Effects of naphthenic acids and acid extractable organic mixtures on development of the frog *Silurana (Xenopus) tropicalis*

Juan Manuel Gutierrez Villagomez

Thesis submitted in partial fulfillment of the requirements for the Ph.D. degree in Biology

with specialization in Chemical and Environmental Toxicology

Ottawa-Carleton Institute of Biology

Faculty of Science, University of Ottawa

University of Ottawa • Universite d'Ottawa

Dedicated to my family, especially M. Rosalia Villagomez Gomez and

Juan Antonio Gutierrez Negrete

Acknowledgments

I would like to express my deepest gratitude to my supervisor, Dr. Vance Trudeau, for five years of continuous guidance and expertise. Thank you for your enormous support, encouragement, and patience.

I also would like to thank Dr. Michael Rene van den Heuvel, Dr. John T. Arnason, Dr. Stacey Robinson, and Dr. Paul White for reviewing this thesis. I am gratefully indebted to their very valuable comments.

I would like to thank the experts who were involved in this research thesis: Dr. Juan Vazquez Martinez. Enrique Ramirez Chavez, Dr. Jorge Molina Torres, Kerry M. Peru, Dr. John V. Headley, Dr. Christopher Martyniuk, and Dr. Valerie Langlois. Without their participation and input, this work could not have been successfully conducted.

I want to acknowledge the help of Dr. Lei Xing, Julie Bilodeau, Maria Vu, Connor Edington, Ailsa Gan, Sarah Richardson, Kim Mitchell, Marilyn Vera Chang, Brooke Cameron, Erin Blake, Sue Zhang and Christine Wong. Also, to all the past and present members of TeamEndo. They made things easier and funnier in the lab.

I would like to thank Dr. Sarah A. Hughes for providing the oil sands process-affected water sample and to Dr. Matthew Pamenter for letting me use his equipment for gene expression analysis. The assistance provided by William Fletcher and Christine Archer during frog injections was greatly appreciated.

Finally, I must express my very profound gratitude to my family and friends for providing me with encouragement throughout my years of study. Friends from near and far, Dario, Tomas, Mayra, Hector, Gen, Rafa, Xico, Clausula, Alondra, Ann, Carlos, Jossua, Juan Ma, Maja, Lety, and the Sirois-Leclerc family.

Abstract

Naphthenic acids (NAs) are oil-derived mixtures of carboxylic acids and are aquatic contaminants of emerging concern. The objective of the research presented in this thesis was to investigate the toxicity of NAs in tadpoles of the frog *Silurana (Xenopus) tropicalis*. Using electrospray ionization high-resolution mass spectrometry (ESI-HRMS), I determined that the proportions of O_2 (presumably carboxylic acid moiety) species were 98.8, 98.9 and 58.6% respectively, for two commercial extracts (S1 and S2), and acid extractable organics (AEOs) from oil sands process-affected water (OSPW). The rank order potency based on the lethal concentration fifty (LC_{50}) and effect concentration fifty (EC_{50}) with and without normalization for the quantity of O_2 species was S1 > S2 > AEO. The main effects observed were reduced body size, edema, and cranial, cardiac, gut and ocular abnormalities. Oligonucleotide microarray technology was used to determine the transcriptomic responses in developing *S. tropicalis* embryos following exposure to S1 and S2 at a sub-lethal concentration of 2 mg/L. Some of the significantly enriched pathways ($p < 0.05$) included metabolism and cell membrane depolarization, and some were related to observed abnormalities including edema, gastrointestinal system, and cartilage differentiation. I established and validated a derivatization method for NAs using pentafluorobenzyl bromide (PFBBr) prior to gas chromatography-electron impact mass spectrometry (GC-EIMS) to increase chromatographic resolution, and sensitivity, compared to boron trifluoride-methanol (BF_3/MeOH) and *N-tert*-Butyldimethylsilyl-*N*-methyltrifluoroacetamide (MTBSTFA). Solid-phase microextraction of volatiles originating from S1, S2, Merichem NAs and an AEO mixture led to the identification of 54, 56, 40 and 4 compounds, respectively. The

compounds identified in the mixtures included aliphatic and cyclic hydrocarbons, carboxylic acids, alkyl-benzenes, phenols, naphthalene and alkyl-naphthalene, and decalin compounds. To determine the chemical nature of the toxic compounds in NA mixtures, the S2 and AEOs preparations were fractionated using open column chromatography. A non-polar and a polar fraction were obtained from S2. Overall, the toxicity of the polar fraction was not significantly different from whole S2 ($p > 0.05$). Six fractions of AEOs were obtained, however because of limited material, only the toxicities of F3 and F4 were assessed. The toxicity of F3 was significantly lower than AEOs ($p < 0.05$) and F4 was not toxic for *S. tropicalis* ($p > 0.05$). These results suggest that during fractionation, toxic compounds were lost or that the toxicity of AEOs results from the combined effects of the compounds present in the whole extract. The toxicological dose descriptors, morphometric, transcriptomic and chemical analysis herein presented may contribute to the development of environmental guidelines for NAs and AEOs.

Résumé

Les acides naphténiques (ANs) sont des mélanges d'acides carboxyliques dérivés du pétrole et sont considérés comme des contaminants aquatiques émergents. L'objectif de la recherche présentée dans cette thèse était d'étudier la toxicité des ANs chez les têtards de la grenouille *Silurana (Xenopus) tropicalis*. Utilisant la spectrométrie de masse à haute résolution par ionisation par électronébuliseur, j'ai déterminé que les proportions d'espèces O_2 (probablement groupement d'acide carboxylique) étaient de 98.8, 98.9 et 58.6% respectivement, pour deux extraits commerciaux (S1 et S2) et organiques extractibles à l'acide (OEA) provenant des eaux affectées par les sable bitumineux. L'ordre de classement basé sur CL_{50} et CE_{50}, avec et sans normalisation pour la quantité d'espèces O_2, était S1 > S2 > AEO. Les principaux effets observés chez *S. tropicalis* ont été la réduction de la taille corporelle, de l'œdème et des anomalies crâniennes, cardiaques, intestinales et oculaires. Des puces à ADN ont été utilisées pour identifier des réponses transcriptomiques chez des embryons de *S. tropicalis* en développement après une exposition à S1 et S2 à une concentration sublétal de 2 mg/L. Certaines des voies significativement enrichies ($p < 0.05$) comprenaient le métabolisme et la dépolarisation de la membrane cellulaire. Certaines des voies enrichies étaient liées aux anomalies observées comprenant l'œdème, le système gastro-intestinal et la différenciation du cartilage. J'ai établi et validé une méthode de dérivatisation pour les ANs utilisant le bromure de pentafluorobenzyle (PFBBr) avant la chromatographie en phase gazeuse couplée à la spectrométrie de masse par ionisation électronique (GC-IEMS) pour augmenter la résolution chromatographique et la sensibilité, comparé au trifluorure de bore-méthanol

(BF_3/MeOH) et N-tert-butyldiméthylsilyl-N-méthyltrifluoroacétamide (MTBSTFA). La micro-extraction sur phase solide des substances volatiles provenant de S1, S2, Merichem et d'un mélange d'OEA a conduit à l'identification de 54, 56, 40 et 4 composés, respectivement. Les composés identifiés dans les mélanges comprenaient des hydrocarbures aliphatiques et cycliques, des acides carboxyliques, des alkylbenzènes, des phénols, des naphtalènes et des alkyl-naphtalènes, et des composés de décaline. Pour déterminer la nature chimique des composés toxiques dans les mélanges d'AN, les préparations de S2 et d'OEA ont été fractionnées en utilisant la chromatographie sur colonne ouverte. Une fraction non polaire et une fraction polaire ont été obtenues à partir de S2. Dans l'ensemble, la toxicité de la fraction polaire n'était pas significativement différente de S2 entière ($p > 0.05$). Six fractions d'OEA ont été obtenues, mais en raison du matériel limité, seulement les niveaux de toxicité de F3 et F4 ont été évaluées. La toxicité de F3 était significativement plus faible que celle des OEA ($p < 0.05$) et F4 n'était pas toxique pour *S. tropicalis* ($p > 0.05$). Ces résultats suggèrent que lors du fractionnement, des composés toxiques ont probablement été perdus ou que la toxicité des OEA résulte des effets combinés des composés présents dans l'extrait entier. Les descripteurs de dose toxicologique, morphométrique, d'expression génique et d'analyse chimique présentés dans cette thèse peuvent contribuer à l'élaboration de directives environnementales pour les AN et les OEA.

Table of contents

List of Figures

List of Tables

Abbreviations

AEOs	Acid extractable organics
ANOVA	Analysis of variance
BTEX	Benzene, toluene, ethylbenzene, and xylene
CCFD	Center composite face design
DBE	Double bond equivalences
DVB/CAR/PDMS	Divinylbenzene/carboxen/polydimethylsiloxane
EIC	Extracted ion chromatogram
ESI-HRMS	Electrospray ionization high-resolution mass spectrometry
FDR	False-discovery rate
GC-EIMS	Chromatography coupled to electron impact mass spectrometry
GO	Gene ontology
hCG	Human chorionic gonadotropin
HPLC	High-performance liquid chromatography
IOD	Interorbital distance
IUPAC	International Union of Pure and Applied Chemistry
m/z	Mass charge ratio
Mcf	Thousand cubic feet
MTBSTFA	*N*-(t-butyldimethylsilyl)-*N*-methyltrifluoroacetamide
NA	Naphthenic acids
OSPW	Oil Sands Process-affected Water
PAGE	Parametric analysis of gene set enrichment
PAHs	Polycyclic aromatic hydrocarbons
PCA	Principal component analysis
PFBBr	Pentafluorobenzyl bromide
pKa	Acid dissociation constant at logarithmic scale
R^2	Coefficient of determination
Ri	Retention indices

RSM	Response surface methodology
RT	Retention times
RT-qPCR	Real-time quantitative polymerase chain reaction
SNEA	Sub-network enrichment analysis
SPME	Solid phase microextraction
SVL	Snout-vent length
TaL	Tail length
TIC	Total ion chromatograms
TL	Total length
VOCs	Volatile organic compounds

Chapter 1

General introduction

1.1. Problem identification and thesis overview

Naphthenic acids (NAs) are complex mixtures of carboxylic acid produced in large volumes from several sources. One of the main sources of NAs are the Canadian oil sands in Alberta, Canada. The disposition of water and the reclamation of land contaminated with these compounds have become a challenge and a public concern. Therefore, it is necessary to establish regulations for the protection of wildlife and ecosystems. The assessment of the dose-dependent effects on physical abnormalities and gene expression in the frog model *Silurana tropicalis*, coupled with the chemical characterization of mixtures of NAs can offer valuable information for alternative wastewater and land management.

The general hypothesis I have tested is that NA mixtures induce detrimental abnormalities on *S. tropicalis* embryos through the alteration of the expression of developmental and metabolic genes. I have attempted to correlate the extent of effects with the chemical compositions of each mixture. The main objective of this work was to determine the lethal, sub-lethal and molecular effects of the exposure to NAs on *S. tropicalis* in relation to mixture composition.

This thesis is a compilation of manuscripts that includes an introductory chapter with literature review, five data chapters (Chapter 2 to 6) and a conclusion chapter. Each chapter aimed to answer a specific question to better understand the complexities of NAs, their toxicity, and their composition. The physiological effects of the exposure to

NAs were explored in Chapter 2 and 3. I used several techniques for this purpose, including bioassays, microarray analysis, and RT-qPCR. In Chapter 2, I report the lethal and sub-lethal effects of the exposure to NAs on *S. tropicalis* together with a chemical analysis by electrospray ionization high-resolution mass spectrometry (ESI-HRMS). In Chapter 3, I explored the molecular mechanism of the toxicity of NAs using a microarray. In Chapter 4, I report the use of a derivatization method for the analysis of NAs using gas chromatography coupled to electron impact mass spectrometry (GC-EIMS). In Chapter 5, I report the volatile organic compound profiles of the mixtures using solid phase microextraction (SPME) coupled to GC-EIMS. In Chapter 6, I report on the toxicity of specific fractions of the NA mixtures in *S. tropicalis*. Each of the chapters is written as single independent manuscripts intended for publication. Chapter 4 was already accepted for publication in October 2016 by the journal Talanta. My original contributions to the field are briefly summarized in chapter 7. I also present the following introduction and review of relevant literature to set the stage for the original research presented in this thesis.

1. 2. Literature review

1.2.1. Oil sands

Oil is an important source of energy, accounting for 32% of the consumption worldwide in 2010 [1]. Canada has the third-largest proven oil reserves in the world, with the oil sands representing 92% of these reserves [2]. There are three oil sands deposit areas in Canada: Athabasca, Peace River and Cold Lake [3]. These sites are the only commercially exploited oil sands deposits in the world, and production has

increased since becoming an economically recoverable source of oil in 2003. On the whole, the deposits cover an area of 142,000 square kilometers [4,5].

According to the United States Energy Information Administration, in 2014, Canada was the world's 5th producer of oil with 4.5 million barrels per day and was the 9th consumer using 2.4 million barrels per day [6]. The volume of the reserves in the oil sands is estimated to be 28.3 billion cubic meters (178 billion barrels), but only 12% is recoverable with current technology [5]. In 2010 the oil sands production represented the 1.9% of the world oil supply [4]. The projections indicate that in 2035 the oil sands will be producing approximately 4.8 millions of barrels every day [7].

The oil sands are non-conventional deposits of petroleum in the form of bitumen. It is believed that the oil sands in Alberta, Canada formed as marine deposits 50 million years ago with subsequent migration of the oil to large areas of sands. Oxidation, evaporation, water erosion, and bacteria have collectively degraded the simplest and heaviest hydrocarbons, resulting in bitumen [4]. Bitumen is a form of petroleum in solid state, generally found mixed with sand and water (Fig. 1.1). The proportions of these components of the oil sands can vary, but generally, it is composed of 75-80% inorganic material, 3-5% water, and 10-12% bitumen [5]. Natural bitumen deposits, defined as oil with viscosity 10 000 centipoises and an API gravity of less than 10º, have been reported in twenty-three countries around the world. The composition of the oil from the bitumen is different than conventional crude oil; it is more abundant in nitrogen, oxygen, heavier hydrocarbons [8], complex sulfur compounds and metals [3], and it is deficient in hydrogen [5].

There are two main methods of extracting petroleum from the oil sands: in-situ production underground and traditional mining. Nearly 80% of the oil sands bitumen is too deep for mining but are recoverable by in-situ technologies, which have a recovery rate between 25 and 50% [4]. The two most common in-situ methods are Steam Assisted Gravity Drainage (SAGD) and Cyclic Steam Stimulation (CSS). To recover bitumen using the SAGD method, steam is injected into the pit using two parallel tubes in order to soften the bitumen, and pumps are used to bring the bitumen to the surface. CSS is a three-phase process. First, high-pressure steam is injected for several weeks to separate the bitumen from sand and reduce its viscosity. Once the reservoir is saturated, the injection of steam stops and the reservoir soaks for several weeks. Finally, the production phase starts when the bitumen is extracted [9,10].

Approximately 20% of Alberta's oil sands can be exploited using the mining method [9]. The area is cleared of vegetation, then large volumes of oil sands are removed, crushed, and mixed with hot water. The mixture is transported to the extraction plant by pipelines for separation. The mixture is combined with hot water, sodium hydroxide and air [11]. And the bitumen which floats to the surface is recovered and transported to the upgrading facilities. The mix of water, sand, bitumen, and chemicals is transferred to the settling ponds [12]. The recovery rate of the mining process is around 82%. The bitumen's high viscosity and sulfur content are not ideal for pipelines, therefore it has to undergo fractional distillation to produce Synthetic Crude Oil (SCO) with 36°API and sulfur content of 0.015% [3].

Oil sands development has resulted in a large environmental footprint. The industrial activity in the oil sands is linked to the emission of carbon dioxide (CO_2),

methane (CH_4), sulfur dioxide (SO_2), nitrogen oxides (NO_X), hydrogen sulfide (H_2S), carbon monoxide (CO), volatile organic compounds (VOCs), ozone (O_3), polycyclic aromatic hydrocarbons (PAHs), particulate matter (PM), reduced sulfur compounds (SC_s) and other trace air contaminants [13,14]. Oil sands greenhouse emissions represent 5% of Canada's total greenhouse emissions and 21% of the total emissions of Alberta [10]. SAGD requires 1 Mcf (thousand cubic feet) of natural gas and CSS requires 1.0 to 1.2 Mcf per barrel of oil to produce high-pressure steam. The process of upgrading requires 0.75 Mcf of natural gas per barrel of oil [5]. In-situ facilities require, on average, half a barrel of fresh water to generate steam, to produce one barrel of oil, to separate bitumen from sand and for hydro transportation, besides the upgrading process. For the mining method, almost two tonnes of soil are used to obtain one barrel of crude oil; in 2009, it was reported that an area of 602 km^2 of land has been disturbed due to oil sands mining operations [10].

The mining process produces large volumes of fine tailings; after the bitumen is extracted, hydrocarbon- and solvent-contaminated water, sand, silt, and clay remain. Currently, for each unit of bitumen produced, approximately 2.5 units of water are required [15]. The main source of water for the oil industry in Alberta is the Athabasca River [16]. The water used in the recovering process becomes contaminated and is therefore kept in tailing ponds. The tailing ponds are a disposal area for the Oil Sands Process-affected Water (OSPW) that has been in contact with bitumen or other contaminants. Some tailings ponds can be as large as 30 km^2 [14] and the migration of pollutants from the tailing ponds has become of public concern. Frank et al. compared NA profiles of samples of groundwater of the Athabasca River area and samples of

locally produced OSPW and found important similarities, suggesting a common source, and concluded that OSPW is being released to the environment [17]. The tailings water contains high concentrations of inorganic salts, heavy metals, PAHs, BTEX compounds (benzene, toluene, ethylbenzene, and xylene), residual bitumen and NAs [14,18]. During the oil extraction, the NAs are solubilized and concentrated in the ponds.

1.2.2. Naphthenic acids

In 1984, carboxylate surfactants were first reported in samples of the Syncrude Mildred Lake oil sands facilities [19] and were later identified as NAs [20]. Being natural components of oil [21], the NA content of petroleum ranges between 0-3% by weight [22]. In the industry, NAs refers to all the carboxylic acids present in crude oil. However, a more scientific definition for NAs would be the mixture of alkyl-substituted cyclic and acyclic aliphatic carboxylic acids, represented by the general formula: $C_nH_{2n-z}O_2$, where *n* indicates the carbon number [22], *z* specifies the deficiency of hydrogen from the formation of rings and groups the NAs in homologous series [23] and relatively low molecular weight (<500 Daltons) [24] (Fig. 1.2). The International Union of Pure and Applied Chemistry (IUPAC) defines NAs as “acids chiefly monocarboxylic, derived from naphthene” [25]. NAs are amphiphilic molecules and they are relatively soluble in water due to their hydrophilic carboxylic acid functional group and low molecular weight. The carbon chain represents the hydrophobic part. The approximate pKa values of carboxylic acids are around 5 [26] and specifically, the pKa value for NAs is 4.9 [27], therefore NAs are weak acids. The NAs are commercially used to make metal salts and are also used as wood preservative. In 1986, NAs were identified as the main toxic component of OSPW [20].

There are many different concentrations of total NAs reported in the OSPW of different tailings ponds. In the literature, NA concentration values vary between 48 to 128 mg/L [17,24,28–30], reflecting the heterogeneity between tailings ponds. The concentration reported in the water and groundwater of the Athabasca River near the oil sands extraction facilities ranges from not detected to 48 mg/L [17]. The concentration of the NAs in the OSPW depends on the age of the pond, and where the sample was taken. Furthermore, the levels of dissolved oxygen, which influence the rate of biodegradation, vary considerably with the depth of the pond [31], so the distribution of NAs is not homogeneous. Also, the ponds are exposed to abiotic factors such as the air, light, temperature, radiation [32] and biotic factors (such as macro- and microorganisms), therefore the composition and proportions of NAs in the water of the ponds change constantly.

1.2.3. Chemical analysis of NAs

The water that has been in contact with bitumen takes up some of the components of the oil, making the OSPW a substance of unknown and variable composition. Due to the complexity of the mixtures, the separation, identification, and quantification of each component of the blend have not yet been achieved, and only a few compounds from the OSPW have been clearly isolated and identified such as adamantane-1-carboxylic acid, adamantane-2-carboxylic acid, and 3-ethyl adamantane-1-carboxylic acid [33]. Also, the resin acid dehydroabietic acid (DHAA) has been isolated from the OSPW [34] and is also found in pulp mill effluents [35]. It is therefore highly unlikely that DHAA comes from petroleum products in oil sands *per se*, but may

be related to forest disturbances during mining. Also, some series of NAs have been tentatively identified [36].

To evaluate the concentration and composition of the NAs from OSPW, extraction, concentration, and purification are necessary. The process of extraction involves acidifying the sample to a pH of 2, adding a solvent (usually dichloromethane), and drying the solution. The residue represents the fraction that contains the NAs [24]. However, the methodologies that have been reported are not specific for NAs. The substances that result from the OSPW acid are called acid-extractable organics (AEOs), which includes NAs [17] along with fulvic acids, humic acids, and metallic ions [37]. Moreover, the presence of aromatic organic acids has been reported in recent articles analyzing mixtures of NAs extracted from the OSPW [34,38]. The classic definition of NAs does not include these aromatic NAs, therefore the establishment of a new definition is required to name all categories of organic acids in the AEO.

With the application of new techniques and the combination of analytical methodologies, it is possible to get a better understanding of complex mixtures. It is important to underline that the accuracy of the determination of concentration and the identification of the components of the NAs will depend on the methodology used to assess it [28]. During my Ph.D. project, I used electrospray ionization high-resolution mass spectrometry (HRMS) and gas chromatography coupled to electron impact mass spectrometry (GC-EIMS) for analysis and characterization of the mixtures. The two analytical techniques are briefly described below.

One of the first applications of mass spectrometry (MS) was the analysis of oil and its byproducts. This approach is based on the formation of charged particles from

the molecules of the analyte. There is an energetic process to charge the molecules, followed by the detection and measurement of the ratio of the mass and charge of the molecules (*m/z*). Depending on the level of the energy of the process, the molecules may be fragmented. When fragmentation does not occur, a mass spectrum representing the molecular weights of the components is obtained. When the fragmentation occurs, a unique fragmentation pattern to each compound under the same conditions of analysis is obtained and allows for identification when a standard is available. There are several sources of ionization energy, such as electrospray ionization (ESI), electronic impact (EI) and chemical ionization (CI). The electronic impact approach has a high ionization energy and the molecules will be fragmented, however, with ESI, the molecules are not fragmented. ESI-MS allows the identification of the molecular weight of the analytes, but it does not provide structural information. Electronic impact ionization gives information about the structure from the analysis of the fragmentation patterns; however, the fragmentation pattern does not always show the molecular ion. For the analysis of individual components in complex mixtures, it is necessary to complement with a separation technique [39,40]. ESI-MS has been applied to the analysis of mixtures of NAs as a fast, semi-quantitative method. It is possible to obtain the relative proportions of the molecular weights with detection limits around 0.01 mg/L [37]. However, this technique does not differentiate NAs from contaminants with similar molecular weight, nor between isomers. Another variation of the mass spectrometer, besides the source of ionization, is the mass analyzer; the resolution of molecular masses relies on this device. Equipment with relatively low resolution can differentiate *m/z* up to 20 ppm, whereas a high resolution allows the

differentiation of ~5 ppm. With higher resolution, there is a superior identification of the components of a sample since few compounds have exactly the same molecular weight [41]. The study of mixtures of NAs with the Orbitrap mass spectrometer allows the distinction of compounds with the identical molecular weight as the NAs and compounds with small variations of molecular weight. However, it is not possible to distinguish between isomers since the mass is exactly the same. Another advantage of this MS variant is the low detection limit and the high linearity to the detection of NAs [42].

Gas chromatography-electron impact mass spectrometry combines the capacity of separation of a GC with the detection of the ionized molecules of an MS. Gas chromatography is a method of separation that is based on the partition of the components between an inert gas in the mobile phase and a stationary phase, generally, a liquid immobilized in a solid support. The separation takes place inside of a capillary column with variable composition, diameter, length and thickness of the layer of the stationary phase depending on the target compounds. The GC is an excellent methodology for the separation of complex mixtures of compounds that are volatile in the temperatures that the columns can resist (40-370 ºC). When the compounds are not volatile in such conditions, GC is not effective, however, many compounds can be chemically transformed to more stable and volatile derivatives under GC operating conditions. After the various compounds have been separated in the GC, they elute from the column to the ionization source in the MS. Once the ions are formed, they are detected and measured in the mass analyzer of the MS. This approach provides a chromatogram with the retention times (RT) and the intensity of the ions for each

component of the mixture. For the same series of compounds, the retention time is proportional to the number of carbons [40]. The analysis of NAs using GC provides useful quantitative information that allows the measurement of the total concentration of NAs in a sample and gives a notion of the number of atoms of carbon for each compound in a sample according to the retention time, with the advantage of a third dimension that shows the fragmentation pattern of each separated compound in an specific retention time [43].

Since NAs are polar and labile compounds, in order to improve the chromatographic characteristics of the analyte, it is necessary to derivatize the carboxylic group for a more stable group to reduce the polarity of the molecule. The hydrogen can be substituted with a methyl group, a silane group or other. One disadvantage of this technique is that it cannot provide information about molecular weight and molecular structure. Therefore, if there are no appropriate standards available for NAs, this is not a useful identification technique [44]. The analysis of mixtures of NAs produces a chromatogram with a characteristic hump in which all the compounds are distributed. Without the capacity to resolve this sample, the mass spectrum represents the average spectrum of all of the components. This analysis is helpful to determine the differences in the composition of NAs mixtures. However, for a proper quantification, a better resolution of the samples is necessary [45,46]. By coupling other types of mass spectrometer, it could be possible to improve the analysis of the NAs samples.

1.2.4. Toxicity of the NAs

After the Exxon Valdez oil spill in March 1989, there has been a growing concern about the environmental effects of chemicals and the activities related to the extraction, production, and transportation of oil products. Research on oil and its byproducts have shown that exposure to crude oil has negative effects in fish, such as reduced growth rates [47], decreased survival, and induction of development of ascites [48]. Interest in the toxicity of OSPW emerged in the 1960s after development began in the Alberta oil sands [28]. The toxicity of OSPW has thus far been mainly attributed to the NAs [20].

Frank et al. [49] proposed narcosis as the mechanism of toxicity of the NAs. Narcosis is believed to be a nonspecific process in which a hydrophobic chemical compound interacts with the lipid bilayer of the cell membrane, modifying its fluidity, tension, and thickness. Narcosis causes the disruption of the cell membrane and can induce cell death [50]. Surfactants are toxic because they cause membrane disruption and protein denaturation [51]. However, other mechanisms of toxicity have been reported. Immunological and endocrine disruption effects from NAs have been reported. The immune system of rainbow trout was depressed after being exposed to OSPW in-situ [52] and in laboratory facilities [53]. Furthermore, in both studies, the response against *Aeromonas salmonicida* infection was lower compared to the unexposed fish. Putative structures somewhat similar to estrone and estradiol have been reported in OSPW [54]. Male fathead minnows exposed to OSPW have been reported to have lower 11-keto-testosterone levels in plasma [55]. *Lithobates sylvaticus* tadpoles raised in young, tailings-affected wetlands fail to reach metamorphosis [56]. In Sprague-Dawley rats, NAs in high concentration have deleterious effects such as reduction of

body weight, and reduction of live-born offspring [57]. It is possible that compounds present in the OSPW act as endocrine disruptors.

1.2.5 *Silurana (Xenopus) tropicalis* as model organism

Amphibians are ectothermic vertebrates and include three lineages, Anura, Urodela, and Caecilia. Their skin is permeable to water and gases [58], making them susceptible to environmental toxicants. Anura represents the largest lineage of amphibians that currently includes ~6806 species around the world [59]. Frogs are key species in food webs as they are both major prey items and predators. They are highly susceptible to pollutants and endocrine disruptors [60]. The frog *S. tropicalis* is a member of the Pipidae family and the only Xenopus-type frog that has a diploid genome, with ten pairs of chromosomes [61]. In S. tropicalis, spawning can be artificially induced year-round [62], allowing easy work with embryos and tadpoles at all stages. Moreover, in 2010 the genome of *S. tropicalis* was published. The *S. tropicalis* genome with about 1.7×10^9 base pairs encodes more than 20,000 protein-coding genes exhibiting homology with many human genes [63].

1.3. Collaboration

The progress of this thesis involved the training and collaboration of colleagues from different institutions, including University of Ottawa, Environment and Climate Change Canada, University of Florida, and CINVESTAV Mexico. The collaborations are stated in each chapter of the thesis. Other contributions are recognized in the acknowledgment section.

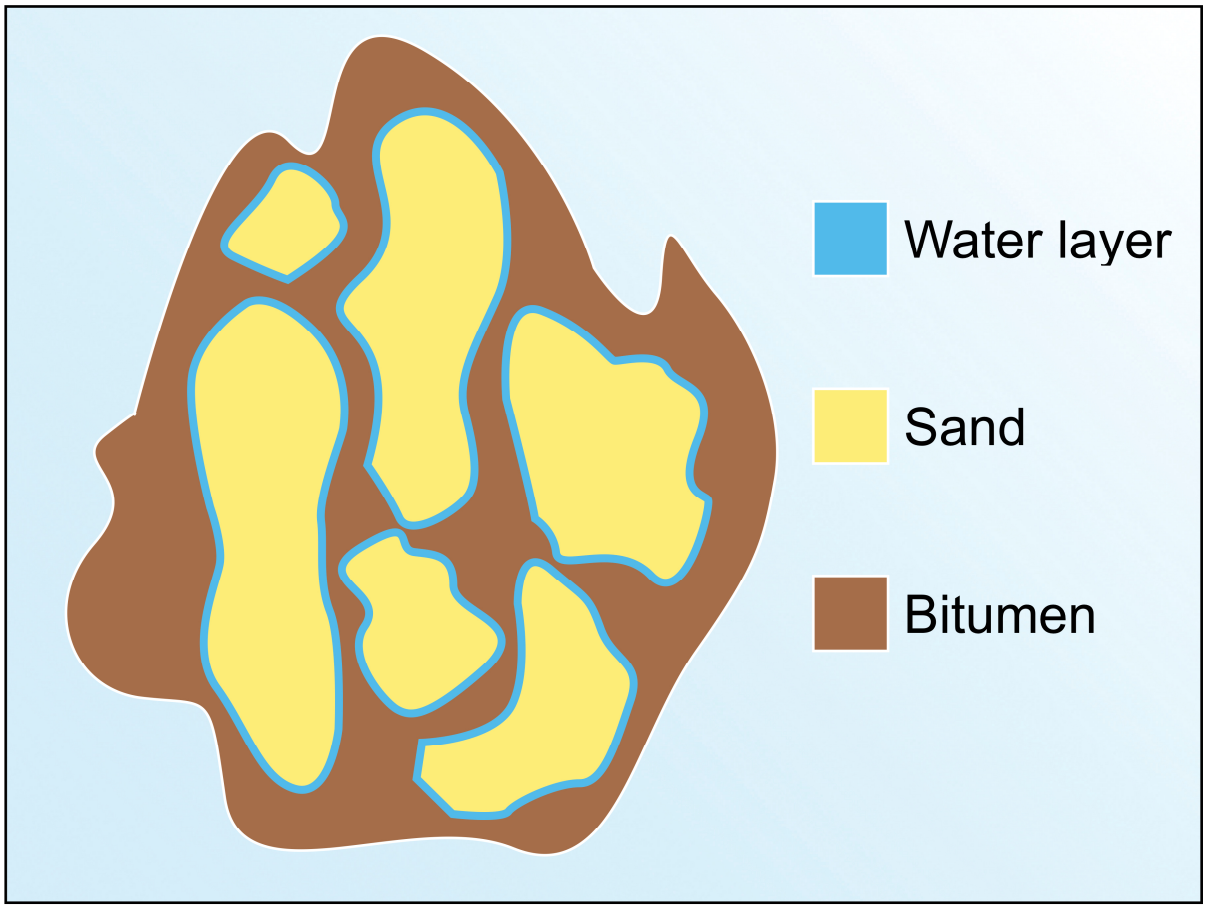


Figure 1.1. Illustration of the composition of bituminous sands.

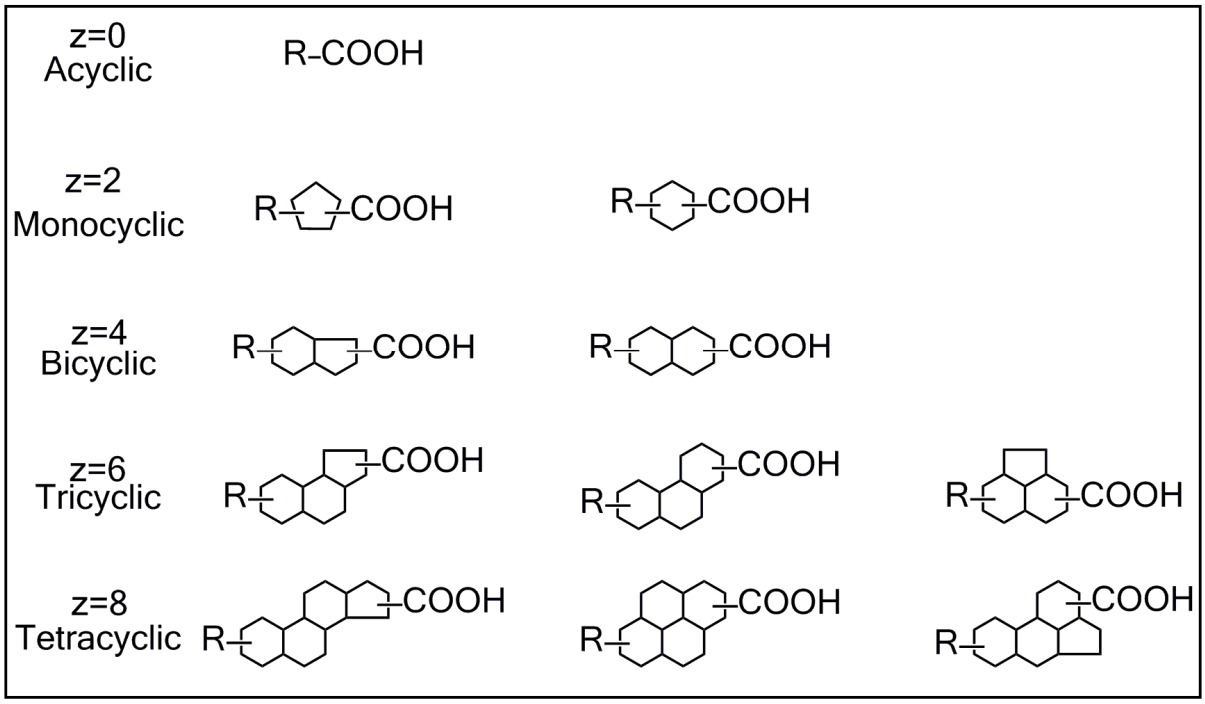


Figure 1.2. Conventional NAs structures grouped in z families. R=alkyl group.

Chapter 2

Naphthenic acid mixtures and acid-extractable organics extract impairs embryonic development of *Silurana (Xenopus) tropicalis*

2.1. Abstract

Naphthenic acids (NAs) are carboxylic acids naturally present in oil. The Canadian Oil Sands is a major source of NAs, where the water used in the extraction process, or oil sands process-affected water (OSPW), is stored in tailing ponds. The NAs have also been detected in sediment after substantial oil spills. In this chapter, I report the lethal and teratogenic effects of the exposure to an acid extractable organic (AEO) mixture from OSPW of the Canadian oil sands region. I used the frog model *Silurana (Xenopus) tropicalis* and the results were compared to the effects of two commercial NA extracts, Sigma 1 (S1) and Sigma 2 (S2). The 96h LC_{50} estimates were 10.4, 11.7 and 52.3 mg/L for S1, S2, and the AEOs, respectively. The 96h EC_{50} based on frequencies of developmental abnormalities was 2.1, 2.6 and 14.2 mg/L for S1, S2, and the AEOs, respectively. The main effects observed were reduced body size, cranial abnormalities, heart abnormalities, gut abnormality, edema, and ocular abnormalities. The severity and frequency of abnormalities increased with concentration ($p < 0.05$). Using electrospray ionization high-resolution mass spectrometry (ESI-HRMS), I determined that the O_2 species proportion were 98.8, 98.9 and 58.6% for S1, S2, and AEOs, respectively. The rank order potency based on LC_{50} and EC_{50} with and without normalization for the quantity of O_2 species was S1 > S2 > AEOs.

2.2. Introduction

Naphthenic acids (NAs) are carboxylic acid mixtures naturally occurring in crude oil and are classified as emergent contaminants [64]. They are of public interest due to their widespread use as wood and fabric preservatives, emulsifiers and surfactants, paint and ink driers. Besides their presence in households, NAs have been found in relatively high concentration in some sediments after oil spills. For example, during the Deepwater Horizon oil spill in the Gulf of Mexico, ~5 million barrels of oil were released [65]. The degradation of the oil hydrocarbons led to their conversion to carboxylic acids [66]. Following an oil spill in the Taean region of South Korea, concentrations of NAs in affected sediments reached as high as 130 mg/kg dry weight [67]. As over 100 countries produce oil [6], the potential toxicity that NAs present is of international concern.

Canada has the third-largest proven oil reserves in the world, with the oil sands representing 98% of these reserves [68]. The oil sands are non-conventional deposits containing bitumen mixed with sand, clay, and water. There are three oil sands deposit areas in Canada: Athabasca, Cold Lake and Peace River [68]. Encompassing an area of 142,000 square kilometers, these sites currently represent the largest commercially exploited oil sands deposits in the world [69]. Bitumen deposits have been reported in 22 other countries [3] and oil sands can be found in Venezuela, the United States, and Russia [69]. Approximately 20% of Alberta's oil sands can be exploited using the surface mining method [68]. The recovery process requires ~2.5 units of water for every 1 unit of bitumen produced [15]. This results in a mixture of contaminated water, sand,

bitumen and various petroleum-based chemicals destined for nearby settling ponds [70].

The tailing ponds are a storage area for the Oil Sands Process-affected Water (OSPW) containing bitumen and other contaminants. The tailings water contains high concentrations of inorganic salts, heavy metals, polycyclic aromatic hydrocarbons (PAHs), benzene, toluene, ethylbenzene, and xylene (BTEX compounds), residual bitumen and NAs [14,18,71]. During the oil extraction, NAs are solubilized and concentrated in the ponds. As of 2013, the area covered by oil sands tailing ponds was 77 square kilometers [70]. Some ponds are in close proximity to the Athabasca River [14], raising concerns regarding possible migration of pollutants to the surrounding and downstream environments. Frank et al. [17] compared NA profiles of groundwater in the Athabasca River area to locally produced OSPW and found important similarities, suggesting a common source, and concluded that tailings pond water is being released to the environment.

