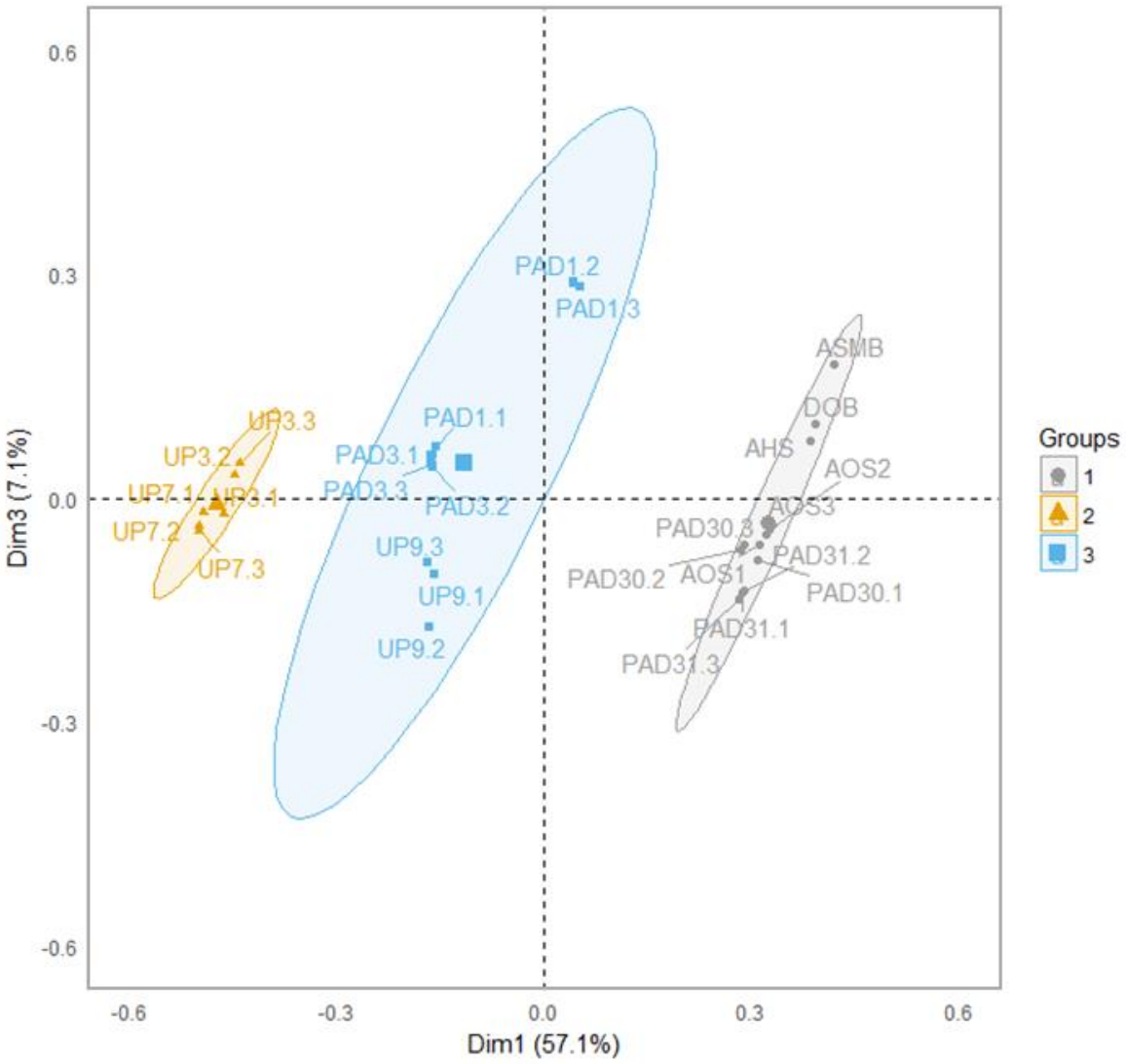

Figure S2.5. Ordination plot showing site scores for principle components ('Dim') 1 and 3 (64.2% of total variance) of biomarker data in sediments samples in 2017 and potential sources oils (AOS1, AOS2, AOS3, AHS, ASMB, and DOB), n=27 samples. 'Groups' represents clusters obtained from k-means partitioning. Confidence ellipses are drawn for the 95% confidence interval.

CHAPTER 3. METAGENOMIC ANALYSIS OF THE HYDROCARBON DEGRADATION POTENTIAL OF MICROBIAL COMMUNITIES IN LAKES OF THE PEACE-ATHABASCA DELTA, AB, CANADA

Michelle-Claire Roy[1], Joshua R. Thienpont[1], Julien Tremblay[2], Jessica Wasserscheid[2], Charles W. Greer[2], and Jules M. Blais[1]

[1]University of Ottawa, Ottawa, Ontario, Canada

[2] Energy, Mining and Environment, National Research Council Canada, Montréal, Québec, Canada

3.1 Abstract

The potential for biodegradation of oil sands-related contaminants in Canadian freshwater systems remains vastly understudied. The Peace-Athabasca Delta (PAD) is a prime location to study biodegradation as it has a long history of exposure to naturally-derived petrogenic hydrocarbons (PHCs) from bituminous deposits. Metagenomes were analyzed in delta and non-delta lakes alongside a target analysis of PHCs (including polycyclic aromatic compounds (PACs) and *n*-alkanes) in sediment samples to evaluate the impact of PHCs on microbial communities and vice-versa. Microbial communities in the PAD are highly diverse and significantly different from other lakes in the region located outside of the delta. A total of 3885 genes involved in the degradation of *n*-alkanes and PACs were identified in sediments from both regions, which were significantly correlated with sediment concentrations of short-chain alkanes, alkyl-PACs, and organic carbon, nitrogen, and sulfur. Genes for anaerobic degradation of PHCs were identified in syntrophic bacteria (Syntrophobacterales), methanogens (Methanomicrobiales), nitrate reducers (Hydrogenophilales), and sulfate reducers (Desulfobacterales and Desulfomonodales), demonstrating the potential for syntrophic degradation of PHCs. This is the first known study to demonstrate the potential for biodegradation of oil sands-related contaminants in the PAD. Further assessment of the transcriptional activity of genes highlighted in this study is warranted.

3.2 Introduction

Petrogenic hydrocarbons (PHCs) can be emitted into the environment through a variety of ways, be it by accidental release or natural seeps of oil into an ecosystem. PHCs encompass four classes of compounds derived from petroleum reservoirs including saturates, aromatics, resins, and asphaltenes (Yang et al., 2011). Biodegradation by microorganisms is the primary mechanism by which PHCs can be removed from the environment. These organisms can use various hydrocarbon compounds as a source of carbon and/or energy (Das & Chandran, 2011). Though biodegradation of PHCs has been thoroughly examined in marine environments (Leahy & Colwell, 1990), much less is known about freshwater microbial communities and their hydrocarbon-degrading mechanisms. The degradation of oil sands-related hydrocarbons may be the most important to consider in a Canadian context, as much of Canada's oil and equivalent resources are in contact with freshwater ecosystems (Kelly et al., 2009). Research on the biodegradation of PHCs in freshwater has focused on microbial communities in the Athabasca Oil Sands Region (AOSR) (An et al., 2013; Hadwin et al., 2006; Yergeau et al., 2012, 2013). Some genera identified as bioindicators of oil sands-related pollution in the AOSR include nitrogen-fixing *Azonexus* and the methanogen *Methanobrevibacter*, which were positively correlated with concentrations of oil sands-derived PHCs (Yergeau et al., 2012). Cyanobacteria on the other hand have been reported to be negatively affected by oil sands-derived PHCs when exposed to contaminated sediments (Yergeau et al., 2012, 2013). A large part of the microbial community identified via 16S rRNA in oil sands tailings was anaerobic (Yergeau et al., 2012, 2013), which may indicate the importance of anaerobic bacteria and archaea to the degradation of oil sands organic contaminants. Additionally, prior-exposure to hydrocarbon contaminants, from either anthropogenic or natural sources, has been shown to increase the hydrocarbon

degradation potential of a microbial community (Leahy & Colwell, 1990). Therefore, it is of value to investigate microbial communities in freshwater systems with a long-term exposure to oil sands-related PHCs and their ability to degrade such hydrocarbons.

The Peace-Athabasca Delta (PAD), located in northeastern Alberta, is the world's largest inland freshwater boreal delta. This valuable ecosystem is habitat for an identified 219 species of birds, 43 species of mammals, and 31 species of fish (Timoney, 2013b), and provides sustenance and spiritual well-being to First Nations and Métis people of the land. The PAD receives PHCs from a variety of sources, both natural (erosion of oil and coal sands deposits) and anthropogenic (AOS mining-related petcoke) (Conly et al., 2002; Jautzy, Ahad, Gobeil, et al., 2015; Jautzy, Ahad, Hall, et al., 2015). Polycyclic aromatic compounds (PACs) – part of the aromatic fraction of PHCs – are of particular concern in the PAD because several PACs are known to be toxic and carcinogenic (Kelly et al., 2009). These compounds have been traced from oil sands-related activities in the AOSR (Jautzy, Ahad, Gobeil, et al., 2015; Kelly et al., 2009), as well as the Athabasca and Peace Oil Sands deposits (Jautzy, Ahad, Hall, et al., 2015). Data from sediment cores suggests that petrogenic PACs are enriched in lakes connected to the PAD, and their distribution is influenced by flooding events (Hall et al., 2012; Jautzy, Ahad, Hall, et al., 2015; Timoney & Lee, 2011). However, it is unclear whether microbial communities harbor the genetic capabilities to degrade these and other PHCs.

The PAD's long history of exposure to naturally-derived petrogenic hydrocarbons, namely PACs (Hall et al., 2012), may have allowed indigenous microbial communities to develop the genetic capabilities to degrade these compounds. This makes the PAD an excellent location to study hydrocarbon degradation in a Canadian freshwater system, where we have much less knowledge compared to marine systems. The main objectives of the study were to: (1)

examine the composition of PHCs (including *n*-alkanes and PACs) in two lakes of the Athabasca Delta and two lakes of the Peace Delta, and compare these to lakes in the boreal uplands that do not receive floodwaters; (2) examine microbial community structure in delta and boreal upland lakes; (3) analyze the metagenome of these lakes and screen for hydrocarbon degradation genes; and (4) evaluate the influence of petrogenic hydrocarbons on the structure and gene composition of these microbial communities. The comparison of these three regions (Athabasca Delta, Peace Delta, and boreal uplands) allows a proper contrast of microbial communities in the PAD, which may receive a significant load of PHCs from floodwaters, to those in the nearby boreal uplands with no connection to the Peace and Athabasca rivers. These objectives were achieved using a metagenomic approach, which enabled us to recover the assemblage of genes and reconstruct taxonomic profiles directly from environmental samples. Applying this approach to the PAD is advancing the current state of knowledge on hydrocarbon degradation in this valuable ecosystem and in Canadian freshwater systems.

3.3 Materials and Methods

3.3.1 Site Selection

Sites were selected with the main goal to maximize contrast between PAD lakes subject to the influence of river floodwaters carrying PHCs and non-delta boreal upland lakes. This allowed us to compare PAD lakes to boreal upland lakes that are nearby and likely receive the same atmospheric deposition, yet do not receive floodwaters from the Peace and Athabasca rivers. We predicted that PAD lakes would consequently have higher concentrations of PHCs than boreal upland lakes. Figure 1 shows the sites that were sampled in July 2017. Delta sites were selected in the Athabasca Delta (lakes PAD 30 and 31), as well as in the Peace Delta (lakes PAD 1 and 3) for their proximity to major rivers. Lakes PAD 30 (210 m above sea level (a.s.l.))

and 31 (213 m a.s.l.) are flood-prone lakes located adjacent to Mamawi Creek. Lakes PAD 1 (212 m a.s.l.) and 3 (213 m a.s.l.) are closed-drainage lakes adjacent to the Rivière des Rochers and are less frequently flooded compared to PAD 30 and 31. Boreal upland lakes sampled to the northeast area of the region (subsequently named 'Uplands') include lakes UP 3 (256 m a.s.l.), UP 7 (254 m a.s.l.), and UP 9 (273 m a.s.l.). All three of these boreal upland lakes are closed-drainage lakes located on the Canadian Shield, underlain by bedrock. A small creek connects UP 9 to a larger lake approximately 3 metres of elevation below.