Naphthenic acids are a group of bioavailable compounds that present a risk for human and environmental well-being. Despite their relatively high concentrations in the environment, the effects of exposure to NAs are still unclear as there has been limited research. Therefore, this chapter aims to assess the toxicity and teratogenic effects of two closely related oil-derived extracts (S1 and S2) and an acid extractable organic (AEO) mixture from OSPW that is known to contain NAs, using western clawed frog *Silurana (Xenopus) tropicalis* embryos. Embryonic development is a window period sensitive for development [72] and could impact long-term phenotypic characteristics in animals [73]. Frogs are important indicator species due to their susceptibility to

pollutants in the water and their integral value in the food webs of habitats they occupy. Little is known about the effects of NAs and AEOs on the development of frogs and only a few articles regarding the toxicity of NAs on frogs have been published [56,74,75]. Moreover, the western clawed frog *S. tropicalis* could be used as a valuable tool to identify potential human developmental toxicants [76]. Furthermore, I used electrospray ionization high-resolution mass spectrometry (ESI-HRMS) for mixture characterization and concentration assessment of S1, S2, and the AEOs.

In September 2016, I sent aqueous solutions of NAs and AEOs to be analyzed by Dr. John V. Headley and Kerry M. Peru at Environment and Climate Change Canada. The chemical analyses were used in this chapter.

2.3. Material and methods

2.3.1. Chemicals

Human chorionic gonadotropin (hCG) was purchased from Millipore (Burlington, Massachusetts, United States). The salts necessary to prepare the FETAX solution (NaCl, $NaHCO_3$, KCl, $CaCl_2$, $CaSO_4$, and $MgSO_4$) were purchased from Fisher Scientific (Waltham, Massachusetts, United States) and Sigma-Aldrich (St. Louis, Missouri, United States). L-cysteine was acquired from Sigma-Aldrich. The two oil extracted NA mixtures, Sigma 1 (S1) and Sigma 2 (S2) were purchased from Sigma-Aldrich with lot number BCBC9959V and BCBK0736V, respectively. A sample of OSPW was collected in June 2015 and the resulting AEOs were extracted from 600 L following established procedures [77]. Shortly after collection, the OSPW was left to settle and the clarified OSPW was acidified to pH 2.5 using sulfuric acid. One L of OSPW was placed in a

funnel, 500 mL of dichloromethane was added and the mixture was manually shaken for a couple of minutes. The organic phase was collected and concentrated using a rotary evaporator set to 40 ºC. 150 mL of 0.1 N NaOH was added to the AEO. The pH was reduced to 10 using sulfuric acid. The solution was filtered using 1000 MW cut-off membranes (Millipore). The AEO mixture was stored at 4 ºC protected from light until usage for the exposures. The concentration of NAs in the AEOs was 8050 mg/L measured using ESI-MS in comparison to a different, previously characterized sample of OSWP-extract [77]. The term AEOs refers to the compounds obtained via acidification and organic extraction of OSPW. In addition to classic naphthenic acids, the AEO mixture likely contains other complex structures, such as diamondoid compounds [33], resin acids [34], and sulfur- and nitrogen-containing compounds [78].

2.3.2. Chemical analysis

The two commercial NA mixtures and the AEO extract were analyzed using ESI-HRMS. The analysis was in collaboration with Dr. Headley and Kerry Peru from the Watershed Hydrology and Ecology Research Division, Water Science and Technology Directorate, Environment and Climate Change Canada. Samples (5 µL) were introduced into the mass spectrometer by loop injection (flow injection analysis) using a Surveyor MS pump (Thermo Fisher Scientific Inc.) and a mobile phase of 50:50 acetonitrile/water containing 0.1% NH_4OH. Mass spectrometry analysis was carried out using a dual pressure linear ion trap–orbitrap mass spectrometer (LTQ Orbitrap Elite, Thermo Fisher Scientific, Bremen, Germany) equipped with an ESI interface operated in negative ion mode. Data was acquired in full scan mode from *m/z* 100 to 600 at a setting of 240,000 resolution (average mass resolving power (m/Δm50%) was 242,000 at *m/z* 400). Mass

accuracy error of <1 ppm was obtained using a lock mass compound (n-butyl benzenesulfonamide) for scan-to-scan mass calibration correction. Xcalibur version 2.2 software (Thermo Fisher Scientific San Jose, CA) was used for data acquisition, instrument operation, and semi-quantitative data analysis. Class distributions were established with acquired accurate mass data and Composer version 1.5.3 (Sierra Analytics, Inc. Modesto, CA) with mass accuracies of <2 ppm.

Samples with nominal concentrations of 0, 1, 5, 10, 50 and 100 mg/L (n = 3) were analyzed by ESI-HRMS as described above. The two commercial mixtures were used as standards for the quantification of NAs in their respective water samples. A previously characterized Athabasca OSPW large volume extract [77] was used as standard for the quantification of NA in the AEO samples. Linear equations forced to zero were built from the data. The coefficient of determination was 0.989, 0.994, and 0.964, for S1, S2, and AEOs, respectively. The nominal concentrations of the exposures were adjusted based on the linear equations. The adjusted values are herein presented as calculated concentrations. There is currently no standard method for the quantification of NAs, therefore the results are semi-quantitative.

2.3.3. Animal husbandry and breeding

All the bioassays were performed at the University of Ottawa in accordance with the guidelines of the Canadian Council of Animal Care and approved by the Animal Care and Veterinary Services. Originally obtained from the US Environmental Protection Agency [79], adult golden strain *S. tropicalis* frogs from our breeding colony were used. Frogs were fed daily with Nasco frog brittle for juvenile Xenopus and kept under a 12 h light/dark cycle and 25 ºC water temperature in a 21020 Tecniplast unit. Embryos of *S.*

tropicalis were obtained by injection of hCG into the posterior lymph sac to induce spawning [62]. They received a priming dose of 12.5 IU hCG and were moved to glass tanks with FETAX solution at 21 ± 1 ºC with males and females separated. After 24 h, they received a boosting dose of 100 IU hCG and breeding pairs were placed in glass tanks with FETAX solution at 21 ± 1 ºC. Embryos were collected and the jelly coat was removed with a two min exposure to 2% L-cysteine [76]. Fifty mL of test solution and 10 embryos at Nieuwkoop-Faber (NF) stage 9-10 were placed in each 10 cm petri dish.

2.3.4. Solution preparation

Test solutions were prepared daily from a stock stored at 4 ºC and protected from light. The stock solutions from S1 and S2 were prepared using ethanol as a solvent. The final concentration of ethanol in all treatments was 0.0025%. Concentrations of ethanol of 0.005% cause no observable effects on embryos of *S. tropicalis* [80]. The AEO extract contained 8050 mg/L NAs in a solution of 0.1 N of NaOH. The final concentration of NaOH in all of the treatments was 0.001 N. The pH of all treatments was 7.8 ± 0.1. The nominal concentrations tested for S1 and S2 were 0, 0.5, 1, 2, 3, 4, 6, 8, 12 and 24 mg/L. The nominal concentrations tested for AEOs were 0, 0.5, 1, 2, 3, 4, 6, 8, 12, 24, 48, and 96 mg/L. Two controls were used: FETAX solution and carrier solvent (ethanol or NaOH).

2.3.5. Toxicity assays

Glass Petri dishes were used as exposure vessels. Plastics were avoided as they are known to leach estrogenic chemicals [81]. Solutions were prepared using glass cylinders. All glassware was washed as follows: soaked for 15 min in hot water and detergent, rinsed twice with reverse osmosis water, later rinsed in a 10% HCl solution,

rinsed twice with reverse osmosis water, rinsed with acetone (Fisher Scientific, Certified ACS grade, 99.8%), rinsed twice with reverse osmosis water and kept at 120 ºC for 3 h. Seven replicates with ten embryos were used for each treatment. Exposures were conducted in a room maintained at 27 ± 1 ºC and 12:12h light:dark cycle. Tadpoles were observed every 24 h, dead embryos were removed, and solutions were fully renewed. Preliminary dose range-finding experiments were carried out to establish the final test concentrations. After exposures, digital images of the embryos were taken under a light stereomicroscope (Nikon SMZ 1500), with a Nikon DS-Fi1 camera and NIS Elements version 3.22.00 software. Images were used to extract morphometric data: tail length (TaL), snout-vent length (SVL), total length (TL) and interorbital distance (IOD). The abnormalities were classified according to the FETAX Atlas of abnormalities [82]. The frequency of the abnormalities was used for the calculation of the EC_{50} and EC_{20}. A tadpole was considered affected by the exposure if at least one type of abnormality was observed. The abnormalities observed were cranial abnormality, edema, heart, ocular and gut abnormalities.

2.3.6. Data processing and statistical analysis

Lethal concentration fifty (LC_{50}), lethal concentration twenty (LC_{20}), effect concentration fifty (EC_{50}), and effect concentration twenty (EC_{20}) were analyzed using a log vs. normalized response-variable slope model (Y=100/(1+10^((LogEC50-X)*HillSlope))) in GraphPad Prism 6. Morphological effects of the different mixtures on the *S. tropicalis* embryos were assessed using ImageJ™ 1.48 (http://imagej.nih.gov/ij/). Frequencies of abnormalities were analyzed as arcsine square root transformed data. The Shapiro-Wilk and Levene's tests were performed for data normality and

homogeneity of variances. To compare the control and solvent carrier, the normally distributed data were analyzed using t-tests and the non-normally distributed data were analyzed using Mann-Whitney tests. Data that met assumptions of normality and homoscedasticity were analyzed using ANOVA and Tukey's post-hoc test. Normally distributed data with unequal were analyzed using ANOVA Welch and Dunnet T3 as the post-hoc test. Non-normally distributed data were analyzed with Kruskal-Wallis tests, followed by the Student-Newman-Keuls (SKN) test. The alpha (α) level was set to 0.05. The statistical analyses were performed using SPSS V21 IBM and Sigma Plot 12.0.

2.4. Results and discussion

2.4.1. Chemical characterization

ESI-HRMS analysis showed that the AEO mixture contains a higher number of components, being more complex than S1 and S2 (Fig. 2.1). The two Sigma extracts (S1 and S2) had similar profiles; however, the proportion of the components between 150-250 *m/z* was different. Moreover, Sigma and Merichem mixtures have been reported to contain in addition to NAs, ~7% non-polar components with a mass spectra characteristic similar to alkanes [83].

The proportion of O_2 species was 98.8, 98.9 and 58.6% for S1, S2, and AEOs, respectively (Fig. 2.2). These O_2 species are represented by classical NAs with one carboxylic acid group denoted by the formula: $C_nH_{2n-z}O_2$ where *n* indicates the carbon number, *z* specifies the deficiency of hydrogen from the formation of rings and groups the NAs in a homologous series [22]. The S1 and S2 extracts are composed largely of the O_2 and O_4 classes. Other reports show a lower proportion of O_2 species in

commercial extracts, ranging from 78.5-86.1% and specifically another Sigma extract at 79.9% [84]. However, the AEO extract (Fig. 2.2) is composed mainly of O_2, O_3, O_4 (89.3%) and other species in a lower proportion (10.7%). The O_2 species proportion in the AEO extract was estimated as 58.6%, which is in the range of other extracts (50-82.5%) reported in the literature [30,84].

The double bond equivalences (DBE) were calculated using the equation DBE ($C_cH_hN_nO_oS_s$) = c-(h/2) + (n/2) + 1 [85]. The data obtained by ESI-HRMS in negative mode were principally deprotonated ions $[M-H]^-$. Therefore, the DBE values are half-integers and 0.5 is subtracted [86]. The DBE can be attributed to unsaturated bonds or cyclic structures. The DBE profile for O_2 species of S1 and S2 are similar. The S1 has a slightly higher proportion of components in the 1.5 group with some differences in the proportions of the groups DBE 2.5, 3.5, and 4.5 (Fig. 2.3). The DBE groups of 3.5 and 4.5 were more abundant in the AEOs. This indicates that S1 is mainly composed of aliphatic NAs followed by DBE 2.5 and 3.5 that are presumably single and double ring NAs. The extract S2 is also mostly composed of aliphatic NAs but followed by 2.5 DBE group that could be attributed to unsaturated bonds or single ring NAs. It has been previously reported that Sigma commercial mixtures are principally composed of aliphatic NAs [83,84]. However, the AEOs are mainly composed of 3.5 and 4.5 DBE, presumably double and triple ring NAs. In 2010, Barrow et al. [78] reported similar results in the analysis of an OSPW sample using atmospheric pressure photoionization (APPI) and electrospray ionization (ESI) Fourier transform ion cyclotron resonance mass spectrometry. In that case, DBE 3, 4, and 7 were the most abundant groups. The most abundant components in S1 and S2 were aliphatic NAs with 19 carbons. The

differences between the S1 and S2 were observed in the proportion of the aliphatic NAs with 9, 10 and 11 carbons (Fig. 2.4). However, the most abundant components in the AEO mixture were tentatively double ring structures with 13 carbons and triple ring structures with 14 carbons (Fig. 2.4).

2.4.2. Acute lethality

Lethal and sub-lethal effects of S1, S2 and the AEO mixture in *S. tropicalis* were assessed with a dose-response approach. The solvent carriers had no significant effect on survival, or frequency of teratogenic endpoints (Fig. 2.5-2.8). However, the exposure to the different extracts reduced survival after four days and induced teratogenesis (Fig. 2.9 a, d, and g). Based on the calculated $LC_{50S,}$ S2 was 1.1 times less toxic than S1, and the AEO was 5 times less toxic than S1 (Table 2.1). While the 96-h LC_{50} values of S1 and S2 are similar, however, the observed mortality pattern varied. For example, 100% mortality was achieved with the 24 mg/L for S1 after 48h and after 72 h 100% mortality was observed in the 12 mg/L treatment. In contrast, nearly 50% mortality was observed after 96 h for the 12 mg/L treatment for tadpoles exposed to S2 (Fig. 2.9 a and d). Similar lethal effects have been demonstrated in other model species. For example, the LC_{50} reported for fathead minnow (*Pimephales promelas*) is 2.5 mg/L for a different Sigma extract and the 96 h LC_{50} was 7.5-18.5 mg/L for different AEO extracts [84].

2.4.3. Sub-lethal effects

S1, S2, and AEO extract exposures significantly increased abnormality rates in embryos of *S. tropicalis* (Fig. 2.9 b, e and h). Based on calculated EC_{50S}, S2 was 1.2 times less toxic than S1, and S1 was 6.7 times more toxic than AEOs (Fig, 2.9, b, e and

h, Table 2.1). In fathead minnow embryos, an EC_{50} for hatching success of 2.5 and 7.5-18.5 mg/L has been reported for a Sigma extract and AEOs, respectively [84]. This indicates that fish and frogs may respond similarly to commercial NAs and AEOs. The EC_{20} analysis showed significant differences between S1 and S2, S1 being the most toxic extract we tested (Table 2.2). The exposure to S1, S2 and the AEO extract significantly decreased the TL, SVL, TaL, and, IOD (Fig. 2.10-2.12 a-d). The main effects observed include cranial abnormalities, heart abnormalities, gut abnormality, edema, and ocular abnormality. Edema, gut abnormalities, and heart abnormalities were particularly evident in the 9.3, 11.3 and 55.5 mg/L treatments of the S1, S2, and AEO extracts, respectively (Fig. 2.9, c, f and i). There was a positive relationship between treatment concentration and severity and frequency of abnormalities (Fig. 2.10-2.12). Furthermore, edema and cardiovascular abnormalities have been observed in zebrafish (*Danio rerio*) larvae after being exposed during seven days to AEOs extracted from oil sands from the Daqing (China) oil exploring area at a concentration of 2.5 mg/L [87]. The results here presented are also consistent with Marentette et al. [84]. They reported cardiovascular deformities in fathead minnows after exposure to naphthenic acid fraction components (NAFCs) until hatching. Marentette et al. [88] suggested oxidative stress as a possible toxic mechanism of naphthenic acid fraction components (NAFC) in walleye (*Sander vitreus*) embryos. Further research is needed to address the mechanism of NA toxicity, especially in relation to edema, heart, and gut development. In Chapter 3, I explored the response of the exposure to NA extracts in the transcriptional profile of *S. tropicalis*.

The HRMS analyses of the two commercial extracts (S1 and S2) indicate differences in carboxylic acid composition. Differences were observed in small MW NAs and in the relative proportions of various components ranging from MW 157-283 (Fig. 2.1 a-b). Slight differences between S1 and S2 were also observed in the DBE distribution (Fig. 2.3 a-b) and in the distribution of carbon numbers (Fig. 2.4 a-b). These differences may be linked to the different effects observed in tadpoles. According to the EC_{50} and EC_{20} of edema, cranial, heart and gut abnormalities, animals exposed to S1 have significantly more cranial defects ($p < 0.05$) (Table 2.3). The mixture S2 induced significantly more edema in the tadpoles compared to S1 and AEOs ($p < 0.05$) (Table 2.4). However, there were no significant differences between S1 and S2 in terms of heart and gut abnormalities ($p > 0.05$) (Table 2.5-2.6). The toxicity variation between the two commercial extracts could reflect variation in environmental samples that might have different origin and weathering history. Marentette et al. compared the profile of OSPW from two different sources, reporting differences in the proportion of 2 and 3 ring structures and small differences in the z family 8, 10, and 12 [84]. This indicates that the oil extraction process plays an important role in the group of compounds detected in the OSPW. In the same fashion, Brunswick et al. [89] reported differences in the percentage response of individual homologs between several Merichem, Acros, and Sigma NA mixtures. Therefore, the use of different extracts is expected to have different toxic effects according to its chemical profile. Commercial mixtures have been used previously as standards for concentration measurements of NAs [30,90], however, these results may be difficult to replicate due to variation in commercial extracts. This

also highlights the challenges with the analysis of toxicity and chemistry of AEO extracts since the origin, history, and resulting composition of samples vary among publications.

A possible solution to remedy the issue of differences in concentration and characterization of NAs and AEO mixtures would be to standardize a method for the extraction and measurement of the NA concentration. Another option would be an adjustment based on the O_2 species concentration. The O_2 family is represented by classical NAs with a single carboxylic acid group. Morandi et al. [91] reported that O_2 species are the main toxic NA species in OSPW. Furthermore, Huang et al. [92] determined that the proportion of O_2 species was a good parameter for the differentiation of the sample source, and normalized the results based on the O_2 species concentration. Based on the calculated concentration, S1 is the most toxic mixture being 1.1 times more toxic than S2 and 5 times more toxic than the AEO extract. However, when the data was adjusted to the O_2 family concentration, the toxicity of AEOs increases and becomes three times less toxic than S1 (Table 2.1, Fig. 2.13). The rank order potency with and without normalization to the content of O_2 species was S1 > S2 > AEO.

The LC_{50}, $EC_{50,}$ and EC_{20} values for the commercial extracts are lower than for the AEOs. This is in agreement with previous reports where fathead minnows embryos were exposed to fresh and aged OSPW, and commercial NAs [84]. The LC_{50}, EC_{50} and EC_{20} values for *S. tropicalis* embryos with and without O_2 species normalization are less than or within the concentrations reported in the tailing ponds and natural groundwater. Concentrations of NAs ranging from of 80-128 mg/L (High-performance liquid chromatography; HPLC) or 93-103 mg/L (Orbitrap MS) in OSPW have been reported in

tailing ponds [30]. The concentration in groundwater has been reported in some cases of up to 48 mg/L and in fresh water of up to 0.7 mg/L [17,28].

2.5. Conclusions

The chemical analysis of two commercial NAs and an AEO was obtained by ESI-HRMS. The spectral analysis revealed a higher complexity in the AEO sample compared to the commercial NAs. Similarities in the speciation, DBE profile, and the carbon distribution were observed in the two commercial NAs. The commercial blends were 99% composed of O_2 species. However, the AEOs O_2 content was ~60%. 4-day exposure of *S. tropicalis* to NAs and an AEO induced mortality (LC_{20}) at concentrations above 9.6, 9.5, and 41.2 mg/L for S1, S2, and AEOs, respectively. Exposure to NAs and AEOs also increased teratogenesis such as cranial abnormalities, heart abnormalities, gut abnormality, edema, and ocular abnormality, and reduced body size in *S. tropicalis*. These were observed (EC_{20}) above concentrations of 0.6, 1.3, 5.5 mg/L, for S1, S2, and AEOs, respectively. The chemical and morphometric analyses herein presented, represent one step towards understanding the effects of NAs, and may contribute to the development of environmental regulations and mitigation strategies for the management of OSPW.

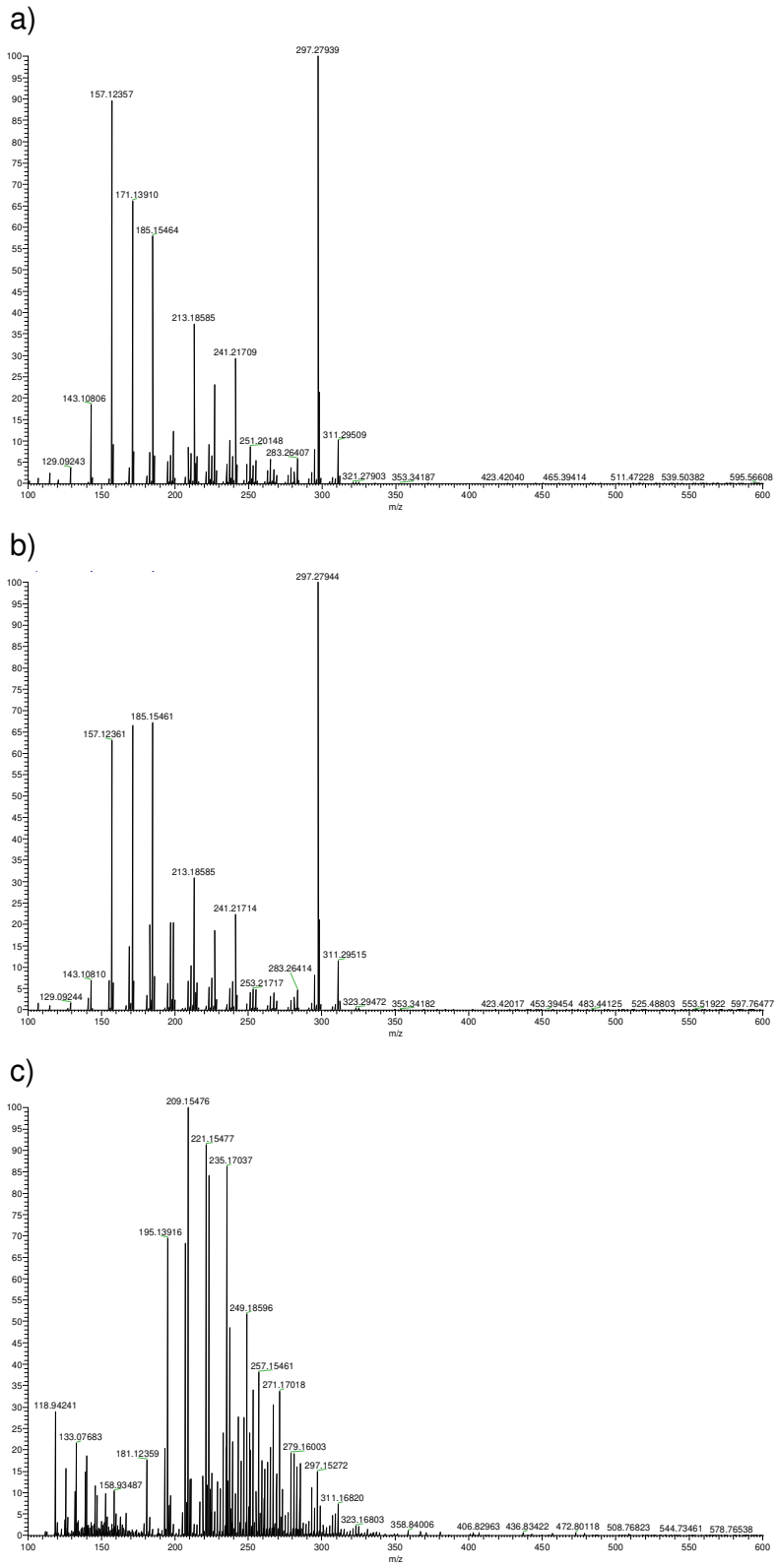


Figure 2.1. ESI-HRMS for S1 (a), S2 (b), and AEOs (c).

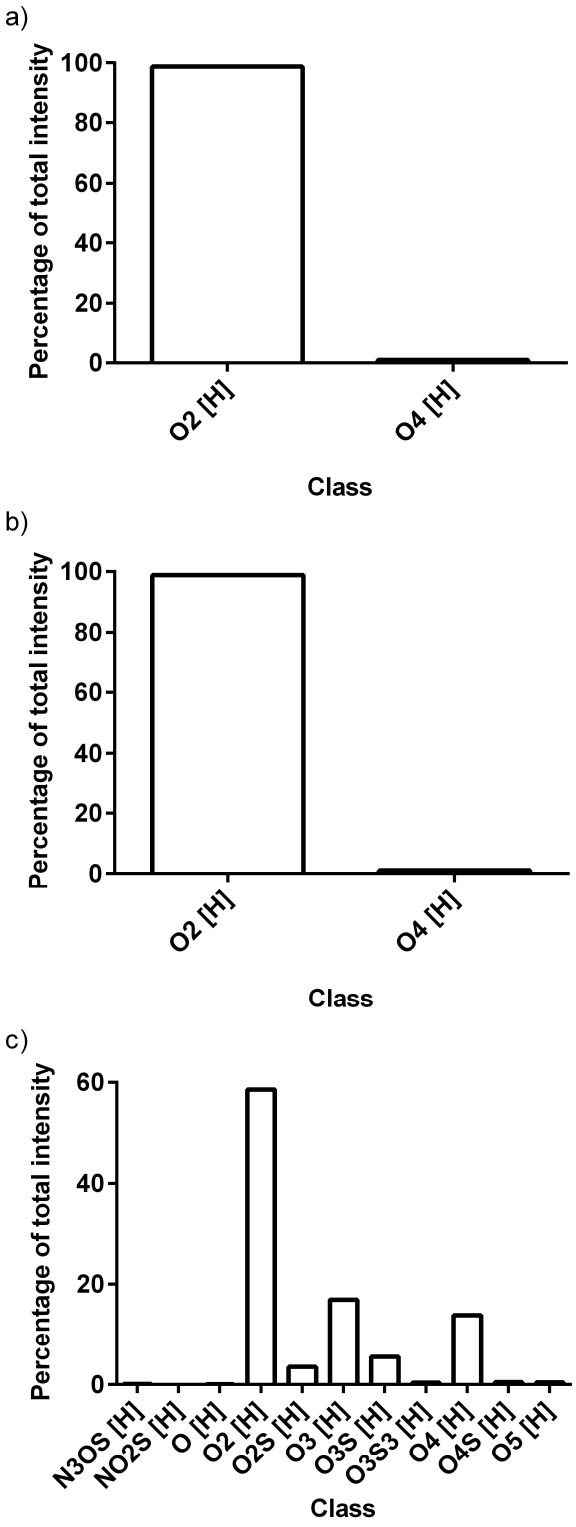


Figure 2.2. ESI-HRMS speciation profiles of S1 (a), S2 (b), and AEOs (c).

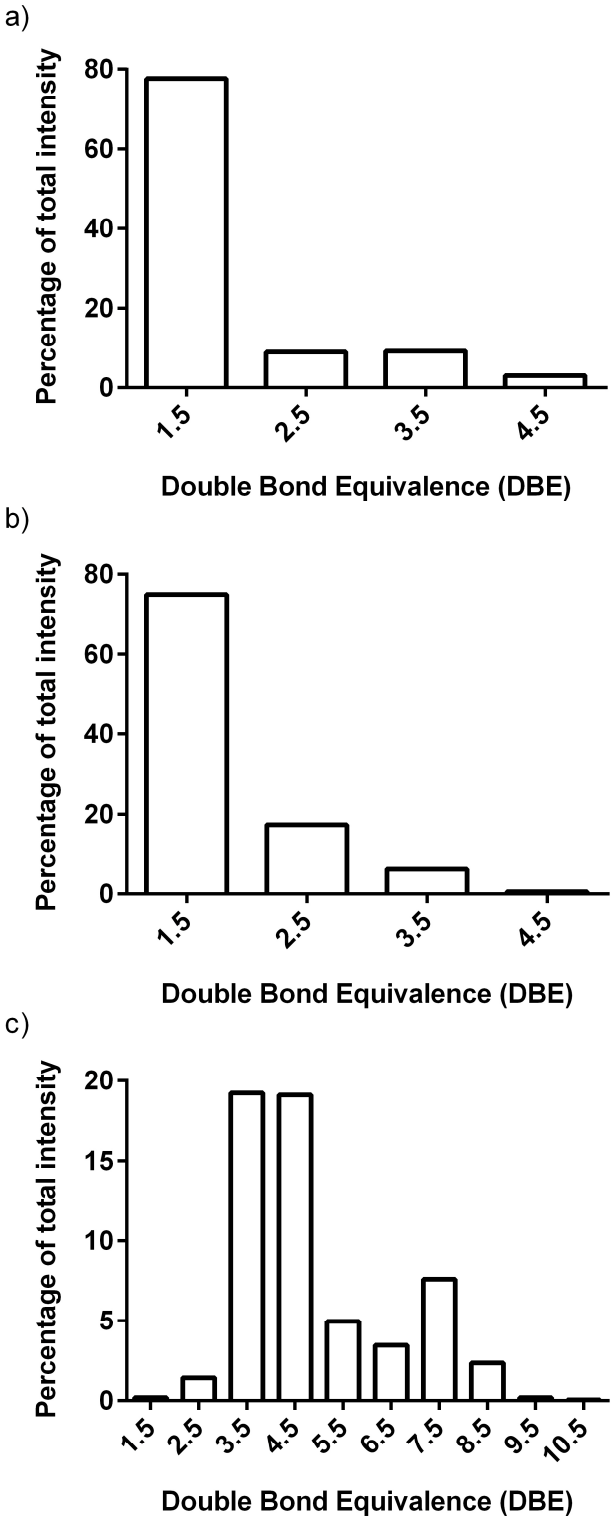


Figure 2.3. DBE distribution is given as a function of abundance for O_2 for S1 (a), S2 (b), and AEOs (c).

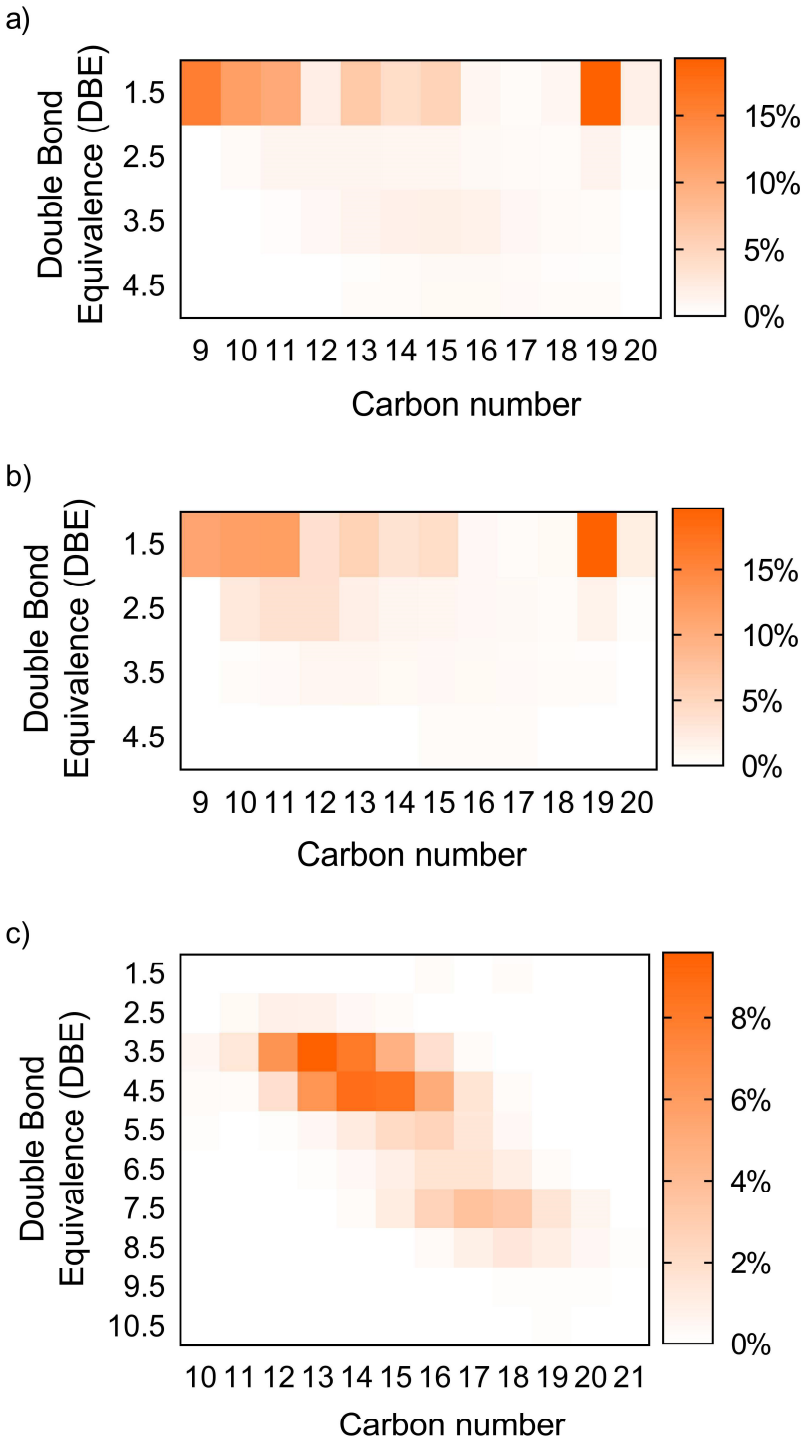


Figure 2.4. Analysis of the distribution of carbon numbers and DBE of NAs in S1 (a), S2 (b), and AEO (c). The values represent the percentage of total signal that accounts for a given carbon number of a given DBE family. The sum of all of the values equals 100%.

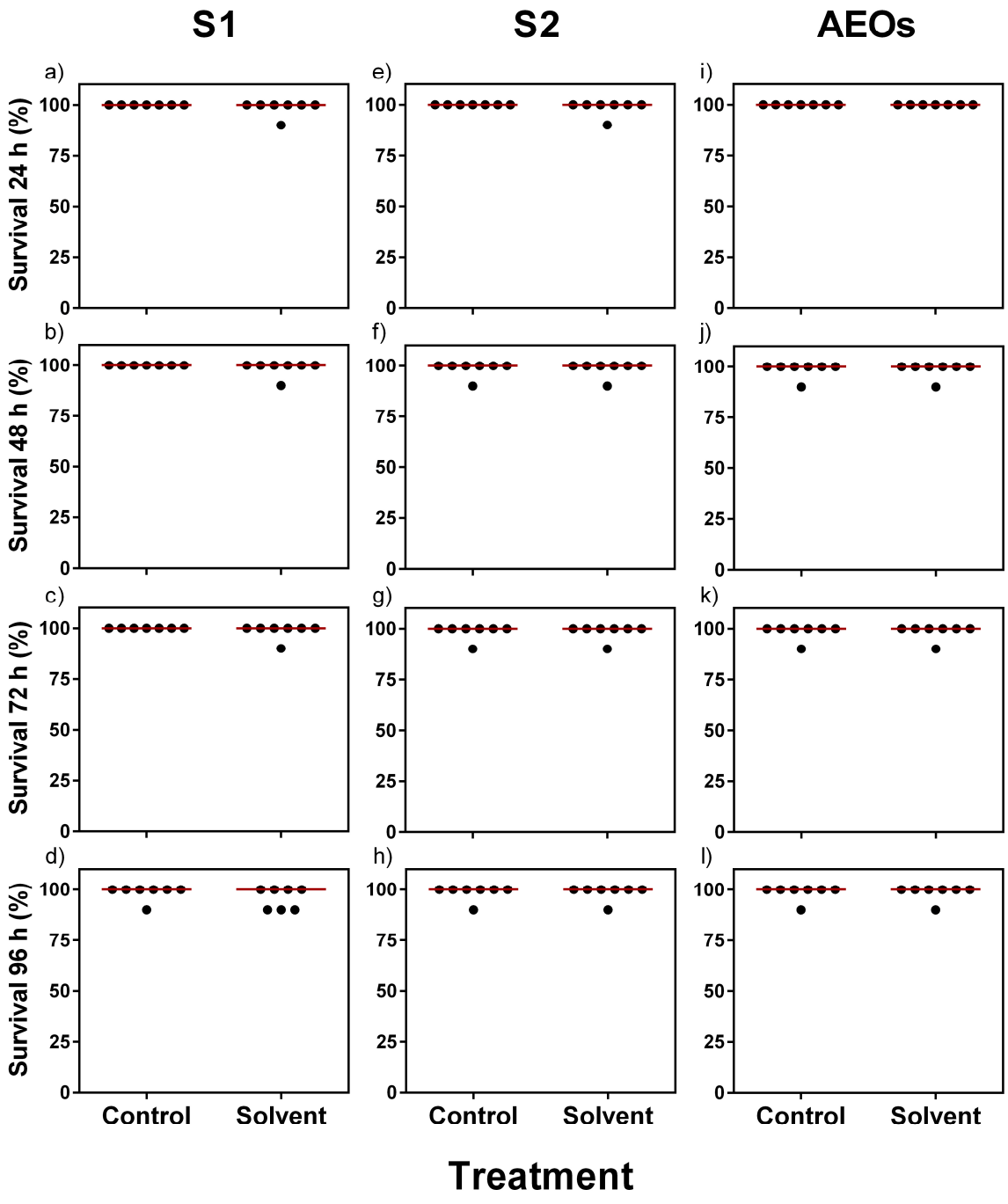


Figure 2.5. Survival rate of control against solvent treatments for 24, 48, 72, and 96 h, for the exposure of *S. tropicalis* embryos to S1 (a-d), S2 (e-h), and an AEOs (i-l). Ethanol at a final concentration of 0.0025% was used as a carrier solvent for S1 and S2. NaOH at a final concentration of 0.001 N was used as carrier solvent for AEOs. Each point represents a replicate and the red lines represent the median.

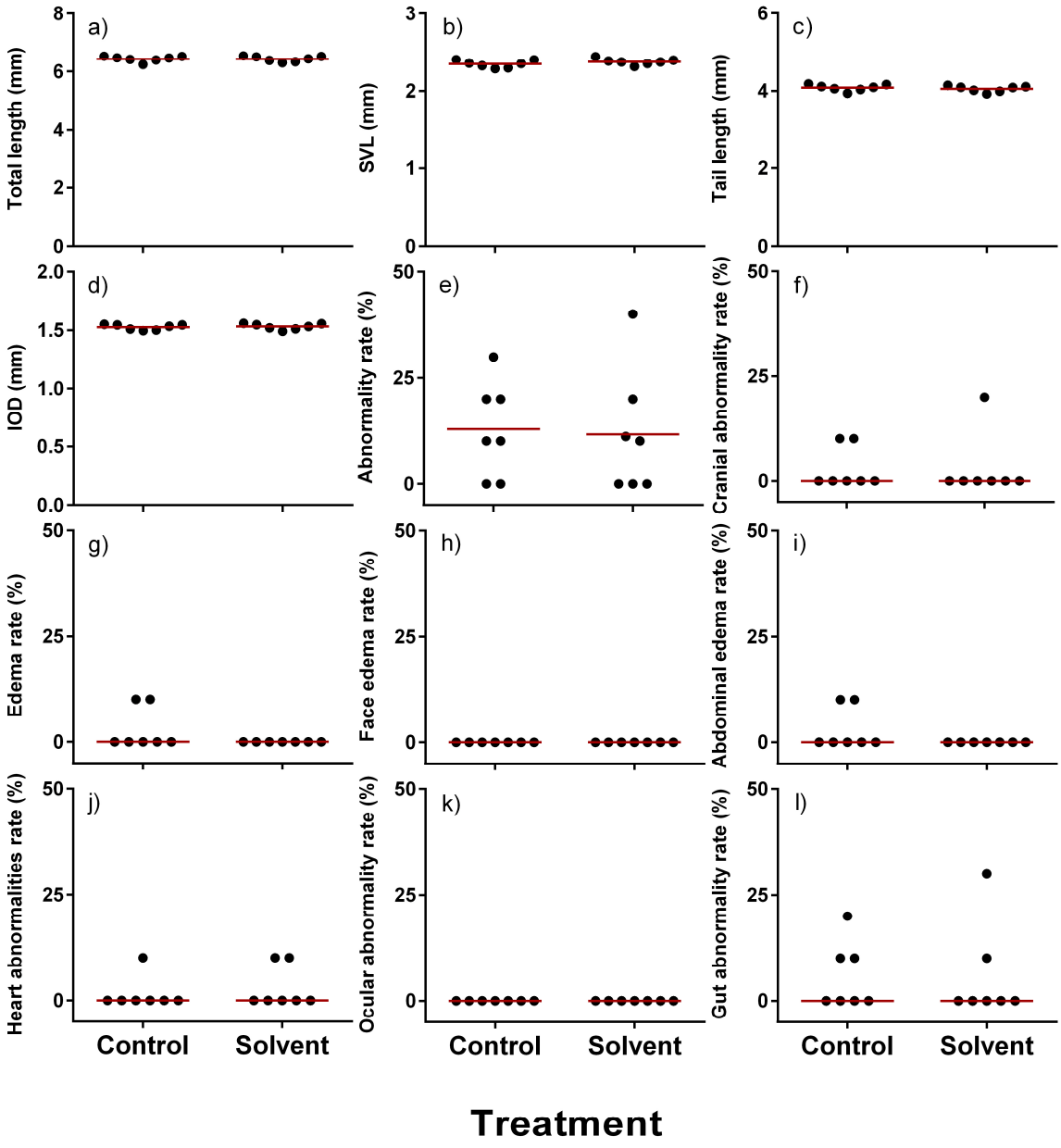


Figure 2.6. Effects of the solvent carrier on teratogenic endpoints in *S. tropicalis* embryos exposed to S1. The teratogenic endpoints were TL (a), SVL (b), TaL (c), IOD (d), abnormality rate (e), cranial abnormality (f), edema (g), face edema (h), abdominal edema (i), heart abnormality (j), ocular abnormality (k), and gut abnormality (l). Ethanol at a final concentration of 0.0025% was used as a carrier solvent. Each point represents a replicate. Red line represents the mean in a-e. Red line represents the median in f-l.

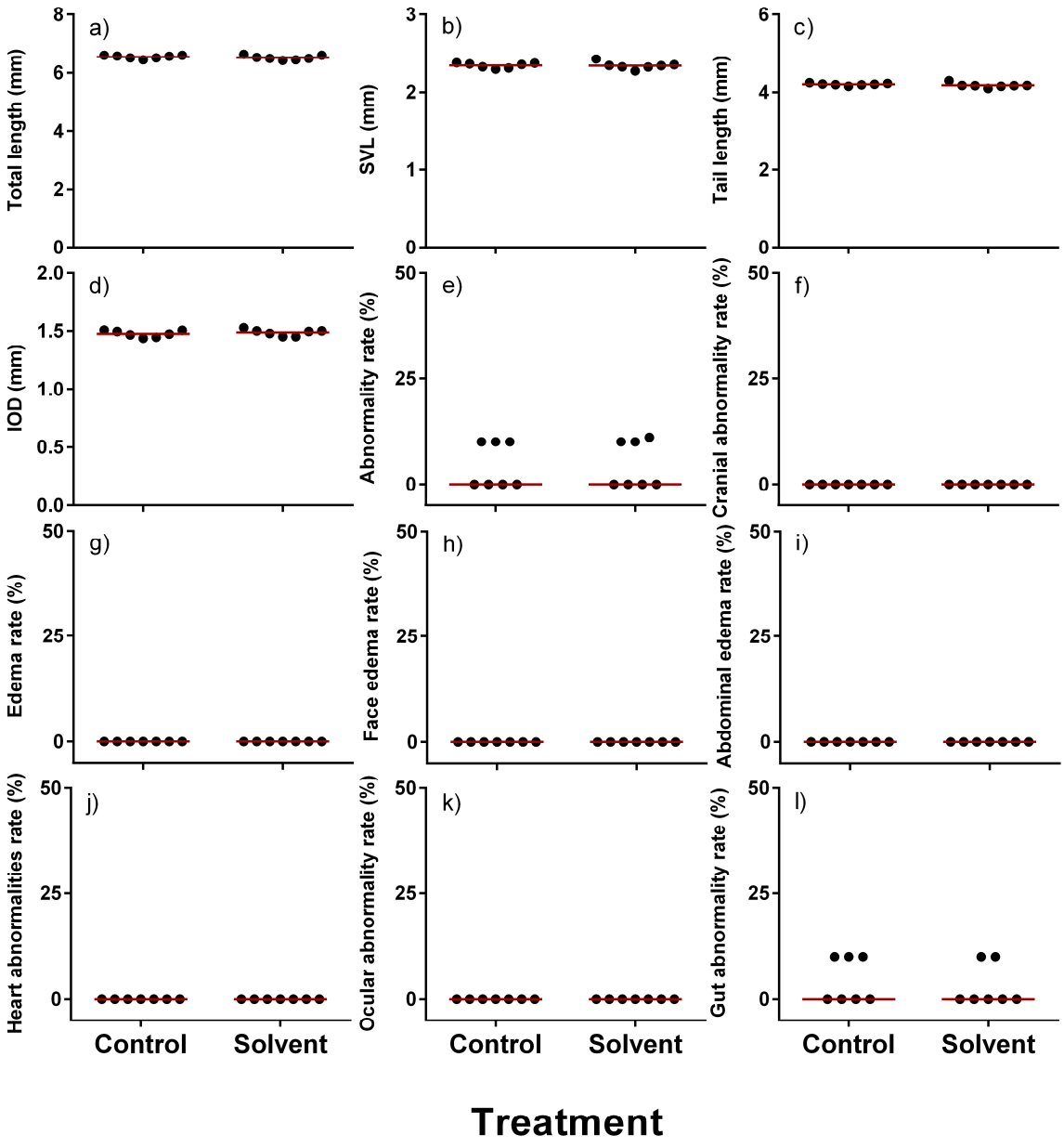


Figure 2.7. Effects of the solvent carrier on teratogenic endpoints in *S. tropicalis* embryos exposed to S2. The teratogenic endpoints were TL (a), SVL (b), TaL (c), IOD (d), abnormality rate (e), cranial abnormality (f), edema (g), face edema (h), abdominal edema (i), heart abnormality (j), ocular abnormality (k), and gut abnormality (l). Ethanol at a final concentration of 0.0025% was used as a carrier solvent. Each point represents a replicate. Red line represents the mean in a-d. Red line represents the median in e-l.

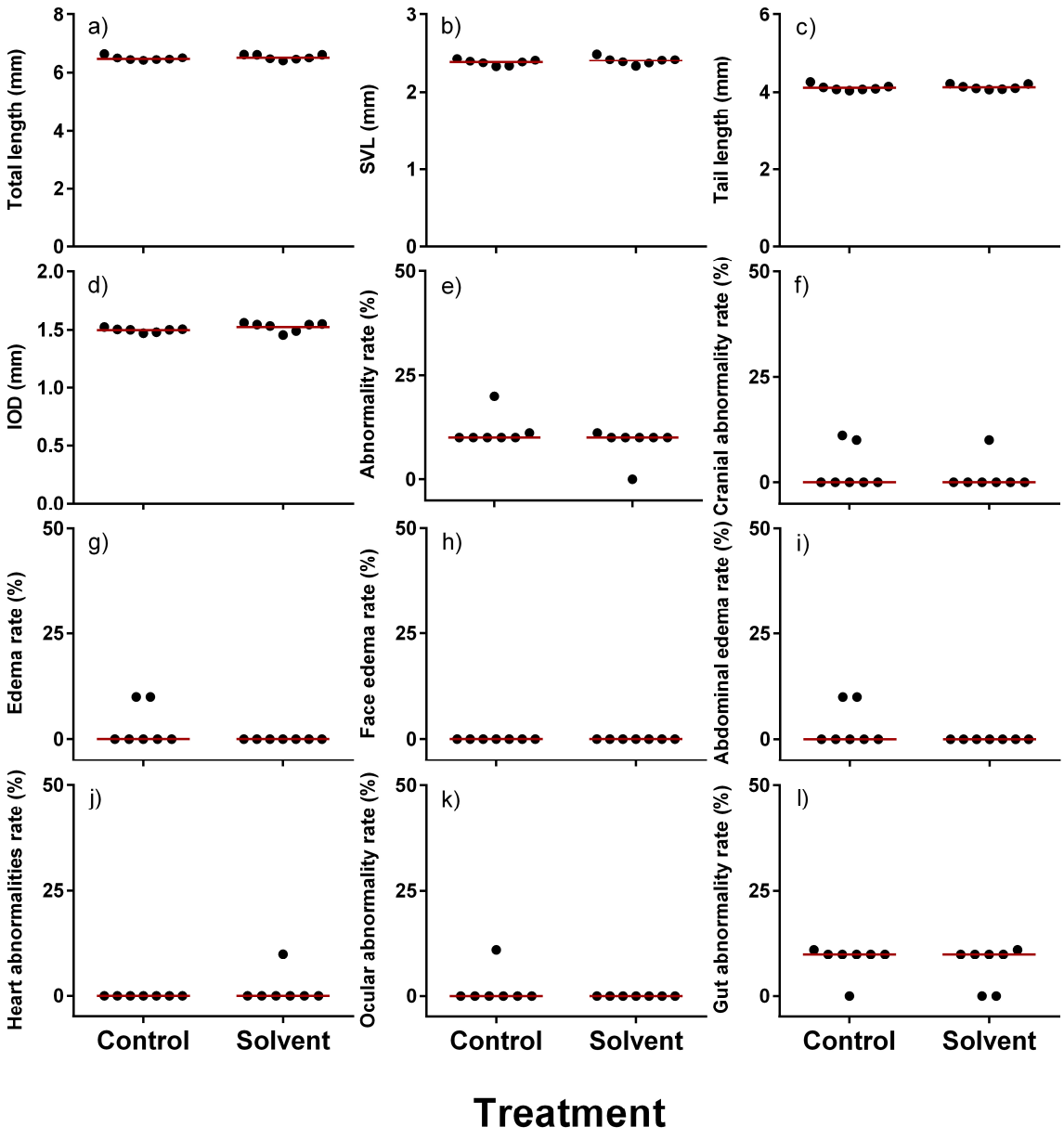


Figure 2.8. Effects of the solvent carrier on teratogenic endpoints in *S. tropicalis* embryos exposed to AEOs. The teratogenic endpoints were TL (a), SVL (b), TaL (c), IOD (d), abnormality rate (e), cranial abnormality (f), edema (g), face edema (h), abdominal edema (i), heart abnormality (j), ocular abnormality (k), and gut abnormality (l). NaOH at a final concentration of 0.001 N was used as carrier solvent. Each point represents a replicate. Red line represents the mean in a-d. Red line represents the median in e-l.

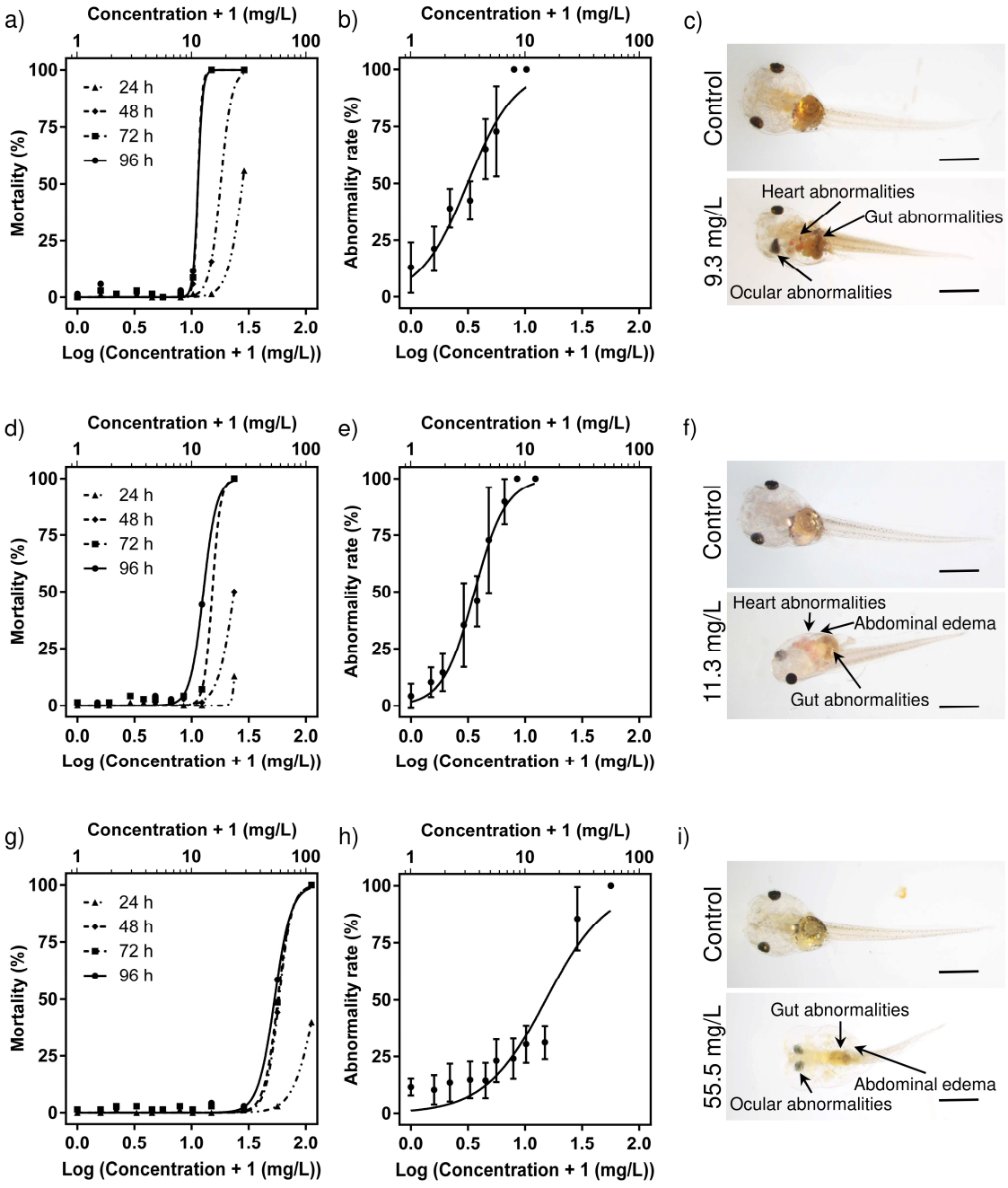


Figure 2.9. Concentration-response curves for mortality, abnormality rate and main teratogenic effects induced by exposure of *S. tropicalis* embryos to different mixtures of NAs, S1 (a-c), S2 (d-f), and AEOs (g-i) based on calculated concentrations. Each point represents the average of seven replicates and bars represent the standard deviation of the mean. The most commonly observed abnormalities were: edema, cranial malformation, gut malformation, heart abnormalities, and ocular malformation, as indicated by arrows. Scale bar is equal to 1 mm.

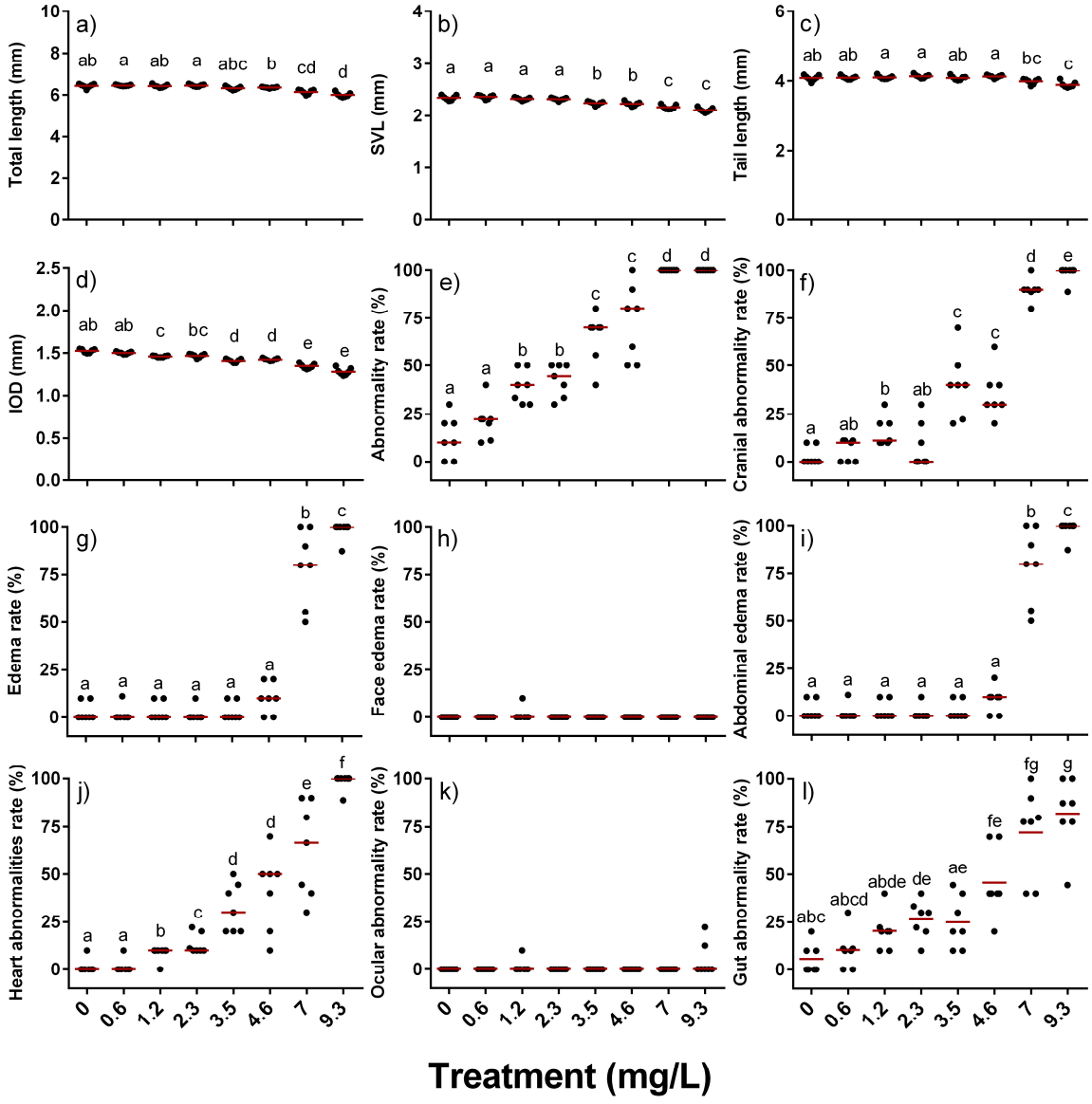


Figure 2.10. Teratogenic effects observed in *S. tropicalis* embryos exposed to S1. The teratogenic endpoints were TL (a), SVL (b), TaL (c), IOD (d), abnormality rate (e), cranial abnormality rate (f), edema rate (g), face edema rate (h), abdominal edema rate (i), heart abnormality rate (j), ocular abnormality rate (k), and gut abnormality rate (l). Different letters indicate significant difference ($p < 0.05$). Each point represents a replicate. Red line indicates the mean in a-d and l. Red line represents the median in e-k.

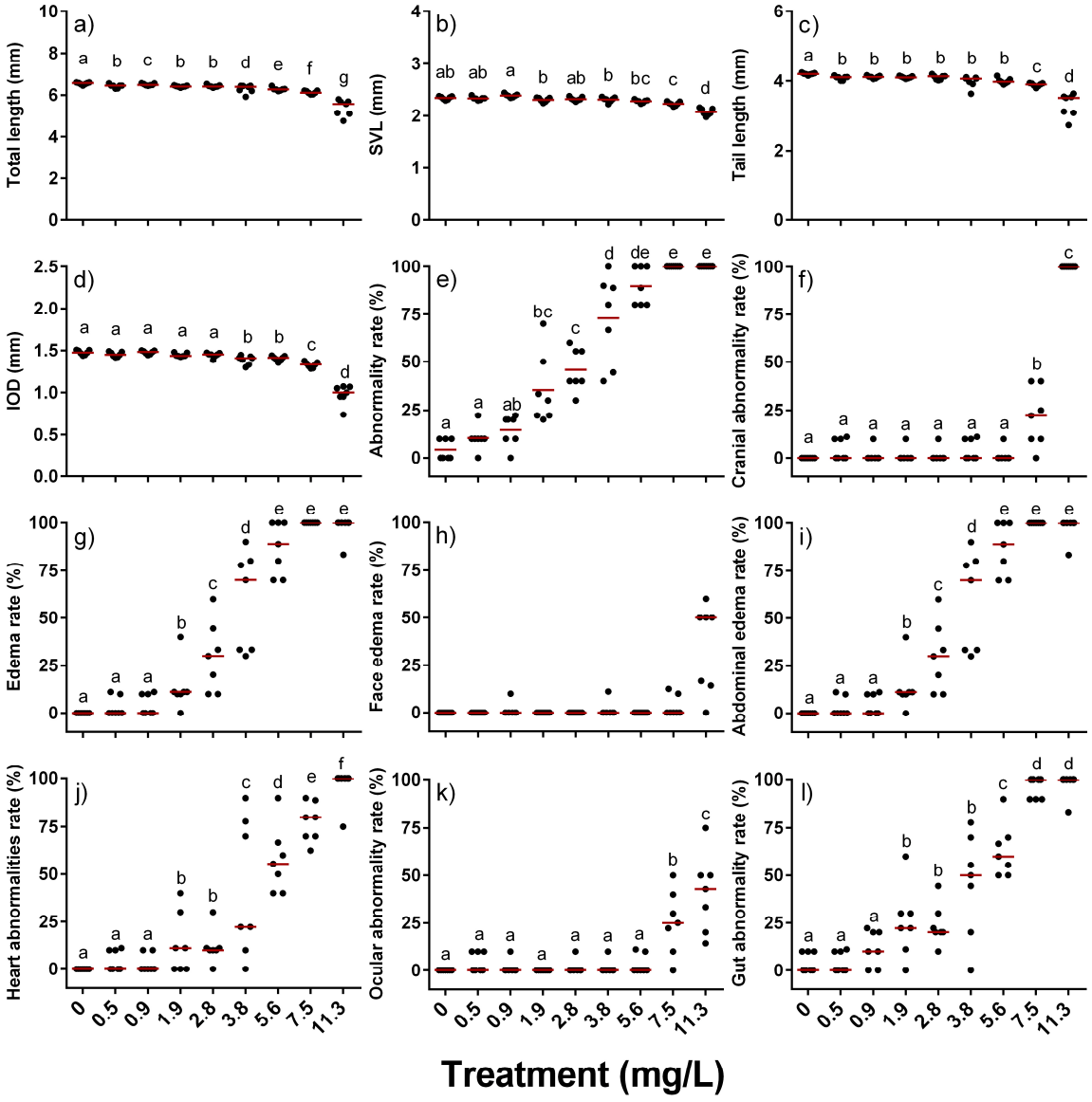


Figure 2.11. Teratogenic effects observed in *S. tropicalis* embryos exposed to S2. The teratogenic endpoints were TL (a), SVL (b), TaL (c), IOD (d), abnormality rate (e), cranial abnormality rate (f), edema rate (g), face edema rate (h), abdominal edema rate (i), heart abnormality rate (j), ocular abnormality rate (k), and gut abnormality rate (l). Different letters indicate significant difference ($p < 0.05$). Each point represents a replicate. Red line represents the mean in b and e. Red line represents the median in a, c, d, and f-l.

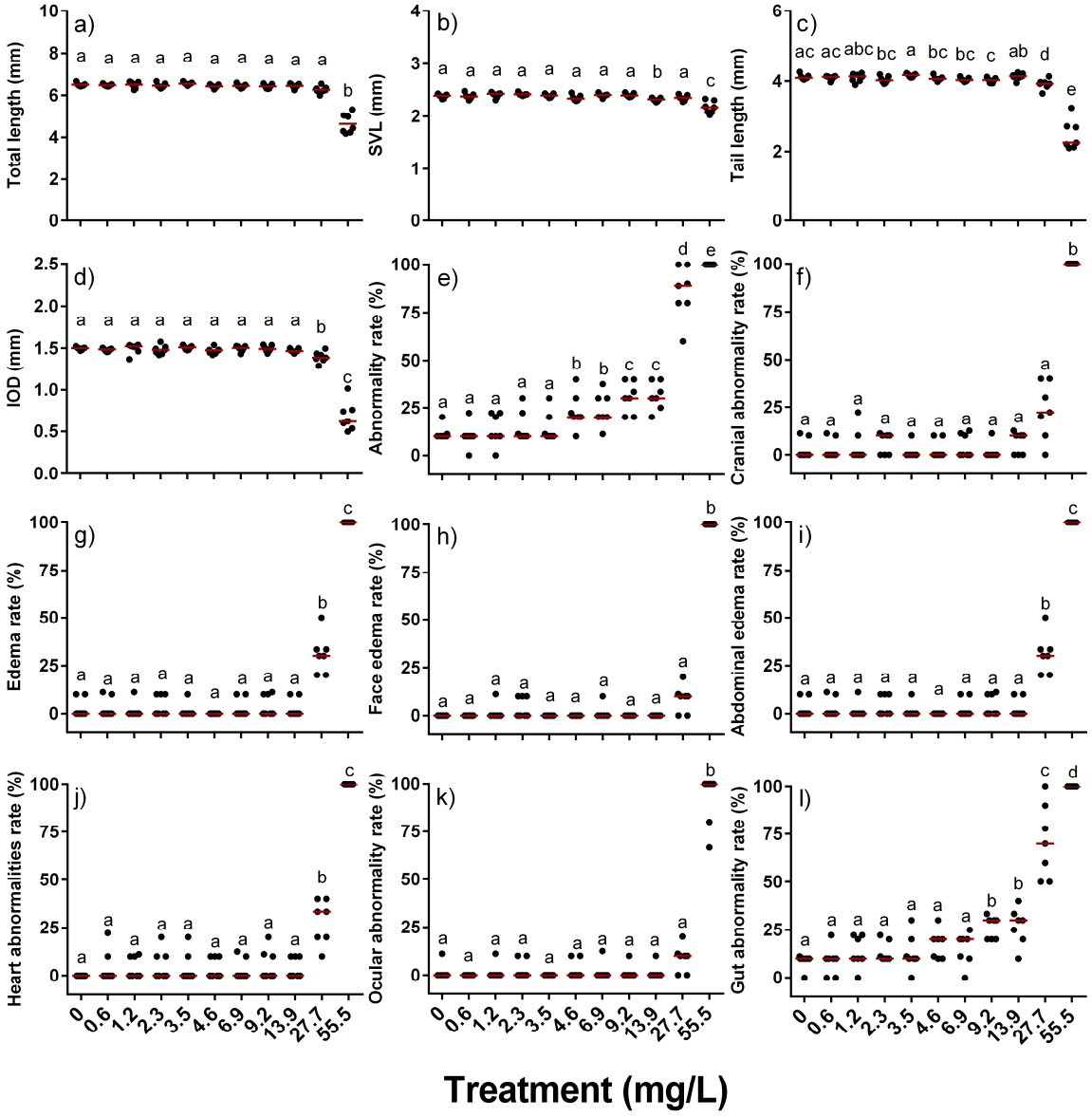


Figure 2.12. Teratogenic effects observed in *S. tropicalis* embryos exposed to AEOs. The teratogenic endpoints were TL (a), SVL (b), TaL (c), IOD (d), abnormality rate (e), cranial abnormality rate (f), edema rate (g), face edema rate (h), abdominal edema rate (i), heart abnormality rate (j), ocular abnormality rate (k), and gut abnormality rate (l). Different letters indicate significant difference ($p < 0.05$). Each point represents a replicate. Red line represents the mean in a. Red line represents the median in b-l.

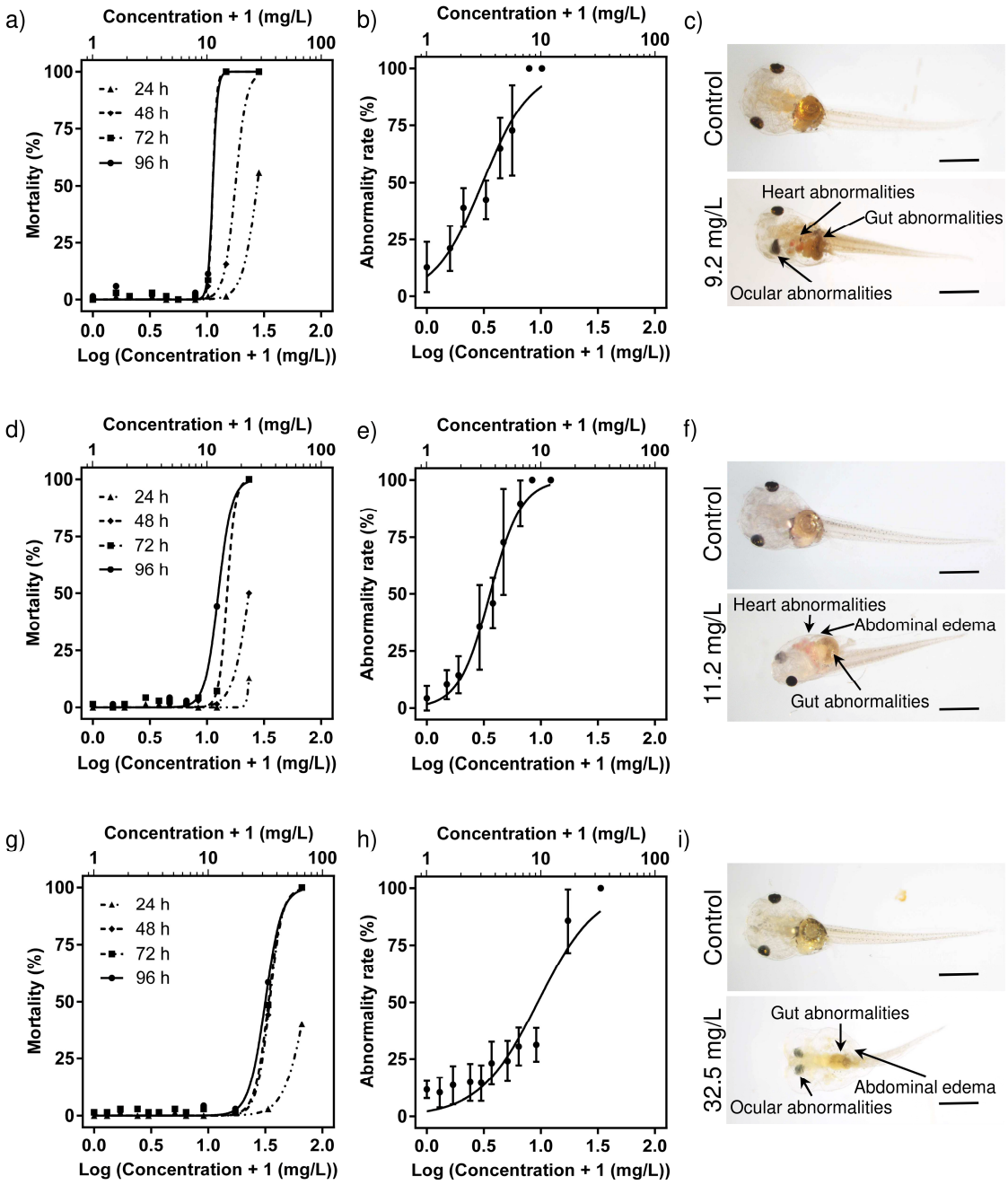


Figure 2.13. Concentration-response curves for mortality, abnormality rate and main teratogenic effects induced by exposure of *S. tropicalis* embryos to different mixtures of NAs, S1 (a-c), S2 (d-f), and AEOs (g-i). Concentrations were corrected based on O_2 family proportion. Each point represents the average of seven replicates and bars represent the standard deviation of the mean. The most commonly observed abnormalities were: edema, head malformation, gut malformation, heart abnormalities, and ocular malformation, as indicated by arrows. Scale bar is equal to 1 mm.

Table 2.1. Comparison of the LC_{50} and EC_{50} of S1 and S2 NA extracts and an AEO extract on *S. tropicalis* after a 4-day exposure, based on the nominal concentration, calculated concentration, and O_2 family concentration. Confidence interval 95% indicated between brackets. The coefficient of determination (R^2) calculated in GraphPad is also indicated. Different letters indicate values that are significantly different ($p < 0.05$).

	S1 LC_{50}	S2 LC_{50}	AEOs LC_{50}	S1 EC_{50}	S2 EC_{50}	AEOs EC_{50}
Nominal concentration (mg/L)	8.9^a	12.4^b	45.2^c	1.8^a	2.8^b	12.3^c
	[7.8, 10.2]	[12.0, 12.8]	[43.6, 46.9]	[1.6, 2.1]	[2.5, 3.0]	[10.8, 14.1]
	$R^2 = 0.987$	$R^2 = 0.961$	$R^2 = 0.966$	$R^2 = 0.859$	$R^2 = 0.875$	$R^2 = 0.846$
Calculated concentration (mg/L)	10.4^a	11.7^a	52.3^b	2.1^a	2.6^b	14.2^c
	[9.1, 11.8]	[11.3, 12.1]	[50.4, 54.2]	[1.9, 2.4]	[2.4, 2.9]	[12.4, 16.2]
	$R^2 = 0.987$	$R^2 = 0.961$	$R^2 = 0.966$	$R^2 = 0.860$	$R^2 = 0.908$	$R^2 = 0.844$
Concentration based on the O_2 families (mg/L)	10.2^a	11.6^a	30.6^b	2.1^a	2.6^b	8.4^c
	[8.9, 11.8]	[11.2, 12.0]	[29.5, 31.8]	[1.8, 2.4]	[2.4, 2.8]	[7.4, 9.5]
	$R^2 = 0.987$	$R^2 = 0.961$	$R^2 = 0.966$	$R^2 = 0.856$	$R^2 = 0.907$	$R^2 = 0.854$

Table 2.2. Comparison of the LC_{20} and EC_{20} of S1 and S2 NA extracts and an AEO extract on *S. tropicalis* after a 4-day exposure, based on the nominal concentration, calculated concentration, and O_2 family concentration. Confidence interval 95% indicated between brackets. The coefficient of determination (R^2) calculated in GraphPad is also indicated. Different letters indicate values that are significantly different ($p < 0.05$).

	S1 LC_{20}	S2 LC_{20}	AEOs LC_{20}	S1 EC_{20}	S2 EC_{20}	AEOs EC_{20}
Nominal concentration (mg/L)	8.3^a	10.2^a	35.6^b	0.5^a	1.4^b	4.7^c
	[8.1, 9.7]	[9.5, 10.7]	[31.4, 39.2]	[0.3, 0.7]	[1.1, 1.7]	[3.9, 5.6]
	$R^2 = 0.987$	$R^2 = 0.961$	$R^2 = 0.966$	$R^2 = 0.859$	$R^2 = 0.875$	$R^2 = 0.846$
Calculated concentration (mg/L)	9.6^a	9.5^a	41.2^b	0.6^a	1.3^b	5.5^c
	[9.4, 11.3]	[8.9, 10]	[36.2, 45.3]	[0.4, 0.9]	[1.1, 1.6]	[4.5, 6.5]
	$R^2 = 0.987$	$R^2 = 0.961$	$R^2 = 0.966$	$R^2 = 0.860$	$R^2 = 0.908$	$R^2 = 0.844$
Concentration based on the O_2 families (mg/L)	9.5^a	9.4^a	24.2^b	0.6^a	1.3^b	3.2^c
	[9.3, 11.2]	[8.8, 9.9]	[21.3, 26.6]	[0.3, 0.8]	[1.1, 1.6]	[2.6, 3.8]
	$R^2 = 0.987$	$R^2 = 0.961$	$R^2 = 0.966$	$R^2 = 0.856$	$R^2 = 0.907$	$R^2 = 0.854$

Table 2.3. Comparison of the EC_{50} and EC_{20} based on cranial abnormality rates of S1 and S2 NA extracts and an AEO extract on *S. tropicalis* after a 4-day exposure, based on the nominal concentration, calculated concentration, and O_2 family concentration. Confidence interval 95% indicated between brackets. The coefficient of determination (R^2) calculated in GraphPad is also indicated. Different letters indicate values that are significantly different ($p < 0.05$).