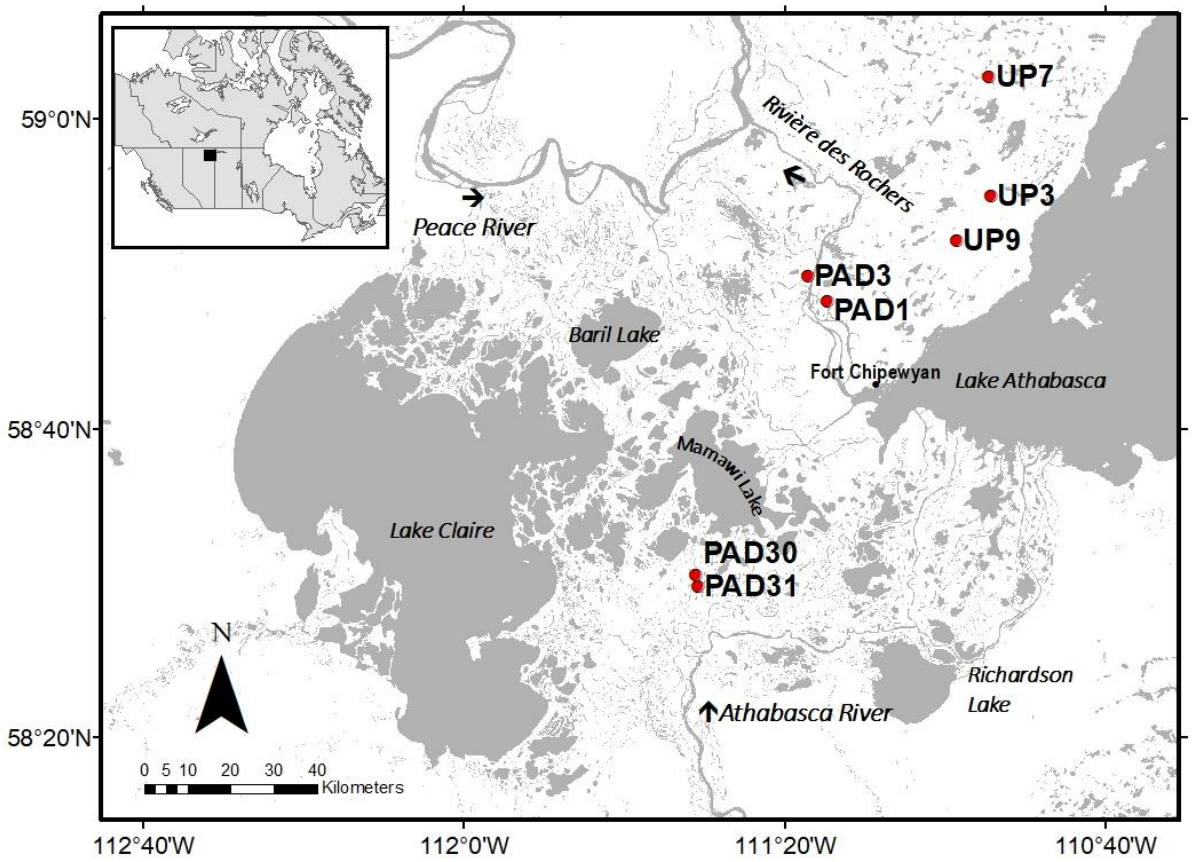

Figure 3.1. Map identifying locations of sampling sites in the Peace-Athabasca Delta: PAD 1 (58°48'N, 111°14'W), PAD 3 (58°49'N, 111°17'W), PAD 30 (58°30'N, 111°31'W), PAD 31 (58°29'N, 111°30'W), and sampling sites outside of the Delta: Uplands 3 (58°55'N, 110°54'W), Uplands 7 (59°02'N, 110°54'W), and Uplands 9 (58°52'N, 110°58'W). Also shown are the Athabasca and Peace rivers and their respective flow directions under normal, non-flood conditions.

3.3.2 Sample Collection

Sample collection was carried out from July 11th to July 13th, 2017. Surface sediment samples (uppermost 5 cm from each core) were collected from each lake in triplicate with a mini-Glew gravity sediment corer. The samples were placed into individual sterile Whirl-Pak® bags and kept on ice (4 °C) in the field. Surface sediment for each individual triplicate was thoroughly homogenized and subsampled into two 50 mL centrifuge tubes – one for DNA extraction, the other for chemical (% carbon, nitrogen, and sulfur) and PHC analysis – and frozen after subsampling. Surfaces and tools used for sediment subsampling were sterilized using 70% v/v isopropyl alcohol wipes between samples. All DNA samples (sediment and filters) were kept frozen and stored at -80 °C upon arrival at the University of Ottawa.

Surface water destined for DNA analysis was sampled in triplicate using sterile 1 L Nalgene® bottles, while water collected for water chemistry analyses was sampled using 4 L carboys at ~10 cm depth below surface. All samples were kept on ice (4 °C) in the field and frozen (-20 °C) within 5 h following subsampling or filtration. Surface water collected in 1 L Nalgene® bottles was filtered through Millipore GSWP filters (0.22 μm pore size, 47 mm diameter), and the filter was collected in sterile aluminum paper and kept frozen. Surfaces and tools used for water filtration were sterilized using 70% v/v isopropyl alcohol wipes between samples. Filtration manifolds were soaked in 6% sodium hypochlorite and rinsed with sterile Milli Q water between each filtration.

3.3.3 Sediment and Water Chemistry

For sediment samples, elemental analysis was performed for total organic carbon, nitrogen, and sulfur (% CNS) at G.G. Hatch Stable Isotope Laboratory (University of Ottawa). A

subsample of sediment was freeze-dried, then digested in an acid desiccator containing hydrochloric acid for 72 h to remove inorganic carbon, followed by three rinses using distilled water, and re-freeze-dried. Dry, digested sediments (5-10 mg) were then transferred to 5 x 8 mm tin capsules along with ~20 mg WO_3, and analyzed with VarioEL cube (Elementar Americas Inc., New York). The remaining sediment samples were used for PHC analysis (below).

Surface water samples collected in 4 L carboys were analyzed for total nitrogen (TN), total phosphorus (TP), total dissolved nitrogen (TDN), total dissolved phosphorus (TDP), combined nitrate+nitrite (NO_3+NO_2), dissolved inorganic carbon (DIC), dissolved organic carbon (DOC), major anion analysis (Cl^-, SO_4^{2-}), and major cation analysis (Ca^{2+}, Mg^{2+}, Na^+, K^+, SiO_2) at the University of Alberta's Biogeochemical Analysis Service Laboratory (BASL). For TN, TP, TDN, TDP, DSi, NO_3+NO_2, DIC, DOC, and major anions (Cl^-, SO_4^{2-}), unfiltered water was subsampled and analyzed at the BASL. For major cations, cellulose-filtered water (0.45 μm) was transferred to 15 mL centrifuge tubes, preserved using 15% nitric acid and analyzed at the BASL.

3.3.4 Extraction and Analysis of PHCs

Sediment samples destined for PHC analysis were processed at Environment and Climate Change Canada (ECCC) (River Road, Ottawa, ON). Water content was determined by placing 2-5 g wet sediment in an aluminium dish and drying at 60 °C oven overnight (~17 h). Water content was calculated from the weight difference between wet and dry samples. Approximately 5 g wet sediment were added to extraction thimbles and combined with sufficient amounts of anhydrous sodium sulphate to absorb water from the sediments. The thimbles were spiked with 100 μL of recovery surrogate (200 μg/mL ortho-terphenyl, 200 μg/mL d_{50}-tetracosane ($C_{24}D_{50}$), and 10 μg/mL of each of the five deuterated PACs (d_8-naphthalene, d_{10}-acenaphthene, d_{10}-

phenanthrene, d_{12}-benz[*a*]anthracene and d_{12}-perylene)) and Soxhlet extracted with 200 mL of 9:1 dichloromethane(DCM):acetone overnight (~17 h). Sample extracts were solvent exchanged with hexane and concentrated to ~2 mL.

Concentrated sample extracts were transferred to 3 g silica gel chromatographic columns for cleanup and fractionation as per Yang et al. (2011). Columns were conditioned with 20 mL hexane before the sample was loaded. The saturated fraction (F1) containing *n*-alkanes was eluted with 12 mL hexane and collected. The aromatic fraction (F2) containing the PACs was eluted with 15 mL of DCM:hexane and collected. The F1 and F2 fractions were evaporated under a gentle stream of nitrogen to 0.9 mL and 0.8 mL, respectively. The concentrated fractions were then spiked with 100 µL of each internal standard (IS), which included 200 µg/mL of 5-α–androstane for F1, and 200 µg/mL of 5-α–androstane and 10 µg/mL d_{14}-terphenyl for F2. The total pre-injection volume of 1 mL was analyzed by gas chromatography coupled with mass spectrometry (GC-MS). Target PHCs included 15 unsubstituted PACs, 7 homologous series of alkylated PACs (alkyl-PACs) (naphthalenes, phenanthrenes, dibenzothiophenes, fluorene, fluoranthenes, benzonaphthothiophenes, and chrysenes), and *n*C9-*n*C40 alkanes (Table S3.1).

3.3.5 DNA Extraction and Metagenomic Sequencing

DNA was extracted from sediments (2 g wet weight) and filters using DNeasy® PowerSoil® and PowerWater® kits (Qiagen, Hilden, Germany). Pre-extraction buffer solution (10 mM EDTA pH 8.0 + 50 mM Tris-HCl pH 8.0 + 50 mM sodium phosphate dibasic heptahydrate ($Na_2HPO_4 \cdot 7H_2O$) pH 8.0) was applied to remove humic substances from sediments and increase DNA yield. DNA extracts were shipped on dry ice (-80 °C) to the National Research Council Canada (NRC, Montreal) for genomic DNA (gDNA) library preparation. Extracts were treated with RNase to remove any RNA from the gDNA and purified with QIAEX II Gel Extraction Kit

(Qiagen, Hilden Germany). Libraries were prepared from 2 ng of gDNA with Nextera® XT DNA Library Prep kit (Illumina, San Diego, California) with Nextera® Index Primers S502, S503, S505-507, N701-707, N711, N712, N714, and N715 (Illumina, San Diego, California). A total of 45 samples (24 from water, 21 from sediments) were submitted to Génome Québec (Montreal, Quebec) for library QC for paired-end sequencing on an Illumina HiSeq 2500 PE (2 x 125 bp).

3.3.6 Genomics and Bioinformatics

Sequencing results were processed using NRC's shotgun metagenomics pipeline (Fortin et al., 2017). In brief, raw reads were quality controlled and co-assembled with Megahit (PMID: 27012178) software v1.1.2 on a 3 Terabyte RAM compute node using iterative kmer lengths of 21, 31, 41, 51, 61, 71, 81 and 91 bp (N50 >= 902 bp, N90 >= 547 bp, GC = 52.45%) with assembly statistics and read count statistics available in supplemental table S3.2 and S3.3 respectively. Gene prediction on each contig was performed with Prodigal v2.6.2, followed by gene annotation following guidelines set by the Joint Genome Institute (Huntemann et al., 2016). Taxonomy and full lineage information were assigned to contigs with the NCBI taxonomy database. Reads of each sequencing library (*i.e.*, sample) were mapped to the contigs of the metagenome assembly (bwa mem v0.7.15) to obtain contig coverage values for each sample (samtools). Gene coverage was computed from contig coverage (*i.e.*, bam files) using bedtools (v2.17.0) with the gene prediction coordinates given by Prodigal. Gene coverages of each sample were merged into one file, which provided an abundance matrix of read counts and further normalized (edgeR) to obtain a count per million (CPM) matrix for each gene of each sample. This matrix and associated taxonomy were used for degradation gene analyses. A total of 220 metagenomic bins were generated using MetaBAT v2.12.1, decontaminated using a majority rule-based decision scheme (Fortin et al., 2017), and incorporated into an abundance matrix. This

metagenome bin abundance matrix was used for downstream community structure and taxonomic analyses.

3.3.7 Statistical Analyses

All statistical analyses were performed using R software (v3.3.2). Statistical significance and differences between chemical data for each replicate was analyzed using the "Kruskal.test" function when assumptions of the t-test could not be met (normality, sample size, or equality of variance). Analysis of alpha and beta diversity were performed using the 'vegan' package. The alpha diversity index was calculated with the "diversity" function (Shannon's H diversity index). Analysis of beta diversity was performed by calculating the ecological distance between sites based on Bray-Curtis dissimilarity of bin abundances ("vegdist" function). Permutational analysis of variance (perMANOVA) was then performed on this distance matrix to test if the composition of binned species in the two sectors (PAD versus boreal upland) were significantly different from each other ("adonis" function). Multivariate dispersion was subsequently calculated using the "betadisper" function and mean distance to group centroid calculated using the "permutest" function. Unconstrained ordination using non-metric multidimensional scaling (NMDS) was performed with the "metaMDS" function to represent the original position of communities in multidimensional space. Confidence ellipses on NMDS plots were generated for each region to serve as visual indicators of correlation (95% confidence interval calculated from the binned species abundance in samples within a region) between sites using the "stat_ellipse" attribute of the 'ggplot2' package in R.

Redundancy analysis (RDA) was performed on degradation genes (normalized CPM) and chemical variables using the "rda" function. Selection of important chemical predictors of change in community matrices was calculated using the "envfit" and "anova" functions. Once

the chemical predictors were selected, variation partitioning was performed using the "varpart" function, effectively splitting the chemical variables into two parts to be considered independently (sediment or water chemistry versus PAC or *n*-alkane concentration). The "anova" function was then used to test the significance of variation in partitions.

3.3.8 Availability of Data

Raw sequence reads of the shotgun metagenomic data were submitted to the NCBI's Sequence Read Archive under accession no. SRP151914 under Bio Project PRJNA479434.

3.4 Results and Discussion

3.4.1 Chemical Analyses and PHCs

Water chemistry analysis (Table S3.4) demonstrated that northern Peace sector lakes PAD 1 and 3 have the highest total nitrogen content (2,210 and 2,210 µg/L, respectively) and highest pH (10.50 and 10.19, respectively). It should be noted that the pH for these lakes lies outside of the optimal range for hydrocarbon degradation (pH 6-9) (Das & Chandran, 2011) and may in fact hinder degradation. Southern Athabasca sector lakes PAD 30 and 31 have noticeably higher sulfate concentrations (14.36 and 27.01 mg/L, respectively). Overall, lakes in the PAD have elevated concentrations of major cations and anions, total phosphorus, and dissolved inorganic carbon compared to lakes in the uplands sector (Kruskal-Wallis: $\chi^2 = 4.5$, $p < 0.05$). This may be expected as lakes in the PAD are shallow, highly productive lakes that receive large quantities of organic matter and inorganic nutrients from the Athabasca and Peace rivers. In contrast, upland lakes are closed-drainage with small catchments, and lie on the Canadian Shield, underlain by more slowly weathering granitic bedrock. Sediment chemical data showed that concentrations of

Table 3.1. Sediment chemical data for lakes sampled in July 2017. PAC and *n*-alkane concentrations are normalized to organic carbon content. Numbers represent the mean of triplicate samples for each lake such that n=3 samples per lake.

Parameters	PAD 1	PAD 3	PAD 30	PAD 31	UP 3	UP 7	UP 9
Elemental Analysis							
Organic Carbon (%)	11.67	13.17	5.70	3.10	14.13	35.10	9.24
Organic Nitrogen (%)	1.19	1.22	0.50	0.21	1.41	3.14	0.30
Organic Sulfur (%)	0.84	1.01	0.15	0.10	0.09	0.18	0.13
Petrogenic Hydrocarbons (µg/g OC)							
Total PACs	5.06	3.77	12.89	31.61	0.08	1.1	2.69
Total alkylated PACs	4.44	3.3	11.51	29.03	0.07	0.98	2.18
Total unsubstituted PACs	0.62	0.47	1.38	2.58	0.01	0.12	0.51
Total *n*-Alkanes	115.51	111.72	118.9	109.4	11.6	141.53	81.11
Total C_9:C_{16} *n*-Alkanes	3.35	2.85	7.48	6.48	0.1	0.56	1.21
Total C_{17}:C_{40} *n*-Alkanes	112.16	108.87	111.42	102.92	11.5	140.97	79.9

short-chain (C_9-C_{16}) *n*-alkanes were significantly different between Peace Delta, Athabasca Delta, and boreal upland lakes (Kruskal-Wallis: $\chi^2 = 13.492$, $p < 0.05$) (Table 3.1). There is no significant difference between concentrations of long-chain (C_{17}-C_{40}) *n*-alkanes in samples from PAD and boreal upland sectors (Kruskal-Wallis: $\chi^2 = 2.85$, $p = 0.24$). The high carbon preference index (CPI) (Wang et al., 2014) for the entire range of *n*-alkanes in all samples (3.1-10.0) as well as the predominance of long-chain alkanes (Table 3.1) indicates a biogenic source of *n*-alkanes in all lakes samples. Athabasca Delta lake sediments also had the highest total PACs, followed by Peace Delta lakes, and lowest PAC concentrations were found in boreal upland lakes. The

concentrations were significantly different between sediments from these three regions for total PACs (Kruskal-Wallis: $\chi^2 = 15.1$, $p < 0.001$), alkyl-PACs (Kruskal-Wallis: $\chi^2 = 15.1$, $p < 0.001$), and the 15 unsubstituted PACs (Kruskal-Wallis: $\chi^2 = 14.23$, $p < 0.001$).

Alkyl-PACs dominate the PAC fraction in sediments from all lakes (81.0-91.8% of total) (Table 3.1). Lakes in the southern Athabasca sector (PAD 30 and 31) had the highest concentrations of alkyl-PACs and included petrogenic compounds such as C0-C3 chrysenes, C0-C3 dibenzothiophenes, and C0-C4 benzonaphthothiophenes (Table S3.5). These lakes receive floodwaters derived from the Athabasca River which is most directly connected to the AOSR and the McMurray formation (Das & Chandran, 2011). The presence of these high molecular weight alkyl-PACs sourced from oil sands deposits in lake PAD 31 has been previously reported (Kelly et al., 2009), yet this is the first time benzonaphthothiophenes (BNTs) have been analyzed and identified in the PAD. BNTs are a type of polycyclic aromatic sulfur heterocycle (PASH) commonly found in crude oil and bitumen (Michel & Hayes, 1999). Although these compounds are thought to arise solely from petroleum reservoirs, BNTs and their methylated counterparts can be formed through the microbial degradation of benzo[*b*]thiophenes via the 4S desulfurization pathway (Mcfall et al., 1984).

Alkyl-PACs were also identified in uplands lake sediments (Table S3.5). Low molecular weight compounds make up the largest proportion of alkyl-PACs in uplands lake sediments, including C0-C1 phenanthrenes, C1 fluorenes, and C0-C4 naphthalenes. These compounds are typically associated with wood combustion (Wang et al., 2014). Therefore, their presence in uplands lakes at low concentrations can likely be explained by atmospheric deposition of PACs sourced from the intense forest fire seasons in 2016 and 2017. Interestingly, traces of C0-C3 dibenzothiophenes and C0-C3 chrysenes typically associated with petrogenic hydrocarbon

sources were also identified at low concentrations in UP 7 (mean of 0.06 and 0.02 μg/g OC, respectively) and UP 9 (mean of 0.13 and 0.02 μg/g OC, respectively) surface sediments, suggesting atmospheric deposition of these compounds has occurred recently.

The degree of alkylation (R-group substitutions) of PACs is higher in PAD samples than in the boreal uplands. PAD sediments sampled in the Athabasca sector (PAD 30 and 31) had high concentrations of 2 to 4 R-group substituted PACs, while sediments sampled from the Peace sector (PAD 1 and 3) had high concentrations of 1 to 3 R-group substituted PACs, followed by uplands sediments which had alkyl-PACs with 1 to 2 R-group substitutions. This has implications for biodegradation since alkyl-substitution may limit the accessibility of dioxygenase enzymes and require additional steps for degradation (Seo et al., 2009). Degradation rates of PACs have been shown to decrease with increasing degree of alkylation (Deshpande et al., 2018; Novaković et al., 2012). However, biodegradation of alkylated phenanthrene, dibenzothiophene, and naphthalene has been reported (Budzinski et al., 1998; Deshpande et al., 2018), and may occur through a variety of different pathways (Mahajan, Phale, & Vaidyanathan, 1994). Overall, the composition of PHCs is distinct in lake sediments of the Athabasca Delta, the Peace Delta, and boreal uplands lakes. Flood-prone lakes of the Athabasca Delta have the highest concentration of highly alkyl-PACs and short-chain alkanes, followed by closed-drainage lakes of the Peace Delta, and significantly lower concentrations of PACs and short-chain alkanes in boreal upland lakes.

3.4.2 Microbial Community Profiles and Diversity

Despite the presence of the aforementioned PHCs, sediment microbial communities in both PAD and boreal upland lake systems are highly diverse (Shannon's H = 4.27-4.58) with no significant difference in the diversity indices. Betaproteobacteria are the dominant taxa in all

sediment samples, followed by Opitutae, Alphaproteobacteria, and Deltaproteobacteria (Figure 2). The class Betaproteobacteria has been identified as a contributor to the degradation of PACs (Martin et al., 2012), including the genus *Thiobacillus* and members of the family Rhodocyclaceae, which were identified in all lakes sampled. Slightly higher abundances of contigs belonging to Methanomicrobia were identified in sediments from the upland lakes, whereas class Dehalococcoidia was prevalent in lakes PAD 1 and 3.

Water microbial communities were more diverse in PAD samples (Shannon's H = 3.49-3.99) than in upland samples (Shannon's H = 2.67-3.90) and the microbial diversity was significantly different in these two groups (ANOVA: $F = 6.994$, $p < 0.05$). Alpha- and Beta-proteobacteria were the dominant classes in water samples from all lakes (Figure 2). Slightly higher proportions of Cyanobacteria were present in water from upland lakes 3 and 7, while classes Flavobacteriia and Actinobacteria were slightly more abundant in all PAD lake water samples. Cyanobacteria are negatively influenced by PHCs including saturates and aromatics (Yergeau et al., 2012, 2013), which may explain their low abundance in PAD lake water and sediment, despite favourable conditions in water chemistry.

A clear separation in microbial community structure between PAD and boreal upland lakes is apparent with NMDS ordination of metagenomic bins (Figures 3). The Athabasca Delta, Peace Delta, and boreal upland lakes differ significantly in their species composition for both sediment and water (perMANOVA: $F = 10.566$, $p = 0.001$ and $F = 9.707$, $p = 0.001$, respectively). Key species responsible for this separation in sediment community composition include *Thiobacillus denitrificans* ($r^2 = 0.71$), *Dechloromonas aromatica* ($r^2 = 0.66$), *Methanoregula formicica* ($r^2 = 0.66$), and *Geobacter sulfurreducens* ($r^2 = 0.70$). Site scores for the Athabasca Delta (grey) and Peace Delta (yellow) and their respective confidence ellipses also

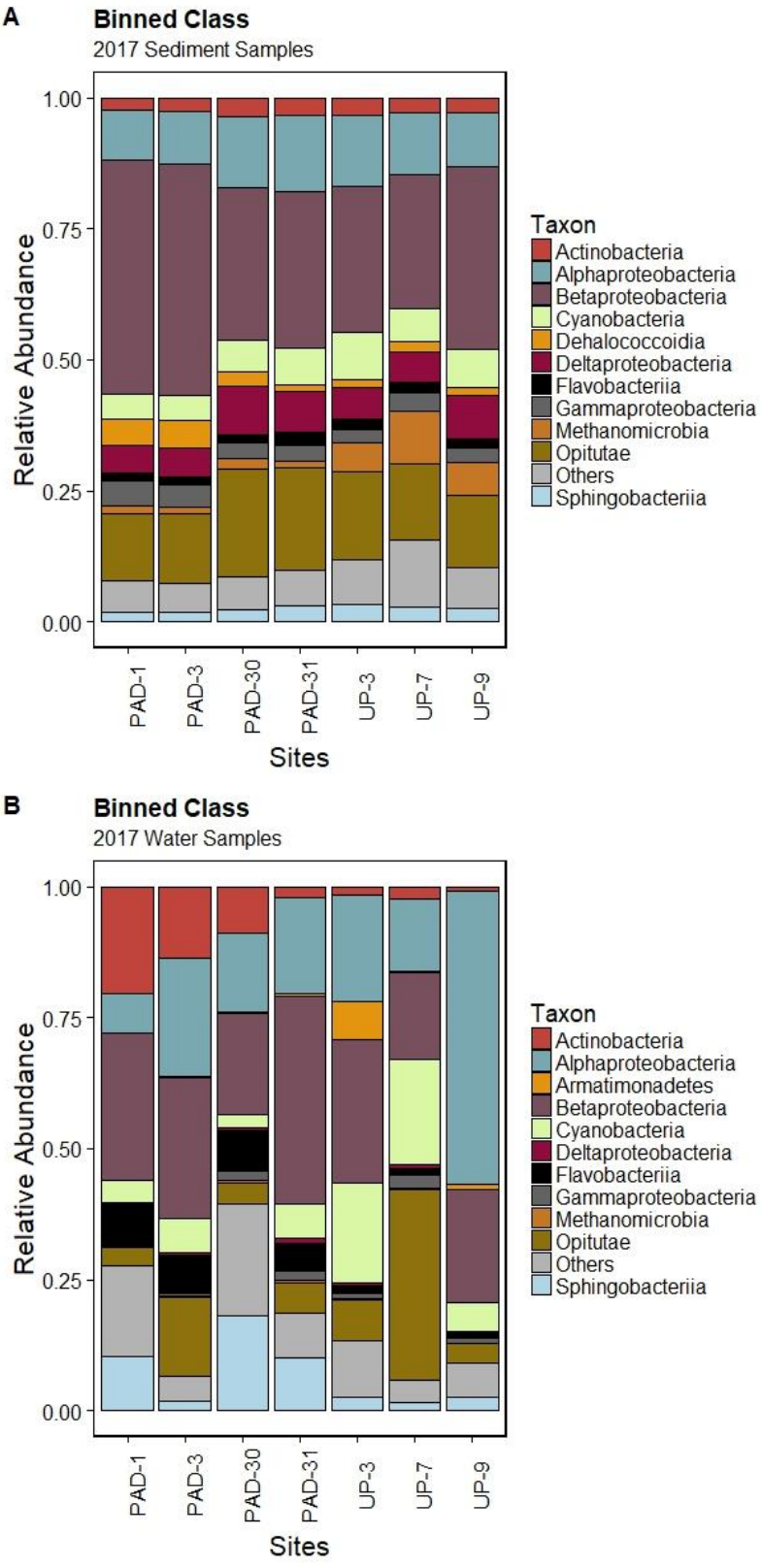

Figure 3.2. Relative abundance of binned contiguous metagenomic sequences at the class level for bacteria and archaea. A) Taxonomic assemblage in sediment samples. 'Others' include *Euryarchaota, Acetothermia, Armatimonadetes, Bacteroidia, Gemmatimonadetes, Planctomycetia, Cytophagia,* as well as bacterial gene sequences that could not be affiliated to a taxon. B) Taxonomic assemblage in water samples. 'Others' include *Euryarchaota, Acetothermia, Armatimonadetes, Bacteroidia, Gemmatimonadetes, Planctomycetia, Cytophagia, Dehalococcoidia* as well as bacterial gene sequences that could not be affiliated to a taxon.

plot in separate locations of the graph, demonstrating that microbial communities in these two regions of the delta are distinct from each other. Metagenomic bins for anaerobes and facultative anaerobes such as *Geobacter anodireducens, Dechloromonas aromatica, Opitutus terrae,* as well as denitrifying bacteria *Thiobacillus denitrificans* and *Steroidobacter denitrificans* contribute to