	S1 EC_{50}	S2 EC_{50}	AEOs EC_{50}	S1 EC_{20}	S2 EC_{20}	AEOs EC_{20}
Nominal concentration (mg/L)	3.9^a	8.7^b	27.5^c	2.5^a	8.0^b	23.5^c
	[3.6, 4.3]	[8.3, 10.0]	[25.4, 35.3]	[2.2, 2.9]	[7.5, 9.5]	[20.6, 31.6]
	$R^2 = 0.872$	$R^2 = 0.960$	$R^2 = 0.935$	$R^2 = 0.872$	$R^2 = 0.960$	$R^2 = 0.935$
Calculated concentration (mg/L)	4.5^a	8.1^b	31.7^c	2.9^a	7.5^b	27.1^c
	[4.2, 4.9]	[7.8, 9.4]	[29.3, 40.7]	[2.5, 3.3]	[7.0, 8.9]	[23.8, 36.4]
	$R^2 = 0.875$	$R^2 = 0.960$	$R^2 = 0.935$	$R^2 = 0.875$	$R^2 = 0.960$	$R^2 = 0.935$
Concentration based on the O_2 families (mg/L)	4.5^a	8.1^b	18.6^c	2.9^a	7.0^b	16.0^c
	[4.1, 4.9]	[7.7, 9.3]	[17.3, 24.0]	[2.5, 3.3]	[7.4, 8.7]	[14, 21.4]
	$R^2 = 0.869$	$R^2 = 0.960$	$R^2 = 0.935$	$R^2 = 0.869$	$R^2 = 0.960$	$R^2 = 0.935$

Table 2.4. Comparison of the EC_{50} and EC_{20} based on edema rates of S1 and S2 NA extracts and an AEO extract on *S. tropicalis* after a 4-day exposure, based on the nominal concentration, calculated concentration, and O_2 family concentration. Confidence interval 95% indicated between brackets. The coefficient of determination (R^2) calculated in GraphPad is also indicated. Different letters indicate values that are significantly different ($p < 0.05$).

	S1 EC_{50}	S2 EC_{50}	AEOs EC_{50}	S1 EC_{20}	S2 EC_{20}	AEOs EC_{20}
Nominal concentration (mg/L)	5.2^b	3.7^a	26.5^c	4.4^b	2.5^a	22.3^c
	[4.9, 5.3]	[3.4, 3.9]	[25.1, 32.2]	[4.1, 4.7]	[2.2, 2.8]	[19.6, 23.3]
	$R^2 = 0.954$	$R^2 = 0.918$	$R^2 = 0.964$	$R^2 = 0.954$	$R^2 = 0.918$	$R^2 = 0.964$
Calculated concentration (mg/L)	6.0^b	3.5^a	30.5^c	5.1^b	2.3^a	25.8^c
	[5.7, 6.2]	[3.2, 3.7]	[29.0, 37.2]	[4.8, 4.4]	[2.1, 2.6]	[22.6, 26.9]
	$R^2 = 0.954$	$R^2 = 0.918$	$R^2 = 0.964$	$R^2 = 0.954$	$R^2 = 0.918$	$R^2 = 0.964$
Concentration based on the O_2 families (mg/L)	5.9^b	3.4^a	18.0^c	5.0^b	2.3^a	15.2^c
	[5.7, 6.2]	[3.2, 3.6]	[17.1, 21.9]	[4.7, 5.3]	[2.1, 2.6]	[13.3, 15.8]
	$R^2 = 0.954$	$R^2 = 0.918$	$R^2 = 0.964$	$R^2 = 0.954$	$R^2 = 0.918$	$R^2 = 0.964$

Table 2.5. Comparison of the EC_{50} and EC_{20} based heart abnormality rates of S1 and S2 NA extracts and an AEO extract on *S. tropicalis* after a 4-day exposure, based on the nominal concentration, calculated concentration, and O_2 family concentration. Confidence interval 95% indicated between brackets. The coefficient of determination (R^2) calculated in GraphPad is also indicated. Different letters indicate values that are significantly different ($p < 0.05$).

	S1 EC_{50}	S2 EC_{50}	AEOs EC_{50}	S1 EC_{20}	S2 EC_{20}	AEOs EC_{20}
Nominal concentration (mg/L)	4.3^a	5.1^a	26.9^b	2.5^a	3.0^a	22.7^b
	[3.9, 4.7]	[4.6, 5.7]	[25.2, 34.2]	[2.1, 2.9]	[2.5, 3.5]	[19.5, 23.7]
	$R^2 = 0.850$	$R^2 = 0.830$	$R^2 = 0.931$	$R^2 = 0.850$	$R^2 = 0.830$	$R^2 = 0.931$
Calculated concentration (mg/L)	5.0^a	4.8^a	31.1^b	2.9^a	2.8^a	26.2^b
	[4.5, 5.5]	[4.3, 5.3]	[29.1, 39.4]	[2.4, 3.4]	[2.3, 3.3]	[22.6, 27.3]
	$R^2 = 0.850$	$R^2 = 0.831$	$R^2 = 0.931$	$R^2 = 0.850$	$R^2 = 0.831$	$R^2 = 0.931$
Concentration based on the O_2 families (mg/L)	4.9^a	4.7^a	18.3^b	2.9^a	2.8^a	15.4^b
	[4.5, 5.4]	[4.3, 5.3]	[17.1, 23.2]	[2.4, 3.3]	[2.3, 3.3]	[13.3, 16.1]
	$R^2 = 0.848$	$R^2 = 0.829$	$R^2 = 0.931$	$R^2 = 0.848$	$R^2 = 0.829$	$R^2 = 0.931$

Table 2.6. Comparison of the EC_{50} and EC_{20} based on gut abnormality rates of S1 and S2 NA extracts and an AEO extract on *S. tropicalis* after a 4-day exposure, based on the nominal concentration, calculated concentration, and O_2 family concentration. Confidence interval 95% indicated between brackets. The coefficient of determination (R^2) calculated in GraphPad is also indicated. Different letters indicate values that are significantly different ($p < 0.05$).

	S1 EC_{50}	S2 EC_{50}	AEOs EC_{50}	S1 EC_{20}	S2 EC_{20}	AEOs EC_{20}
Nominal concentration (mg/L)	4.1^a	4.2^a	15.2^b	1.7^a	2.3^a	6.4^b
	[3.6, 4.7]	[3.8, 4.7]	[13.4, 17.4]	[1.2, 2.2]	[1.9, 2.7]	[5.3, 7.6]
	$R^2 = 0.740$	$R^2 = 0.852$	$R^2 = 0.842$	$R^2 = 0.740$	$R^2 = 0.852$	$R^2 = 0.842$
Calculated concentration (mg/L)	4.7^a	4.0^a	17.5^b	2.0^a	2.1^a	7.4^b
	[4.1, 5.5]	[3.6, 4.4]	[15.5, 20.1]	[1.4, 2.5]	[1.7, 2.6]	[6.2, 8.7]
	$R^2 = 0.739$	$R^2 = 0.853$	$R^2 = 0.840$	$R^2 = 0.739$	$R^2 = 0.853$	$R^2 = 0.840$
Concentration based on the O_2 families (mg/L)	4.7^a	3.9^a	10.3^b	1.9^a	2.1^a	4.3^b
	[4.1, 5.4]	[3.5, 4.4]	[9.1, 11.8]	[1.4, 2.5]	[1.7, 2.5]	[3.6, 5.0]
	$R^2 = 0.739$	$R^2 = 0.852$	$R^2 = 0.848$	$R^2 = 0.739$	$R^2 = 0.852$	$R^2 = 0.848$

Chapter 3

Transcriptome analysis reveals that naphthenic acids perturb gene networks related to metabolic processes, membrane function, and gut function in *Silurana (Xenopus) tropicalis* embryos

3.1. Abstract

This chapter aims to identify potential mechanisms of toxicity of naphthenic acids (NA) mixtures in *Silurana (Xenopus) tropicalis* embryos. The sublethal transcriptomic responses following exposure to 2 mg/L oil-derived Sigma commercial NA extracts (S1 and S2) were assessed using a custom oligonucleotide microarray. Both NA mixtures induced embryonic abnormalities including edema and cardiac and gut abnormalities. Exposure to NAs also decreased morphometric parameters, such as total length, tail length, and interorbital distance. Microarray analysis using a custom 4 x 44 K Agilent platform showed that 3,656 common genes were differentially expressed ($p < 0.05$) in embryos exposed to S1 and S2 compared to the control group. Gene ontology analysis revealed that 18 biological processes, five cellular components, and 19 molecular functions were significantly enriched after both S1 and S2 exposures ($p < 0.05$). Pathways enriched by sub-network enrichment analysis (SNEA) were related to the observed phenotypic abnormalities such as gut function, edema, and cartilage differentiation. Other notable networks affected by NAs included metabolism, cell membrane function, and metabolism. In a separate dose-response experiment, the expression of key genes identified by microarray (*cyp4b1*, *abcg2*, *slc26a6*, *eprs*, and *slc5a1*) was measured by real-time PCR in *S. tropicalis* embryos exposed to NAs. In general, the RT-qPCR data supported the microarray data. In *S. tropicalis* embryos

exposed to acid-extractable organics from a tailings pond from Canada's oil sands, the expression levels of *eprs* (*bifunctional glutamate/proline-tRNA ligase*) and *slcs5a1* (sodium/glucose cotransporter 1) were significantly decreased compared to the controls. These changes are likely indicative of increased edema and disrupted gut function, respectively. These data suggest that NAs have multiple modes of action to induce developmental toxicity in amphibians. In addition, some mechanisms of toxicity may be shared between commercial NAs and AEOs.

3.2. Introduction

Interest in the toxicity of the oil sands process-affected water (OSPW) emerged in the 1960s after the initiation of the oil extraction activities in the oil sands region of Alberta, Canada [28]. Canada has the third-largest proven oil reserves in the world [93], with oil sands representing 98% of these resources [94]. In oil sands, oil is present in the form of bitumen mixed with inorganic material [94]. The oil extraction processes require up to 2.5 barrels of water to produce one barrel of crude oil [15]. Since 1967, almost 895 km^2 had been disturbed by bitumen extraction activities [95]. Moreover, the OSPW is kept in separate tailing ponds under a zero-discharge policy. As of 2013, tailings ponds covered 77 km^2 [70]. In 2014, Frank et al. [17] found important similarities between the chemical profile of the groundwater of the Athabasca River to OSPW and, concluding that OSPW is contaminating the environment.

The toxicity of the OSPW has mainly been attributed to the presence of naphthenic acids (NAs) [20]. The term NAs refers to all the carboxylic acids present in crude oil and these are now classified as water contaminants of emerging concern [64].

Naphthenic acids are used in the production of metal salts and as wood preservative [22]. They have been also detected in sediments after oil spills [66,67]. In 2008, Frank et al. [49] proposed narcosis as the mechanism of toxicity of the NAs. In 2017, Marentette et al. [88] suggested oxidative stress as a possible toxic mechanism of naphthenic acid fraction components (NAFC) of OSPW affecting walleye (*Sander vitreus*) embryos. There are few published articles on the mechanisms of toxicity of NAs in aquatic organisms. Therefore, it is necessary to assess in detail the effects of this group of compounds.

Frogs are a valuable test organism due to their critical role in the food web, their susceptibility to aquatic pollutants and endocrine disruptors, and as a model for human development. Specifically, the Western clawed frog *Silurana (Xenopus) tropicalis* is a diploid model species commonly used in developmental biology and toxicology. A sequenced genome and relatively high number of conserved genes to humans, in particular, those involved with development and disease [63] make *S. tropicalis* of interest to this study. Moreover, early stages of development are susceptible to chemical disruptors, and studies have shown that stress in larval frogs affects future body size, locomotor ability, and survivorship [73]. This study investigates the response of *S. tropicalis* embryos to NAs to determine potential mechanisms of toxicity using a transcriptomic approach.

In July 2016, I visited Dr. Christopher Martyniuk's laboratory at the University of Florida to conduct gene ontology and pathway analysis of microarray data. During my stay, Dr. Martyniuk trained me with the use of the software and helped me solve problems.

3.3. Methods

3.3.1. Chemicals

The two commonly available commercial NA mixtures were purchased from Sigma-Aldrich (lot number BCBC9959V and BCBK0736V), hereafter called Sigma 1 (S1) and Sigma 2 (S2), respectively. Human Chorionic Gonadotropin (hCG) was purchased from Millipore. The salts used in the preparation of the FETAX solution (NaCl, $NaHCO_3$, KCl, $CaCl_2$, $CaSO_4$, and $MgSO_4$) were purchased from Fisher Scientific. L-cysteine was obtained from Sigma-Aldrich.

3.3.2. Animals and exposure

Embryos of *S. tropicalis* were obtained by injection of hCG into the posterior lymph sac of adult *S. tropicalis*, as previously described [62]. Adult frogs were from the University of Ottawa *S. tropicalis* breeding colony. Frogs were fed daily with Nasco pellets and kept under a 12:12 h light:dark cycle. Briefly, the frogs were injected with a priming dose of 12.5 IU of hCG into the posterior lymph sac and move to glass tanks with FETAX solution with a temperature of 21.4 ºC and pH of 5.9. After 24 h, the breeding pairs were injected with a boosting dose of 100 IU of hCG and placed in glass tanks with FETAX solution at a temperature of 21.6 ºC and pH of 5.8. Embryos were collected and the jelly coat was removed by gently mixing in a 2% w/v L-cysteine solution prepared with FETAX solution for two min [76]. Clean embryos were then moved to a tank with FETAX solution. Embryos at stage 9-10 Nieuwkoop and Faber (NF) staging system [96] were individually selected under a light microscope. Four replicates with 10 embryos were used for each treatment and were placed in 10 cm Petri dishes containing 50 mL of the NA solutions. All the glassware used in the

exposure was washed as previously described in Chapter 2. The embryos were exposed to the two NA mixtures at 2 mg/L. This concentration was selected as it induced sub-lethal effects in *S. tropicalis* embryos as shown in Chapter 2. Control embryos were placed in FETAX solution. Treatment solutions were prepared daily from a respective stock solution that was kept at 4 ºC and covered from light. Every 24 h the embryos were observed, dead embryos were removed, and the solutions were replaced to ensure concentration and water quality. The exposure was performed in a controlled light/temperature room with a 12:12 light:dark cycle. The water temperature was 25 ± 1 ºC. After 96 h of exposure, tadpoles were sampled to measure gene expression profiles. They were anesthetized with a solution of 0.5 g/L Tricaine-S (MS-222, TMS, tricaine methanesulfonate; Western, Chemical Inc.) buffered with sodium bicarbonate. Digital images of the embryos were taken under a light stereomicroscope (Nikon SMZ 1500), with a Nikon DS-Fi1 camera and NIS Elements version 3.22.00 software. Images were used for morphometric measurements. The abnormalities were classified according to the FETAX Atlas of abnormalities [82]. Four samples per treatment were used for microarray analysis (n = 4). Each sample consisted of a pool of 10 larvae to ensure adequate quantity of RNA. The care and treatment of animals used in this study were in accordance with the guidelines of the Animal Care Committee of the University of Ottawa and the Canadian Council on Animal Care.

3.3.3. Microarray analysis

Total RNA extraction was done using RNeasy Mini Kit (Qiagen) with on-column RNase-free DNase treatment (Qiagen). RNA concentrations were determined using NanoDrop-2000 spectrophotometer (Thermo Scientific). Total RNA integrity was

assessed using RNA 6000 Nano Assay Kit with the 2100 Bioanalyzer (Agilent Technologies). The RNA integrity number (RIN) for the *S. tropicalis* samples ranged from 8.0-8.8.

A custom 4 x 44 K Agilent microarray developed for *S. tropicalis* was used to identify transcripts affected by the exposure to NA mixtures. The microarray platform is deposited in Gene Expression Omnibus (GPL15626) at NCBI and comprised 32,899 unique probes and 1417 Agilent control features [97]. The microarray platform was kindly provided by our collaborators Valérie Langlois (Institut national de la recherche scientifique, Quebec City, QC, Canada) and Christopher Martyniuk (University of Florida, FL, USA). Microarray hybridization was performed according to the Agilent One-Color Microarray-Based Gene Expression Analysis protocol using Cyanine 3 (Cy3). For the production of complementary DNA (cDNA), 1 µg total RNA per sample was used. Labeling was conducted according to the manufacturer's protocol. Microarrays were hybridized for 17 h and washed according to the Agilent protocol. The Agilent G2600D Microarray Scanner was used for microarray scanning at 5 µm. The Agilent Feature Extraction Software (v. 11.0.1.1) was used to extract the signal intensities from the TIFF microarray images. Intensity data were imported into JMP® Genomics (Version 6.0) and data were normalized using quantile normalization. The microarray limit of detection was 2.5, according to the lowest point in the standard curve and the Agilent negative controls. Thus, probes with signal intensity < 2.5 were filtered to a value of 2.5. Following the removal of all the control spots, the differentially expressed genes were identified using a one-way ANOVA and a false-discovery rate of 5% (FDR = 0.05).

3.3.4. Bioinformatics

Any probe that significantly different from the control ($p < 0.05$) prior to FDR adjustment in either treatment was used to broaden the analysis of the transcriptomic responses. Clustering analysis was used to assess separation in transcriptomes among treatments. Two-way hierarchical clustering of differentially expressed probes was performed using the Fast Ward algorithm. Rows were centered to a mean of zero prior to clustering and were also scaled to a variance of one.

Gene set enrichment on gene ontology (GO) terms were determined using the parametric analysis of gene set enrichment (PAGE) algorithm, which is a two-sided z-score for gene ontology categories [98]. Sub-network enrichment analysis (SNEA) was performed using Pathway Studio 9.0 (Ariadne, Rockville, MD, USA) and ResNet 9.0. SNEA uses known relationships based on expression, binding and common pathways between genes to build networks focused around gene hubs. The GeneBank ID was then used for mapping human homologs in Pathway Studio; 4572 genes were successfully mapped. Enrichment *p*-value cut-off was set at $p < 0.05$. The analysis used the function “highest fold change, best *p*-value” for duplicate probes. This bioinformatics method leverages the entire dataset regardless of *p*-value and builds a distribution based on fold change to statistically test for enrichment of processes.

3.3.5. RNA extraction, cDNA synthesis, and RT-qPCR for microarray validation

The samples obtained in Chapter 2 were used for RT-qPCR validation of the microarray. RNA was isolated using the RNeasy Mini kit (Qiagen) according to the manufacturer’s protocol with on-column RNase-free DNase treatment (Qiagen). Before

cDNA synthesis, the concentration and quality of all samples were assessed using NanoDrop ND-2000 spectrophotometer (Thermo Scientific) and an agarose gel. The quality of the RNA was indicated by the presence of two defined bands. The band represents 28S ribosomal RNA (rRNA) and the second band represents the 18S rRNA [99]. Total cDNA was prepared using Maxima First Strand cDNA Synthesis Kit for RT-qPCR (Thermo Scientific). The cDNA samples were synthesized in parallel 3 different batches, S1, S2, and Acid-extractable organics (AEOs). The AEO mixture contains all the organic compounds from the acid extraction of an OSPW sample (see details in Chapter 2).

Real-time quantitative polymerase chain reaction (RT-qPCR) with SYBR green dye technology was used to validate relative gene expression. Based on the microarray results six genes were selected for further analysis (Table 3.8). The *ribosomal protein L8 (rpl8) gene* was selected because it did not significantly change with the exposure to NAs (Table 3.8) and is a ribosomal gene often used for normalization [100]. Primers were designed using Primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) and synthesized by Integrated DNA Technologies (Table 3.1). The PCR products were cloned into the pGEM®-T Easy vector (Promega, Madison, WI, USA) and sequenced using 3730 DNA Analyzer (Applied Biosystems) to confirm the amplification of the regions of interest. The Rotor-Gene SYBR Green kit (Qiagen) and Rotor-Gene Q (Qiagen) were used to amplify and detect the transcripts of interest. The thermal cycling parameters were as suggested by the manufacturer. An activation step at 95 ºC for 5 min, followed by 40 cycles of 95 ºC denaturation step for 5 s and one primer annealing temperature 60 ºC for 10 s. After 40 cycles, a melt curve was performed over a range of

60-95 ºC with increments of 1 ºC to ensure a single amplified product. The concentration of each primer in the RT-qPCR reactions was 1 µM. The efficiency of all RT-qPCR reactions was 94-107% (100.7 ± 3.4%) and coefficient of determination (R^2) was ≥ 0.99 (0.995 ± 0.004). Data were analyzed using the Rotor-Gene Q Series software (Qiagen). Analysis of outliers was performed using the ROUT method in GraphPad Prism 6 [101]. The relative standard curve method was used to calculate relative mRNA abundance between samples, normalized using NORMA-GENE algorithm [102] and then presented as fold change of gene expression from replicates (n = 5-6; assayed in triplicate) for each group.

3.3.6. Statistical analysis

Frequencies of abnormalities were arcsine square root transformed. Data normality and homogeneity of variance were assessed by conducting Shapiro–Wilk's test and Levene's test respectively. Data that violated these assumptions were log10 transformed. One-way ANOVA and Tukey post-hoc analyses were performed on normally distributed data. Non-parametric Kruskal–Wallis test followed by Student–Newman–Keuls (SNK) test were performed on not normally distributed data. The significance level was set at $\alpha = 0.05$. Statistical analysis was performed using SPSS V21 IBM and Sigma Plot 12.0.

3.4. Results and discussion

3.4.1. Effects of exposure to commercial NAs

The exposures of *S. tropicalis* embryos to NAs at 2 mg/L resulted in a significant reduction in total length (TL), snout-vent length (SVL), tail length (TaL), and interorbital

distance (IOD) ($p < 0.05$) (Fig. 3.1-3.2). There was also a significant increase in all abnormalities assessed. The most common were cranial abnormalities, edema, gut abnormalities, heart abnormalities, and eye abnormalities (Fig. 3.1-3.2). These teratogenic effects have been reported previously in *S. tropicalis* exposed to commercial extracts and AEOs in a dose-dependent response (see Chapter 2). These types of abnormalities have also been described in other vertebrates, such as larval zebrafish (*Danio rerio*) exposed to AEOs extracted from oil sands from the Daqing oil exploring area [87] and larval fathead minnows (*Pimephales promelas*) exposed to NA fraction components (NAFCs) from OSPW from Alberta, Canada [84].

3.4.2. Gene expression profiling following NA exposure

The microarray analysis performed on embryos of *S. tropicalis* exposed to S1 revealed 6074 differentially expressed genes relative to control ($p < 0.05$), of which 2873 were up-regulated and 3201 were down-regulated. A total of 894 genes passed the correction for multiple testing (FDR < 0.05). This includes 526 up-regulated and 368 down-regulated genes. Furthermore, the analysis revealed 6789 differentially expressed genes versus control for S2 ($p < 0.05$). Of these, 3395 were up-regulated and 3394 were down-regulated by the exposure to S2. A total of 1177 genes passed the FDR test (FDR < 0.05). Of these, 620 up-regulated and 557 down-regulated. There were 3656 common genes affected by S1 and S2 ($p < 0.05$). Of these, 564 genes passed the FDR test (FDR < 0.05). More genes were significantly affected ($p < 0.05$) by S2 compared to S1, indicating that S2 has a higher impact at the transcriptomic level of *S. tropicalis*. However, the EC_{50} and EC_{20} of S1 are significantly lower than S2 (see Chapter 2).

3.4.3. Hierarchical analysis

The hierarchical clustering analysis of the differential expression ($p < 0.05$) data showed two main clades that mark the clear separation of the control samples and the two NA treatments (S1 and S2). The second clade is composed only of expression profiles from embryos exposed to the NA extracts. Treatments S1 and S2 are sister groups, indicating that these two blends affect the transcriptome of *S. tropicalis* similarly. The different replicates are located in their respective cluster and there are not mixed between groups, indicating a relatively strong transcriptome response to S1 and S2 (Fig 3.3).

3.4.4. Gene set enrichment on gene ontology (GO) terms

In total, 64 biological processes, 25 cellular components, and 52 molecular functions were significantly enriched after S1 exposure ($p < 0.05$) (Table 3.2). Moreover, 77 biological processes, 29 cellular components, and 57 molecular functions were significantly enriched after S2 exposure ($p < 0.05$) (Table 3.3). The number of significantly enriched processes common to both S1 and S2 for each of biological process, cellular components, and molecular functions was 18, 5 and 19, respectively ($p < 0.05$) (Table 3.4). The differences between the number of GO terms significantly enriched with the exposure to S1 and S2 is proposed to be a consequence of the chemical differences of the two mixtures (see Chapter 2).

The significant enrichment of GO terms such as respiratory electron transport chain, respiratory chain, adenylylsulfate kinase activity and pyruvate kinase activity indicates that NAs affect metabolism and that this could be a mechanism of toxicity of NAs (Table 3.4). Protein import into mitochondrial inner membrane and mitochondrial intermembrane GO terms were also significantly enriched ($p < 0.05$). These results

support the hypothesis of Frank et al. [49] of narcosis as part of the toxic mechanism of NAs. Furthermore, visual perception, structural constituent of eye lens, actin cytoskeleton and cytoskeleton organization are GO terms significantly enriched ($p < 0.05$) after the exposure to S1 and S2. This may be associated with eye and morphometric abnormalities in *S. tropicalis* induced by S1 and S2 (Fig. 3.2, Chapter 2).

3.4.5. Sub-network enrichment analysis (SNEA)

Using SNEA, a total of 75 cell process and 42 disease pathways were affected by S1 exposure ($p < 0.05$; Table 3.5). Furthermore, 66 cell process pathways and 46 disease pathways were perturbed in the embryos following S2 exposure ($p < 0.05$; Table 3.6). There were 38 common cell process pathways and 14 common diseases pathways perturbed by S1 and S2 ($p < 0.05$; Table 3.7). Some pathways that were significantly affected by the exposure to S1 and S2 included networks related to metabolism (Fig. 3.4 a-b), such as liver metabolism, cholesterol metabolism, retinoid metabolism, retinoic acid metabolism, steroid metabolism, arachidonic acid metabolism, drug metabolism, and xenobiotic clearance. This corroborates the GO analysis which also indicated that metabolic processes were significantly enriched following exposure. Other pathways significantly enriched, were related to membrane function (Fig. 3.4 c-d), such as membrane depolarization, Cl- transport, Na+ influx co-transport, ion transport, anion transport, fluid transport, and HCO3- transport. This further supports the hypothesis that toxicity of NAs is at least partially through narcosis, defined as the disruption cell membrane function [103]. The morphological analysis showed that NAs significantly induce gut abnormalities (Fig. 3.1j) on *S. tropicalis* embryos, and the lack of gut coiling was observed in some of the tadpoles (Fig. 3.2). Importantly, gene networks

related to gut function (Fig. 3.4 e-f), such as digestion, intestinal absorption, acid secretion, gastrointestinal system digestion, bile acid secretion, biliary flow, were significantly disrupted by the exposure to NAs. Other pathways related to the abnormalities observed on *S. tropicalis* (Fig. 3.1-3.2, see Chapter 2) include cartilage differentiation and edema (Table 3.7).

3.4.6. RT-qPCR validation of differentially expressed genes

Five genes were selected for microarray validation and for further study based on the following criteria: (1) their expression was altered by both S1 and S2; (2) the transcripts were annotated and had an identified human homolog; and (3) the transcripts appeared in the pathway analysis (Table 3.8). The cytochrome *P450 4B1b* (*cyp4b1*) gene encodes a monooxygenase enzyme that catalyzes reactions involved in the metabolism of cholesterol, steroids, and xenobiotics [104]. This gene forms part of the metabolism networks that was significantly disrupted by NAs, specifically arachidonic acid metabolism (Fig. 3.4 a-b). Arachidonic acid is a fatty acid involved in eye and brain development [105]. The *ATP-binding cassette sub-family G member 2* (*abcg2*) gene encodes a membrane transporter [106] involved in the gene network of gut function, specifically in the pathway of intestinal absorption (Fig. 3.4 e-f). The *solute carrier family 26 member 6* (*slc26a6*) gene encodes an anion transporter protein and regulates acid-base homeostasis [107,108]. This gene forms part of networks that were disturbed by the exposure to NAs, in particular, intestinal absorption and Cl^- transport (Fig. 3.4 c-f). Furthermore, Slc26a6-null mice have defective intestinal oxalate secretion and a higher incidence of stony concretions in the urinary tract [109]. The *bifunctional glutamate/proline--tRNA ligase* (*eprs*) gene encodes the enzyme that charges tRNA

with glutamate and proline [110], was selected as the SNEA showed that this gene is related to edema in *S. tropicalis* embryos (Table 3.7). The *sodium/glucose cotransporter 1* (*slc5a1*) gene encodes a sodium-dependent glucose transporter [111] and it is involved in several networks affected by NAs, such as liver metabolism, membrane depolarization, Na^+ influx co-transport, and intestinal absorption (Fig. 3.4 a-f). In 1994, Lee et al. [112] reported that *slc5a1* is involved in glucose uptake and is highly and moderately expressed in rats' intestine and liver respectively [113]. Also, the mutation of *slc5a1* gene can induce severe diarrhea and dehydration in humans [114]. These five genes appear to be associated with the significant increase ($p < 0.05$) of phenotypic abnormalities in *S. tropicalis* embryos, such as gut abnormalities, eye abnormalities, and edema (Fig. 3.1-3.2, see also Chapter 2). The *myod1* gene was selected because it changed significantly with both S1 and S2 treatment and is linked to myogenesis and muscle differentiation [115]. Therefore, it may be associated with the significant decrease of TL, SVL, TaL, and IOD of *S. tropicalis* exposed to NAs (Fig. 3.1-3.2).

The *S. tropicalis* embryos obtained in Chapter 2 were used for validation and to determine if there was a dose-response effect in the expression of these genes. The doses used for RT-qPCR analysis were 2, 4, 6 and 8 mg/L for S1, and 2, 6, 8, and 12 mg/L for S2. In general, the expression obtained by microarray analysis and RT-qPCR are similar in direction of change (Fig. 3.5, column S1 and S2 and Table 3.8). The genes *cyp4b1*, *abcg2*, *slc26a6*, *eprs*, and *slc5a1*, were affected in a dose-response fashion and the fold changes were comparable between S1 and S2 (Fig. 3.5). However, some discrepancies were noted for *myod1*. The microarray analysis showed a

significant down-regulation of *myod1* by the exposure to S1 and S2 (Table 3.8) while the RT-qPCR indicated an upregulation (Fig. 3.5 s-t).

The chemical analysis presented in Chapter 2 showed differences in the chemical profiles of commercial extracts versus the AEOs. However, we had hypothesized that the toxic mechanisms would be similar. To test this hypothesis, samples of the *S. tropicalis* exposed to 2, 12, 24 and 48 mg/L AEOs (Chapter 2) were analyzed by RT-qPCR for the genes described for microarray validation (Table 3.8). The expression of *cyp4b1*, *abcg2*, and *slc26a6* was not affected in a dose-response manner by the exposure to AEOs (Fig. 3.5, AEO column). The RT-qPCR results also showed that NaOH at 0.001 N significantly impacted the expression of all of the genes here analyzed (Fig. 3.5, column AEO). After extraction, the AEOs were dissolved in a solution of 0.1 N NaOH. The solvent control did not have a significant effect on the parameters measured in Chapter 2, such as morphometric and frequency of abnormalities. Likewise, in 2015, Marentette et al. reported no effect on the frequency of abnormalities up to concentrations of 50 mM NaOH in fathead minnow embryos [84]. However, it is important to use and report the effects of NaOH with the toxicity assessment of AEOs. Moreover, the expression level of *eprs* and *slc5a1* significantly decreased at 48 mg/L compared to the control and NaOH control ($p < 0.05$) (Fig. 3.5 l and o). The *eprs* and *slc5a1* are part of networks related to edema, membrane function, metabolism, and gut function (Table 3.8). The frequency of edema and gut abnormalities was significantly higher in samples exposed to AEOs (see Chapter 2) and there may be a relationship between the phenotypic abnormalities in *S. tropicalis* embryos and *eprs* and *slc5a1.*

3.5. Conclusion

In this chapter, I report the first analysis of the effects of NAs on the transcriptomic profile of a vertebrate, specifically early embryonic stages of *S. tropicalis*. A diverse number of GO terms and pathways were significantly enriched by exposures to S1 and S2, suggesting multiple mechanisms of toxicity. The results obtained in the microarray analysis between the samples of *S. tropicalis* exposed to S1 and S2 are highly similar, suggesting differences in the composition. This is supported by chemical characterization using electrospray ionization high-resolution mass spectrometry (ESI-HRMS) in Chapter 2. Metabolism and membrane function GO terms and pathways were significantly enriched, indicating that NAs disrupt metabolism and cell membrane function in *S. tropicalis*. These data support the hypothesis that narcosis is a mechanism of toxicity of NAs. Other GO terms and pathways significantly enriched were related to the several abnormalities observed in the *S. tropicalis* larvae after exposure to NAs. These included gut function, cartilage differentiation, actin cytoskeleton, cytoskeleton organization, and edema. Future research should focus on the abnormalities induced by NAs in specific tissues, considering that gut abnormalities were one of the most relevant phenotypic abnormalities observed during exposures and in the transcriptomic analysis. The link between transcriptomic and phenotypic changes should be specifically addressed.

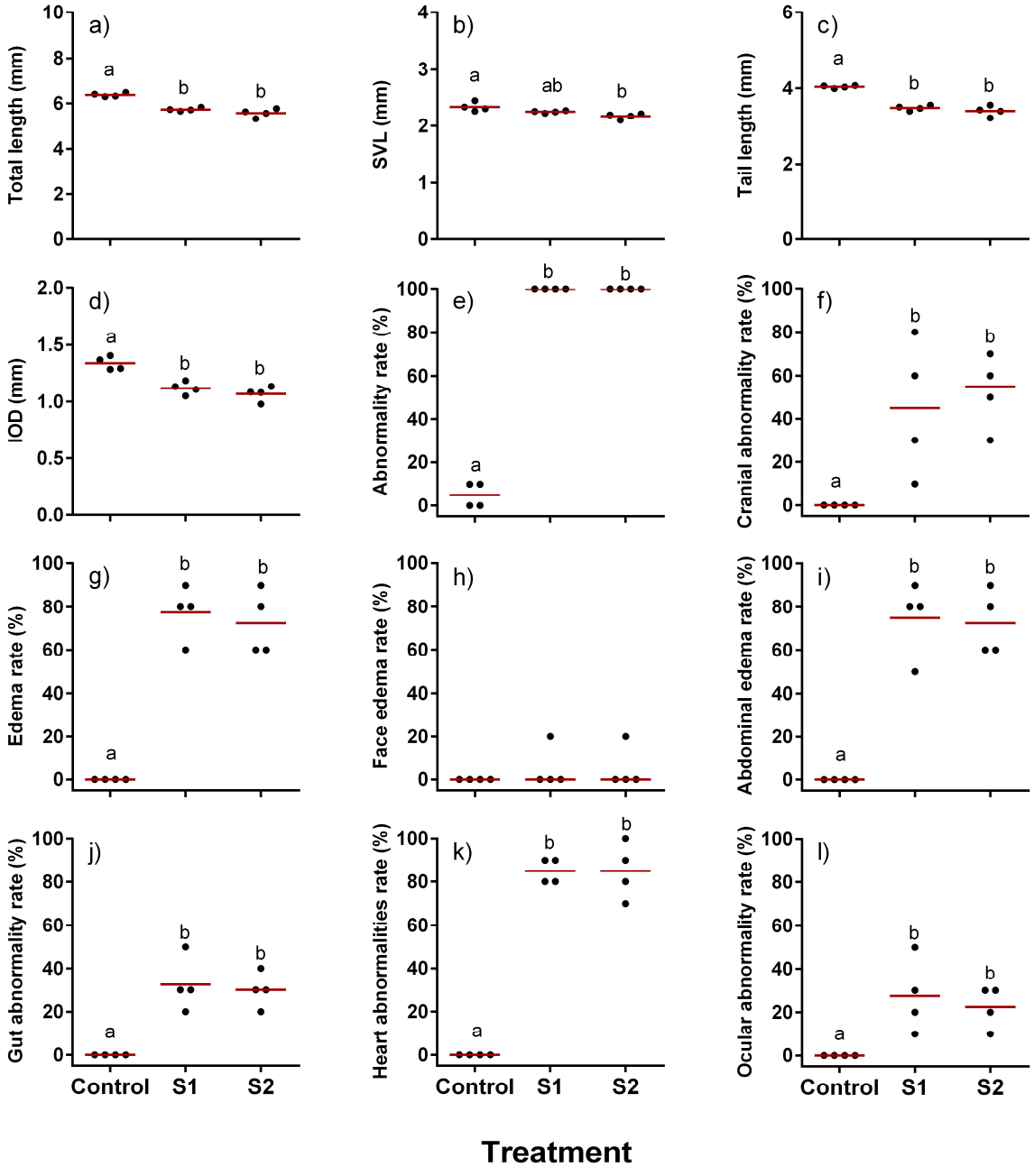


Figure 3.1. Teratogenic effects observed in *S. tropicalis* embryos exposed to S1 and S2. TL (a), SVL (b), TaL (c), IOD (d), abnormality rate (e), cranial abnormality rate (f), edema rate (g), face edema rate (h), abdominal edema rate (i), gut abnormality rate (j), heart abnormality rate (k), and ocular abnormality rate (l). Different letters indicate significant difference ($p < 0.05$). Each point represents a replicate. Red line indicates the mean in a-d, g, i, j, and l. Red line indicates the median in e, f, h, and k.

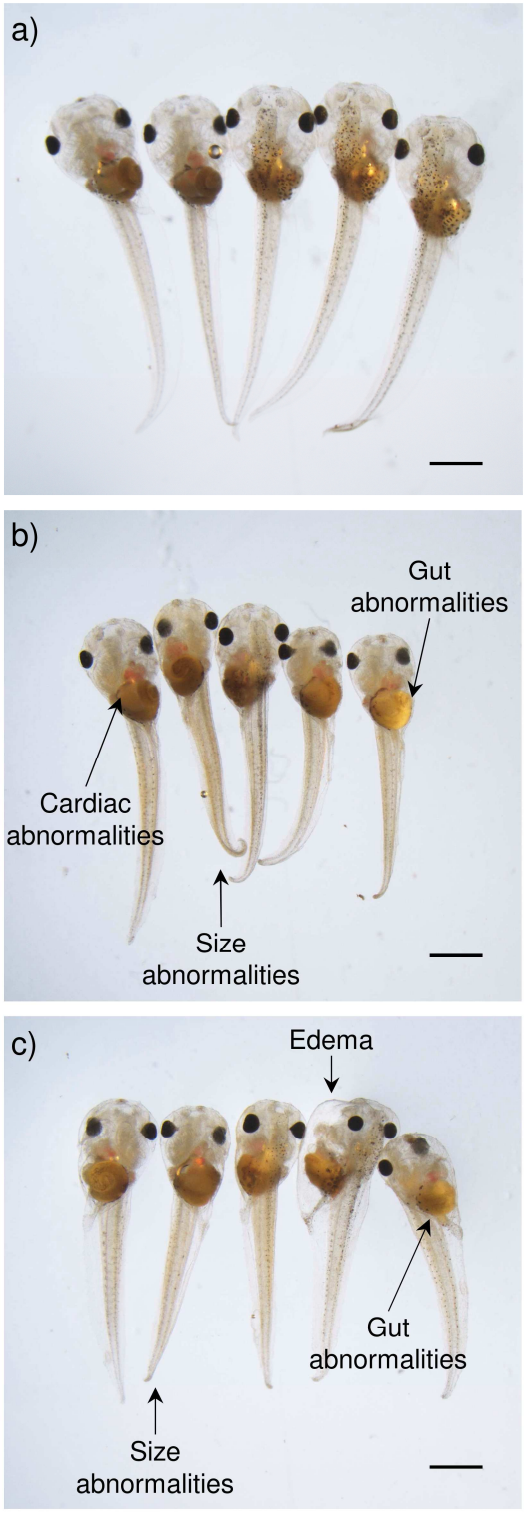


Figure 3.2. Photographs of *S. tropicalis* indicating the main teratogenic effects induced by exposure to NAs, control (a), S1 (b) and S2 (c). Scale bar is equal to 1 mm. Some of the observed abnormalities are indicated with arrows.