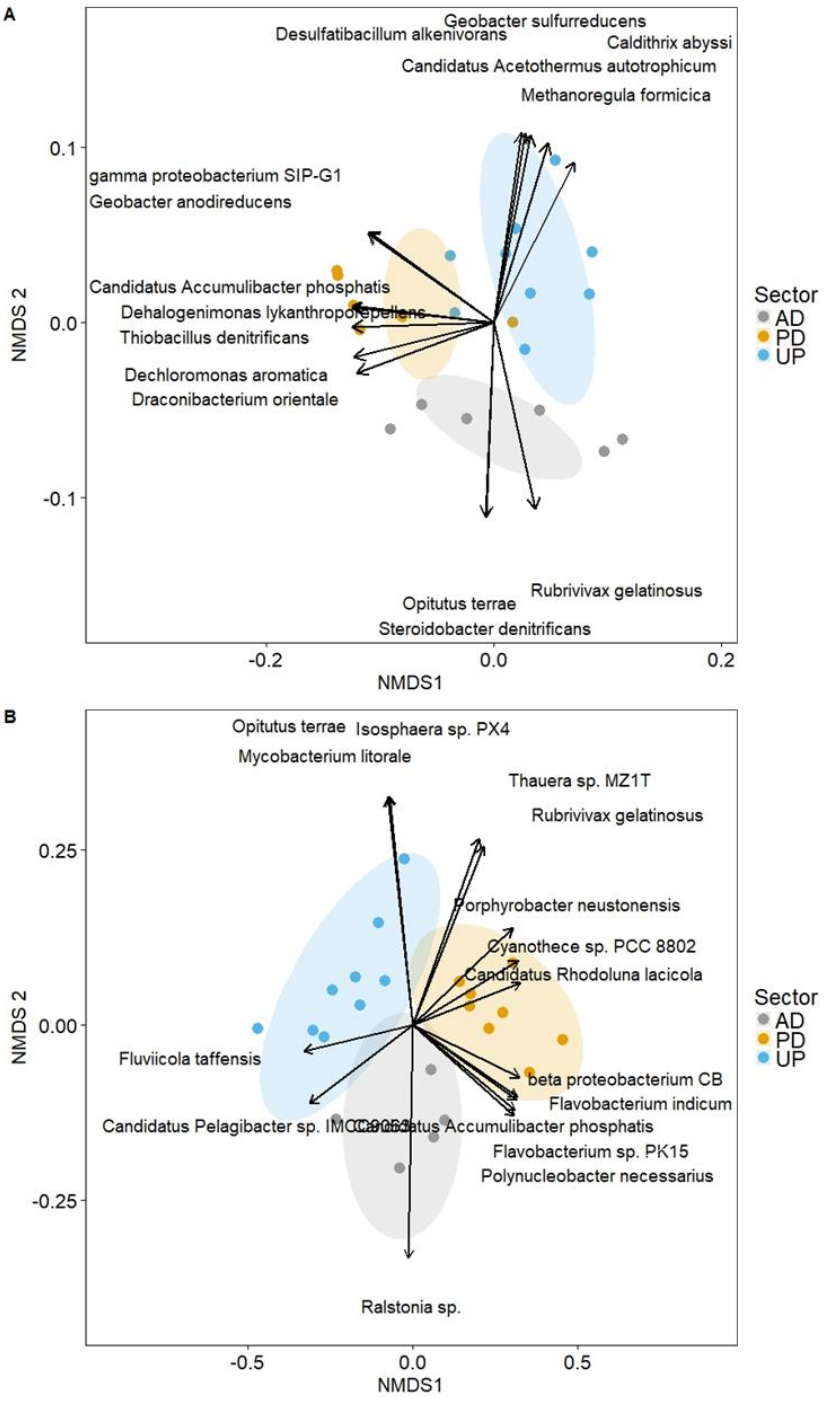

Figure 3.3. Non-metric multidimensional scaling (NMDS) plot showing the ecological distance of the binned species in: A) sediment samples (perMANOVA: n=21, F=10.566, p=0.001), and B) water samples (perMANOVA: n=21, F=9.707, p=0.001). Bray-Curtis dissimilarity index (abundance weighted) was used to calculate the ecological resemblance between samples. Stress (similarity of observed distance to ordination distance) = 0.05771 for panel A and 0.07960 for panel B, confirming acceptable fit for both ordinations. Binned species vectors in sediment are shown for bins that fit to the ordination space with a p-value = 0.001 and $r^2 >$ 0.65. Binned species vectors are shown for bins that fit to the ordination space with a p-value = 0.001 and r2 > 0.35. Confidence ellipses calculated for the Peace Delta (PD), Athabasca Delta (AD) and boreal uplands (UP), assuming a multivariate 't' distribution.

their separation. The ecological separation between microbial communities in PAD and boreal upland water was attributable to *Rubrivirax gelatinosus* ($r^2 = 0.75$), *Thauera* sp. MZ1T ($r^2 = 0.74$), *Flavobacterium indicum* ($r^2 = 0.60$), *Cyanothece* sp. PCC 8802 ($r^2 = 0.60$), and *Polynucleobacter necessaries* ($r^2 = 0.50$), among others. Site loadings and confidence ellipses for water samples plot closer together and are less distinct from each other than in sediments. However, site scores from the Peace Delta, the Athabasca Delta, and boreal upland surface water do plot in separate locations of the NDMS ordination, demonstrating the ecological distance between microbial communities in these three regions. As demonstrated in water and sediment chemistry, differences in concentrations of essential nutrients (nitrogen, phosphorus, and carbon), physicochemical parameters (*i.e.* pH), and PHCs may contribute to the ecological distance between Peace and Athabasca Delta microbial composition.

3.4.3 Hydrocarbon Degradation Gene Analysis

A total of 3885 genes involved in the degradation of short-chain alkanes, PACs, and thiophene-containing PACs, were identified in the dataset based on the 69 gene queries searched using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Table S3.6). Due in part to the semi-quantitative nature of metagenomic data, there is no significant difference in the relative abundance of these genes in PAD versus upland samples. However, the percent composition of degradation genes is different between water and sediment samples (Figure 4). Genes involved in the aerobic degradation of PACs and *n*-alkanes dominate water and sediment samples (91.1% and 76.3%, respectively). The prevalence of genes involved in the mono- and di-hydroxylation pathway of aerobic PAC degradation may be influenced by the higher number of genes entered in the query search in this category (Table S3.6). However, it is important to note that these genes are present alongside significant concentrations of a variety of PAC congeners,

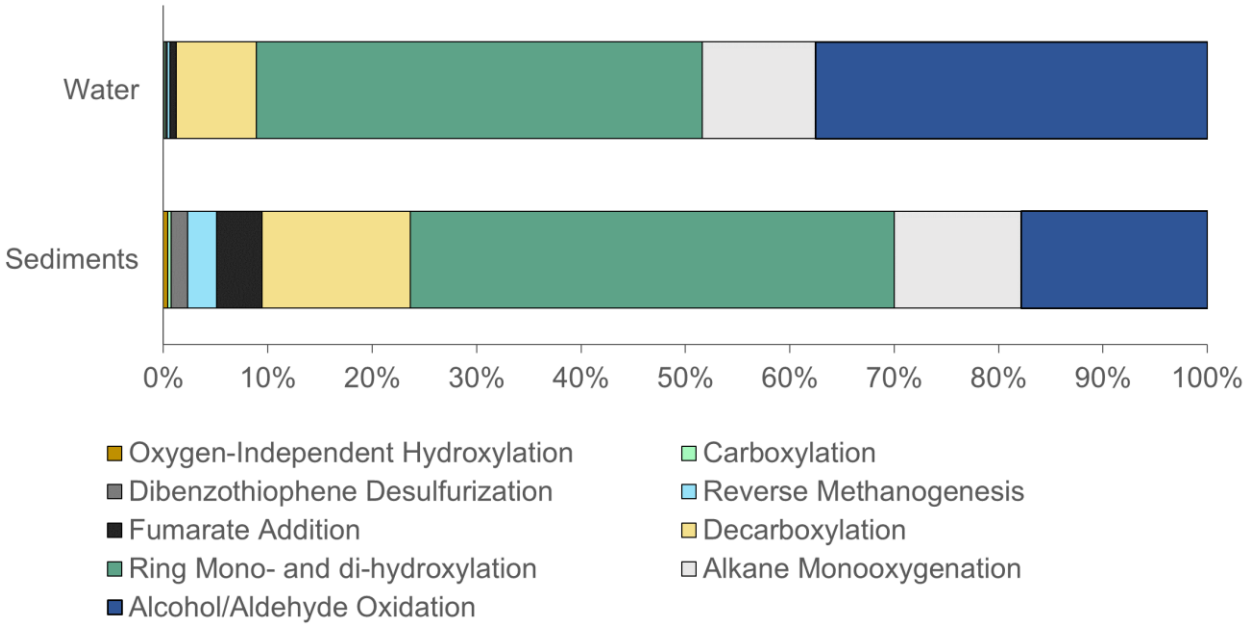

Figure 3.4. Percent composition of PAC and alkane degradation genes in 2017 water (n=21) and sediment (n=21) samples . Genes categories that function aerobically include alkane monooxygenation, alcohol/aldehyde oxidation, and ring mono and di-hydroxylation. Gene categories that function anaerobically include fumarate addition and decarboxylation, carboxylation, oxygen-independent hydroxylation, dibenzothiophene desulfurization, and reverse methanogenesis.

which demonstrates the degradation potential in these sediments. Of the 117 genera containing these genes, Betaproteobacteria are the dominant class, followed by Alphaproteobacteria and Gammaproteobacteria.

Anaerobic degradation may play an important role in the degradation of PHCs in PAD and upland lake sediments. The percent composition of degradation genes in water and sediment showed that the relative abundance of genes involved in anaerobic degradation of PACs and *n*-alkanes (fumarate addition, decarboxylation, oxygen-independent hydroxylation, carboxylation, and reverse methanogenesis) is higher in sediment samples than in water samples (23% and 4%, respectively). Genes coding for fumarate addition and decarboxylation (benzyl succinate

synthase, benzylsuccinate CoA-transferase, benzylsuccinyl-CoA, benzylsuccinyl-CoA dehydrogenase, phenylitaconyl-CoA hydratase, 4-hydroxy-3-polyprenylbenzoate decarboxylase, flavin phenyltransferase) dominate the set of anaerobic degradation genes in both sediment and water samples (18.6% and 8.3%, respectively). Hydrocarbon activation via fumarate addition has been shown with toluene, xylene, ethylbenzene, methylnaphthalene, alkanes, and alicyclic alkanes (Abbasian & Lockington, 2016). This reaction can be performed anaerobically by denitrifying, sulphate-reducing, methanogenic, and metal-reducing bacteria (Sierra-Garcia & de Oliveira, 2013) all of which are present in the dataset, especially sediment samples. A total of 23 genera contained genes for fumarate addition including *Geobacter, Desulfobacula,* and *Methylomirabilis.* Reverse methanogenesis (8.1% and 1.0% in sediment and water, respectively), a process where hydrocarbon compounds such as aliphatics and aromatics are converted to methane by methanogenic consortia, has been demonstrated on hexadecane, ethylbenzene and naphthalene (Siegert et al., 2011). A total of 7 archaeal genera contained genes for reverse methanogenesis including *Methanoregula, Methanosaeta, Methanobacterium,* and *Methanocella.* One gene coding for the desulfurization of thiophene compounds (*e.g.* dibenzothiophene) was also identified in sediment and water samples. 2'-hydroxybiphenyl-2-sulfinate desulfinase (*dszB*), an enzyme involved in the intermediate step of dibenzothiophene degradation responsible for the release of sulfur from 2-(2-hydroxyphenyl)-benzene sulfonate (Abbasian & Lockington, 2016), accounted for 1.6% of degradation genes in sediment and 0.1% of degradation genes in water. The following genera contained *dszB*: *Methanosarcina, Brucella, Ralstonia,* and *Nitrosomonas.* In addition to dibenzothiophene desulfurization, the 4S desulfurization pathway is also used by bacteria to degrade benzo[*b*]thiophene leading to the formation of the aforementioned BNTs. The dominance of BNTs in lakes flooded by the

Athabasca Rivers (14% and 38% of total PAC concentration in PAD 30 and 31, respectively) together with the high concentrations of SO_4^{2-} (14.36 and 27.01 mg/L, respectively) provides evidence that these BNTs may be formed in part by bacteria via the 4S desulfurization pathway. Further investigation on the potential formation of BNTs in the Athabasca Delta watershed is warranted.

The composition of hydrocarbon degradation genes in lake sediments was influenced by the chemical nature of the sediments and the concentration of PHCs. PAC degradation genes correlated significantly with total PAC (envfit: $r^2 = 0.6076$, $p < 0.001$), alkyl-PAC concentration (envfit: $r^2 = 0.5698$, $p < 0.001$), percent organic carbon (envfit: $r^2 = 0.4462$, $p < 0.01$), nitrogen (envfit: $r^2 = 0.4716$, $p < 0.01$), and sulfur (envfit: $r^2 = 0.8860$, $p < 0.001$) (Figure 5). *n*-Alkane degradation genes varied significantly with short-chain (C_9-C_{16}) alkane concentrations (envfit: $r^2 = 0.5943$, $p < 0.01$), carbon (envfit: $r^2 = 0.4118$, $p < 0.05$), nitrogen (envfit: $r^2 = 0.4378$, $p < 0.01$), and sulfur (envfit: $r^2 = 0.9087$, $p < 0.001$). When partitioning the variance of these environmental variables into two parts (i.e chemical nature, %CNS versus PHC, alkyl-PAC, total PAC, and short-chain alkane concentrations), %CNS better explained the variance in both PAC (ANOVA: $F = 1.7177$, $p < 0.001$) and *n*-alkane (ANOVA: $F = 1.7747$, $p < 0.01$) degradation genes across samples than PHC concentrations. The pattern of separation in the RDA between Athabasca Delta, Peace Delta, and boreal upland samples based on degradation genes is similar to the pattern shown in Figure 3 based on ecological distance of binned species. The composition of degradation genes for PACs and *n*-alkanes in PAD 30 and 31 plotted in a unique location relative to those of lakes PAD 1 and 3. This further supports that microbial communities in lakes of the Athabasca Delta are distinct from those in the Peace Delta and may harbor degradation genes that reflect the composition of PACs and *n*-alkanes in these sediments. Further research

should address whether the microbial community is actively taking part in the degradation of alkyl-PACs and short-chain alkanes in PAD sediments.

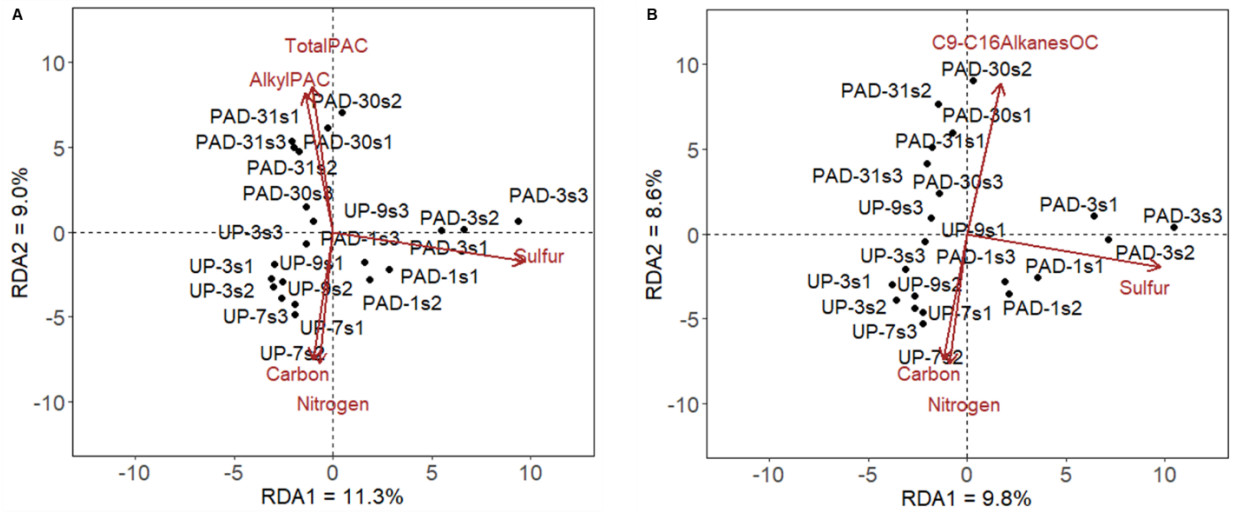

Figure 3.5. Redundancy analysis (RDA) of degradation genes for A) PACs and B) *n*-alkanes in sediment samples (scores represented, n=21) with vectors showing chemical variables that vary significantly alongside degradation genes. Total variance explained in the ordination space is 20.3% for PACs and 18.4% for *n*-alkanes. Vectors are significantly related to degradation genes in both PAC (ANOVA test: F=1.5246, p<0.001) and *n*-alkane (ANOVA test: F=1.4856, p<0.001) RDAs.

3.4.4 Genetic Potential for Syntrophic Degradation of PHCs

Lakes of the PAD sampled in this study support highly diverse microbial communities. These communities also harbour a diverse consortium of genes related to the degradation of PHCs shared amongst hundreds of species of bacteria and archaea. This illustrates that there may be multiple groups of microorganisms working as a consortium to degrade hydrocarbons. The results provide evidence of syntrophic degradation, a mode of anaerobic degradation in which multiple groups of microorganisms including methanogens, nitrate and sulfate reducers actively degrade hydrocarbon compounds together (Gieg, Fowler, & Berdugo-Clavijo, 2014). Syntrophic

oxidation was demonstrated in the anaerobic degradation of long-chain *n*-alkanes in oil sands tailings (Siddique, Penner, Semple, & Foght, 2011).

Genes for anaerobic hydrocarbon degradation were identified in syntrophic bacteria of the order Syntrophobacterales in addition to methanogens of the order Methanomicrobiales, nitrate reducers of the order Hydrogenophilales, and sulfate reducers of the orders Desulfobacterales and Desulfomonodales. These bacteria and archaea were present in Peace Delta, Athabasca Delta, and boreal uplands sediments, yet their composition was significantly different (perMANOVA: $F = 10.566$, $p < 0.001$). Chemical variables in sediment that significantly correlated to this difference were percent carbon (envfit: $r^2 = 0.4118$, $p < 0.01$), percent nitrogen (envfit: $r^2 = 0.4378$, $p < 0.01$), percent sulfur (envfit: $r^2 = 0.9087$, $p < 0.001$), short-chain alkanes concentration (envfit: $r^2 = 0.5943$, $p < 0.01$), alkyl-PAC concentration (envfit: $r^2 = 0.4525$, $p < 0.01$), unsubstituted PAC concentration (envfit: $r^2 = 0.4874$, $p < 0.05$), and total PAC concentration (envfit: $r^2 = 0.5943$, $p < 0.01$).

The presence of methanogens and genes involved in anaerobic degradation of PHCs in boreal upland lakes is likely influenced by high concentrations of long-chain alkanes (Gieg et al., 2014) from biogenic sources as well as higher carbon and nitrogen content (Figure S1). On the other hand, higher concentrations of PACs, short-chain alkanes, and higher sulfur content (especially in lakes PAD 1 and 3) may influence the presence of methanogens, nitrate and sulfate reducers with the genetic potential to degrade these compounds. These findings demonstrate the potential for syntrophic degradation in both systems and provide insight into how hydrocarbon degradation may occur in freshwater systems. Additional research is necessary to determine whether these genes are active in the degradation of PHCs identified in this study. As such,

freshwater systems like the PAD may be favourable environments for the biodegradation of PHCs via syntrophic degradation.

3.5 References

Abbasian, F., & Lockington, R. (2016). A Review on the Genetics of Aliphatic and Aromatic Hydrocarbon Degradation, 224–250. http://doi.org/10.1007/s12010-015-1881-y

An, D., Brown, D., Chatterjee, I., Dong, X., Ramos-Padron, E., Wilson, S., … Voordouw, G. (2013). Microbial community and potential functional gene diversity involved in anaerobic hydrocarbon degradation and methanogenesis in an oil sands tailings pond. Genome, 56(10), 612–618. http://doi.org/10.1139/gen-2013-0083

Budzinski, H., Raymond, N., Nadalig, T., Gilewicz, M., Garrigues, P., Bertrand, J. C., & Caumette, P. (1998). Aerobic biodegradation of alkylated aromatic hydrocarbons by a bacterial community. Organic Geochemistry, 28(5), 337–348. http://doi.org/10.1016/S0146-6380(98)00002-3

Conly, F. M., Crosley, R. W., & Headley, J. (2002). Characterizing sediment sources and natural hydrocarbon inputs in the lower Athabasca River, Canada. Journal of Environmental Engineering and Science, 1, 187–199. http://doi.org/10.1139/s02-013

Das, N., & Chandran, P. (2011). Microbial degradation of petroleum hydrocarbon contaminants: an overview. Biotechnology Research International, 2011, 941810. http://doi.org/10.4061/2011/941810

Deshpande, R. S., Sundaravadivelu, D., Techtmann, S., Conmy, R. N., Santo Domingo, J. W., & Campo, P. (2018). Microbial degradation of Cold Lake Blend and Western Canadian select dilbits by freshwater enrichments. Journal of Hazardous Materials, 352(July 2017), 111–120. http://doi.org/10.1016/j.jhazmat.2018.03.030

Fortin, N., Tremblay, J., Yergeau, E., Fortin, N., Cobanli, S., & Elias, M. (2017). Chemical dispersants enhance the activity of oil- and gas condensate-degrading marine bacteria Chemical dispersants enhance the activity of oil- and gas condensate-degrading marine bacteria, (August). http://doi.org/10.1038/ismej.2017.129

Gieg, L. M., Fowler, S. J., & Berdugo-Clavijo, C. (2014). Syntrophic biodegradation of hydrocarbon contaminants. Current Opinion in Biotechnology, 27, 21–29. http://doi.org/10.1016/j.copbio.2013.09.002

Hadwin, A. K. M., Del Rio, L. F., Pinto, L. J., Painter, M., Routledge, R., & Moore, M. M. (2006). Microbial communities in wetlands of the Athabasca oil sands: Genetic and metabolic characterization. FEMS Microbiology Ecology, 55(1), 68–78. http://doi.org/10.1111/j.1574-6941.2005.00009.x

Hall, R. I., Wolfe, B. B., Wiklund, J. A., Edwards, T. W. D., Farwell, A. J., & Dixon, D. G. (2012). Has Alberta Oil Sands Development Altered Delivery of Polycyclic Aromatic Compounds to the Peace-Athabasca Delta? PLoS ONE, 7(9), e46089. http://doi.org/10.1371/journal.pone.0046089

Huntemann, M., Ivanova, N. N., Mavromatis, K., Tripp, H. J., Paez-Espino, D., Tennessen, K., … Kyrpides, N. C. (2016). The standard operating procedure of the DOE-JGI

Metagenome Annotation Pipeline (MAP v.4). Standards in Genomic Sciences, 11(1), 1–5. http://doi.org/10.1186/s40793-016-0138-x

Jautzy, J. J., Ahad, J. M. E., Gobeil, C., Smirnoff, A., Barst, B. D., & Savard, M. M. (2015). Isotopic Evidence for Oil Sands Petroleum Coke in the Peace-Athabasca Delta. Environmental Science and Technology, 49(20), 12062–12070. http://doi.org/10.1021/acs.est.5b03232

Jautzy, J. J., Ahad, J. M. E., Hall, R. I., Wiklund, J. A., Wolfe, B. B., Gobeil, C., & Savard, M. M. (2015). Source Apportionment of Background PAHs in the Peace-Athabasca Delta (Alberta, Canada) Using Molecular Level Radiocarbon Analysis. Environmental Science and Technology, 49(15), 9056–9063. http://doi.org/10.1021/acs.est.5b01490

Kelly, E. N., Short, J. W., Schindler, D. W., Hodson, P. V, Ma, M., Kwan, A. K., & Fortin, B. L. (2009). Oil sands development contributes polycyclic aromatic compounds to the Athabasca River and its tributaries. Proceedings of the National Academy of Sciences of the United States of America, 106(52), 22346–51. http://doi.org/10.1073/pnas.0912050106

Leahy, J. G., & Colwell, R. R. (1990). Microbial degradation of hydrocarbons in the environment. Microbiological Reviews, 54(3), 305–315.

Mahajan, M. C., Phale, P. S., & Vaidyanathan, C. S. (1994). Evidence for the involvement of multiple pathways in the biodegradation of 1- and 2-methylnaphthalene by Pseudomonas putida CSV86. Archives of Microbiology, 161, 425–433. http://doi.org/10.1007/s002030050077

Martin, F., Torelli, S., Le Paslier, D., Barbance, A., Martin-Laurent, F., Bru, D., … Jouanneau, Y. (2012). Betaproteobacteria dominance and diversity shifts in the bacterial community of a PAH-contaminated soil exposed to phenanthrene. Environmental Pollution, 162, 345–353. http://doi.org/10.1016/j.envpol.2011.11.032

Mcfall, T., Booth, G. M., Lee, M. L., Tominaga, Y., Pratap, R., Tedjamulia, M., & Castle, R. N. (1984). Mutagenic activity of methyl-substituted tri- and tetracyclic aromatic sulfur heterocycles heterocycles. Mutation Research, 135, 97–103.

Michel, J., & Hayes, M. O. (1999). Weathering patterns of oil residueseight years after the Exxon Valdez oil spill. Marine Pollution Bulletin, 38(10), 855–863.

Novaković, M., Ramadan, M. M. A., Knudsen, T. Š., Antić, M., Beškoski, V., Gojgić-Cvijović, G., … Jovančićević, B. (2012). Degradation of methyl-phenanthrene isomers during bioremediation of soil contaminated by residual fuel oil. Environmental Chemistry Letters, 10(3), 287–294. http://doi.org/10.1007/s10311-012-0354-6

Seo, J. S., Keum, Y. S., & Li, Q. X. (2009). Bacterial degradation of aromatic compounds. International Journal of Environmental Research and Public Health, 6(1), 278–309. http://doi.org/10.3390/ijerph6010278

Siddique, T., Penner, T., Semple, K., & Foght, J. M. (2011). Anaerobic biodegradation of longer-chain n-alkanes coupled to methane production in oil sands tailings. Environmental Science and Technology, 45(13), 5892–5899. http://doi.org/10.1021/es200649t

Siegert, M., Cichocka, D., Herrmann, S., Gründger, F., Feisthauer, S., Richnow, H. H., … Krüger, M. (2011). Accelerated methanogenesis from aliphatic and aromatic hydrocarbons under iron- and sulfate-reducing conditions. FEMS Microbiology Letters, 315(1), 6–16. http://doi.org/10.1111/j.1574-6968.2010.02165.x

Sierra-Garcia, I. N., & de Oliveira, V. M. (2013). Microbial Hydrocarbon Degradation : Efforts to Understand Biodegradation in Petroleum Reservoirs. In Biodegradation - Engineering and Technology (pp. 47–72). IntechOpen. http://doi.org/10.5772/55920

Timoney, K. P. (2013). Natural History: Fauna and Wildlife. In The Peace-Athabasca Delta: Portrait of a Dynamic Ecosystem (pp. 119–149). Edmonton: the University of Alberta Press.

Timoney, K. P., & Lee, P. (2011). Polycyclic aromatic hydrocarbons increase in athabasca river delta sediment: Temporal trends and environmental correlates. Environmental Science and Technology, 45(10), 4278–4284. http://doi.org/10.1021/es104375d

Wang, Z., Yang, C., Parrott, J. L., Frank, R. A., Yang, Z., Brown, C. E., … Hewitt, L. M. (2014). Forensic source differentiation of petrogenic, pyrogenic, and biogenic hydrocarbons in Canadian oil sands environmental samples. Journal of Hazardous Materials, 271, 166–177. http://doi.org/10.1016/j.jhazmat.2014.02.021

Yang, C., Wang, Z., Yang, Z., Hollebone, B., Brown, C. E., Landriault, M., & Fieldhouse, B. (2011). Chemical Fingerprints of Alberta Oil Sands and Related Petroleum Products. Environmental Forensics, 12(2), 173–188. http://doi.org/10.1080/15275922.2011.574312

Yergeau, E., Lawrence, J. R., Sanschagrin, S., Roy, J. L., Swerhone, G. D. W., Korber, D. R., & Greera, C. W. (2013). Aerobic biofilms grown from athabasca watershed sediments are inhibited by increasing concentrations of bituminous compounds. Applied and Environmental Microbiology, 79(23), 7398–7412. http://doi.org/10.1128/AEM.02216-13

Yergeau, E., Lawrence, J. R., Sanschagrin, S., Waiser, M. J., Korber, D. R., & Greer, C. W. (2012). Next-generation sequencing of microbial communities in the athabasca river and its tributaries in relation to oil sands mining activities. Applied and Environmental Microbiology, 78(21), 7626–7637. http://doi.org/10.1128/AEM.02036-12

3.6 Supplemental Information

Table S3.1. PACs and *n*-alkanes targeted for petrogenic hydrocarbon analysis.