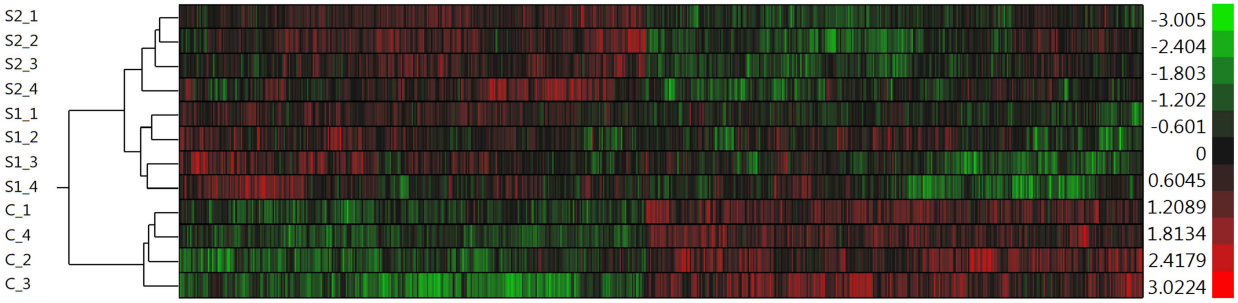


Figure 3.3. Hierarchical clustering of gene expression data using gene probes that were differentially expressed ($p < 0.05$) in *S. tropicalis* embryos exposed to two mixtures of NAs (S1 and S2) during early embryonic development versus control samples (C).

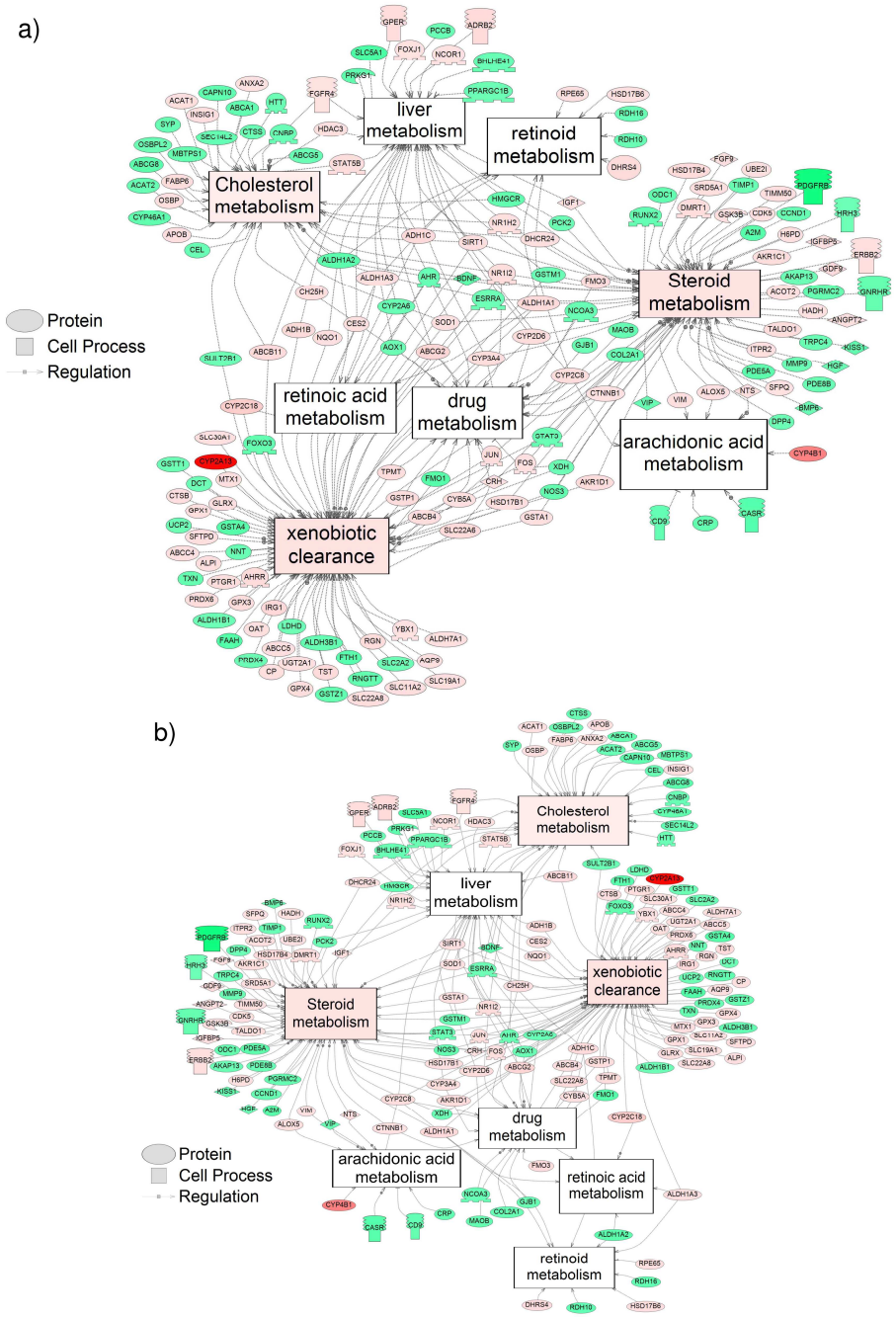


Figure 3.4. Networks associated with metabolism (a-b), membrane function (c-d), and gut function (e-f), indicating altered regulation of genes following exposure of *S. tropicalis* to 2 mg/L of S1 (a, c, and e) and S2 (b, d, and f). Sub-network enrichment analysis (SNEA) was performed using Pathway Studio 9.0. Red indicates that the transcript abundance was increased and green indicates that the transcript was decreased ($p < 0.05$).

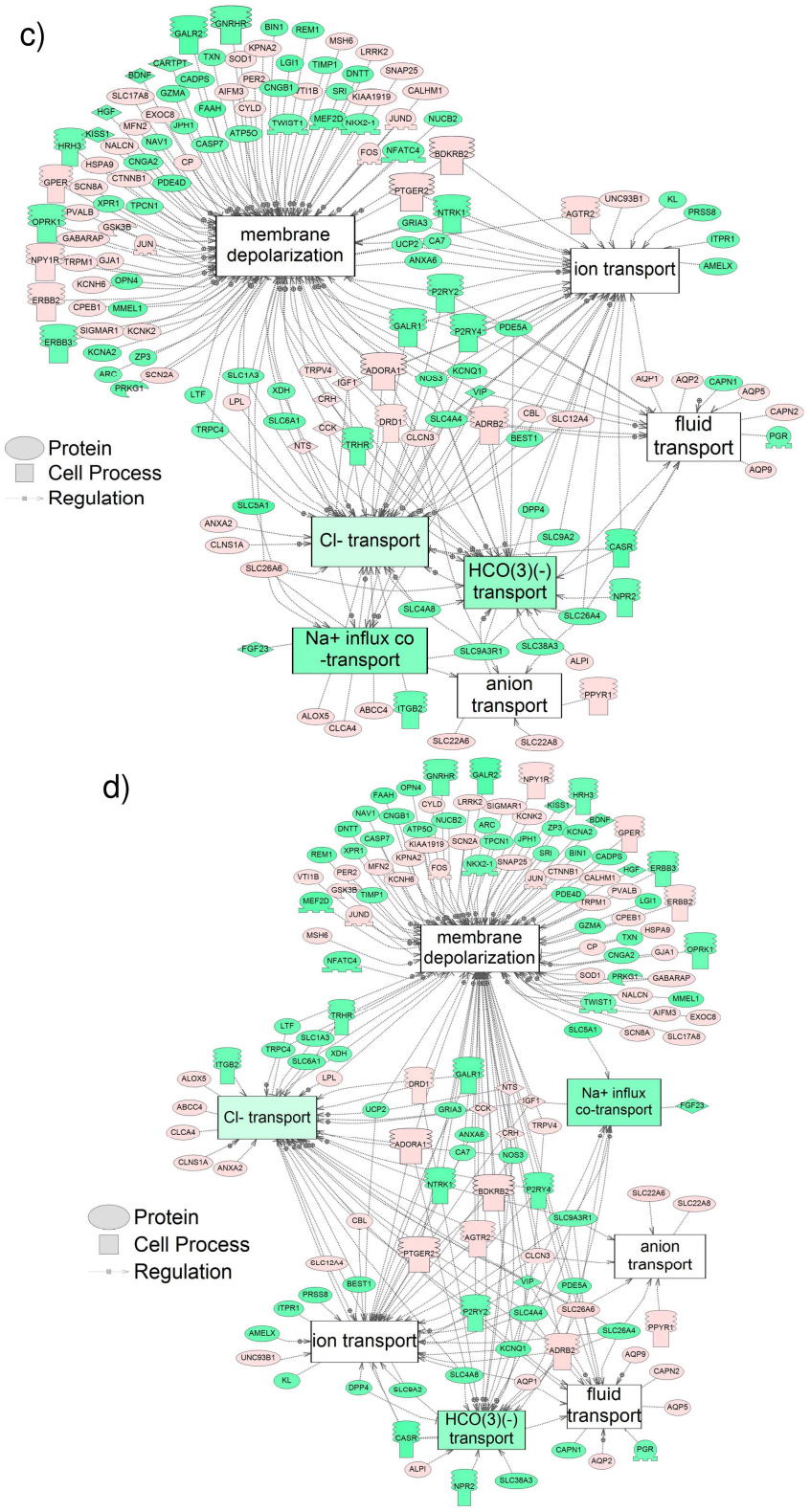


Fig. 3.4. Continued

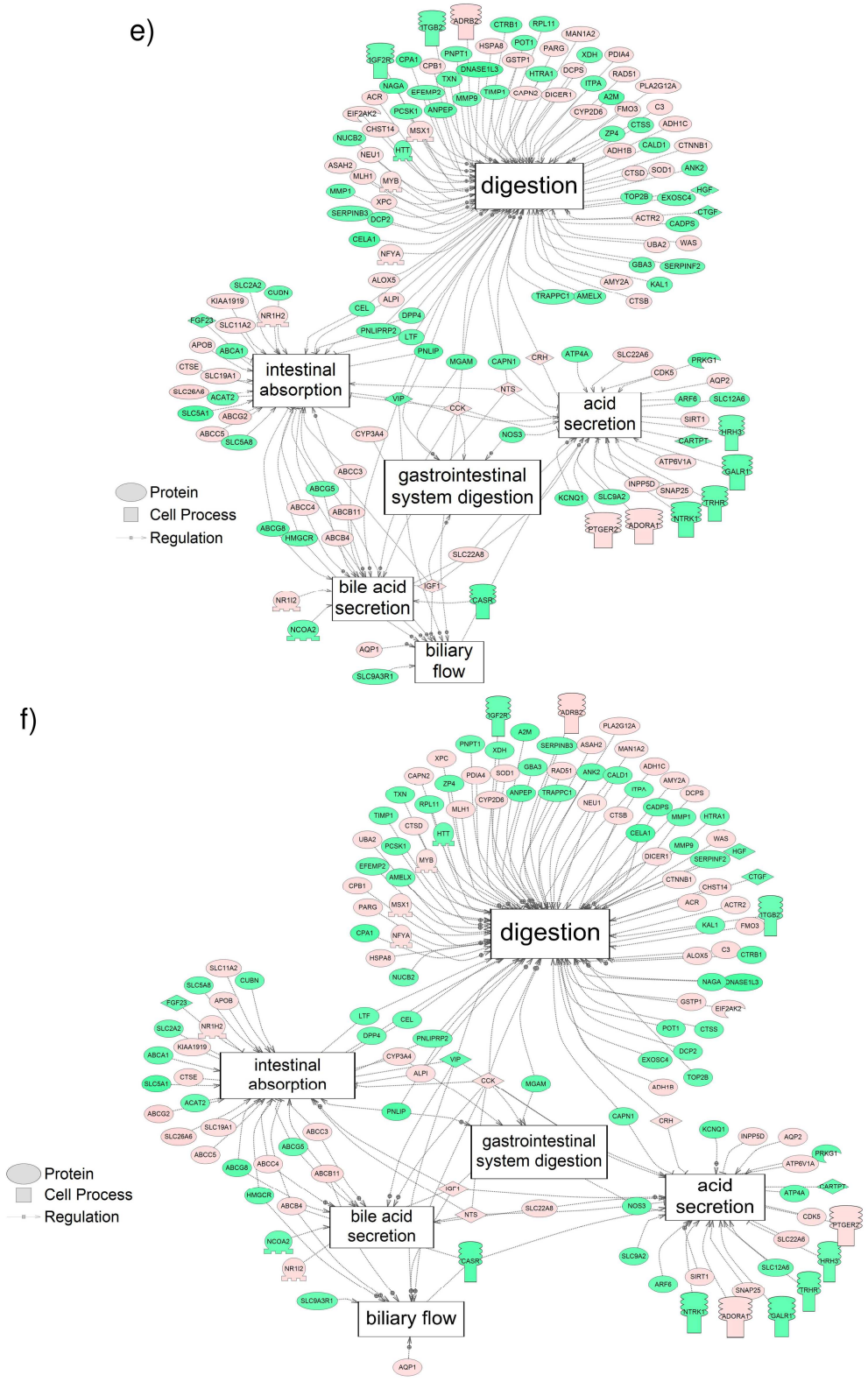


Fig. 3.4. Continued

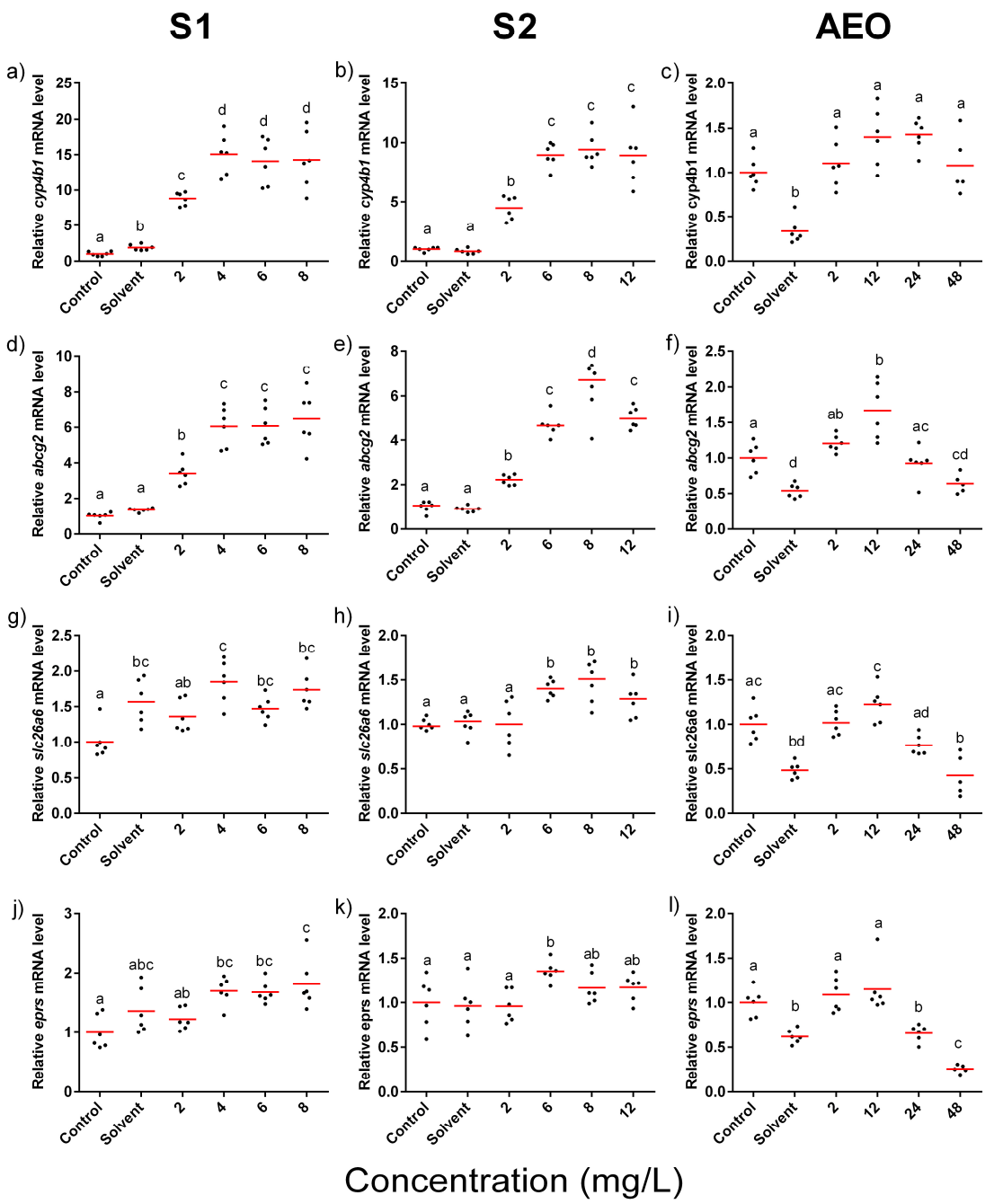


Figure 3.5. RT-qPCR analysis for (a-c) *cyp4b1*, (d-f) *abcg2*, (g-i) *slc26a6*, (j-l) *eprs,* (m-o) *slc5a1,* (p-r) *rpl8*, and (s-u) *myod1* mRNA in *S. tropicalis* exposed to different concentrations of S1 (first column), S2 (second column) and AEOs (third column). Data were normalized and defined as fold change relative to control. Each point represents a replicate. Different letters indicate significant difference ($p < 0.05$). Red line represents the mean in a-d, f, g, i-m, and o-u. Red line represents the median in e, h, and n.

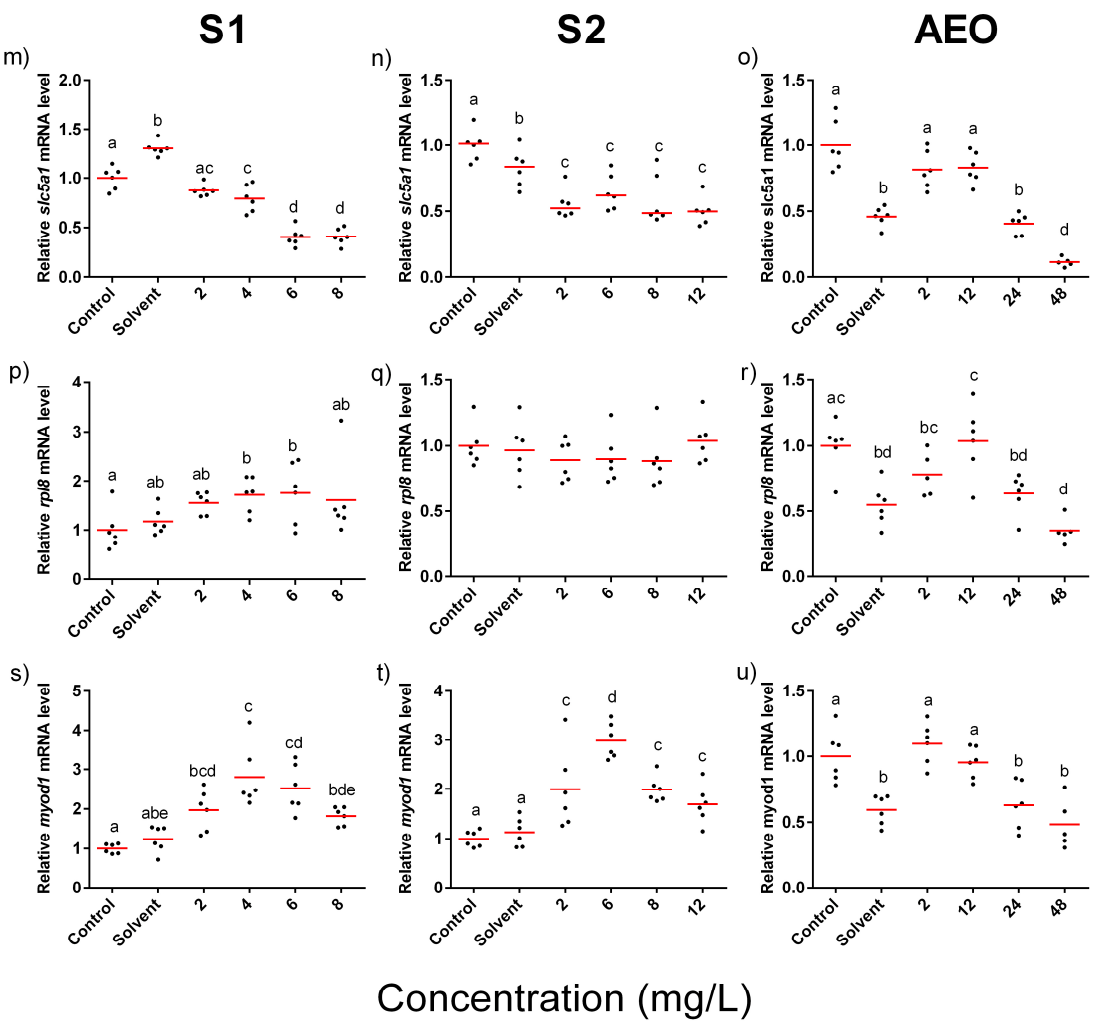


Figure 3.5. Continued

Table 3.1. Primer sets validated and used for real-time qPCR validation of the microarray data. F = forward; R = reverse.

Gene	Product size	Forward primer (5'-3')	Reverse primer (5'-3')	Accession number
abcg2	138	TACCATCCAAGGCAATCCCAA	ATCATGCAAGCCAAAGCCAA	NM_001045762.1
eprs	165	CAGTGGTAAACATGGAGTGGGA	CTGCGTTCTTTGGGTGCTTG	XM_002936017.3
slc26a6	189	TGCTGTTGTGGTCTATGCCTTT	TGCCCTCCTGTGCTTTCTTG	NM_001079448.2
slc5a1	123	TCTGATGCTCTCGGTGATGTT	TCCTTCTCGCTTGCTGTCTT	XM_002931917.4
myod1	179	ACCTCGGACATGAGCTTCTTTG	GCGTTGGTGGTCTTCCTCTT	NM_203641.1
rpl8	124	CACCGTTATCTCCCACAATCCT	AATACGACCACCACCAGCAAC	NM_203594.1
cyp4b1	188	TCACTGGCTGCTCGGTAATG	TGCTGCTCTTTGGTTCTTCTCT	NM_001126601.1

Table 3.2. List of gene ontology terms significantly altered after the exposure of *S. tropicalis* to S1 using Parametric Analysis of Gene Set Enrichment (PAGE) analysis ($p < 0.05$).

GO term biological process	Frequency	Page z-Score	Page Raw *p*-value	Page False Discovery Rate *p*-value
GO:0050909; p:sensory perception of taste;	79	5.07	4.0E-07	2.5E-04
GO:0007186; p:g-protein coupled receptor protein signaling pathway;	61	-4.15	3.3E-05	1.0E-02
GO:0007601; p:visual perception;	76	-3.62	2.9E-04	4.6E-02
GO:0000103; p:sulfate assimilation;	10	-3.55	3.9E-04	4.9E-02
GO:0030163; p:protein catabolic process;	23	3.34	8.3E-04	8.7E-02
GO:0048793; p:pronephros development;	19	-3.27	1.1E-03	9.9E-02
GO:0007369; p:gastrulation;	23	3.15	1.7E-03	1.2E-01
GO:0045039; p:protein import into mitochondrial inner membrane;	11	-3.13	1.7E-03	1.2E-01
GO:0001501; p:skeletal system development;	10	3.10	1.9E-03	1.2E-01
GO:0006672; p:ceramide metabolic process;	20	3.06	2.2E-03	1.2E-01
GO:0042777; p:plasma membrane atp synthesis coupled proton transpor	9	-3.05	2.3E-03	1.2E-01
GO:0032862; p:activation of rho gtpase activity;	5	3.04	2.4E-03	1.2E-01
GO:0006032; p:chitin catabolic process;	10	-3.00	2.7E-03	1.2E-01
GO:0022904; p:respiratory electron transport chain;	16	2.85	4.4E-03	1.8E-01
GO:0009435; p:nad biosynthetic process;	9	-2.84	4.5E-03	1.8E-01
GO:0006364; p:rrna processing;	45	-2.71	6.7E-03	2.5E-01
GO:0006744; p:ubiquinone biosynthetic process;	11	2.67	7.7E-03	2.7E-01
GO:0008156; p:negative regulation of dna replication;	7	2.58	9.8E-03	3.3E-01
GO:0006096; p:glycolysis;	42	2.54	1.1E-02	3.5E-01
GO:0016579; p:protein deubiquitination;	6	2.49	1.3E-02	3.7E-01
GO:0006546; p:glycine catabolic process;	6	2.45	1.4E-02	3.7E-01
GO:0031124; p:mrna 3'-end processing;	7	-2.43	1.5E-02	3.7E-01
GO:0006370; p:mrna capping;	10	-2.42	1.6E-02	3.7E-01
GO:0043161; p:proteasomal ubiquitin-dependent protein catabolic pro	14	2.42	1.6E-02	3.7E-01

GO:0006493; p:protein o-linked glycosylation;	7	2.41	1.6E-02	3.7E-01
GO:0006607; p:nls-bearing substrate import into nucleus;	6	2.40	1.7E-02	3.7E-01
GO:0001707; p:mesoderm formation;	14	-2.39	1.7E-02	3.7E-01
GO:0034641; p:cellular nitrogen compound metabolic process;	6	-2.38	1.7E-02	3.7E-01
GO:0050427; p:3'-phosphoadenosine 5'-phosphosulfate metabolic proce	7	2.37	1.8E-02	3.7E-01
GO:0006414; p:translational elongation;	13	-2.36	1.8E-02	3.7E-01
GO:0046330; p:positive regulation of jnk cascade;	5	-2.34	1.9E-02	3.7E-01
GO:0030036; p:actin cytoskeleton organization;	48	-2.33	2.0E-02	3.7E-01
GO:0051923; p:sulfation;	6	2.32	2.0E-02	3.7E-01
GO:0007179; p:transforming growth factor beta receptor signaling pa	18	-2.31	2.1E-02	3.7E-01
GO:0045765; p:regulation of angiogenesis;	5	2.31	2.1E-02	3.7E-01
GO:0050821; p:protein stabilization;	6	2.30	2.2E-02	3.7E-01
GO:0018105; p:peptidyl-serine phosphorylation;	6	2.29	2.2E-02	3.7E-01
GO:0050852; p:t cell receptor signaling pathway;	9	2.27	2.3E-02	3.8E-01
GO:0007229; p:integrin-mediated signaling pathway;	33	2.26	2.4E-02	3.8E-01
GO:0048260; p:positive regulation of receptor-mediated endocytosis;	5	2.24	2.5E-02	3.8E-01
GO:0060215; p:primitive hemopoiesis;	5	-2.23	2.6E-02	3.8E-01
GO:0009607; p:response to biotic stimulus;	8	-2.22	2.6E-02	3.8E-01
GO:0030168; p:platelet activation;	47	-2.21	2.7E-02	3.8E-01
GO:0007050; p:cell cycle arrest;	24	2.21	2.7E-02	3.8E-01
GO:0005513; p:detection of calcium ion;	16	2.20	2.8E-02	3.8E-01
GO:0032088; p:negative regulation of nf-kappab transcription factor	14	2.16	3.1E-02	4.1E-01
GO:0051225; p:spindle assembly;	8	2.16	3.1E-02	4.1E-01
GO:0001522; p:pseudouridine synthesis;	12	-2.14	3.2E-02	4.2E-01
GO:0007269; p:neurotransmitter secretion;	15	-2.13	3.3E-02	4.2E-01
GO:0008033; p:trna processing;	36	-2.13	3.3E-02	4.2E-01
GO:0001503; p:ossification;	21	2.10	3.6E-02	4.4E-01
GO:0009116; p:nucleoside metabolic process;	14	-2.08	3.8E-02	4.4E-01
GO:0008277; p:regulation of g-protein coupled receptor protein sign	11	-2.08	3.8E-02	4.4E-01
GO:0033504; p:floor plate development;	5	2.07	3.8E-02	4.4E-01

GO:0007596; p:blood coagulation;	56	2.06	3.9E-02	4.4E-01
GO:0051260; p:protein homooligomerization;	22	-2.06	3.9E-02	4.4E-01
GO:0006545; p:glycine biosynthetic process;	5	2.05	4.0E-02	4.4E-01
GO:0007169; p:transmembrane receptor protein tyrosine kinase signal	19	-2.03	4.2E-02	4.4E-01
GO:0040007; p:growth;	28	2.03	4.2E-02	4.4E-01
GO:0009952; p:anterior/posterior pattern formation;	18	-2.01	4.5E-02	4.6E-01
GO:0006069; p:ethanol oxidation;	12	-2.01	4.5E-02	4.6E-01
GO:0030866; p:cortical actin cytoskeleton organization;	5	-1.99	4.7E-02	4.7E-01
GO:0006406; p:mrna export from nucleus;	17	-1.97	4.8E-02	4.7E-01
GO:0006913; p:nucleocytoplasmic transport;	6	1.97	4.8E-02	4.7E-01

GO term cellular component	Frequency	Page z-Score	Page Raw *p*-value	Page False Discovery Rate *p*-value
GO:0035098; c:esc/e(z) complex;	7	3.52	4.3E-04	9.3E-02
GO:0042719; c:mitochondrial intermembrane space protein transporter	11	-3.12	1.8E-03	1.7E-01
GO:0030141; c:stored secretory granule;	9	-2.91	3.6E-03	1.7E-01
GO:0000786; c:nucleosome;	97	2.89	3.9E-03	1.7E-01
GO:0030173; c:integral to golgi membrane;	41	-2.85	4.4E-03	1.7E-01
GO:0005875; c:microtubule associated complex;	16	-2.82	4.7E-03	1.7E-01
GO:0070469; c:respiratory chain;	13	2.69	7.2E-03	2.2E-01
GO:0005604; c:basement membrane;	10	2.59	9.6E-03	2.2E-01
GO:0005938; c:cell cortex;	22	2.57	1.0E-02	2.2E-01
GO:0000784; c:nuclear chromosome, telomeric region;	6	-2.53	1.1E-02	2.2E-01
GO:0015629; c:actin cytoskeleton;	32	-2.53	1.1E-02	2.2E-01
GO:0017059; c:serine c-palmitoyltransferase complex;	5	-2.50	1.3E-02	2.2E-01
GO:0005871; c:kinesin complex;	7	-2.47	1.3E-02	2.2E-01
GO:0005643; c:nuclear pore;	31	2.38	1.7E-02	2.7E-01
GO:0030126; c:copi vesicle coat;	8	-2.33	2.0E-02	2.9E-01
GO:0019005; c:scf ubiquitin ligase complex;	8	2.15	3.1E-02	4.1E-01

GO:0005882; c:intermediate filament;	53	-2.15	3.1E-02	4.1E-01
GO:0032133; c:chromosome passenger complex;	6	2.07	3.9E-02	4.3E-01
GO:0005778; c:peroxisomal membrane;	20	-2.06	3.9E-02	4.3E-01
GO:0030425; c:dendrite;	10	2.04	4.1E-02	4.3E-01
GO:0005921; c:gap junction;	6	2.04	4.2E-02	4.3E-01
GO:0030496; c:midbody;	14	2.00	4.5E-02	4.3E-01
GO:0008305; c:integrin complex;	15	1.99	4.6E-02	4.3E-01
GO:0005923; c:tight junction;	51	-1.97	4.8E-02	4.3E-01
GO:0005777; c:peroxisome;	22	1.97	4.9E-02	4.3E-01

GO term molecular function	Frequency	Page z-Score	Page Raw *p*-value	Page False Discovery Rate *p*-value
GO:0008234; f:cysteine-type peptidase activity;	41	3.83	1.3E-04	4.9E-02
GO:0004692; f:cgmp-dependent protein kinase activity;	9	-3.64	2.7E-04	4.9E-02
GO:0004020; f:adenylylsulfate kinase activity;	10	-3.51	4.4E-04	4.9E-02
GO:0004781; f:sulfate adenylyltransferase (atp) activity;	10	-3.51	4.4E-04	4.9E-02
GO:0042626; f:atpase activity, coupled to transmembrane movement of	19	3.27	1.1E-03	7.6E-02
GO:0042605; f:peptide antigen binding;	8	3.26	1.1E-03	7.6E-02
GO:0030955; f:potassium ion binding;	9	3.21	1.3E-03	7.6E-02
GO:0070330; f:aromatase activity;	14	-3.18	1.5E-03	7.6E-02
GO:0003995; f:acyl-coa dehydrogenase activity;	22	-3.17	1.5E-03	7.6E-02
GO:0004568; f:chitinase activity;	10	-2.96	3.1E-03	1.4E-01
GO:0042826; f:histone deacetylase binding;	13	2.93	3.3E-03	1.4E-01
GO:0010181; f:fmn binding;	12	2.90	3.8E-03	1.4E-01
GO:0004383; f:guanylate cyclase activity;	13	-2.86	4.2E-03	1.4E-01
GO:0017040; f:ceramidase activity;	11	2.79	5.3E-03	1.7E-01
GO:0004743; f:pyruvate kinase activity;	7	2.77	5.7E-03	1.7E-01
GO:0008061; f:chitin binding;	8	-2.71	6.8E-03	1.8E-01
GO:0005506; f:iron ion binding;	68	2.71	6.8E-03	1.8E-01
GO:0004017; f:adenylate kinase activity;	13	-2.66	7.9E-03	1.9E-01

GO:0050253; f:retinyl-palmitate esterase activity;	5	2.63	8.4E-03	2.0E-01
GO:0016757; f:transferase activity, transferring glycosyl groups;	58	2.59	9.6E-03	2.1E-01
GO:0048038; f:quinone binding;	5	2.56	1.0E-02	2.2E-01
GO:0001567; f:cholesterol 25-hydroxylase activity;	9	2.54	1.1E-02	2.2E-01
GO:0043425; f:bhlh transcription factor binding;	9	-2.51	1.2E-02	2.3E-01
GO:0003755; f:peptidyl-prolyl cis-trans isomerase activity;	35	-2.50	1.2E-02	2.3E-01
GO:0016780; f:phosphotransferase activity, for other substituted ph	12	2.47	1.3E-02	2.4E-01
GO:0004601; f:peroxidase activity;	11	-2.46	1.4E-02	2.4E-01
GO:0042802; f:identical protein binding;	37	2.42	1.5E-02	2.5E-01
GO:0004613; f:phosphoenolpyruvate carboxykinase (gtp) activity;	5	2.42	1.6E-02	2.5E-01
GO:0008373; f:sialyltransferase activity;	29	-2.41	1.6E-02	2.5E-01
GO:0019903; f:protein phosphatase binding;	5	2.36	1.8E-02	2.7E-01
GO:0004735; f:pyrroline-5-carboxylate reductase activity;	5	2.35	1.9E-02	2.7E-01
GO:0046914; f:transition metal ion binding;	14	2.27	2.3E-02	3.2E-01
GO:0008308; f:voltage-gated anion channel activity;	9	-2.23	2.6E-02	3.3E-01
GO:0051539; f:4 iron, 4 sulfur cluster binding;	21	2.22	2.6E-02	3.3E-01
GO:0003887; f:dna-directed dna polymerase activity;	23	2.22	2.7E-02	3.3E-01
GO:0030332; f:cyclin binding;	5	-2.21	2.7E-02	3.3E-01
GO:0008121; f:ubiquinol-cytochrome-c reductase activity;	8	2.20	2.8E-02	3.3E-01
GO:0034701; f:tripeptidase activity;	6	-2.19	2.8E-02	3.3E-01
GO:0005544; f:calcium-dependent phospholipid binding;	29	-2.16	3.1E-02	3.4E-01
GO:0004864; f:protein phosphatase inhibitor activity;	13	-2.16	3.1E-02	3.4E-01
GO:0004012; f:phospholipid-translocating atpase activity;	9	2.16	3.1E-02	3.4E-01
GO:0016705; f:oxidoreductase activity, acting on paired donors, wit	63	-2.11	3.5E-02	3.7E-01
GO:0004146; f:dihydrofolate reductase activity;	5	2.09	3.6E-02	3.8E-01
GO:0005212; f:structural constituent of eye lens;	24	-2.08	3.7E-02	3.8E-01
GO:0004707; f:map kinase activity;	22	2.06	4.0E-02	3.9E-01
GO:0051864; f:histone demethylase activity (h3-k36 specific);	5	2.04	4.1E-02	4.0E-01
GO:0004497; f:monooxygenase activity;	54	-2.02	4.4E-02	4.0E-01
GO:0004402; f:histone acetyltransferase activity;	9	-2.02	4.4E-02	4.0E-01
GO:0004221; f:ubiquitin thiolesterase activity;	49	2.01	4.4E-02	4.0E-01

GO:0001948; f:glycoprotein binding;	6	2.01	4.5E-02	4.0E-01
GO:0005523; f:tropomyosin binding;	11	-2.00	4.6E-02	4.0E-01
GO:0005085; f:guanyl-nucleotide exchange factor activity;	38	1.96	5.0E-02	4.2E-01

Table 3.3. List of gene ontology terms significantly altered after the exposure of *S. tropicalis* to S2 using Parametric Analysis of Gene Set Enrichment (PAGE) analysis ($p < 0.05$).