15 unsubstituted PACs	Alkylated PACs	*n*-Alkanes
Biphenyl	C0-Naphthalene	*n*-C_9
Acenaphthylene	2-Methylnaphthalene	*n*-C_{10}
Acenaphthene	1-Methylnaphthalene	*n*-C_{11}
Anthracene	C2-Naphthalene	*n*-C_{12}
Fluoranthene	C3-Naphthalene	*n*-C_{13}
Pyrene	C4-Naphthalene	2,6,10-trimethyldodecane
Benz[*a*]anthracene	C0-Phenanthrene	*n*-C14
Benzo[*b*]fluoranthene	C1-Phenanthrene	*n*-C_{15}
Benzo[*k*]fluoranthene	C2-Phenanthrene	*n*-C_{16}
Benzo[*e*]pyrene	C3-Phenanthrene	2,6,10-trimethylpentadecane
Benzo[*a*]pyrene	C4-Phenanthrene	*n*-C_{17}
Perylene	C0-Dibenzothiphene	Pristane
Indeno[1,2,3-*cd*]pyrene	C1-Dibenzothiphene	*n*-C_{18}
Dibenz[*ah*]anthracene	C2-Dibenzothiphene	Phytane
Benzo[*ghi*]perylene	C3-Dibenzothiphene	*n*-C_{19}
	C0-Fluorene	*n*-C_{20}
	C1-Fluorene	*n*-C_{21}
	C2-Fluorene	*n*-C_{22}
	C3-Fluorene	*n*-C_{23}
	C0-Fluoranthene	*n*-C_{24}
	C1-Fluoranthene	*n*-C_{25}
	C2-Fluoranthene	*n*-C_{26}
	C3-Fluoranthene	*n*-C_{27}
	C4-Fluoranthene	*n*-C_{28}
	C0-BNT	*n*-C_{29}
	C1-BNT	*n*-C_{30}
	C2-BNT	*n*-C_{31}
	C3-BNT	*n*-C_{32}
	C4-BNT	*n*-C_{33}
	C0-Chrysene	*n*-C_{34}
	C1-Chrysene	*n*-C_{35}
	C2-Chrysene	*n*-C_{36}
	C3-Chrysene	*n*-C_{37}
	(3-+2-)-Methylphenanthrene	*n*-C_{38}
	(9-/4-+1-)-Methylphenanthrene	*n*-C_{39}
	4-Methyldibenzothiphene	*n*-C_{40}
	2-/3-Methyldibenzothiphene	
	1-Methyldibenzothiphene	

Table S3.2. Assembly statistics from de novo metagenome assembly using Ray software v2.3.1.

Assembly Criteria	Assembly Statistic
contigs greater than 0.5kb	727,906
contigs greater than 1kb	165,077
contigs greater than 2kb	36,135
contigs greater than 5kb	4,390
contigs greater than 10kb	702
contigs greater than 20kb	118
contigs greater than 40kb	15
contigs greater than 80kb	1
contigs greater than 160kb	0
contigs greater than 320kb	0
contigs greater than 640kb	0
contigs greater than 1Mb	0
contigs greater than 2Mb	0
contigs greater than 4Mb	0
contigs greater than 6Mb	0
contigs greater than 8Mb	0
contigs greater than 10Mb	0
Total contigs	727,906
Total bases in contigs (bp)	669,074,241
Minimum contigs length (bp)	500
Maximum contigs length (bp)	119,402
GC content (%)	52.45
N25 - 25% of total sequence length is contained in the 5,295 sequence(s) having a length >=	1,591 bp
N50 - 50% of total sequence length is contained in the 22,597 sequence(s) having a length >=	902 bp
N75 - 75% of total sequence length is contained in the 59,255 sequence(s) having a length >=	639 bp
N90 - 90% of total sequence length is contained in the 94,356 sequence(s) having a length >=	547 bp

Table S3.3. Quality control statistics for metagenomics reads from July 2017 PAD samples.

Sample Name	Raw Fragments	Surviving Fragments	Surviving Fragments (%)	Surviving Single	Total Reads QCed	Mapped	Mapped (%)	Properly Paired	Properly Paired (%)
MCR_PAD-1-Rep-1_2017_July_PAD-1_Sed	3,544,767	3,497,076	98%	31,214	6,344,264	2,978,158	46%	2,010,470	31%
MCR_PAD-1-Rep-1_2017_July_PAD-1_Water	2,799,207	2,761,681	98%	19,463	4,683,060	3,296,334	70%	2,405,622	51%
MCR_PAD-1-Rep-2_2017_July_PAD-1_Sed	2,851,236	2,808,311	98%	30,693	5,175,666	2,373,204	45%	1,626,904	31%
MCR_PAD-1-Rep-2_2017_July_PAD-1_Water	6,553,271	6,471,477	98%	49,525	11,382,898	8,179,148	71%	6,154,594	54%
MCR_PAD-1-Rep-3_2017_July_PAD-1_Sed	3,182,482	3,141,094	98%	27,464	5,793,834	2,695,908	46%	1,846,114	31%
MCR_PAD-1-Rep-3_2017_July_PAD-1_Water	3,177,027	3,140,531	98%	21,145	5,511,302	4,122,748	74%	3,118,896	56%
MCR_PAD-3-Rep-1_2017_July_PAD-3_Sed	11,031,126	10,875,113	98%	102,892	19,773,966	8,938,815	45%	6,075,444	30%
MCR_PAD-3-Rep-1_2017_July_PAD-3_Water	4,401,567	4,346,349	98%	33,018	7,614,204	4,543,157	59%	3,205,844	42%
MCR_PAD-3-Rep-2F2_2017_July_PAD-3_Water	2,940,895	2,899,421	98%	25,924	5,098,406	3,093,344	60%	2,179,016	42%
MCR_PAD-3-Rep-2_2017_July_PAD-3_Sed	9,730,209	9,610,513	98%	67,881	17,335,438	8,415,044	48%	5,750,750	33%
MCR_PAD-3-Rep-2_2017_July_PAD-3_Water	3,854,605	3,803,214	98%	34,567	6,879,658	4,513,597	65%	3,423,328	49%
MCR_PAD-3-Rep-3F2_2017_July_PAD-3_Water	3,154,008	3,110,496	98%	26,683	5,405,928	3,339,871	61%	2,360,066	43%
MCR_PAD-3-Rep-3_2017_July_PAD-3_Sed	13,817,712	13,647,736	98%	104,950	24,936,876	11,600,830	46%	7,924,970	31%
MCR_PAD-3-Rep-3_2017_July_PAD-3_Water	10,068,187	9,934,551	98%	76,526	17,024,040	11,094,920	65%	7,820,188	45%
MCR_PAD-30-Rep-1_2017_July_PAD-30_Sed	15,417,471	15,183,540	98%	158,453	28,133,890	11,483,175	40%	7,706,496	27%
MCR_PAD-30-Rep-1_2017_July_PAD-30_Water	2,886,937	2,854,239	98%	21,528	5,166,670	2,984,967	57%	2,139,522	41%
MCR_PAD-30-Rep-2_2017_July_PAD-30_Sed	16,160,830	15,908,479	98%	159,875	28,796,854	12,831,284	44%	8,430,672	29%
MCR_PAD-30-Rep-2_2017_July_PAD-30_Water	16,475,111	16,299,160	98%	99,749	29,006,996	18,211,948	62%	13,170,354	45%
MCR_PAD-30-Rep-3_2017_July_PAD-30_Sed	4,051,031	3,988,248	98%	33,545	7,090,094	3,188,183	44%	2,093,212	29%
MCR_PAD-30-Rep-3_2017_July_PAD-30_Water	2,617,664	2,585,488	98%	20,915	4,657,872	2,796,799	60%	2,013,674	43%
MCR_PAD-31-Rep-1_2017_July_PAD-31_Sed	10,071,510	9,873,609	98%	39,948	13,503,738	8,451,158	62%	3,671,988	27%
MCR_PAD-31-Rep-1_2017_July_PAD-31_Water	7,033	5,886	83%	971	8,116	5,219	64%	2,930	36%
MCR_PAD-31-Rep-2_2017_July_PAD-31_Sed	12,549,678	12,353,382	98%	102,638	21,504,816	9,503,642	44%	5,832,410	27%
MCR_PAD-31-Rep-2_2017_July_PAD-31_Water	2,719,230	2,683,759	98%	23,112	4,757,990	2,833,006	59%	1,984,676	41%
MCR_PAD-31-Rep-3_2017_July_PAD-31_Sed	9,287,878	9,135,903	98%	96,178	16,600,042	6,765,319	40%	4,445,262	26%
MCR_PAD-31-Rep-3_2017_July_PAD-31_Water	3,450,725	3,408,803	98%	25,238	5,961,190	3,588,703	60%	2,507,040	42%

Table S3.3 (cont'd). Quality control statistics for metagenomics reads from July 2017 uplands (UP) samples.

Sample Name	Raw Fragments	Surviving Fragments	Surviving Fragments (%)	Surviving Single	Total Reads QCed	Mapped	Mapped (%)	Properly Paired	Properly Paired (%)
MCR_UP-3-Rep-1_2017_July_UP-3_Sed	5,800,001	5,717,787	98%	55,184	10,402,216	3,976,647	38%	2,615,070	25%
MCR_UP-3-Rep-1_2017_July_UP-3_Water	3,027,328	2,984,299	98%	28,979	5,374,664	2,857,591	53%	2,050,792	38%
MCR_UP-3-Rep-2_2017_July_UP-3_Sed	6,530,480	6,435,343	98%	71,026	11,753,530	4,742,915	40%	3,110,120	26%
MCR_UP-3-Rep-2_2017_July_UP-3_Water	6,684,200	6,608,622	98%	31,626	11,037,834	6,737,267	61%	4,524,938	40%
MCR_UP-3-Rep-3_2017_July_UP-3_Sed	3,156,256	3,115,857	98%	28,371	5,696,242	2,375,549	41%	1,569,816	27%
MCR_UP-3-Rep-3_2017_July_UP-3_Water	2,718,190	2,684,670	98%	19,973	4,713,376	2,797,532	59%	1,958,690	41%
MCR_UP-7-Rep-1_2017_July_UP-7_Sed	8,596,250	8,485,538	98%	80,016	15,553,332	6,286,663	40%	4,239,954	27%
MCR_UP-7-Rep-1_2017_July_UP-7_Water	2,753,051	2,707,854	98%	31,223	4,889,478	2,783,277	56%	1,979,208	40%
MCR_UP-7-Rep-2_2017_July_UP-7_Sed	3,775,620	3,731,512	98%	30,885	6,871,062	2,760,786	40%	1,874,284	27%
MCR_UP-7-Rep-2_2017_July_UP-7_Water	9,937,939	9,790,444	98%	69,810	16,050,592	10,520,032	65%	6,984,316	43%
MCR_UP-7-Rep-3_2017_July_UP-7_Sed	4,228,529	4,169,467	98%	42,682	7,694,382	2,979,468	38%	1,996,036	25%
MCR_UP-7-Rep-3_2017_July_UP-7_Water	5,933,777	5,854,721	98%	36,971	9,964,632	5,511,801	55%	3,749,280	37%
MCR_UP-9-Rep-1_2017_July_UP-9_Sed	15,682,558	15,426,949	98%	183,753	28,407,040	11,163,787	39%	7,505,764	26%
MCR_UP-9-Rep-1_2017_July_UP-9_Water	4,684,018	4,626,073	98%	36,414	8,300,992	5,494,464	66%	4,054,320	48%
MCR_UP-9-Rep-2_2017_July_UP-9_Sed	1,881,408	1,852,827	98%	18,755	3,321,708	1,470,629	44%	971,620	29%
MCR_UP-9-Rep-2_2017_July_UP-9_Water	2,465,907	2,436,472	98%	17,055	4,270,876	2,909,563	68%	2,134,632	49%
MCR_UP-9-Rep-3_2017_July_UP-9_Sed	12,600,830	12,420,105	98%	121,813	22,653,918	9,204,072	40%	6,198,900	27%
MCR_UP-9-Rep-3_2017_July_UP-9_Water	1,733,827	1,702,337	98%	26,979	3,374,236	1,819,598	53%	1,360,968	40%

Table S3.4. Water chemistry results from July 2017 surface water samples including total nitrogen (TN), total phosphorus (TP), total dissolved nitrogen (TDN), total dissolved phosphorus (TDP), combined nitrate+nitrite (NO_3+NO_2), dissolved inorganic carbon (DIC), dissolved organic carbon (DOC), major anion analysis (Cl^-, SO_4^{2-}), and major cation analysis (Ca^{2+}, Mg^{2+}, Na^+, K^+,SiO_2).

	PAD 1	PAD 3	PAD 30	PAD 31	UP 3	UP 7	UP 9
Nitrogen							
NO2+NO3 (N µg/L)	<MDL	<MDL	2	<MDL	6	<MDL	<MDL
TDN (N µg/L)	2,030	1,890	715	638	729	1,280	618
TN (N µg/L)	2,210	2,210	776	883	916	1,610	637
Phosphorus							
TP (P µg/L)	47	69	100	48	28	25	8
TDP (P µg/L)	23	30	74	12	4	7	4
Carbon							
DOC(C mg/L)	47.8	34.2	14.1	14.3	15.4	31.6	19.1
DIC(C mg/L)	25.8	26.1	37.3	30.6	2.2	9.6	4.9
Major anions							
Cl (mg/L)	13.87	8.90	5.61	5.50	0.66	0.43	0.41
SO4 (mg/L)	2.30	2.90	14.36	27.01	0.96	0.16	1.00
Major cations							
Na (mg/L)	24.17	14.92	10.41	10.19	1.22	1.66	1.30
K (mg/L)	7.99	22.11	2.28	1.62	0.55	1.07	0.43
Ca (mg/L)	19.98	14.62	51.38	45.45	1.81	13.37	7.04
Mg (mg/L)	20.74	19.08	12.30	11.96	0.87	6.02	1.85
SiO2 (Si mg/L)	0.10	0.49	1.00	1.60	0.11	0.65	0.36
Total Gran Alk (mg/L as CaCO3)	147.82	138.68	176.47	138.90	7.69	51.23	21.74
Endpoint pH	4.47	4.53	4.62	4.61	4.89	4.60	4.73
pH	10.50	10.19	7.78	8.65	7.14	7.96	7.65
Turb (NTU)	0.87	2.67	4.06	2.82	0.75	1.18	0.45

Table S3.5. Mean concentration of alkylated PAC concentrations for 7 homologous series in µg/g OC for each lake. N=3 samples per lake.

Compounds	Abbreviations	PAD 1	PAD 3	PAD 30	PAD 31	UP 3	UP 7	UP 9
Alkylated PACs (µg/g OC)								
Naphthalene								
C0-Naphthalene	C0-N	0.23	0.14	0.38	0.41	ND	0.01	0.05
C1-Naphthalene	C1-N	0.48	0.27	0.49	0.63	ND	0.03	0.08
C2-Naphthalene	C2-N	0.60	0.35	0.56	0.78	$1.48x10^{-3}$*	0.03	0.10
C3-Naphthalene	C3-N	0.39	0.24	0.43	0.63	$3.39x10^{-3}$	0.04	0.14
C4-Naphthalene	C4-N	0.19	0.10	0.25	0.45	$8.69x10^{-4}$*	0.04	0.10
Phenanthrene								
C0-Phenanthrene	C0-P	0.29	0.20	0.34	0.49	0.03	0.14	0.30
C1-Phenanthrene	C1-P	0.37	0.28	0.48	0.79	$8.70x10^{-3}$	0.14	0.27
C2-Phenanthrene	C2-P	0.24	0.19	0.49	0.93	$3.18x10^{-3}$	0.09	0.18
C3-Phenanthrene	C3-P	0.14	0.16	0.50	1.43	$1.56x10^{-3}$*	0.05	0.09
C4-Phenanthrene	C4-P	0.04	0.05	0.38	1.17	$3.07x10^{-4}$*	0.01	0.01
Dibenzothiophene								
C0-Dibenzothiophene	C0-D	0.05	0.04	0.05	0.09	$7.87x10^{-4}$*	0.01	0.02
C1-Dibenzothiophene	C1-D	0.06	0.05	0.13	0.25	$1.01x10^{-3}$*	0.02	0.04
C2-Dibenzothiophene	C2-D	0.06	0.06	0.31	0.71	ND	0.02	0.05
C3-Dibenzothiophene	C3-D	0.03	0.04	0.58	1.81	ND	0.01	0.02
Fluorene								
C0-Fluorene	C0-F	0.19	0.13	0.14	0.23	$4.78x10^{-3}$	0.03	0.05
C1-Fluorene	C1-F	0.15	0.11	0.24	0.31	$2.65x10^{-3}$	0.11	0.34
C2-Fluorene	C2-F	0.11	0.10	0.31	0.41	$3.94x10^{-3}$	0.05	0.09
C3-Fluorene	C3-F	0.08	0.07	0.31	0.70	$2.47x10^{-3}$*	0.03	0.06
Fluoranthene								
C0-Fluoranthene	C0-Fl	0.06	0.05	0.11	0.14	$5.19x10^{-3}$	0.05	0.09
C1-Fluoranthene	C1-Fl	0.11	0.10	0.24	0.54	$9.52x10^{-5}$*	0.03	0.04
C2-Fluoranthene	C2-Fl	0.10	0.09	0.32	0.85	$6.89x10^{-5}$*	0.01	0.02
C3-Fluoranthene	C3-Fl	0.04	0.06	0.26	0.75	ND	$2.66x10^{-3}$	$2.36x10^{-3}$*
C4-Fluoranthene	C4-Fl	0.02	0.03	0.14	0.39	ND	$1.43x10^{-3}$	ND
Benzonaphthothiophene								
C0-Benzonaphthothiophene	C0-B	0.02	0.02	0.07	0.19	$5.24x10^{-5}$*	$3.47x10^{-3}$	$5.43x10^{-3}$
C1-Benzonaphthothiophene	C1-B	0.10	0.08	0.53	1.67	ND	ND	ND
C2-Benzonaphthothiophene	C2-B	0.05	0.08	0.81	3.34	ND	ND	ND
C3-Benzonaphthothiophene	C3-B	ND	ND	1.34	4.55	ND	ND	ND
C4-Benzonaphthothiophene	C4-B	ND	ND	0.50	2.39	ND	ND	ND
Chrysene								
C0-Chrysene	C0-C	0.06	0.05	0.13	0.25	$9.53x10^{-4}$*	0.01	0.02
C1-Chrysene	C1-C	0.10	0.08	0.25	0.59	$7.55x10^{-5}$*	$3.73x10^{-3}$	$8.08x10^{-3}$
C2-Chrysene	C2-C	0.06	0.06	0.32	0.84	ND	$4.38x10^{-3}$	ND
C3-Chrysene	C3-C	0.02	0.02	0.12	0.32	ND	ND	ND
Total alkylated PACs		4.44	3.30	11.51	29.03	0.07	0.98	2.18

* Below upper range of LOQ (1.3 ng/g dw)

ND = not detected

Table S3.6. KEGG Orthology, including the definition and ortholog ID (KEGG Entry) for genes related to the aerobic and anaerobic degradation of short- and long-chain *n*-alkanes, PACs, and PASHs. Gene counts provided for the number of genes identified in the metagenome that match the KEGG Entry.

KEGG Orthology				
KEGG Definition	KEGG Entry	Aerobic	Anaerobic	Gene Count
Short-chain n-alkanes		**1502**		**1502**
alkane 1-monooxygenase	K07425	70		70
alkane 1-monooxygenase	K00496	0		0
alkane 1-monooxygenase	K17687	0		0
alkane 1-monooxygenase	K17688	0		0
long-chain fatty acid omega-monooxygenase	K07425	0		0
alkanesulfonate monooxygenase	K04091	88		88
3-phenylpropionate/trans-cinnamate dioxygenase ferredoxin reductase subunit	K00529	134		134
rubredoxin-NAD+ reductase *needed by alkane 1-monooxygenase	K05297	102		102
alcohol dehydrogenase	K13954	99		99
alcohol dehydrogenase, propanol-preferring	K12953	107		107
S-(hydroxymethyl)glutathione dehydrogenase/alcohol dehydrogenase	K00121	123		123
acetaldehyde dehydrogenase / alcohol dehydrogenase	K04072	31		31
alcohol dehydrogenase (cyt C)	K00114	74		74
quinohemoprotein ethanol dehydrogenase	K17760	55		55
aldehyde dehydrogenase (NAD+)	K00128	607		607
aldehyde dehydrogenase (NAD(P)+)	K00129	8		8
energy-converting hydrogenase A subunit D	K14095	4		4
Long-chain n-alkanes (C>18)		**0**		**0**
long-chain alkane monooxygenase	K20938	0		0
midchain alkane hydroxylase	K15405	0		0
PAC		**1024**	**468**	**1492**
phenol- 2-monooxygenase	K03380	13		13
cyclohexanone monooxygenase	K03379	165		165
1,2-dihydroxynaphthalene dioxygenase	K14583	0		0
2-hydroxymuconate-semialdehyde hydrolase	K10216	68		68
2-oxo-3-hexenedioate decarboxylase	K01617	27		27
2-oxopent-4-enoate/cis-2-oxohex-4-enoate hydratase	K18364	4		4
4-hydroxy 2-oxovalerate aldolase	K01666	9		9
4-hydroxy-3-polyprenylbenzoate decarboxylase	K03182		339	339
flavin phenyltransferase	K03186		123	123
4-oxalocrotonate tautomerase	K01821	70		70
acetaldehyde dehydrogenase	K04073	7		7
acetaldehyde/propanal dehydrogenase	K18366	3		3
benzaldehyde dehydrogenase (NAD)	K00141	18		18

Table S3.6 (cont'd). KEGG Orthology, including the definition and ortholog ID (KEGG Entry) for genes related to the aerobic and anaerobic degradation of short- and long-chain *n*-alkanes, PACs, and PASHs. Gene counts provided for the number of genes identified in the metagenome that match the KEGG Entry.

KEGG Definition	KEGG Entry	Aerobic	Anaerobic	Gene Count
coniferyl-aldehyde dehydrogenase	K00154	69		69
coniferyl-aldehyde dehydrogenase	K12355	1		1
aryl-alcohol dehydrogenase (NADP+)	K05882	233		233
aryl-alcohol dehydrogenase	K00055	17		17
benzoate/toluate 1,2-dioxygenase alpha subunit	K05549	3		3
benzoate/toluate 1,2-dioxygenase beta subunit	K05550	4		4
benzoate/toluate 1,2-dioxygenase reductase subunit	K05784	9		9
biphenyl 2,3-dioxygenase alpha subunit	K08689	1		1
catechol 1,2-dioxygenase	K03381	17		17
catechol 2,3-dioxygenase	K00446	43		43
catechol 2,3-dioxygenase	K07104	45		45
complex iron-sulfur molybdoenzyme family reductase subunit alpha	K17050		3	3
complex iron-sulfur molybdoenzyme family reductase subunit beta	K17051		1	1
complex iron-sulfur molybdoenzyme family reductase subunit gamma	K17052		2	2
ethylbenzene hydroxylase subunit alpha	K10700		0	0
extradiol dioxygenase	K11945	1		1
ferredoxin-NAD(P)+ reductase (naphthalene dioxygenase ferredoxin-specific)	K14581	5		5
hydratase-aldolase	K11946	0		0
monooxygenase	K10215	17		17
naphthalene 1,2-dioxygenase system ferredoxin subunit	K14578	13		13
protocatechuate 4,5-dioxygenase, beta chain	K04101	31		31
salicylate hydroxylase	K00480	131		131
trans-o-hydroxybenzylidenepyruvate hydratase-aldolase	K14585	0		0
PASH		**0**	**16**	**16**
2'-hydroxybiphenyl-2-sulfinate desulfinase	K05977		16	16
dibenzothiophene monooxygenase	K22219	0		0
dibenzothiophene sulfone monooxygenase	K22220	0		0
PAC and n-alkane		**0**	**875**	**875**
(R)-benzylsuccinyl-CoA dehydrogenase	K07545		1	1
2,4-dienoyl-CoA reductase (NADPH2)	K00219		345	345
acetone carboxylase, alpha subunit	K10854		3	3
acetone carboxylase, beta subunit	K10855		1	1
acetone carboxylase, gamma subunit	K10856		3	3
benzoylsuccinyl-CoA thiolase BbsB subunit	K07550		9	9
benzylsuccinate CoA-transferase BbsE subunit	K07543		24	24
benzylsuccinate CoA-transferase BbsF subunit	K07544		39	39
benzylsuccinate synthase	K07540		1	1
E-phenylitaconyl-CoA hydratase	K07546		6	6
methyl-coenzyme M reductase beta subunit	K00401		26	26
methyl-coenzyme M reductase subunit C	K03421		13	13
methyl-coenzyme M reductase subunit D	K03422		14	14
pyruvate formate lyase activating enzyme	K04069		390	390
Gene Count		**2526**	**1359**	**3885**

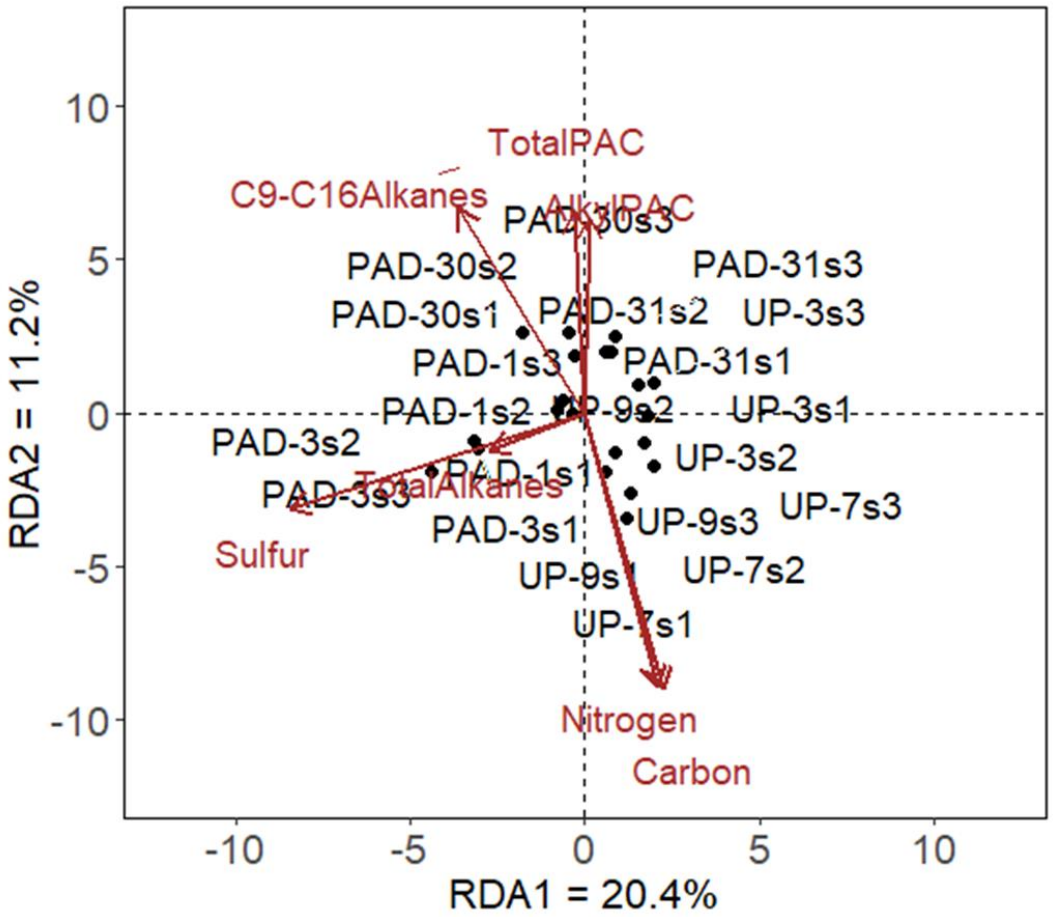

Figure S3.1. Redundancy analysis (RDA) of the abundance of species containing anaerobic degradation genes for PACs and *n*-alkanes in all sediment samples (n=21) with vectors showing chemical variables. Total variance explained in the ordination space is 31.6%. Vectors are significantly related to species abundance (ANOVA test: $F=1.9608$, $p < 0.001$).