GO term biological process	Frequency	Page z-Score	Page Raw *p*-value	Page False Discovery Rate *p*-value
GO:0050909; p:sensory perception of taste;	79	6.21	5.3E-10	3.4E-07
GO:0022904; p:respiratory electron transport chain;	16	3.83	1.3E-04	2.4E-02
GO:0050852; p:t cell receptor signaling pathway;	9	3.82	1.3E-04	2.4E-02
GO:0015992; p:proton transport;	10	3.79	1.5E-04	2.4E-02
GO:0006310; p:dna recombination;	32	3.63	2.9E-04	3.5E-02
GO:0015991; p:atp hydrolysis coupled proton transport;	36	-3.57	3.5E-04	3.5E-02
GO:0006367; p:transcription initiation from rna polymerase ii promo	36	3.51	4.5E-04	3.5E-02
GO:0000103; p:sulfate assimilation;	10	-3.49	4.8E-04	3.5E-02
GO:0051056; p:regulation of small gtpase mediated signal transducti	43	3.49	4.9E-04	3.5E-02
GO:0051225; p:spindle assembly;	8	3.21	1.3E-03	8.0E-02
GO:0007059; p:chromosome segregation;	25	3.14	1.7E-03	8.0E-02
GO:0001522; p:pseudouridine synthesis;	12	-3.12	1.8E-03	8.0E-02
GO:0006744; p:ubiquinone biosynthetic process;	11	3.06	2.2E-03	9.4E-02
GO:0007346; p:regulation of mitotic cell cycle;	5	2.99	2.8E-03	1.1E-01
GO:0006261; p:dna-dependent dna replication;	5	2.96	3.1E-03	1.2E-01
GO:0006364; p:rrna processing;	45	-2.93	3.4E-03	1.2E-01
GO:0007601; p:visual perception;	76	-2.90	3.7E-03	1.2E-01
GO:0002922; p:positive regulation of humoral immune response;	5	2.84	4.6E-03	1.3E-01
GO:0042102; p:positive regulation of t cell proliferation;	5	2.84	4.6E-03	1.3E-01
GO:0045860; p:positive regulation of protein kinase activity;	5	2.84	4.6E-03	1.3E-01
GO:2000360; p:negative regulation of binding of sperm to zona pellu	5	2.84	4.6E-03	1.3E-01
GO:0006825; p:copper ion transport;	5	2.74	6.1E-03	1.5E-01
GO:0007602; p:phototransduction;	15	-2.73	6.4E-03	1.5E-01
GO:0006099; p:tricarboxylic acid cycle;	29	-2.72	6.6E-03	1.5E-01

GO:0045666; p:positive regulation of neuron differentiation;	9	-2.72	6.6E-03	1.5E-01
GO:0043588; p:skin development;	8	2.72	6.6E-03	1.5E-01
GO:0006829; p:zinc ion transport;	16	-2.63	8.5E-03	1.8E-01
GO:0006974; p:response to dna damage stimulus;	15	-2.62	8.8E-03	1.8E-01
GO:0007596; p:blood coagulation;	56	2.62	8.9E-03	1.8E-01
GO:0042777; p:plasma membrane atp synthesis coupled proton transpor	9	-2.60	9.4E-03	1.8E-01
GO:0048511; p:rhythmic process;	8	-2.52	1.2E-02	2.2E-01
GO:0045665; p:negative regulation of neuron differentiation;	12	-2.47	1.3E-02	2.4E-01
GO:0045765; p:regulation of angiogenesis;	5	2.37	1.8E-02	3.0E-01
GO:2000344; p:positive regulation of acrosome reaction;	7	2.35	1.9E-02	3.0E-01
GO:0006414; p:translational elongation;	13	-2.35	1.9E-02	3.0E-01
GO:0006805; p:xenobiotic metabolic process;	56	-2.35	1.9E-02	3.0E-01
GO:0034329; p:cell junction assembly;	9	2.33	2.0E-02	3.0E-01
GO:0048665; p:neuron fate specification;	6	-2.31	2.1E-02	3.0E-01
GO:0060441; p:epithelial tube branching involved in lung morphogene	6	-2.31	2.1E-02	3.0E-01
GO:0060487; p:lung epithelial cell differentiation;	6	-2.31	2.1E-02	3.0E-01
GO:0030521; p:androgen receptor signaling pathway;	9	2.30	2.1E-02	3.1E-01
GO:0006108; p:malate metabolic process;	11	-2.27	2.3E-02	3.1E-01
GO:0043410; p:positive regulation of mapkkk cascade;	6	-2.27	2.3E-02	3.1E-01
GO:0045039; p:protein import into mitochondrial inner membrane;	11	-2.25	2.4E-02	3.1E-01
GO:0000070; p:mitotic sister chromatid segregation;	5	2.24	2.5E-02	3.1E-01
GO:0006541; p:glutamine metabolic process;	8	-2.23	2.6E-02	3.1E-01
GO:0007010; p:cytoskeleton organization;	26	-2.23	2.6E-02	3.1E-01
GO:0016567; p:protein ubiquitination;	15	2.23	2.6E-02	3.1E-01
GO:0006471; p:protein adp-ribosylation;	18	-2.23	2.6E-02	3.1E-01
GO:0006869; p:lipid transport;	28	-2.23	2.6E-02	3.1E-01
GO:0009117; p:nucleotide metabolic process;	8	-2.21	2.7E-02	3.1E-01
GO:0033504; p:floor plate development;	5	2.20	2.8E-02	3.1E-01
GO:0007032; p:endosome organization;	8	-2.20	2.8E-02	3.1E-01

GO:0018108; p:peptidyl-tyrosine phosphorylation;	7	-2.18	3.0E-02	3.2E-01
GO:0006357; p:regulation of transcription from rna polymerase ii pr	48	-2.17	3.0E-02	3.2E-01
GO:0043392; p:negative regulation of dna binding;	9	-2.16	3.1E-02	3.2E-01
GO:0010458; p:exit from mitosis;	5	2.15	3.1E-02	3.2E-01
GO:0070528; p:protein kinase c signaling cascade;	6	2.14	3.2E-02	3.2E-01
GO:0007420; p:brain development;	30	-2.14	3.2E-02	3.2E-01
GO:0050900; p:leukocyte migration;	14	2.14	3.3E-02	3.2E-01
GO:0048260; p:positive regulation of receptor-mediated endocytosis;	5	2.13	3.3E-02	3.2E-01
GO:0048793; p:pronephros development;	19	-2.08	3.8E-02	3.6E-01
GO:0035024; p:negative regulation of rho protein signal transductio	5	2.06	3.9E-02	3.6E-01
GO:0009168; p:purine ribonucleoside monophosphate biosynthetic proc	6	-2.05	4.0E-02	3.6E-01
GO:0030335; p:positive regulation of cell migration;	7	-2.05	4.0E-02	3.6E-01
GO:0006605; p:protein targeting;	6	-2.05	4.0E-02	3.6E-01
GO:0000087; p:m phase of mitotic cell cycle;	8	-2.05	4.0E-02	3.6E-01
GO:0051726; p:regulation of cell cycle;	14	-2.05	4.1E-02	3.6E-01
GO:0045806; p:negative regulation of endocytosis;	5	-2.03	4.2E-02	3.7E-01
GO:0006816; p:calcium ion transport;	6	2.03	4.2E-02	3.7E-01
GO:0008624; p:induction of apoptosis by extracellular signals;	15	-2.01	4.4E-02	3.7E-01
GO:0031295; p:t cell costimulation;	10	2.01	4.4E-02	3.7E-01
GO:0046209; p:nitric oxide metabolic process;	5	-2.00	4.5E-02	3.7E-01
GO:0030866; p:cortical actin cytoskeleton organization;	5	-2.00	4.5E-02	3.7E-01
GO:0009950; p:dorsal/ventral axis specification;	9	-1.99	4.7E-02	3.7E-01
GO:0008202; p:steroid metabolic process;	10	-1.99	4.7E-02	3.7E-01
GO:0018298; p:protein-chromophore linkage;	22	-1.98	4.8E-02	3.8E-01

GO term cellular component	Frequency	Page z-Score	Page Raw *p*-value	Page False Discovery Rate *p*-value
GO:0005792; c:microsome;	55	-3.65	2.7E-04	5.8E-02
GO:0005694; c:chromosome;	33	2.91	3.6E-03	2.2E-01

GO:0033179; c:proton-transporting v-type atpase, v0 domain;	7	-2.76	5.7E-03	2.2E-01
GO:0031463; c:cul3-ring ubiquitin ligase complex;	5	2.76	5.8E-03	2.2E-01
GO:0030131; c:clathrin adaptor complex;	21	2.74	6.2E-03	2.2E-01
GO:0045121; c:membrane raft;	8	2.73	6.3E-03	2.2E-01
GO:0000784; c:nuclear chromosome, telomeric region;	6	-2.68	7.5E-03	2.2E-01
GO:0032133; c:chromosome passenger complex;	6	2.63	8.5E-03	2.2E-01
GO:0042470; c:melanosome;	17	-2.54	1.1E-02	2.2E-01
GO:0005782; c:peroxisomal matrix;	8	-2.51	1.2E-02	2.2E-01
GO:0005578; c:proteinaceous extracellular matrix;	80	2.50	1.2E-02	2.2E-01
GO:0005876; c:spindle microtubule;	9	2.49	1.3E-02	2.2E-01
GO:0001726; c:ruffle;	11	-2.46	1.4E-02	2.2E-01
GO:0070469; c:respiratory chain;	13	2.46	1.4E-02	2.2E-01
GO:0005697; c:telomerase holoenzyme complex;	6	-2.43	1.5E-02	2.2E-01
GO:0030017; c:sarcomere;	10	2.38	1.7E-02	2.3E-01
GO:0033180; c:proton-transporting v-type atpase, v1 domain;	9	-2.36	1.8E-02	2.3E-01
GO:0042719; c:mitochondrial intermembrane space protein transporter	11	-2.30	2.2E-02	2.6E-01
GO:0005744; c:mitochondrial inner membrane presequence translocase	6	-2.23	2.6E-02	2.9E-01
GO:0042555; c:mcm complex;	10	-2.20	2.8E-02	2.9E-01
GO:0005813; c:centrosome;	52	2.19	2.9E-02	2.9E-01
GO:0045261; c:proton-transporting atp synthase complex, catalytic c	6	-2.18	3.0E-02	2.9E-01
GO:0009897; c:external side of plasma membrane;	5	-2.15	3.1E-02	3.0E-01
GO:0030665; c:clathrin coated vesicle membrane;	5	2.11	3.4E-02	3.1E-01
GO:0031901; c:early endosome membrane;	19	-2.10	3.6E-02	3.1E-01
GO:0042383; c:sarcolemma;	5	-2.03	4.2E-02	3.6E-01
GO:0005815; c:microtubule organizing center;	41	-2.00	4.6E-02	3.6E-01
GO:0015629; c:actin cytoskeleton;	32	-1.99	4.6E-02	3.6E-01
GO:0015030; c:cajal body;	18	-1.98	4.7E-02	3.6E-01
GO term molecular function	Frequency	Page z-	Page	Page False

		Score	Raw *p*-value	Discovery Rate *p*-value
GO:0004887; f:thyroid hormone receptor activity;	11	3.84	1.2E-04	3.0E-02
GO:0070330; f:aromatase activity;	14	-3.76	1.7E-04	3.0E-02
GO:0005085; f:guanyl-nucleotide exchange factor activity;	38	3.72	2.0E-04	3.0E-02
GO:0004020; f:adenylylsulfate kinase activity;	10	-3.47	5.3E-04	4.1E-02
GO:0004781; f:sulfate adenylyltransferase (atp) activity;	10	-3.47	5.3E-04	4.1E-02
GO:0030955; f:potassium ion binding;	9	3.46	5.5E-04	4.1E-02
GO:0001104; f:rna polymerase ii transcription cofactor activity;	6	3.31	9.4E-04	5.5E-02
GO:0004864; f:protein phosphatase inhibitor activity;	13	-3.29	1.0E-03	5.5E-02
GO:0042626; f:atpase activity, coupled to transmembrane movement of	19	3.22	1.3E-03	6.2E-02
GO:0004692; f:cgmp-dependent protein kinase activity;	9	-3.13	1.7E-03	7.7E-02
GO:0043425; f:bhlh transcription factor binding;	9	-2.95	3.2E-03	1.3E-01
GO:0030374; f:ligand-dependent nuclear receptor transcription coact	7	2.87	4.1E-03	1.4E-01
GO:0004743; f:pyruvate kinase activity;	7	2.85	4.3E-03	1.4E-01
GO:0032190; f:acrosin binding;	5	2.84	4.5E-03	1.4E-01
GO:0030971; f:receptor tyrosine kinase binding;	9	2.74	6.1E-03	1.8E-01
GO:0004843; f:ubiquitin-specific protease activity;	16	-2.66	7.7E-03	2.1E-01
GO:0005201; f:extracellular matrix structural constituent;	16	-2.59	9.5E-03	2.1E-01
GO:0004519; f:endonuclease activity;	29	2.58	9.9E-03	2.1E-01
GO:0008080; f:n-acetyltransferase activity;	28	-2.57	1.0E-02	2.1E-01
GO:0015078; f:hydrogen ion transmembrane transporter activity;	17	-2.57	1.0E-02	2.1E-01
GO:0008641; f:small protein activating enzyme activity;	6	2.56	1.0E-02	2.1E-01
GO:0017176; f:phosphatidylinositol n-acetylglucosaminyltransferase	5	2.56	1.1E-02	2.1E-01
GO:0005212; f:structural constituent of eye lens;	24	-2.54	1.1E-02	2.1E-01
GO:0042605; f:peptide antigen binding;	8	2.48	1.3E-02	2.4E-01
GO:0004013; f:adenosylhomocysteinase activity;	5	2.47	1.4E-02	2.4E-01
GO:0003725; f:double-stranded rna binding;	37	2.44	1.5E-02	2.4E-01

GO:0004499; f:flavin-containing monooxygenase activity;	10	-2.44	1.5E-02	2.4E-01
GO:0046933; f:hydrogen ion transporting atp synthase activity, rota	23	-2.37	1.8E-02	2.8E-01
GO:0015250; f:water channel activity;	9	2.34	2.0E-02	3.0E-01
GO:0042809; f:vitamin d receptor binding;	5	2.30	2.2E-02	3.2E-01
GO:0009982; f:pseudouridine synthase activity;	15	-2.28	2.3E-02	3.2E-01
GO:0004012; f:phospholipid-translocating atpase activity;	9	2.25	2.4E-02	3.4E-01
GO:0050661; f:nadp binding;	26	-2.24	2.5E-02	3.4E-01
GO:0051864; f:histone demethylase activity (h3-k36 specific);	5	2.22	2.6E-02	3.4E-01
GO:0005487; f:nucleocytoplasmic transporter activity;	5	2.22	2.7E-02	3.4E-01
GO:0046982; f:protein heterodimerization activity;	24	2.21	2.7E-02	3.4E-01
GO:0003887; f:dna-directed dna polymerase activity;	23	2.18	2.9E-02	3.4E-01
GO:0004806; f:triglyceride lipase activity;	15	-2.17	3.0E-02	3.4E-01
GO:0004402; f:histone acetyltransferase activity;	9	-2.17	3.0E-02	3.4E-01
GO:0001948; f:glycoprotein binding;	6	2.15	3.2E-02	3.5E-01
GO:0003956; f:nad(p)+-protein-arginine adp-ribosyltransferase activ	7	-2.14	3.3E-02	3.5E-01
GO:0015631; f:tubulin binding;	11	-2.12	3.4E-02	3.6E-01
GO:0016788; f:hydrolase activity, acting on ester bonds;	24	-2.10	3.5E-02	3.7E-01
GO:0015293; f:symporter activity;	20	-2.07	3.8E-02	3.7E-01
GO:0010181; f:fmn binding;	12	2.07	3.8E-02	3.7E-01
GO:0003697; f:single-stranded dna binding;	20	2.07	3.9E-02	3.7E-01
GO:0001567; f:cholesterol 25-hydroxylase activity;	9	2.05	4.0E-02	3.8E-01
GO:0046961; f:proton-transporting atpase activity, rotational mecha	16	-2.04	4.2E-02	3.8E-01
GO:0017056; f:structural constituent of nuclear pore;	5	-2.03	4.2E-02	3.8E-01
GO:0005391; f:sodium:potassium-exchanging atpase activity;	11	-2.03	4.2E-02	3.8E-01
GO:0004748; f:ribonucleoside-diphosphate reductase activity;	7	-2.01	4.4E-02	3.8E-01
GO:0016651; f:oxidoreductase activity, acting on nadh or nadph;	10	-1.99	4.7E-02	3.8E-01
GO:0070412; f:r-smad binding;	5	-1.99	4.7E-02	3.8E-01
GO:0016769; f:transferase activity, transferring nitrogenous groups	11	-1.98	4.7E-02	3.8E-01
GO:0004470; f:malic enzyme activity;	6	-1.98	4.8E-02	3.8E-01

GO:0005024; f:transforming growth factor beta receptor activity;	14	1.97	4.9E-02	3.9E-01
GO:0003873; f:6-phosphofructo-2-kinase activity;	8	1.96	4.9E-02	3.9E-01

Table 3.4. List of gene ontology terms altered after the exposure of *S. tropicalis* to both S1 and S2 using Parametric Analysis of Gene Set Enrichment (PAGE) analysis ($p < 0.05$).

GO term biological process	S1				S2			
	Frequency	Page z-Score	Page Raw *p*-value	Page False Discovery Rate *p*-value	Frequency	Page z-Score	Page Raw *p*-value	Page False Discovery Rate *p*-value
GO:0050909; p:sensory perception of taste;	79	5.07	4.0E-07	2.5E-04	79	6.21	5.3E-10	3.4E-07
GO:0007601; p:visual perception;	76	-3.62	2.9E-04	4.6E-02	76	-2.90	3.7E-03	1.2E-01
GO:0000103; p:sulfate assimilation;	10	-3.55	3.9E-04	4.9E-02	10	-3.49	4.8E-04	3.5E-02
GO:0048793; p:pronephros development;	19	-3.27	1.1E-03	9.9E-02	19	-2.08	3.8E-02	3.6E-01
GO:0045039; p:protein import into mitochondrial inner membrane;	11	-3.13	1.7E-03	1.2E-01	11	-2.25	2.4E-02	3.1E-01
GO:0042777; p:plasma membrane atp synthesis coupled proton transpor	9	-3.05	2.3E-03	1.2E-01	9	-2.60	9.4E-03	1.8E-01
GO:0022904; p:respiratory electron transport chain;	16	2.85	4.4E-03	1.8E-01	16	3.83	1.3E-04	2.4E-02
GO:0006364;	45	-2.71	6.7E-03	2.5E-01	45	-2.93	3.4E-03	1.2E-01

p:rrna processing; GO:0006744; p:ubiquinone biosynthetic process;	11	2.67	7.7E-03	2.7E-01	11	3.06	2.2E-03	9.4E-02
GO:0006414; p:translational elongation;	13	-2.36	1.8E-02	3.7E-01	13	-2.35	1.9E-02	3.0E-01
GO:0045765; p:regulation of angiogenesis;	5	2.31	2.1E-02	3.7E-01	5	2.37	1.8E-02	3.0E-01
GO:0050852; p:t cell receptor signaling pathway;	9	2.27	2.3E-02	3.8E-01	9	3.82	1.3E-04	2.4E-02
GO:0048260; p:positive regulation of receptor-mediated endocytosis;	5	2.24	2.5E-02	3.8E-01	5	2.13	3.3E-02	3.2E-01
GO:0051225; p:spindle assembly;	8	2.16	3.1E-02	4.1E-01	8	3.21	1.3E-03	8.0E-02
GO:0001522; p:pseudouridine synthesis;	12	-2.14	3.2E-02	4.2E-01	12	-3.12	1.8E-03	8.0E-02
GO:0033504; p:floor plate development;	5	2.07	3.8E-02	4.4E-01	5	2.20	2.8E-02	3.1E-01
GO:0007596; p:blood coagulation;	56	2.06	3.9E-02	4.4E-01	56	2.62	8.9E-03	1.8E-01
GO:0030866; p:cortical actin	5	-1.99	4.7E-02	4.7E-01	5	-2.00	4.5E-02	3.7E-01

cytoskeleton organization;

GO term	S1				S2			
	Frequency	Page z-Score	Page Raw *p*-value	Page False Discovery Rate *p*-value	Frequency	Page z-Score	Page Raw *p*-value	Page False Discovery Rate *p*-value
GO:0042719; c:mitochondrial intermembrane space protein transporter	11	-3.12	1.8E-03	1.7E-01	11	-2.30	2.2E-02	2.6E-01
GO:0070469; c:respiratory chain;	13	2.69	7.2E-03	2.2E-01	13	2.46	1.4E-02	2.2E-01
GO:0000784; c:nuclear chromosome, telomeric region;	6	-2.53	1.1E-02	2.2E-01	6	-2.68	7.5E-03	2.2E-01
GO:0015629; c:actin cytoskeleton;	32	-2.53	1.1E-02	2.2E-01	32	-1.99	4.6E-02	3.6E-01
GO:0032133; c:chromosome passenger complex;	6	2.07	3.9E-02	4.3E-01	6	2.63	8.5E-03	2.2E-01

GO term	S1				S2			
	Frequency	PAGE Z-Score	PAGE Raw *p*-value	PAGE False Discovery Rate *p*-value	Frequency	PAGE Z-Score	PAGE Raw *p*-value	PAGE False Discovery Rate *p*-value
GO:0004692; f:cgmp-dependent protein kinase activity;	9	-3.64	2.7E-04	4.9E-02	9	-3.13	1.7E-03	7.7E-02

GO:0004020; f:adenylylsulfate kinase activity;	10	-3.51	4.4E-04	4.9E-02	10	-3.47	5.3E-04	4.1E-02
GO:0004781; f:sulfate adenylyltransferase (atp) activity;	10	-3.51	4.4E-04	4.9E-02	10	-3.47	5.3E-04	4.1E-02
GO:0042626; f:atpase activity, coupled to transmembrane movement of	19	3.27	1.1E-03	7.6E-02	19	3.22	1.3E-03	6.2E-02
GO:0042605; f:peptide antigen binding;	8	3.26	1.1E-03	7.6E-02	8	2.48	1.3E-02	2.4E-01
GO:0030955; f:potassium ion binding;	9	3.21	1.3E-03	7.6E-02	9	3.46	5.5E-04	4.1E-02
GO:0070330; f:aromatase activity;	14	-3.18	1.5E-03	7.6E-02	14	-3.76	1.7E-04	3.0E-02
GO:0010181; f:fmn binding;	12	2.90	3.8E-03	1.4E-01	12	2.07	3.8E-02	3.7E-01
GO:0004743; f:pyruvate kinase activity;	7	2.77	5.7E-03	1.7E-01	7	2.85	4.3E-03	1.4E-01
GO:0001567; f:cholesterol 25-hydroxylase activity;	9	2.54	1.1E-02	2.2E-01	9	2.05	4.0E-02	3.8E-01
GO:0043425; f:bhlh transcription factor binding;	9	-2.51	1.2E-02	2.3E-01	9	-2.95	3.2E-03	1.3E-01

GO:0003887; f:dna-directed dna polymerase activity;	23	2.22	2.7E-02	3.3E-01	23	2.18	2.9E-02	3.4E-01
GO:0004864; f:protein phosphatase inhibitor activity;	13	-2.16	3.1E-02	3.4E-01	13	-3.29	1.0E-03	5.5E-02
GO:0004012; f:phospholipid-translocating atpase activity;	9	2.16	3.1E-02	3.4E-01	9	2.25	2.4E-02	3.4E-01
GO:0005212; f:structural constituent of eye lens;	24	-2.08	3.7E-02	3.8E-01	24	-2.54	1.1E-02	2.1E-01
GO:0051864; f:histone demethylase activity (h3-k36 specific);	5	2.04	4.1E-02	4.0E-01	5	2.22	2.6E-02	3.4E-01
GO:0004402; f:histone acetyltransferase activity;	9	-2.02	4.4E-02	4.0E-01	9	-2.17	3.0E-02	3.4E-01
GO:0001948; f:glycoprotein binding;	6	2.01	4.5E-02	4.0E-01	6	2.15	3.2E-02	3.5E-01
GO:0005085; f:guanyl-nucleotide exchange factor activity;	38	1.96	5.0E-02	4.2E-01	38	3.72	2.0E-04	3.0E-02

Table 3.5. Sub-network enrichment analysis (SNEA) for cellular processes and diseases in *S. tropicalis* exposed to S1 mixture of NAs. All gene set seeds that had >10% overlap with the annotated pathways are listed ($p < 0.05$). The total number of neighbors refers to the total number of known entities in a sub-network.

Gene Set Seed cellular processes	Total # of Neighbors	# of Measured Neighbors	Median change	*p*-value
xenobiotic clearance	370	80	1.16	2.2E-05
liver uptake	86	15	1.24	4.4E-04
natriuresis	103	20	-1.06	8.1E-04
intestinal absorption	135	34	1.07	8.6E-04
drug metabolism	108	22	1.17	9.4E-04
vascular smooth muscle relaxation	54	9	-1.18	1.6E-03
hepatobiliary excretion	18	6	1.31	3.1E-03
endothelial cell production	45	6	1.11	3.8E-03
pregnancy	1106	158	1.01	4.1E-03
heart polarization	21	5	-1.29	4.9E-03
retinoid metabolism	33	10	1.14	5.0E-03
ion transport	211	34	-1.11	5.7E-03
vitamin A metabolism	19	8	1.05	5.8E-03
digestion	367	87	-1.04	6.3E-03
drug transport	57	13	1.19	6.7E-03
arachidonic acid metabolism	87	9	1.06	7.1E-03
renin-angiotensin system	123	25	1.11	7.2E-03
gallstone formation	41	9	1.19	7.8E-03
T-cell homeostases	103	11	-1.06	8.0E-03
vasoconstriction	320	42	-1.02	8.5E-03
Cl- transport	228	36	-1.06	8.7E-03
Fatty acids import	101	15	1.16	9.8E-03
anion transport	38	7	1.24	1.1E-02
retinoic acid metabolism	24	5	1.13	1.2E-02
liver metabolism	116	31	1.08	1.4E-02
membrane depolarization	654	108	-1.00	1.5E-02
translation termination	42	11	-1.13	1.5E-02
Na+ influx co-transport	44	8	-1.11	1.5E-02
RNA modification	56	7	1.28	1.6E-02
lipid transport	368	64	1.03	1.6E-02
female gonad development	74	16	-1.09	1.6E-02
pore formation	91	15	1.11	1.7E-02
lipoprotein metabolism	121	20	-1.10	1.7E-02
intracellular signaling cascade	195	31	-1.11	1.8E-02
blood vessel contraction	111	11	-1.08	1.9E-02

placenta development	293	42	-1.10	1.9E-02
vasodilation	406	58	1.02	2.2E-02
ethanol metabolism	27	8	1.31	2.4E-02
fluid transport	60	17	-1.11	2.4E-02
bile acid metabolism	52	10	1.29	2.5E-02
uterus blood flow	23	6	1.11	2.6E-02
immune response	1638	204	1.01	2.8E-02
chromatin silencing	24	6	1.32	2.8E-02
bile acid secretion	74	14	1.11	2.8E-02
neuromuscular synaptic transmission	52	7	-1.18	2.9E-02
cholesterol export	200	29	-1.00	3.0E-02
hormone metabolism	39	12	1.21	3.0E-02
acid secretion	156	29	-1.02	3.0E-02
hemato-encephalic barrier	326	52	-1.01	3.0E-02
ROS generation	1172	199	1.01	3.1E-02
HCO(3)(-) transport	116	21	-1.11	3.1E-02
protein-protein cross-linking via L-cystine	109	24	1.11	3.2E-02
intracellular pH	129	18	-1.06	3.3E-02
coronary flow	73	14	1.19	3.3E-02
gastrointestinal system digestion	26	6	1.07	3.6E-02
action potential duration	107	18	1.06	3.6E-02
nucleotide biosynthesis	35	5	-1.04	3.9E-02
histone ubiquitination	22	5	-1.11	4.0E-02
T-helper lymphocyte response	154	7	-1.18	4.0E-02
syncytium formation	58	9	1.02	4.1E-02
Cholesterol metabolism	185	35	-1.05	4.1E-02
fluid secretion	102	19	1.04	4.1E-02
sex maturation	75	11	1.11	4.2E-02
mRNA modification	19	6	1.19	4.2E-02
retinol metabolism	18	5	1.26	4.3E-02
perception of pain	304	48	1.02	4.4E-02
cell communication	145	24	-1.09	4.5E-02
antiport	33	7	1.17	4.6E-02
cartilage differentiation	48	9	-1.20	4.7E-02
synaptic target recognition	7	5	-1.10	4.7E-02
kidney filtration	80	13	1.11	4.8E-02
Steroid metabolism	392	66	1.06	4.9E-02
biliary flow	56	9	1.07	4.9E-02
free radical scavenging	35	10	1.07	5.0E-02
receptor clustering	95	11	1.11	5.0E-02

Gene Set Seed disease	Total # of Neighbors	# of Measured Neighbors	Median change	*p*-value

Neutropenia	90	9	1.48	6.5E-04
Hepatitis	207	27	-1.14	1.4E-03
airway hyperreactivity	274	32	1.05	1.5E-03
liver toxicity	627	110	1.06	5.7E-03
Cholelithiasis	52	13	1.07	5.8E-03
Testicular Neoplasms	49	5	1.27	5.9E-03
Constipation	36	6	-1.36	6.2E-03
Smoking	93	22	-1.04	7.1E-03
liver clearance	53	10	1.31	7.2E-03
Cystic Fibrosis	125	27	-1.06	7.2E-03
Thrombosis	338	50	0.00	7.8E-03
Colitis	503	60	1.01	8.2E-03
Meningitis, Bacterial	33	5	-1.10	8.8E-03
Edema	429	67	-1.02	1.3E-02
vascular reactivity	94	15	-1.12	1.3E-02
Dyslipidemias	147	26	0.00	1.4E-02
inflammatory lesions	73	8	-1.19	1.6E-02
Drug Toxicity	37	9	1.12	1.7E-02
Bronchial Spasm	37	5	1.20	1.8E-02
Dehydration	86	16	1.21	1.9E-02
Malabsorption Syndromes	41	11	-1.13	2.0E-02
Hyperuricemia	22	5	1.87	2.1E-02
Cerebrovascular Disorders	48	8	-1.18	2.3E-02
Hyperlipidemias	137	24	1.04	2.4E-02
Dyskinesias	73	17	-1.02	2.5E-02
tachyphylaxis	44	7	1.11	2.7E-02
Mycoses	56	9	-1.14	2.9E-02
Hypoalgesia	39	8	-1.11	3.0E-02
Glomerulonephritis, Membranoproliferative	30	6	1.06	3.1E-02
cardiovascular risk	123	21	-1.06	3.4E-02
Amyloidosis	82	11	-1.11	3.5E-02
Cardiovascular Diseases	465	73	-1.00	3.6E-02
Tetraploidy	40	7	-1.13	3.6E-02
Wounds and Injuries	1585	254	1.02	4.2E-02
Adverse Drug Reaction	39	9	1.29	4.2E-02
Emphysema	110	24	1.05	4.4E-02
Seizures, Febrile	34	5	-1.10	4.5E-02
Hyperhomocysteinemia	50	5	1.19	4.5E-02
Arterial stiffness	43	7	-1.25	4.5E-02
Cold Ischemia	34	5	-1.11	4.5E-02
Intestinal Diseases	116	16	-1.14	4.6E-02
Bronchitis, Chronic	28	6	-1.04	4.8E-02

Table 3.6. Sub-network enrichment analysis (SNEA) for cellular processes and diseases in *S. tropicalis* exposed to S2 mixture of NAs. All gene set seeds that had >10% overlap with the annotated pathways are listed ($p < 0.05$). The total number of neighbors refers to the total number of known entities in a sub-network.

Gene Set Seed cellular process	Total # of Neighbors	# of Measured Neighbors	Median change	*p*-value
intestinal absorption	135	34	1.16	1.9E-06
xenobiotic clearance	370	80	1.16	2.4E-05
retinoid metabolism	33	10	1.76	2.4E-04
fluid secretion	102	19	1.09	6.2E-04
liver metabolism	116	31	1.14	1.7E-03
drug metabolism	108	22	1.19	1.7E-03
gallstone formation	41	9	1.23	2.3E-03
vascular smooth muscle relaxation	54	9	-1.09	2.6E-03
blood-retinal barrier	61	7	-1.10	2.7E-03
cell communication	145	24	-1.07	2.9E-03
liver uptake	86	15	1.16	3.4E-03
blood vessel contraction	111	11	-1.13	3.4E-03
fluid transport	60	17	-1.01	4.8E-03
lipid transport	368	64	1.05	5.1E-03
retinoic acid metabolism	24	5	1.21	5.5E-03
anion transport	38	7	1.20	7.0E-03
ion transport	211	34	-1.09	7.8E-03
kidney filtration	80	13	1.16	7.8E-03
digestion	367	87	-1.04	8.6E-03
Na+ influx co-transport	44	8	1.15	8.9E-03
Steroid metabolism	392	66	1.05	9.3E-03
LDL oxidation	89	18	1.14	9.5E-03
podocyte function	48	8	1.23	1.1E-02
water transport	201	43	-1.07	1.1E-02
ethanol metabolism	27	8	1.19	1.2E-02
protein-protein cross-linking via L-cystine	109	24	1.09	1.3E-02
bile acid secretion	74	14	1.16	1.5E-02
arachidonic acid metabolism	87	9	1.16	1.5E-02
nucleotide biosynthesis	35	5	-1.08	1.6E-02
muscle cell development	21	8	-1.26	1.6E-02
endothelial cell production	45	6	1.28	1.6E-02
HCO(3)(-) transport	116	21	-1.09	1.9E-02
muscle relaxation	119	19	-1.09	1.9E-02
mammary gland involution	44	6	1.19	2.0E-02

Triacylglycerol mobilization	14	5	-1.17	2.0E-02
Sterol biosynthesis	30	6	1.15	2.1E-02
Cl- transport	228	36	1.05	2.1E-02
male meiosis	36	8	-1.09	2.1E-02
necrotic cell death	127	24	1.12	2.1E-02
tissue maintenance	189	34	1.05	2.3E-02
neutrophil differentiation	53	7	1.19	2.3E-02
biliary flow	56	9	-1.09	2.3E-02
acid secretion	156	29	-1.10	2.3E-02
hepatobiliary excretion	18	6	1.27	2.4E-02
lipid peroxidation	293	58	1.05	2.5E-02
macrophage differentiation	183	27	1.03	2.7E-02
skin barrier	97	12	1.13	3.1E-02
phosphate import	68	10	1.17	3.1E-02
Cholesterol metabolism	185	35	1.02	3.1E-02
antigen processing and presentation	417	46	1.03	3.1E-02
cell redox homeostasis	83	19	1.16	3.2E-02
vasoconstriction	320	42	-1.02	3.4E-02
blood circulation	417	63	1.05	3.5E-02
gastrointestinal system digestion	26	6	-1.17	3.6E-02
blastocyst development	46	9	1.08	3.8E-02
muscle contraction	191	20	1.03	3.9E-02
membrane depolarization	654	108	-1.02	3.9E-02
histone ubiquitination	22	5	-1.18	4.5E-02
triglyceride storage	51	5	-1.17	4.5E-02
hemato-encephalic barrier	326	52	-1.04	4.7E-02
sulfate import	26	5	-1.16	4.7E-02
phosphate metabolism	41	9	1.04	4.7E-02
T-cell response	563	60	1.04	4.8E-02
mitotic spindle orientation	31	9	1.01	4.8E-02
cartilage differentiation	48	9	-1.18	4.9E-02
muscle regeneration	178	30	-1.03	5.0E-02

Gene Set Seed disease	Total # of Neighbors	# of Measured Neighbors	Median change	*p*-value
Malabsorption Syndromes	41	11	-1.19	6.0E-05
liver toxicity	627	110	1.09	2.4E-03
Hepatitis	207	27	-1.07	3.4E-03
Cystic Fibrosis	125	27	-1.09	4.0E-03
Neoplastic Processes	46	8	-1.10	4.5E-03
Protein Deficiency	19	5	-1.19	5.8E-03
Pulmonary Disease, Chronic Obstructive	200	36	-1.09	8.6E-03
Constipation	36	6	-1.19	1.0E-02

Albuminuria	134	23	-1.05	1.1E-02
Heart Septal Defects, Ventricular	37	7	1.12	1.3E-02
Emphysema	110	24	1.05	1.3E-02
tachyphylaxis	44	7	1.13	1.4E-02
Cold Ischemia	34	5	-1.16	1.4E-02
mdx dystrophy	66	13	-1.18	1.5E-02
Neutropenia	90	9	1.35	1.6E-02
Lens Opacities	18	8	1.23	1.8E-02
Neoplasms, Intraepithelial	41	6	1.19	2.0E-02
Bronchitis, Chronic	28	6	-1.09	2.0E-02
Glaucoma, Open-Angle	38	6	-1.16	2.0E-02
Flushing	43	6	1.38	2.1E-02
Seizures, Febrile	34	5	-1.26	2.2E-02
stem cell cancer	17	6	1.19	2.3E-02
Shock	54	7	-1.19	2.3E-02
artery stenosis	73	14	-1.05	2.4E-02
Hypercalcemia	56	8	1.29	2.4E-02
Astrocytoma	149	37	1.00	2.5E-02
eye vascularization	20	5	-1.24	2.6E-02
Retinal Neovascularization	127	19	1.14	2.9E-02
Cicatrix, Hypertrophic	45	10	-1.16	2.9E-02
visual acuity	22	5	-1.24	2.9E-02
Meningitis, Bacterial	33	5	-1.07	3.1E-02
Hyperhomocysteinemia	50	5	1.10	3.2E-02
Airway Obstruction	69	10	-1.18	3.5E-02
bone marrow toxicity	58	9	1.19	3.6E-02
capillary density	95	21	1.06	3.7E-02
Acute Coronary Syndrome	88	11	0.00	3.8E-02
Arteriosclerosis	103	15	1.05	3.8E-02
Liver Cirrhosis, Biliary	63	10	1.16	3.8E-02
hematotoxicity	28	6	1.21	3.9E-02
teratogenicity	49	8	1.21	3.9E-02
Hyperalgesia	365	51	1.04	4.4E-02
Glomerulosclerosis, Focal Segmental	142	20	1.05	4.6E-02
Aortic Aneurysm	52	9	-1.33	4.8E-02
fibrillation	44	8	-1.21	4.8E-02
Stomach Ulcer	108	17	-1.06	4.8E-02
Edema	429	67	1.05	5.0E-02

Table 3.7. Sub-network enrichment analysis (SNEA) for common cellular processes and diseases in *S. tropicalis* exposed to S1 and S2 mixtures of NAs. All gene set seeds that had >10% overlap with the annotated pathways are listed ($p < 0.05$). The total number of neighbors refers to the total number of known entities in a sub-network.

	S1				S2			
Gene Set Seed cellular process	Total # of Neighbors	# of Measured Neighbors	Median change	*p*-value	Total # of Neighbors	# of Measured Neighbors	Median change	*p*-value
xenobiotic clearance	370	80	1.16	2.2E-05	370	80	1.16	2.4E-05
liver uptake	86	15	1.24	4.4E-04	86	15	1.16	3.4E-03
intestinal absorption	135	34	1.07	8.6E-04	135	34	1.16	1.9E-06
drug metabolism	108	22	1.17	9.4E-04	108	22	1.19	1.7E-03
vascular smooth muscle relaxation	54	9	-1.18	1.6E-03	54	9	-1.09	2.6E-03
hepatobiliary excretion	18	6	1.31	3.1E-03	18	6	1.27	2.4E-02
endothelial cell production	45	6	1.11	3.8E-03	45	6	1.28	1.6E-02
retinoid metabolism	33	10	1.14	5.0E-03	33	10	1.76	2.4E-04
ion transport	211	34	-1.11	5.7E-03	211	34	-1.09	7.8E-03
digestion	367	87	-1.04	6.3E-03	367	87	-1.04	8.6E-03
arachidonic acid metabolism	87	9	1.06	7.1E-03	87	9	1.16	1.5E-02
gallstone formation	41	9	1.19	7.8E-03	41	9	1.23	2.3E-03
vasoconstriction	320	42	-1.02	8.5E-03	320	42	-1.02	3.4E-02
Cl- transport	228	36	-1.06	8.7E-03	228	36	1.05	2.1E-02
anion transport	38	7	1.24	1.1E-02	38	7	1.20	7.0E-03
retinoic acid metabolism	24	5	1.13	1.2E-02	24	5	1.21	5.5E-03
liver metabolism	116	31	1.08	1.4E-02	116	31	1.14	1.7E-03
membrane depolarization	654	108	-1.00	1.5E-02	654	108	-1.02	3.9E-02
Na+ influx co-transport	44	8	-1.11	1.5E-02	44	8	1.15	8.9E-03
lipid transport	368	64	1.03	1.6E-02	368	64	1.05	5.1E-03

blood vessel contraction	111	11	-1.08	1.9E-02	111	11	-1.13	3.4E-03
ethanol metabolism	27	8	1.31	2.4E-02	27	8	1.19	1.2E-02
fluid transport	60	17	-1.11	2.4E-02	60	17	-1.01	4.8E-03
bile acid secretion	74	14	1.11	2.8E-02	74	14	1.16	1.5E-02
acid secretion	156	29	-1.02	3.0E-02	156	29	-1.10	2.3E-02
hemato-encephalic barrier	326	52	-1.01	3.0E-02	326	52	-1.04	4.7E-02
HCO(3)(-) transport	116	21	-1.11	3.1E-02	116	21	-1.09	1.9E-02
protein-protein cross-linking via L-cystine	109	24	1.11	3.2E-02	109	24	1.09	1.3E-02
gastrointestinal system digestion	26	6	1.07	3.6E-02	26	6	-1.17	3.6E-02
nucleotide biosynthesis	35	5	-1.04	3.9E-02	35	5	-1.08	1.6E-02
histone ubiquitination	22	5	-1.11	4.0E-02	22	5	-1.18	4.5E-02
Cholesterol metabolism	185	35	-1.05	4.1E-02	185	35	1.02	3.1E-02
fluid secretion	102	19	1.04	4.1E-02	102	19	1.09	6.2E-04
cell communication	145	24	-1.09	4.5E-02	145	24	-1.07	2.9E-03
cartilage differentiation	48	9	-1.20	4.7E-02	48	9	-1.18	4.9E-02
kidney filtration	80	13	1.11	4.8E-02	80	13	1.16	7.8E-03
Steroid metabolism	392	66	1.06	4.9E-02	392	66	1.05	9.3E-03
biliary flow	56	9	1.07	4.9E-02	56	9	-1.09	2.3E-02

Gene Set Seed disease	S1				S2			
	Total # of Neighbors	# of Measured Neighbors	Median change	*p*-value	Total # of Neighbors	# of Measured Neighbors	Median change	*p*-value
Neutropenia	90	9	1.48	6.5E-04	90	9	1.35	1.6E-02
Hepatitis	207	27	-1.14	1.4E-03	207	27	-1.07	3.4E-03
liver toxicity	627	110	1.06	5.7E-03	627	110	1.09	2.4E-03
Constipation	36	6	-1.36	6.2E-03	36	6	-1.19	1.0E-02
Cystic Fibrosis	125	27	-1.06	7.2E-03	125	27	-1.09	4.0E-03
Meningitis, Bacterial	33	5	-1.10	8.8E-03	33	5	-1.07	3.1E-02
Edema	429	67	-1.02	1.3E-02	429	67	1.05	5.0E-02

Malabsorption	41	11	-1.13	2.0E-02	41	11	-1.19	6.0E-05
Syndromes tachyphylaxis	44	7	1.11	2.7E-02	44	7	1.13	1.4E-02
Emphysema	110	24	1.05	4.4E-02	110	24	1.05	1.3E-02
Seizures, Febrile	34	5	-1.10	4.5E-02	34	5	-1.26	2.2E-02
Hyperhomocysteinemia	50	5	1.19	4.5E-02	50	5	1.10	3.2E-02
Cold Ischemia	34	5	-1.11	4.5E-02	34	5	-1.16	1.4E-02
Bronchitis, Chronic	28	6	-1.04	4.8E-02	28	6	-1.09	2.0E-02

Table 3.8. List of differentially expressed genes in *Silurana (Xenopus) tropicalis* exposed to NA S1 and S2 (2 mg/L) from microarray data and selected for RT-qPCR validation.

Gene symbol	Name	Human Protein ID	Gene network	Fold change S1	*p*-value S1	FDR adjusted *p*-value S1	Fold change S2	*p*-value S2	FDR adjusted *p*-value S2
cyp4b1	*cytochrome P450 4B1*	P13584	• Metabolism	110.38	2.42E-10	5.0E-06	105.21	2.66E-10	5.0E-06
abcg2	*ATP-binding cassette sub-family G member 2*	Q9UNQ0	• Metabolism • Gut function	2.83	8.44E-07	5.4E-04	2.62	1.66E-06	8.8E-04
slc26a6	*solute carrier family 26 member 6*	Q9BXS9	• Membrane function • Gut function	2.27	5.19E-03	9.6E-02	2.23	5.90E-03	1.0E-01
eprs	*bifunctional glutamate/proline-tRNA ligase*	P07814	• Edema	2.36	1.59E-03	5.4E-02	3.54	9.79E-05	1.3E-02
slcs5a1	*sodium/glucose cotransporter 1*	P13866	• Membrane function • Metabolism • Gut function	-2.05	3.02E-04	2.4E-02	-1.78	1.34E-03	4.9E-02
myod1	*myoblast determination protein 1*	P15172	-	-4.64	4.15E-02	2.6E-01	-5.65	2.50E-02	2.1E-01
rpl8	*ribosomal protein L8*	P62917	-	-1.01	8.37E-01	9.8E-01	-1.01	9.23E-01	1.0E+00

Chapter 4

Analysis of naphthenic acid mixtures as pentafluorobenzyl derivatives by gas chromatography-electron impact mass spectrometry

Chapter adapted from

Juan Manuel Gutierrez-Villagomez[1], Juan Vázquez-Martínez[2], Enrique Ramírez-Chávez[3], Jorge Molina-Torres[4], Vance L. Trudeau[5]. Talanta Volume 162, 1 January 2017, Pages 440-452 https://doi.org/10.1016/j.talanta.2016.10.057

Contributions:

1. Contributed to original ideas, conducted experimental design, performed experiments, data analysis and manuscript writing and submission.

2. Contributed to original ideas, expertise and revised the manuscript. Departamento de Biotecnología y Bioquímica, CINVESTAV Unidad Irapuato, Guanajuato, Mexico.

3. Expertise and revised the manuscript. Departamento de Biotecnología y Bioquímica, CINVESTAV Unidad Irapuato, Guanajuato, Mexico.

4. Expertise, provided material and revised the manuscript. Departamento de Biotecnología y Bioquímica, CINVESTAV Unidad Irapuato, Guanajuato, Mexico.

5. Financial support and scientific guidance. Department of Biology, University of Ottawa, Ottawa, Ontario K1N 6N5, Canada.

4.1. Abstract

In this chapter, the efficiency of pentafluorobenzyl bromide (PFBBr) for naphthenic acid (NA) mixtures derivatization, and the comparison in the optimal conditions to the most common NAs derivatization reagents, BF_3/MeOH and *N*-(t-butyldimethylsilyl)-*N*-methyltrifluoroacetamide (MTBSTFA) is reported for the first time. Naphthenic acids are carboxylic acid mixtures of petrochemical origin. These compounds are important for the oil industry because of their corrosive properties, which can damage oil distillation infrastructure. Moreover, NAs are commercially used in a wide range of products such as paint and ink driers, wood and fabric preservatives, fuel additives, emulsifiers, and surfactants. Naphthenic acids have also been found in sediments after major oil spills in the United States and South Korea. Furthermore, the toxicity of the oil sands process-affected water (OSPW), product of the oil sands extraction activities in Canada's oil sands, has largely been attributed to NAs. One of the main challenges for the chromatographic analysis of these mixtures is the resolution of the components. The derivatization optimization was achieved using surface response analysis with molar ratio and time as factors for derivatization signal yield. After gas chromatography-electron impact mass spectrometry (GC-EIMS) analysis of a mixture of NA standards, it was found that the signal produced by PFB-derivatives was 2.3 and 1.4 times higher than the signal produced by methylated and MTBS-derivatives, respectively. The pentafluorobenzyl derivatives have a characteristic fragment ion at 181 *m*/*z* that is diagnostic for the differentiation of carboxylic and non-carboxylic acid components within mixtures. In the analysis of a Sigma and a Merichem derivatized oil extract NA mixtures, it was found that some peaks lack the characteristic fragment ion;

therefore they are not carboxylic acids. Open column chromatography was used to obtain a hexane and a methanol fraction of the Sigma and Merichem mixtures. The components in the hexane fraction, presumably hydrocarbons that did not react with PFBBr were ~7% by weight. The effectiveness of PFBBr was confirmed when the two NA oil extracts were spiked with 8 distinct NA standards and identified by GC-EIMS in the methanol fraction. The retention indices of the methyl, MTBS and PFB derivatives of these 8 NAs are also reported. The use of PFBBr increases sensitivity, chromatographic resolution, and identification accuracy for the analysis of standards and mixtures of NAs compared to MTBSTFA and BF_3/MeOH. This methodology will have wide applications in the elucidation of NA mixtures.

4.2. Introduction

The IUPAC defines naphthenic acids (NAs) as “acids chiefly monocarboxylic, derived from naphthene” [25]. However, the oil industry has adopted the term NAs to encompass all the carboxylic acids present in crude oil which includes a mixture of alkyl-substituted cyclic and aliphatic carboxylic acids (Fig. 1.2) [22]. Not all of these are naphthalene derivatives, thus the industry definition will be used here. NAs are represented by the general formula: $C_nH_{2n-z}O_2$, where *n* indicates the carbon number, *z* specifies the deficiency of hydrogen from the formation of rings and groups the NAs in a homologous series (Fig. 1.2) [22].

Although NAs correspond to 0–3% of oil by weight [22], they are important for the oil industry because of their corrosive properties [116,117], which can damage oil distillation towers [118]. Moreover, NAs are also used commercially in a wide range of

products such as paint and ink driers, wood and fabric preservatives, fuel additives, emulsifiers, and surfactants. NAs have been found in sediments after major oil spills in the United States [65,66] and South Korea [67]. The toxicity of the oil process-affected water (OSPW), product of the oil sands extraction activities in Canada's oil sands, has largely been attributed to the presence of NAs [20,119].

The chemical composition of NA mixtures is highly diverse and may also contain non-classical NAs that do not fit the general formula of NAs [120], nitrogen or sulfur containing compounds [78], alcohols, ketones, and ethers [121]. The analysis of such mixtures has been widely explored [122–127], however, the main challenge for their characterization is the resolution of the components in the chromatographic phase. In this regard, the use of derivatization techniques to improve the chromatographic characteristics of the mixtures of NAs by GC-MS has been reported previously. Despite the problems summarized by Walker Christie [128] and the superior results offered by silylation methods [129], the alkylation reaction of NAs with BF_3/MeOH remains widely used [33,34,36,44,54,130,131]. In the methylation reaction, the labile hydrogen of the carboxylic group is substituted with a methyl group [132]. Silylation is also commonly used for the analysis of NA mixtures [24,27,45,133–135], wherein the case of *N*-(t-butyldimethylsilyl)-*N*-methyltrifluoroacetamide (MTBSTFA) the labile hydrogens of carboxylic acids, alcohols, phenols, and amines, among others, are replaced with a *t*-butyldimethylsilyl (*t*-BDMS) group [132].

This chapter describes the optimization and comparison of three derivatization reagents in a mixture of NA standards and two oil extracted NA mixtures. The comparison was performed using optimal conditions of molar ratio (derivatization

reagent/NA) and time. The derivatization reagents were BF_3/MeOH, MTBSTFA, and an alkylation reagent not previously reported for the analysis of NA mixtures: *alpha*-bromo-2,3,4,5,6-pentafluorotoluene (pentafluorobenzyl bromide or PFBBr). In the reaction with PFBBr, the labile hydrogen of the carboxylic group is substituted with a pentafluorobenzyl group [132]. PFBBr can be used for the derivatization of phenols, thiols and carboxylic acids [136,137]. However, the reaction is selective to carboxylic acids in the presence of a weak base, e.g., KOAc, $KHCO_3$, KCNO or organic bases [138,139]. PFBBr has been used successfully for the analysis of mixtures of compounds with carboxylic groups, for example, carboxylic acids and phenols in air [140], hydroxy acids in wine [141] and fatty acids in blood plasma [142]. In this part of the thesis, it is shown that PFBBr derivatization increases sensitivity, chromatographic resolution and NAs identification accuracy for NA mixtures analysis by GC-EIMS.

In January 2016, I visited Dr. Jorge Molina Torres' laboratory at CINVESTAV IPN Mexico to conduct the experiments for this chapter. I conducted all these experiments under the supervision of Dr. Juan Vazquez Martinez and M. Sc. Enrique Ramirez Chavez, who helped me refine the experimental design and solving problems.

4.3. Experimental

4.3.1. Reagents and chemicals

Eight single NA standards were purchased from two suppliers, as described in Table 4.1. One of the NA mixtures was purchased from Sigma-Aldrich Co. Ltd. (St. Louis, United States), with lot number BCBC9959V. The Merichem NA mixture was a gift from Dr. John Headley (Environment Canada, Saskatoon, SK). All solvents were

HPLC-grade (Sigma-Aldrich Co. Ltd, St. Louis, United States). BF_3/MeOH (14% w/v), MTBSTFA, PFBBr were used as purchased (Sigma-Aldrich Chemie, Steinheim, Germany). NA standards were selected to cover a variety of chemical structures and molecular weights. A C7 to C40 alkane calibration standard (Sigma-Aldrich Chemie, Steinheim, Germany) was used to calculate the retention indices (Ri).

4.3.2. Derivatization

For the three derivatization reactions, an equimolar solution of all of the standards was prepared in methanol. From this solution, a defined volume containing 1 µmol of each standard was transferred to vials and dried under nitrogen flux, to later proceed with the derivatizations. Sigma and Merichem extract stock solutions were prepared in methanol. For these derivatizations, an average molecular weight of 184 amu was assumed based on preliminary experiments. From these solutions, a defined volume containing 5 µmol was transferred to vials and dry under nitrogen flux to later proceed with the derivatizations.

To obtain the methyl esters of the NAs, the derivatization was performed by adding BF_3/MeOH in different volumes depending on the reaction molar ratio (Table 4.2) and the reaction was adjusted to 500 µL with acetonitrile as organic solvent. Samples were then heated to 60 °C and mixed at 1000 rpm. To recover the methyl esters, 500 µL of hexane and 500 µL of water were added and the resulting mix was vortexed. The mix was then centrifuged for 2 min at 6000 rpm to separate the two phases. An aliquot of the hexane phase was taken and mixed with isooctane to obtain a final concentration of 0.3 mM for each methylated standard and 6 µg/µL for the oil extracted NA mixtures, assuming 100% methyl esterification.

The silyl derivatives of the mixture of NA standards and the oil extracted NA mixtures were obtained with MTBSTFA (> 97% purity), and 20 µL of pyridine (99% purity) in each reaction as a catalyst. MTBSTFA was added according to Table 4.2. The reaction volume was adjusted to 200 µL with isooctane as organic solvent. Samples were then heated to 80 °C and mixed at 1000 rpm. Once the reaction finalized, the solvent, the MTBSTFA, and the pyridine excess were evaporated to dryness under nitrogen flux, and the residue was re-dissolved in isooctane to a final concentration of 0.3 mM for each standard and a concentration of 6 µg/µL for the oil extracted NA mixtures.

In addition to BF_3/MeOH and MTBSTFA, we propose pentafluorobenzyl bromide (PFBBr) for the derivatization and analysis of NAs. This idea emerged from the necessity to improve NA resolution during the chromatography phase to analyze the components within mixtures. The pentafluorobenzyl group is a relatively stable ion that herein is found as characteristic ion for PFB-carboxylic acid derivatives. The derivatization reaction with PFBBr can be selective for NAs and other carboxylic acids if the reaction is done using a weak base [138,139]. The weak base used for these reactions was *N,N*-diisopropylethylamine. PFBBr was added according to Table 4.2 and 50 µL of *N,N*-diisopropylethylamine were added in each reaction as a catalyst and in order to make the reaction selective to carboxylic acids. The reaction volume was adjusted to 200 µL with chloroform as organic solvent. The samples were then heated to 60 °C and mixed at 1000 rpm. At the end of the reaction the solvent, the weak base, and the PFBBr excess were evaporated to dryness under nitrogen flux, the residue was re-dissolved in a mixture of chloroform/methanol 1:1 to a final concentration of 0.3 mM

for each standard and a concentration of 6 µg/µL for the oil extracted NA mixtures. Additionally, estradiol-17β (Sigma-Aldrich, ≥ 98%) was used as negative control for the selective derivatization of carboxylic acids with PFBBr using a weak base.

4.3.3. Response surface analysis

In order to find the optimum conditions for the three derivatization reactions, response surface methodology (RSM) with center composite face design (CCFD) was used to evaluate the effect of the molar ratio of each derivatization reagent/NA and the reaction time. Each factor had three levels –1, 0 and 1. The conditions of the reactions are shown in Table 4.2. This methodology has been widely used in analytical chemistry [143]. The derivatization performance was evaluated according to the signal yielded or area under the peak by the derivatized standards measured by the GC-EIMS and analyzed in the chromatograms using the methodology previously reported by Shepherd et al. [129]. Briefly, 10 experiments were carried out per derivatization to obtain 10 chromatograms with 8 peaks in each chromatogram. The signal produced by these 80 peaks correspond to 100% of the signal. The percentage of the total signal that is represented by each out of the 80 peaks was obtained with the formula $NA_{signal\text{-}yield} = (peak\ area)/(total\ area)*100$. The signal produced by the derivatized standards according to the conditions set for each of the 10 experiments was calculated using the formula $Experiment\ signal\ yield = \sum NA_{signal\text{-}yield}$. The same principle was applied for the comparison of the 3 different derivatizations. Briefly, the mixture of standards was derivatized under optimized conditions as determined by RSM analysis. Three chromatograms were obtained with 8 peaks in each chromatogram. The area of these 24 peaks corresponds to 100% of the signal produced by the derivatives. The

percentage of the signal that is represented by each peak was calculated using the formula $NA_{signal\text{-}yield} = (peak\ area)/(total\ area)*100$ and a summation was used to know the percentage of the whole signal that is represented by each type of derivatization.

4.3.4. Open column chromatography

The chromatographic columns were prepared by loading the 0.7 cm × 10 cm glass columns with silica gel 60 (Merck KGaA, Darmstadt, Germany) suspended in cold hexane so that the column packing was ~8 cm long. Each column was flushed with 3 mL of cold hexane, then loaded with the respective oil extracted NA mixture (0.1 g) as a “neat oil” and it was allowed to pass through the column by gravity. Once the oil extracted NA mixture was fully loaded, hexane was added until 2 mL was recovered from the column to collect the hexane fraction. This was followed by 1 mL of ethyl acetate and 2 mL of methanol until all the fraction passed through the column. This process was done for the two oil extracted NA mixtures. Additionally, the two oil extracted NA mixtures were spiked with 1.3 µmol of each of the 8 NA standards and separated by open column chromatography to confirm the separation of carboxylic acids (methanol fraction) from hexane fraction in the column. The 2 fractions obtained in each of the column experiments (methanol and hexane fraction Sigma extract, methanol and hexane fraction Merichem extract, methanol and hexane fraction Sigma extract spiked with standards, methanol and hexane fraction Merichem extract spiked with standards) were dried under nitrogen flux, weighted, later derivatized with PFBBr and analyzed by GC-EIMS.

4.3.5. Instrumentation and software

A gas chromatograph (Agilent Technologies model 7890A GC System) coupled with an electron impact ionization mass spectrometer (Hewlett Packard model 5973 Mass Selective Detector) was used for the analysis. The data obtained by the GC-EIMS was collected with the software MassHunter Workstation version B.06.00 (Agilent Technologies, Inc.). Agilent MassHunter Qualitative Analysis version B.06.00 was used for the ion distribution analysis of the commercial mixtures. The software Automated Mass Spectral Deconvolution and Identification System "AMDIS" (http://www.amdis.net/) was used for the determination of the retention time and mass spectrum for each component of the chromatograms of the NA standards and oil extracted NA mixtures. MSD ChemStation Data Analysis version E.02.01.1177 was used to calculate the linearity of the response of the derivatized standards. Minitab 17 was used for the analysis of the surface response analysis. The conditions and parameters of the chromatography phase were adjusted to obtain the best mixture components resolution with the lowest interference of the noise signal. All the parameters of the MS were optimized to obtain the best signal/noise ratio of the analytes.

Pulsed splitless injection (1 µL) was used. The injector temperature was set to 250 °C. The separation of the components was performed in a capillary column Zebron ZB-1MS (60 m × 320 µm × 1 µm) and helium used as carrier gas at a constant flow rate of 1 mL/min. The GC oven program began at an initial temperature of 50 °C, held for the first minute, and then was increased at a rate of 10 °C/min to a final temperature of 300 °C, held for 35 min. The transfer line temperature was set at 260 °C.

Electron impact mass spectra were obtained at 70 eV of electron energy. Measurements were performed in SCAN mode with *m*/*z* range set to 40-550. The ion source and quadrupole analyzer temperature were 230 °C and 150 °C respectively and operated at 2.9 scans per second. A different solvent delay was selected for each type of derivatization to avoid damage to the MS filament, without disrupting the measurement of low molecular weight NAs. The solvent delay for BF_3/MeOH was 11 min and 15 min for MTBSTFA and PFBBr derivatives.

4.3.6. Data processing

The Ri of the derivatives from the NA standards were calculated by alkane linear retention indices (C7 to C40) following the methodology of Sun and Stremple [144]. Using the optimal conditions for the three derivatizations, a concentration series of the three derivatized mixtures of NA standards ranging from 0.0006 to 0.3 mM were analyzed in order to determine linearity, limits of detection and limits of quantification. Calibration curves were generated to assess the linearity of the GC-EIMS measurements using ChemStation. The linear ranges varied from 0.001 to 0.3 mM. Limits of detection and quantification were calculated using AMDIS for each of the eight standards on the basis of a signal to noise ratio of 3 and 10, respectively [145]. To generate the ion distribution of the oil extract mixtures, a custom database of possible NA formulas was created spanning n = 4 - 27 and z = 0–10 starting from the formula of classical NAs ($C_nH_{2n-z}O_2$), using the methodology reported by Clemente et al. [146]. Only peaks with areas greater than 1% compared to the greatest peak were considered for analysis. The resolution of adjacent peaks was calculated according to the IUPAC formula for resolution in GC [25].

4.4. Results and discussion

4.4.1 Response surface methodology

The relationship between the derivatization signal yield, the molar ratio and derivatization time was determined using RSM. Ten experiments were carried out per derivatization and the results are shown in Table 4.2. Analysis of variance (ANOVA) with alpha (α) level of 0.05 showed the significance of all of the quadratic models. After a stepwise selection of terms where the terms with alpha level greater than 0.15 were removed, the equations obtained were: $z = -1.19+0.01558x+0.1050y-0.000005x^2$ ($p < 0.001$) for BF_3/MeOH, $z = 12.06-0.0534x+0.0619y-0.002293y^2+0.000913xy$ ($p = 0.021$) for MTBSTFA and $z = 4.09+0.2769x+0.1069y-0.00250x^2-0.003062xy$ ($p = 0.025$) for PFBBr (z = signal yield, x = molar ratio and y = time). The unexplained variance for each model is 5.84%, 13.56%, and 14.51%, for BF_3/MeOH, MTBSTFA, and PFBBr, respectively. The significance of each of the terms of the models is shown in Table 4.3. The cross-validation analysis indicated the predicted R^2 for each of the models of 81.9%, 10.6%, and 3.7%, for BF_3/MeOH, MTBSTFA, and PFBBr, respectively. Thus, there are likely other factors affecting the derivatization of NAs with MTBSTFA and PFBBr. The predicted R^2 value of the BF_3/MeOH model indicates a greater predictive ability than the MTBSTFA and PFBBr models.

The contour and 3D surface plots for the model of each derivatization reagent show the responses of the molar ratio of derivatization reagent/NA and the time of reaction (Fig. 4.1). The Figs. 4.1 a-b show a maximum derivatization signal yield for the conditions here tested for BF_3/MeOH at 1250 M ratio and 50 min of reaction. The

optimal conditions for MTBSTFA (Figs. 4.1 c-d) were observed at 25 M ratio and 18.5 min. PFBBr plots (Figs. 4.1 e-f) show optimal conditions at 49 M ratio and 10 min. It was observed that there is an increase in the derivatization signal yield when the molar ratio and the time of reaction increase with BF_3/MeOH (Figs. 4.1 a-b). The derivatization using BF_3/MeOH is disadvantageous compared to MTBSTFA and PFBBr because it takes more than 50 min for the reaction to reach maximum signal yield and requires a large molar excess of reagents. With MTBSTFA, the area for optimal derivatization was ~25–30 M ratio and 10–25 min of reaction, however beyond the optimal conditions, the derivatization signal yield decreased (Figs. 4.1 c-d). For PFBBr derivatization there is a wide area of optimal signal yield with the interaction of the time and the molar ratio (Figs. 4.1 e-f). The optimal conditions were selected to compare the results of the three derivatization reagents according to their derivatization signal yield.

4.4.2. GC analysis of mixtures of standards

The Fig. 4.2 shows the GC-EIMS total ion chromatograms (TIC) of the derivatized NA standards mixture as listed in Table 4.1. In these chromatograms, it is possible to observe the elution order and retention time for NA standards in relation to the derivatization reagent used. The methylated-NAs had shorter retention times, followed by the MTBS-NAs. The derivatization with PFBBr increased the retention time for the corresponding standards. This is a result of the increases in molecular weight of the compounds: 14 amu are added for each labile hydrogen in the NAs through methylation, whereas 114 amu are added to NAs using MTBSTFA, and 180 amu using PFBBr.

The results showed that PFBBr is the most effective among the three derivatization reagents when considering their derivatization signal yields (Fig. 4.2, Table 4.4). The signal produced by PFB-derivatives is 2.3 times higher than the signal produced from using the methylation reagent and 1.4 times higher to the silylation reagent. This could be due to the addition of 180 amu to the model NAs, increasing the sensitivity of the GC-EIMS [147]. PFBBr might also have a greater chemical derivatization yield, an increase in the volatility of the PFB-derivatives or increase in ionization efficiency for PFBBr derivatives. However, this is speculative. Furthermore, MTBSTFA has higher signal yields compared to BF3/MeOH (Table 4.4).

The span of the retention times for the first and last standards eluted for methylated NAs, MTBS-NAs, and PFB-NAs were 19.8, 19.5, and 21.3 min, respectively. This is a result of the increase of molecular weights according to the derivatization reagent employed. It is important to note that reaction with BF_3/MeOH and MTBSTFA give an aliphatic group, however, PFBBr gives a halogenated-aromatic group, altering the interaction of the compounds with the column. On average, PFB-derivatives show 54% and 140% more chromatographic resolution compared to methylated and silyl derivatives, respectively (Fig. 4.2). Due to the complexity of the mixtures, only a few NAs have been clearly identified previously [54]. However, the use of PFBBr as derivatization reagent for NA mixtures may improve the analysis and identification of NAs, since PFBBr increases the span of elution time and resolution of NAs with 6-20 carbons. Furthermore, the preparation of PFB-derivatives is more efficient since the derivatization reaction with PFBBr takes 10 min producing a higher signal. In

comparison the methylation and silylation reactions take 50 and 18.5 min, respectively, producing a lower derivatization signal yield (Table 4.4, Fig. 4.2).

4.4.3. Mass spectral analysis of NA standards

The mass spectral analysis of the NA standards mixture revealed several advantages of the PFBBr derivatization. All of the PFB derivative mass spectra show a fragment at 181*m/z* as the base peak or as the second most abundant ion, which corresponds to the pentafluorobenzyl radical fragment (Fig. 4.3 q-y). When the derivatization reaction is performed using a weak base, PFBBr reacts selectively with NAs [138,139], yielding a fragment ion at 181*m/z*. Estradiol-17β was used as a negative control for the selective derivatization of carboxylic acids with PFBBr using a weak base. Estradiol-17β has two hydroxyl groups, one of them forming a phenolic moiety (hydroxyl group bonded to an aromatic hydrocarbon group). Using the reaction and conditions herein proposed for derivatization of NAs with PFBBr, the peak of estradiol-17β-PFB was not detected in the GC-EIMS chromatogram (data not shown), however, a peak of non-derivatized estradiol-17β was detected at 42.5 min. This is especially useful because the differentiation of carboxylic acids from other components in complex mixtures of NAs is possible with PFBBr as the derivatization reagent. It was not possible to identify a general characteristic fragment of the derivatized standards using BF_3/MeOH or MTBSTFA. Nevertheless, MTBS-derivatives present a 75*m/z* ion corresponding to $[C_2H_6OSi+H]^+$, which appears in one case as base peak (Fig 4.3 i-p). Another important advantage of using PFBBr is that the molecular ion $[M]^+$ of all PFB derivatives can be detected. This was not possible in all of the cases using MTBSTFA (Table 4.5). When MTBSTFA is used as derivatization reagent in the mixture of NA

standards the molecular ion was not detected in 6 out of the 8 standards. In these 6 cases, the major fragmentations peaks of the MTBS derivatives were $[M\text{-}CH_3]^+$ due to the loss of a methyl group [148–150].

The signature fragment ion of PFB-NA standards at 181*m*/*z* is one of the 10 most abundant in the respective mass spectrum (Fig. 4.3 q-y). The relative proportion of the fragment ion at 181*m*/*z* ranged from 33% to 99% (Table 4.5). It was predicted and determined that the commercial NA mixture mass spectra components also had the characteristic fragment ion at 181*m*/*z*. The variation in the proportion of the 181*m*/*z* ion is related to variations in the chemical stability of the derivatized NA to the electron impact ionization. For example, the mass spectrum of PFB-adamantane carboxylic acid has a characteristic fragment ion at 135*m/z* (in addition to the fragment ion at 181*m*/*z*) that corresponds to the polycyclic ring, which forms the core of the adamantanes (Fig. 4.3 u-v). The observation that characteristic ions are conserved provides information that helps with the identification of such compounds (Fig. 4.3).

In order to identify any compound using its mass spectrum data is necessary to obtain a “pure” mass spectrum, but when two or more components of a complex mixture cannot be resolved in the chromatographic phase, the mass spectrum of the unresolved peak contains the combined fragments of the coeluted components. In a complex NA mixture in which there is no information on the number of components and their respective molecular weights, there is a possibility that each fragment of an unresolved peak corresponds to the molecular weight of a compound. In the PFB derivatives, the possibility to assign an ion to the molecular weight of a component decreases considerably. The minimum expected molecular weight of the smallest PFB-NA is 268,

whereas for methylated NAs is 102 and for silyl NAs is 202. Consequently, only an ion with a higher *m*/*z* than the molecular weight of the smallest derivatized NA can be the molecular ion. For MTBS derivatives, the loss of a methyl (*m*/*z* = 15) was observed in some cases, therefore the expected molecular weight of the smallest MTBS-NA is 187, making the ion distribution analysis more difficult. The results obtained from the PFBBr derivatization lead to a more accurate NA identification and carboxylic acid characterization.

In three out of the eight standards, the limit of quantification was improved using MTBSTFA. In four out of the eight standards, the limit of quantification was improved using PFBBr. The limit of quantification for 1-adamantanecarboxylic acid was the same using MTBSTFA and PFBBr. A similar trend was observed in the calculation of the limits of detection. The results are listed in Table 4.6. Overall, the derivatization with PFBBr provided the best results for the analysis of model NAs.

4.4.4. Analysis of the oil extracted NA mixtures

The three derivatization procedures were evaluated to study two complex oil extract NA mixtures (Sigma and Merichem). The time of reaction and molar ratio of derivatization reagent/NAs were as found to be optimal in the experiments with the NA standards. The reaction times for BF_3/MeOH, MTBSTFA, and PFBBr were 50, 18.5 and 10 min, respectively. The molar ratios of derivatization reagent/NAs for BF_3/MeOH, MTBSTFA, and PFBBr were 1250, 25 and 49, respectively.

From the analysis of a derivatized NA mixture with BF_3/MeOH, it is not possible to obtain useful information for the identification of its components because the chromatogram shows a characteristic hump with unresolved peaks (Fig. 4.4 a and d),

similar to results that have been previously reported [33,34,36,44,54,121]. The chromatogram associated with the NA mixture derivatized with MTBSTFA (Fig. 4.4 b and e) exhibits better-resolved peaks than the methylation technique, but not to the same degree as with PFB derivatives (Fig. 4.4 c and f). In an unknown sample it is important to obtain the molecular ions of the components, however using MTBSTFA in 6 out of the 8 standards the molecular ion was not detected, giving incomplete mass spectra. Therefore, for precise molecular weight determinations, a more complex spectral analysis is required.

In the analysis of the Sigma mixture derivatized with PFBBr, 206 peaks were detected, of which 75% present the fragment ion at 181*m*/*z*. For the Merichem mixture, 176 peaks were detected of which 68% present the fragment ion at 181*m*/*z*. As determined by a lack of the fragment ion at 181*m*/*z* several components in the mixtures were not derivatized. This was approximately 25% in the Sigma mixture and 32% in the Merichem mixture (Fig. 4.4 c and f) so there are compounds in the mixtures that are not carboxylic acids since they do not react with PFBBr. The analysis of the peaks that did not have the fragment ion at 181*m*/*z* in the 2 oil extract mixtures showed “ski-slope” fragmentation patterns with peaks separated by 14*m*/*z*, which correspond to CH_2 units. These mass spectra might correspond to branched or unbranched saturated hydrocarbons [151]. This outcome shows an important advantage of derivatizing complex NA mixtures with PFBBr. The detection of these non-derivatizable components has not been reported with the other derivatization methods, perhaps since complex mixtures of NAs have not yet been entirely resolved. In addition, the differentiation

between derivatized and non-derivatized components in mixtures of NAs is not possible using BF_3/MeOH and MTBSTFA.

The mass spectrum of the peaks at the point of highest signal/noise ratio was analyzed to determine the ions that could be assigned to molecular weights of NAs. The peaks that did not have the fragment ion at 181 *m*/*z* associated with PFFBr derivatization were discarded because these cannot be attributed to NAs. All of the detected peaks were used for the analysis of the mixtures using BF_3/MeOH and MTBSTFA. A database was built to classify the probable molecular ions. The limitation of this analysis is that only ions that fit the classic formula of NAs can be included. The methylated NA, MTBS-NA and PFB-NA standards retention indices (Table 4.5) were used to obtain more accurate results of ion distribution. According to the retention indices (Ri), it is expected that components of the PFB-NA mixture elute at a specific retention time range according to the carbon number.

According to the ion distribution of the methylated Sigma mixture, ~42% of the ions may be related to aliphatic NAs. Approximately 16% of the ions belong to diaromatic NAs and the aliphatic NAs with 7 carbons being the most abundant (Fig.4. 5a). With the silyl ion distribution for the Sigma mixture, it was found that most of the ions are classified as aliphatic NAs (~36%), and among these NAs with 11 carbons were the most abundant group (Fig. 4.5b). However, using PFBBr as the derivatization reagent it was observed that ~75% of the compounds analyzed belong to the z-0 family (i.e., aliphatic naphthenic acids) with 4–18 carbons; among these, the NAs with 6 carbons were the most abundant (Fig. 4.5c). With the analysis of the Sigma NA mixture with the three reagents, it was observed that aliphatic NAs are the most abundant group

of NAs in this mixture, confirming previous results using MTBSTFA in a Sigma extract [135]. The possible presence of NAs of the z-2 family, as well as aromatic compounds, was also observed.

According to the ion distribution of the methylated Merichem mixture, ~24% of the ions may be related to z-4 NAs and monoaromatic NAs with 7 carbons being the most abundant (Fig. 4.5d). With the ion distribution of the silyl Merichem mixture, it was found that most of the ions are classified as z-6 NAs (~23%), and among these NAs with 12 carbons were the most abundant group (Fig. 4.5e). However, using PFBBr as derivatization reagent it was observed that ~33% of the compounds analyzed belong to the z-6 family with 12–19 carbons; among these, the NAs with 16 carbons were the most abundant (Fig. 4.5f). Correspondingly, the ion distributions depicted in Fig. 4.5 c and f for NA mixtures derivatized with PFBBr are different from those using the methylation and silylation techniques (Fig. 4.5 a, b, d, and e) and also different from others previously reported [135,146].

The 2 oil NA extracts (Sigma and Merichem) were spiked with the 8 NA standards to confirm that PFBBr was reacting with all of the NAs present in the mixtures (Fig. 4.6). The 8 NA standards were identified in the chromatogram, the mass spectra of the 8 NA standards derivatives were detected and the spectral analysis showed the presence of the molecular ion and the fragment ion at 181 m/z in all of the standards extracted mass spectra (Fig. 4.6-4.8).

Based on the mass spectral analysis of the mixtures and the presumable presence of saturated hydrocarbons an open column chromatography approach was used to separate the components of the mixtures according to their polarity and to

determine an approximate proportion of non-NA components. From the column, a hexane and a methanol fraction of each oil NA extract were obtained. The NAs are expected to be detected in the methanol fraction and the saturated hydrocarbons are expected to be detected in the hexane fraction. The components in the methanol fraction corresponded to 93% and 92% by weight of the Sigma and Merichem extracts, respectively. A sample of each of the fractions was derivatized with PFBBr. For the methanol fraction 96% and 94% of the peaks have the fragment ion at 181*m*/*z* of the Sigma and Merichem mixtures, respectively, therefore they are carboxylic acids. However, the fragment ion at 181*m*/*z* was not detected in the hexane fraction (Fig. 4.9). The components in the hexane fraction have “ski-slope” fragmentation patterns with peaks separated by 14*m*/*z*, which correspond to CH_2 units, which is characteristic of branched or unbranched saturated hydrocarbons [151], so they might be present in both mixtures. For example, a peak detected at 26.9 min in the non-polar fraction of the Sigma extract (AMDIS, purity = 50, s/n = 68) was found to have a high match in the NIST Library with hexadecane (Match = 840, R Match = 840) (Fig. 4.10). The derivatization methodologies reported here focus on the analysis of carboxylic acids and their derivatives. An alternative and specific methodology would need to be developed for the analysis of trace non-carboxylic acid compounds in NA mixtures, such as non-volatile components, alcohols, ketones, and ethers that could also be present [121]. Moreover, low molecular weight alcohols, ketones, and ethers are very volatile and may be lost during sample processing. Thus, other components such as alcohols, ketones, and ethers, are not observed perhaps because they are not present in these specific mixtures, or they may be in low concentration compared to the carboxylic acids.