CHAPTER 4. CONCLUSION

To my knowledge, the results presented herein provide the first metagenomic analysis of indigenous microbial communities in a freshwater system like the PAD with a history of exposure to PHCs. The aim of this thesis was to evaluate the hydrocarbon-degrading potential of microbial communities in restricted- and closed-drainage lakes of the PAD by way of comparison to closed-drainage lakes in the elevated boreal uplands. This was achieved by

(1) measuring the concentration and distribution of PHCs (including *n*-alkanes, PACs) in surface sediments in addition to characterizing their potential petrogenic origins using petroleum biomarkers (*i.e.* Chapter 2); and
(2) analyzing the metagenomes of surface sediments and water to reconstruct microbial community structure and screen for hydrocarbon degradation genes related to the PHCs measured in Chapter 2 (*i.e.* Chapter 3).

Thus, I have taken the preliminary steps to characterize the genetic potential for hydrocarbon degradation in the PAD and determine whether these were significantly different from freshwater lakes in the boreal upland region.

4.1 Study Outcomes and Evaluations

Chapter 2 provides an analysis of PHCs (*n*-alkanes, PACs, and petroleum biomarkers of terpanes, hopanes, and steranes) in surface sediments of strategically selected lakes in three regions; the Athabasca Delta, the Peace Delta, and the elevated boreal uplands. Results show that *n*-alkanes in all lakes sampled were predominately biogenic, sourced from a mixture of macrophytes (both submerged and emergent) as well as some terrestrial organic matter. However, significant differences between these three regions were found upon analysis of PACs

and petroleum biomarkers. The results of the study confirmed the initial prediction that PACs in restricted-drainage lakes of the Athabasca Delta were higher in concentration and dominated by compounds abundant in bitumen (*e.g.* dibenzothiophenes, fluoranthenes, and benzonaphthothiophenes), whereas PACs in closed-drainage lakes of the Peace Delta were likely derived from a mixture of fossil-fuels and wood combustion. PACs in the boreal uplands were much less concentrated and consisted mainly of atmospherically deposited hydrocarbons possibly sourced from forest fires (*e.g.* dominance of parent and C1-alkylated phenanthrenes and fluorenes) or petrogenic PACs known to associate with dust (*e.g.* low concentrations of dibenzothiophenes and C0-C1 alkylated chrysenes). As predicted in Chapter 1 of this thesis, lakes that were hydrologically connected to the AOSR (PAD 30 and 31) had the highest concentrations of PHCs, namely petrogenic PACs, as confirmed using petroleum biomarkers. To our knowledge, this is the first environmental forensic study to use petroleum biomarkers to confirm long-range fluvial transport of bituminous materials, potentially hundreds of kilometers from its source. The results from this study show that the full suite of petroleum biomarker compounds can in fact be identified in restricted-drainage lakes of the Athabasca Delta. Some lighter terpanes and hopanes can also be identified in closed-drainage Peace Delta and boreal upland lakes, which may be explained by atmospheric deposition. Taking the analysis one step further, multivariate statistics were used to identify raw Alberta oil sands as a potential source of petrogenic materials in Athabasca Delta lakes. This study goes further than previous work in characterizing a larger suite of PHCs to identify which ones dominate the mixture and their potential sources. The hydrological and geographical differences used in this study suggest a shift in PHC source from petrogenic to non-petrogenic that occurs with decreasing influence of the Athabasca River. As such, future work in the PAD should further investigate the occurrence

and distribution of petroleum biomarkers throughout the delta. Obtaining and characterizing petroleum biomarkers from suspected petrogenic sources (*i.e.* McMurray formation, Gates and Gething lower Cretaceous formation, Athabasca Oil Sands, Peace Oil Sands, tailings ponds, petcoke dust, etc.) should be a priority for proper identification of potential sources relevant to the PAD. Being that petroleum biomarkers have a higher resistance to weathering processes and higher specificity in identifying petrogenic inputs than PACs, future PHC investigations in the PAD and elsewhere should include this class of compounds in their analyses.

Chapter 3 provided a metagenomic analysis of the microbial community in the same lakes considered in Chapter 2, screening for genes related to the degradation of *n*-alkanes and PACs. Though the microbial diversity did not significantly vary among lakes, microbial communities in both sediment and water from the Athabasca Delta (PAD 30 and 31), Peace Delta (PAD 1 and 3), and boreal upland (UP 3, 7, and 9) lakes were significantly different from each other, as determined by a permutational ANOVA. The genetic potential for both aerobic and anaerobic PHC degradation was identified in sediments and water, related to various pathways previously identified in the literature (Abbasian & Lockington, 2016; Sierra-Garcia & de Oliveira, 2013). Some examples include aerobic degradation of *n*-alkanes and PACs via monooxygenation, and ring hydroxylation, respectively. Genetic potential for anaerobic degradation of *n*-alkanes and PACs was more prominent in lake sediments than water, and included: fumarate addition, reverse methanogenesis, carboxylation, and oxygen-independent hydroxylation. What's more, a gene (2'-hydroxybiphenyl-2-sulfinate desulfinase (*dszB*)) involved in an intermediate step of desulfurization of thiophene compounds (*e.g.* dibenzothiophene and benzonaphthothiophene) was identified in sediments. Many of the genes coding for anaerobic degradation of PHCs were identified in: syntrophic bacteria of the order

Syntrophobacterales, methanogens of the order Methanomicrobiales, nitrate reducers of the order Hydrogenophilales, and sulfate reducers of the orders Desulfobacterales and Desulfomonodales. The results of the study suggest that the presence of these bacteria and archaea as well as conditions in PAD lakes (*i.e.* higher concentrations of PACs, short-chain alkanes, and sulfur) may allow degradation to occur syntrophically between these microbial consortia. As predicted in Chapter 1 of this thesis, the assemblage of genes coding for the degradation of *n*-alkanes and PACs varies alongside the concentrations of PHCs (more specifically, PACs and short-chain *n*-alkanes), as well as organic carbon, nitrogen, and sulfur in surface sediments. However, the relationship between variations in degradation gene abundance and these chemical variables in sediment was poor (20.7% and 17.7% variation explained for PAC and *n*-alkane degradation genes, respectively), indicating that there may be other important environmental factors that were not included in this analysis. Therefore, the hypothesis of this thesis – that lakes in the PAD that have been exposed to PHCs have significantly different microbial communities and gene assemblages from lakes in the boreal uplands – can be accepted while acknowledging that PHC exposure may not be the main driver of this difference. To my knowledge, this study provides the first metagenomic assessment of the functional and taxonomic diversity of microbial communities in PAD lakes, and their metabolic potential for hydrocarbon degradation of bitumen-associated compounds. Though the potential for hydrocarbon degradation was identified, metagenomics is limited by its use of DNA and semi-quantitative nature. Further research is necessary to determine whether the hydrocarbon degradation genes identified in this study are transcriptionally active and taking part in the degradation of PHCs (*i.e.* via metatranscriptomics or microcosms).

4.2 Concluding Remarks

The health of the PAD and its inhabitants is subject to various stressors such as climate change and upstream industrial development on both its major rivers including mining, river regulation, forestry, and agriculture. Concerns have been mounting about major resource extraction in the AOSR less than 200 km south of the delta, which has been shown to emit PACs and PASHs to the Athabasca River, its tributaries, and the atmosphere (Evans et al., 2016; Kelly et al., 2009; Kurek et al., 2013). Though this thesis does not attempt to distinguish between anthropogenic and natural sources of PHCs, the results show that PACs were in fact higher in PAD lakes than in the nearby boreal upland region, and that restricted-drainage lakes adjacent to Mamawi Creek in the Athabasca Delta (lakes PAD 30 and 31) received the highest concentration of petrogenic PACs with a hopanoid/steroid distribution matching raw AOS samples. In accordance with previous studies (Hall et al., 2012; Jautzy, Ahad, Gobeil, et al., 2015; Jautzy, Ahad, Hall, et al., 2015; Timoney & Lee, 2011), these findings highlight the role that the Athabasca River watershed plays in delivering PHCs to the PAD and the importance of understanding the distribution of these compounds according to the delta's hydrology.

By applying techniques used in oil spill forensics, this thesis illustrates the importance of informed interpretations of PHC sources using a multiple proxy approach. In fact, petroleum biomarkers of terpanes, hopanes, and steranes were validated as a tool to accurately identify the input and potential origin of PHCs in the PAD. Since these compounds are uniquely sourced from petroleum reservoirs and more resistant to weathering processes than PACs and *n*-alkanes, their potential uses in future work are numerous. Uses for petroleum biomarkers in the PAD include, but are not limited to; paleolimnological assessments, petrogenic source identification, and monitoring PHC input. The continuation of PHC monitoring efforts in the PAD is of utmost

importance to the preservation of aquatic life and traditional land use practices by First Nations and Métis. As this thesis shows, these compounds continue to be present in lake surface sediments, some of which, surpass interim sediment quality guidelines set by the Canadian Environmental Quality Guidelines (Canadian Council of Ministers of the Environment, 1999).

This thesis provides the first metagenomic reconstruction of microbial communities in the PAD and their genetic potential for hydrocarbon degradation. In fact, metagenomes in lakes of the PAD and boreal uplands reveal that these microbial communities contain genes involved in the initial activation and further metabolism of PHCs, under both aerobic and anaerobic conditions. Despite the presence of PACs, microbial communities in PAD lakes were found to be as diverse as those in the boreal upland region. The results also demonstrate the potential for anaerobic degradation of PHCs via syntrophic interactions between syntrophic bacteria, methanogens, nitrate reducers, and sulfate reducers. Therefore, this thesis has taken the preliminary steps to advance our understanding of how hydrocarbon degradation occurs in Canadian freshwater systems such as the PAD. Since the PAD is downstream to major deposits of bituminous sands and industrial mine tailings, determining whether microbial communities are capable of degrading bituminous compounds in a freshwater context is important since biodegradation is the primary mechanism by which PHCs can be removed from the environment. The present thesis determines that microbial communities in the PAD have the genetic potential for hydrocarbon degradation, validating future research into whether these microbes can actively degrade PHCs.

4.3 References

Abbasian, F., & Lockington, R. (2016). A Review on the Genetics of Aliphatic and Aromatic Hydrocarbon Degradation, 224–250. http://doi.org/10.1007/s12010-015-1881-y

Canadian Council of Ministers of the Environment. (1999). Canadian Sediment Quality Guidelines for the Protection of Aquatic Life: Polycyclic Aromatic Hydrocarbons (PAHs). Canadian Environmental Quality Guidelines. Winnipeg.

Evans, M., Davies, M., Janzen, K., Muir, D., Hazewinkel, R., Kirk, J., & De Boer, D. (2016). PAH distributions in sediments in the oil sands monitoring area and western Lake Athabasca: Concentration, composition and diagnostic ratios. Environmental Pollution, 213, 671–687. http://doi.org/10.1016/j.envpol.2016.03.014

Hall, R. I., Wolfe, B. B., Wiklund, J. A., Edwards, T. W. D., Farwell, A. J., & Dixon, D. G. (2012). Has Alberta Oil Sands Development Altered Delivery of Polycyclic Aromatic Compounds to the Peace-Athabasca Delta? PLoS ONE, 7(9), e46089. http://doi.org/10.1371/journal.pone.0046089

Jautzy, J. J., Ahad, J. M. E., Gobeil, C., Smirnoff, A., Barst, B. D., & Savard, M. M. (2015). Isotopic Evidence for Oil Sands Petroleum Coke in the Peace-Athabasca Delta. Environmental Science and Technology, 49(20), 12062–12070. http://doi.org/10.1021/acs.est.5b03232

Jautzy, J. J., Ahad, J. M. E., Hall, R. I., Wiklund, J. A., Wolfe, B. B., Gobeil, C., & Savard, M. M. (2015). Source Apportionment of Background PAHs in the Peace-Athabasca Delta (Alberta, Canada) Using Molecular Level Radiocarbon Analysis. Environmental Science and Technology, 49(15), 9056–9063. http://doi.org/10.1021/acs.est.5b01490

Kelly, E. N., Short, J. W., Schindler, D. W., Hodson, P. V, Ma, M., Kwan, A. K., & Fortin, B. L. (2009). Oil sands development contributes polycyclic aromatic compounds to the Athabasca River and its tributaries. Proceedings of the National Academy of Sciences of the United States of America, 106(52), 22346–51. http://doi.org/10.1073/pnas.0912050106

Kurek, J., Kirk, J. L., Muir, D. C. G., Wang, X., Evans, M. S., & Smol, J. P. (2013). Legacy of a half century of Athabasca oil sands development recorded by lake ecosystems. Proceedings of the National Academy of Sciences of the United States of America, 110(5), 1761–6. http://doi.org/10.1073/pnas.1217675110

Sierra-Garcia, I. N., & de Oliveira, V. M. (2013). Microbial Hydrocarbon Degradation : Efforts to Understand Biodegradation in Petroleum Reservoirs. In Biodegradation - Engineering and Technology (pp. 47–72). IntechOpen. http://doi.org/10.5772/55920

Timoney, K. P., & Lee, P. (2011). Polycyclic aromatic hydrocarbons increase in athabasca river delta sediment: Temporal trends and environmental correlates. Environmental Science and Technology, 45(10), 4278–4284. http://doi.org/10.1021/es104375d