The 2 mixtures (Sigma and Merichem) were spiked with the 8 standards to confirm that NAs were effectively separated by column chromatography in the methanol fraction. A sample of each of the two fractions per extract was taken and derivatized using PFBBr. The 8 standards were detected in the methanol fractions, confirming the effectiveness of this approach to separate NAs from other components. The 8 exhibited the characteristic fragment ion at 181*m*/*z* as expected. The standards were not detected in the hexane fraction, nor was the fragment ion at 181*m*/*z (Fig. 4.11).* On average the open column recovery was 97±2%. Sigma and Merichem mixtures have been used as reference mixtures for validation, calibration, and quantification of NAs in water samples assuming a high concentration of NAs [90,124,152]. Our results showed that these standards contain ~7% by weight of other components. Herein, we have found that separation by column chromatography in combination with derivatization using PFBBr, offers a more accurate characterization of carboxylic acids in NA mixtures. Significantly, these non-NA components in the mixtures were evident because they did not contain the fragment ion at 181*m*/*z* that resulted from the use of the PFBBr derivatization reagent.

4.5. Conclusion

This chapter reports on the effectiveness of PFBBr as a derivatization reagent for the analysis of NAs. Using PFBBr prior to GC-EIMS analysis of NA standards and mixtures yielded better results compared to using BF_3/MeOH or MTBSTFA. The derivatization of NAs with PFBBr offers several advantages, including increased derivatization signal yields, resolution, and sensitivity. In the mass spectral analysis of

PFB derivatives, a characteristic fragment from the ion pentafluorobenzyl was identified at 181*m*/*z*. In the analysis of the two oil extract NA mixtures derivatized with PFBBr, several peaks that lacked the fragment ion at 181*m*/*z* were detected; therefore, these peaks are not NAs. Together, our results indicate that PFBBr derivatization allows the differentiation of compounds with no labile hydrogens, greatly increasing the accuracy of analyzing complex NA mixtures. The use of PFBBr increases sensitivity, chromatographic resolution, and identification accuracy for the analysis of standards and mixtures of NAs.

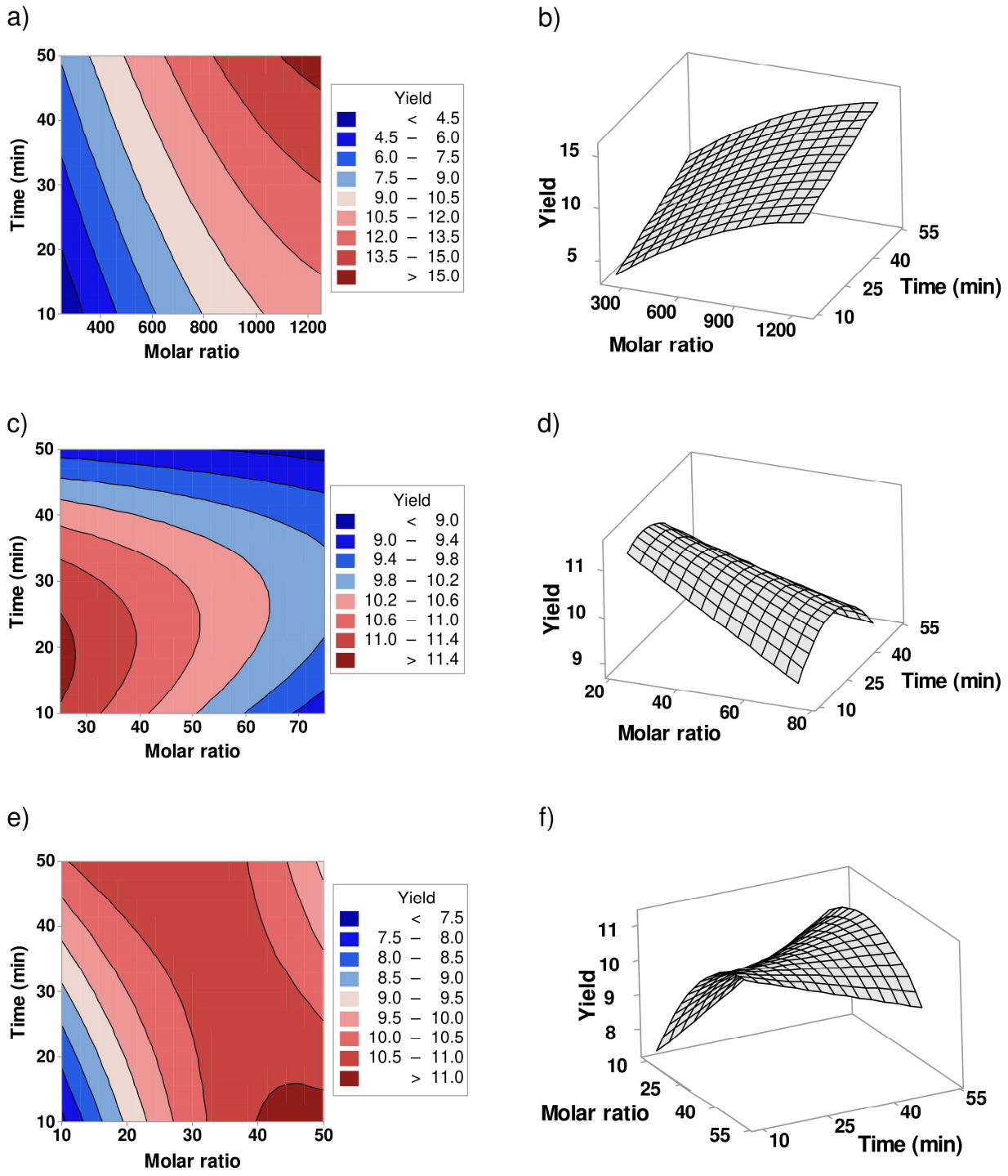


Figure 4.1. Contour and surface plots of derivatization signal yield vs molar ratio and reaction time of the three derivatization reagents and the mixture of NA standards. BF_3/MeOH (a-b; $R^2 = 94.16$), for MTBSTFA (c-d; $R^2 = 86.44$) and for PFBBr (e-f; $R^2 = 85.49$).

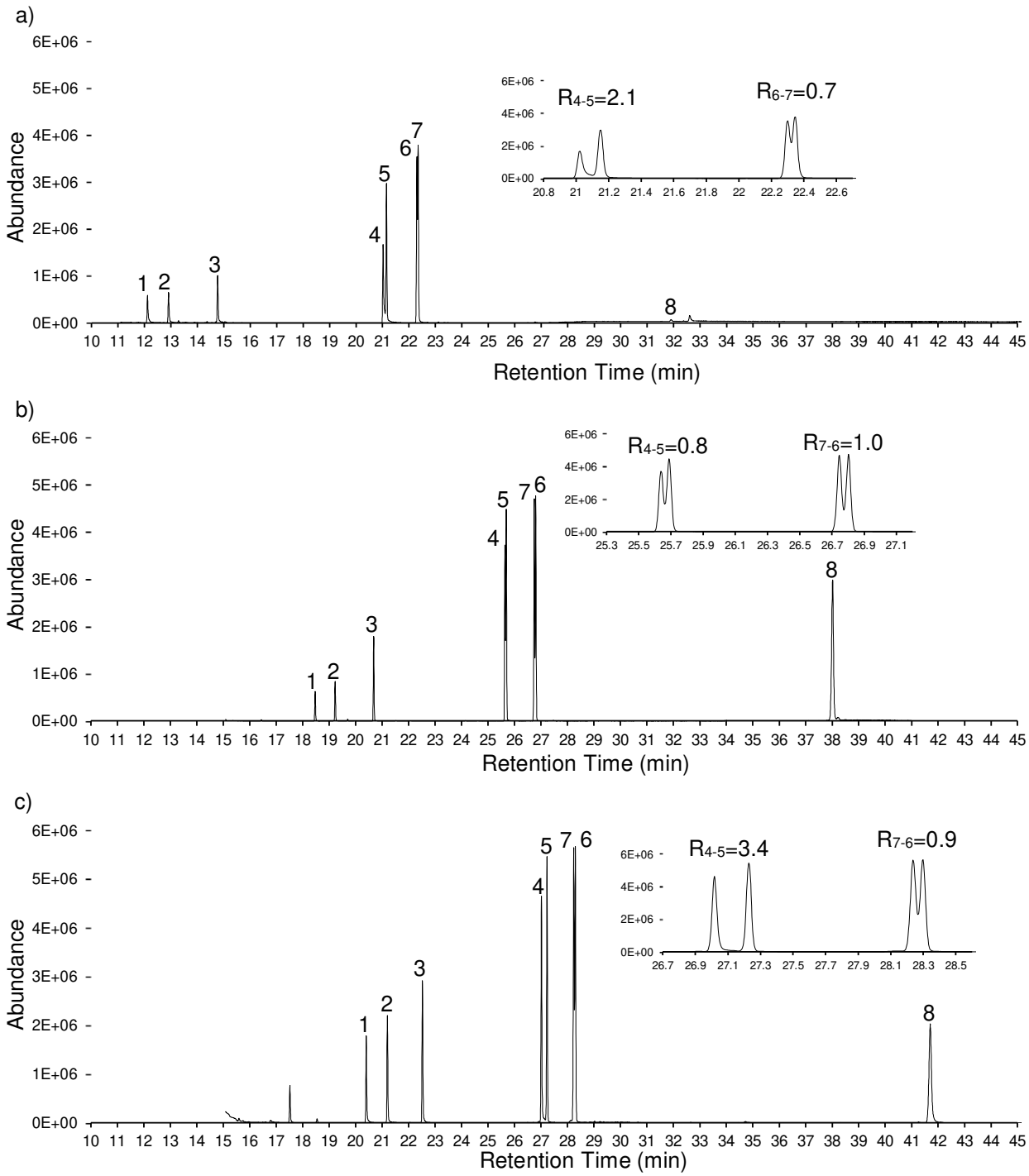


Figure 4.2. GC-EIMS TICs for a derivatized mixture of NA standards. (a) Methylated NAs, (b) MTBS-NAs and (c) PFB-NAs. Numbers correspond to NAs derivatives following the codes of Table 1. R represents the resolution of two adjacent peaks calculated according to the IUPAC resolution equation for chromatography.

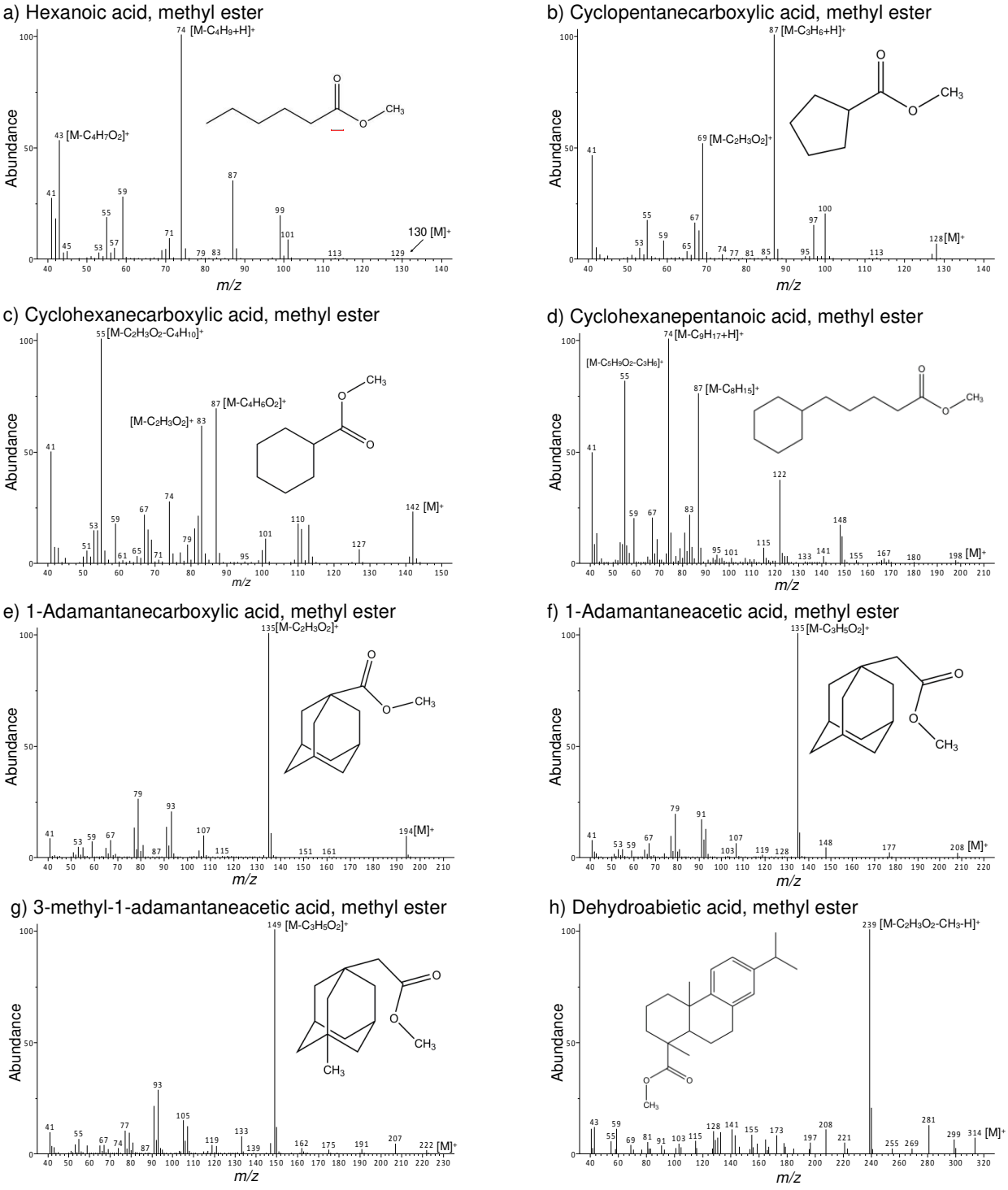


Figure 4.3. Mass spectra of methylated NAs (a-h), MTBS-NAs (i-p), and PFB-NAs (q-y), and their corresponding structures.

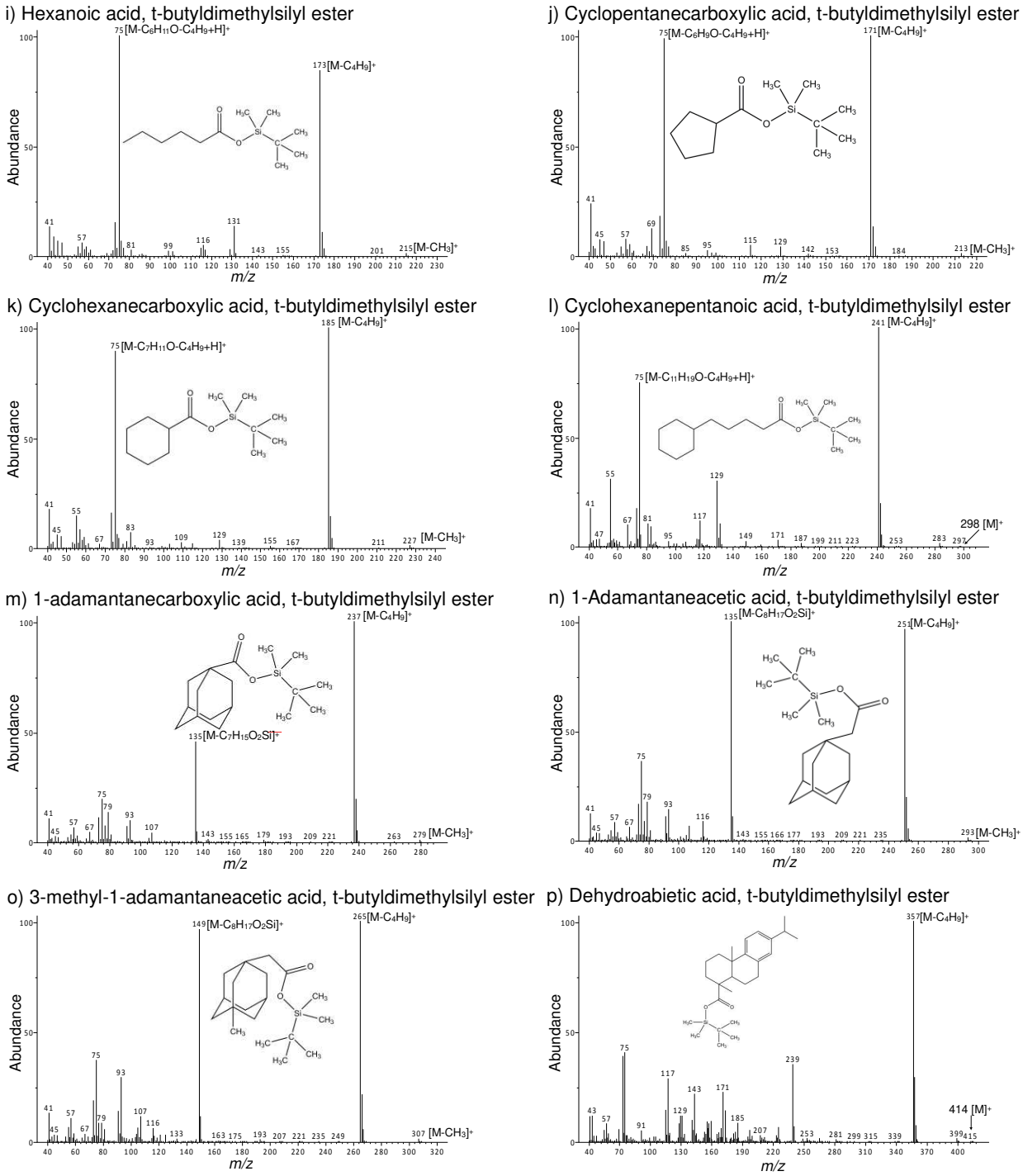
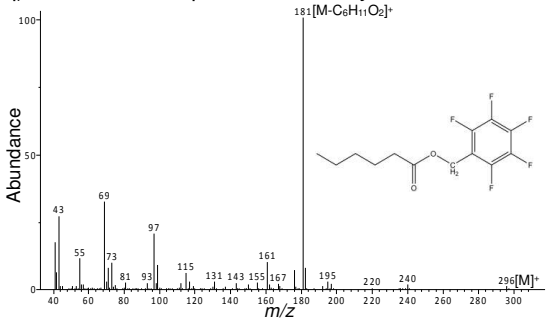
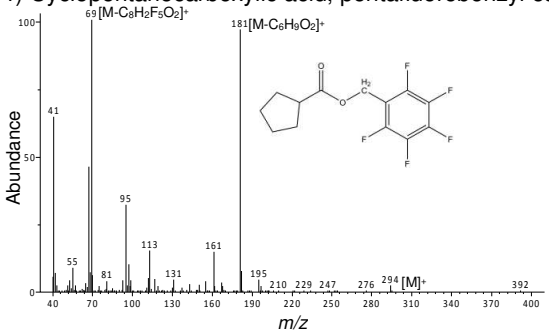


Figure 4.3. Continued.

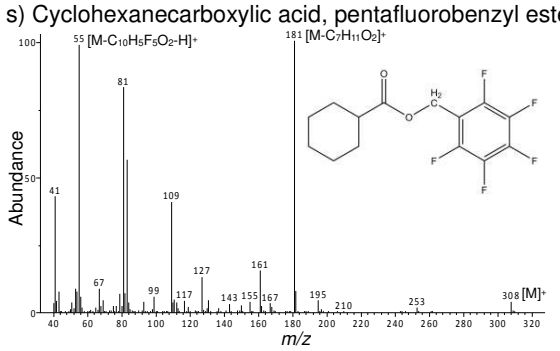
q) Hexanoic acid, pentafluorobenzyl ester



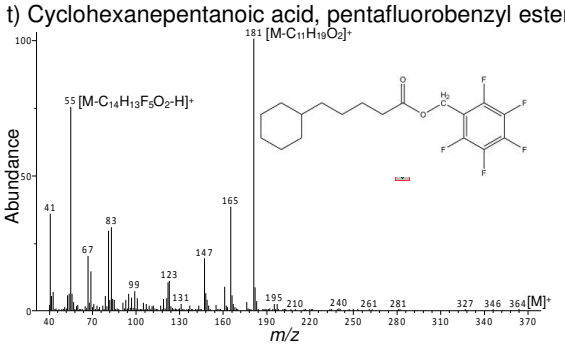
r) Cyclopentanecarboxylic acid, pentafluorobenzyl ester



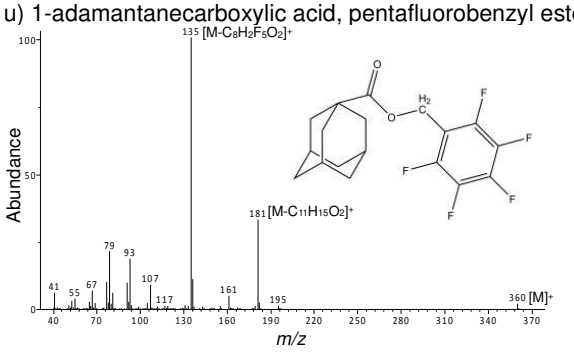
s) Cyclohexanecarboxylic acid, pentafluorobenzyl ester



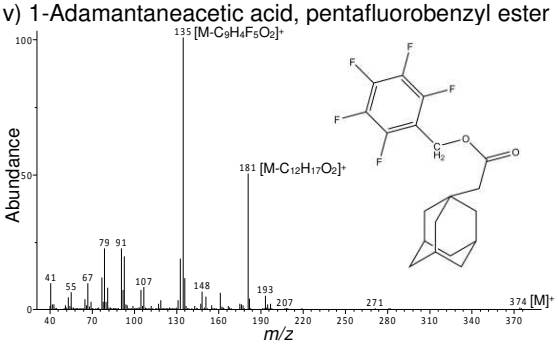
t) Cyclohexanepentanoic acid, pentafluorobenzyl ester



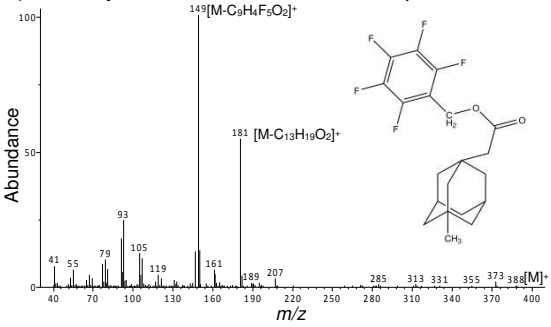
u) 1-adamantanecarboxylic acid, pentafluorobenzyl ester



v) 1-Adamantaneacetic acid, pentafluorobenzyl ester



w) 3-methyl-1-adamantaneacetic acid, pentafluorobenzyl ester



x) Dehydroabietic acid, pentafluorobenzyl ester

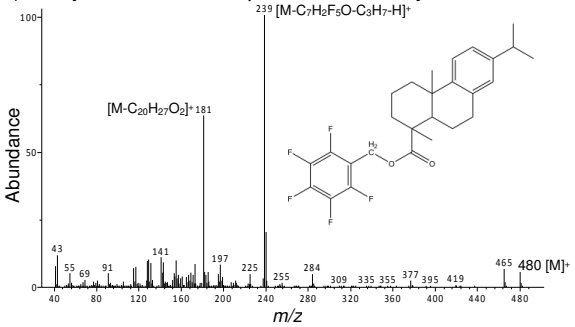


Figure 4.3. Continued.

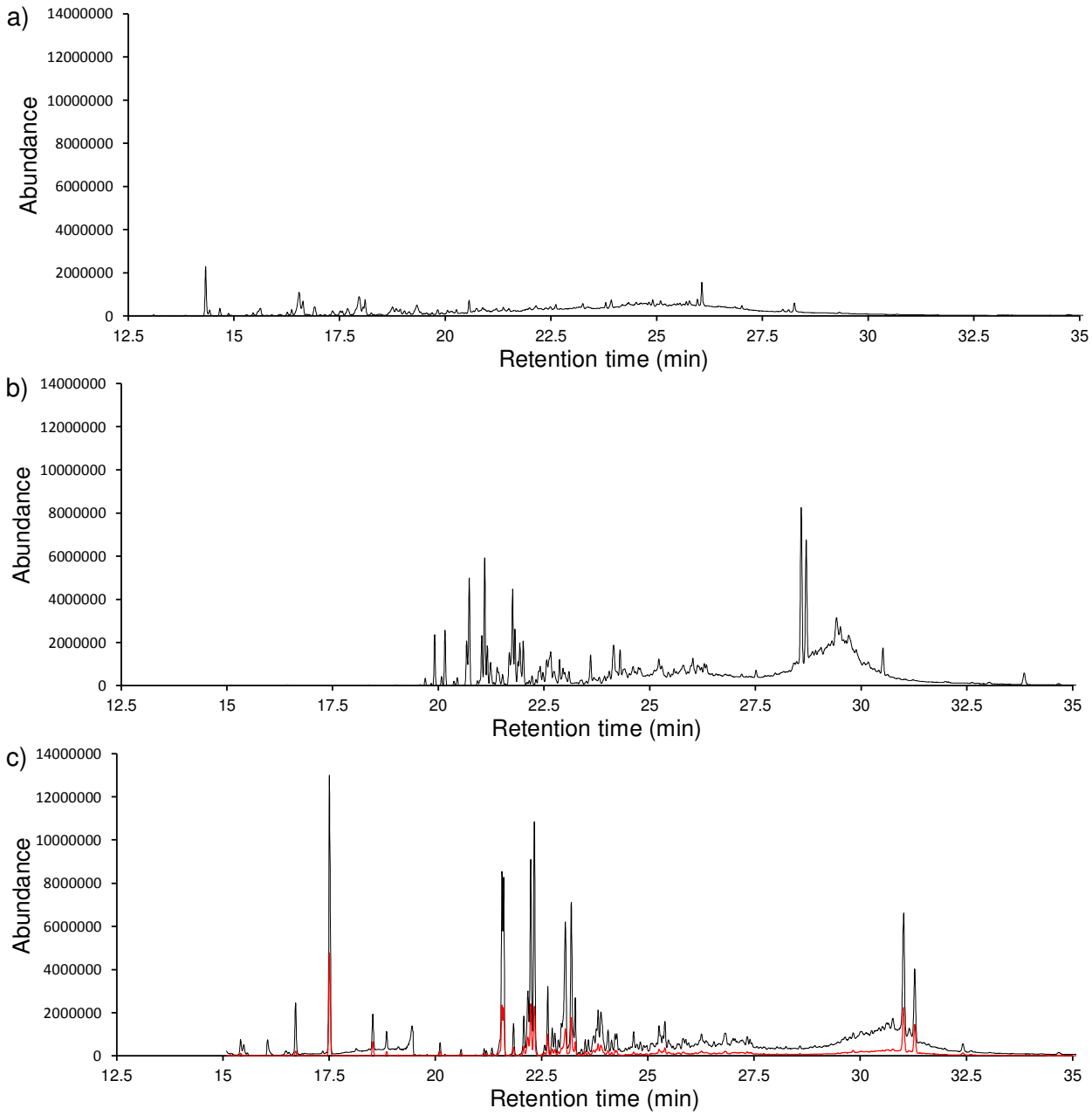


Figure 4.4. GC-EIMS TIC of a Sigma NA mixture derivatized with BF_3/MeOH (a), MTBSTFA (b) and PFBBr (c), shown in black. GC-EIMS TIC of a Merichem NA mixture derivatized with BF_3/MeOH (d), MTBSTFA (e) and PFBBr (f), shown in black. GC-EIMS EIC (181*m*/*z*) from a Sigma and a Merichem NA mixture derivatized with PFBBr in c and f, shown in red.

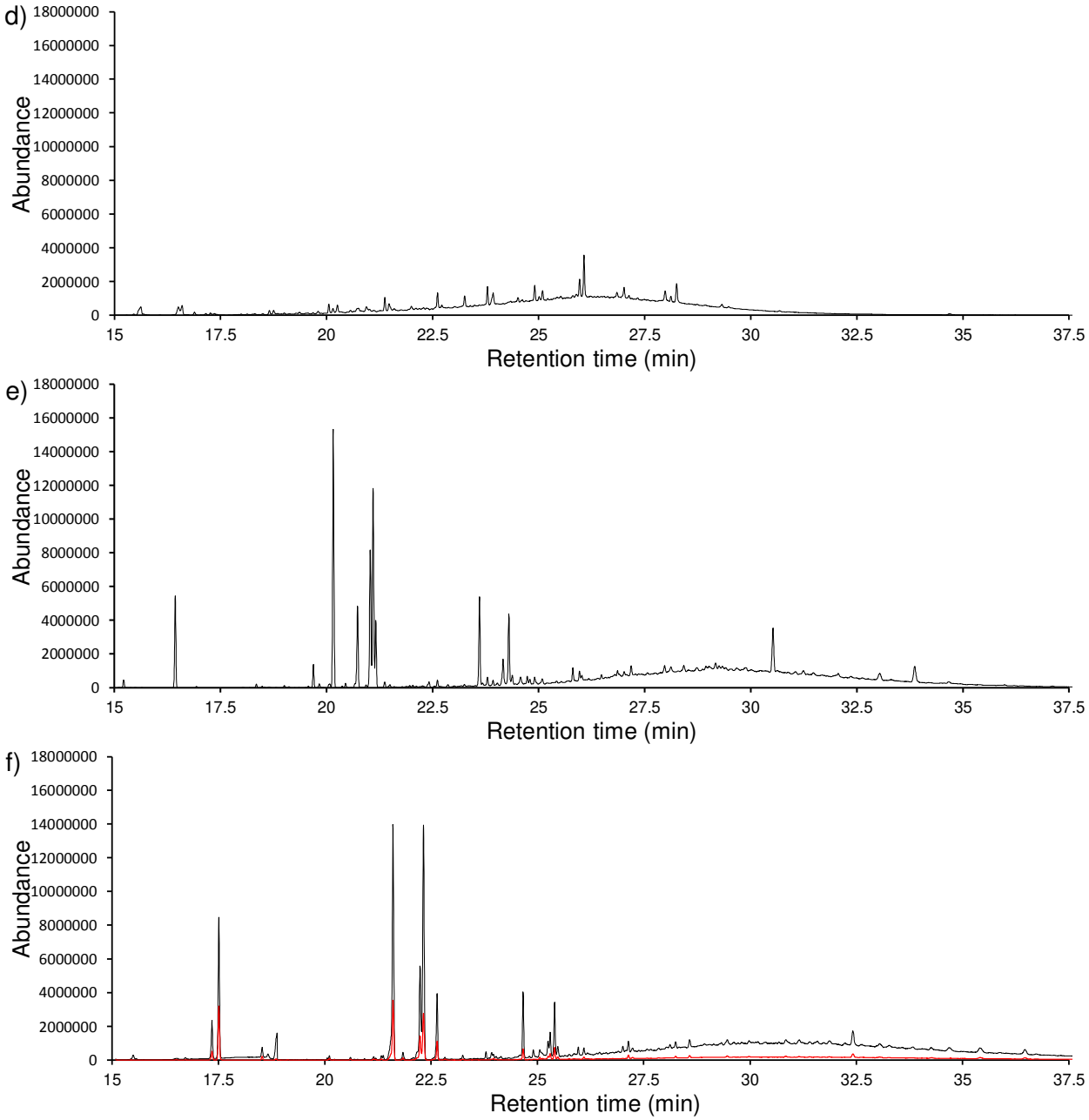


Figure 4.4. Continued.

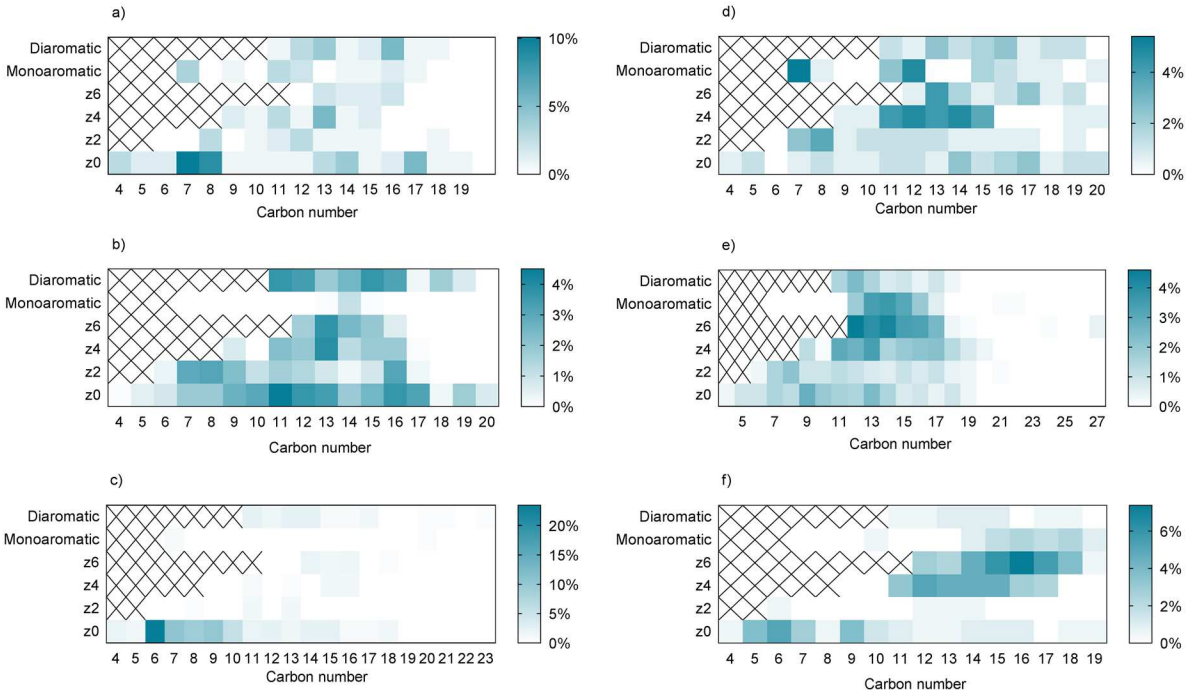


Figure 4.5. Ion distribution of a Sigma (a-c) and a Merichem (d-f) NA extract using BF_3/MeOH (a and d), MTBSTFA (b and e) and PFBBr (c and f) as derivatization reagents. Values correspond to various carbon numbers and z families in the NA commercial mixtures. The color intensity represents the percentage, by number of ions of NAs in the mixture that account for a given carbon number in a given z family, corresponding to specific *m*/*z* values from GC–EIMS analysis. The sum of all the values equals 100%.

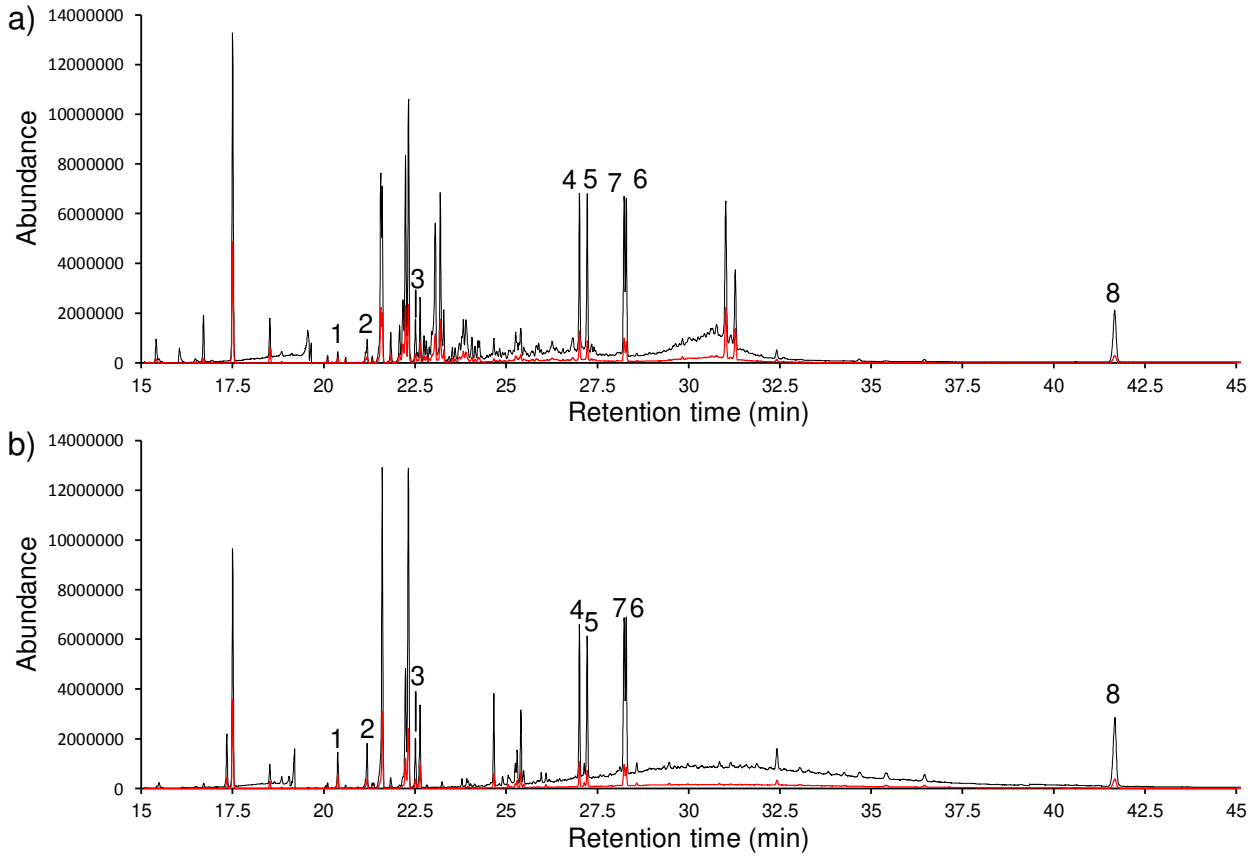


Figure 4.6. GC-EIMS TIC of a Sigma (a) and a Merichem (b) NA mixture spiked with eight NA standards and derivatized with PFBBr, shown in black. GC-EIMS EIC (181 *m/z*) from a Sigma and a Merichem NA mixture derivatized with PFBBr, shown in red. Numbers correspond to the eight NA standards following the codes of Table 1.

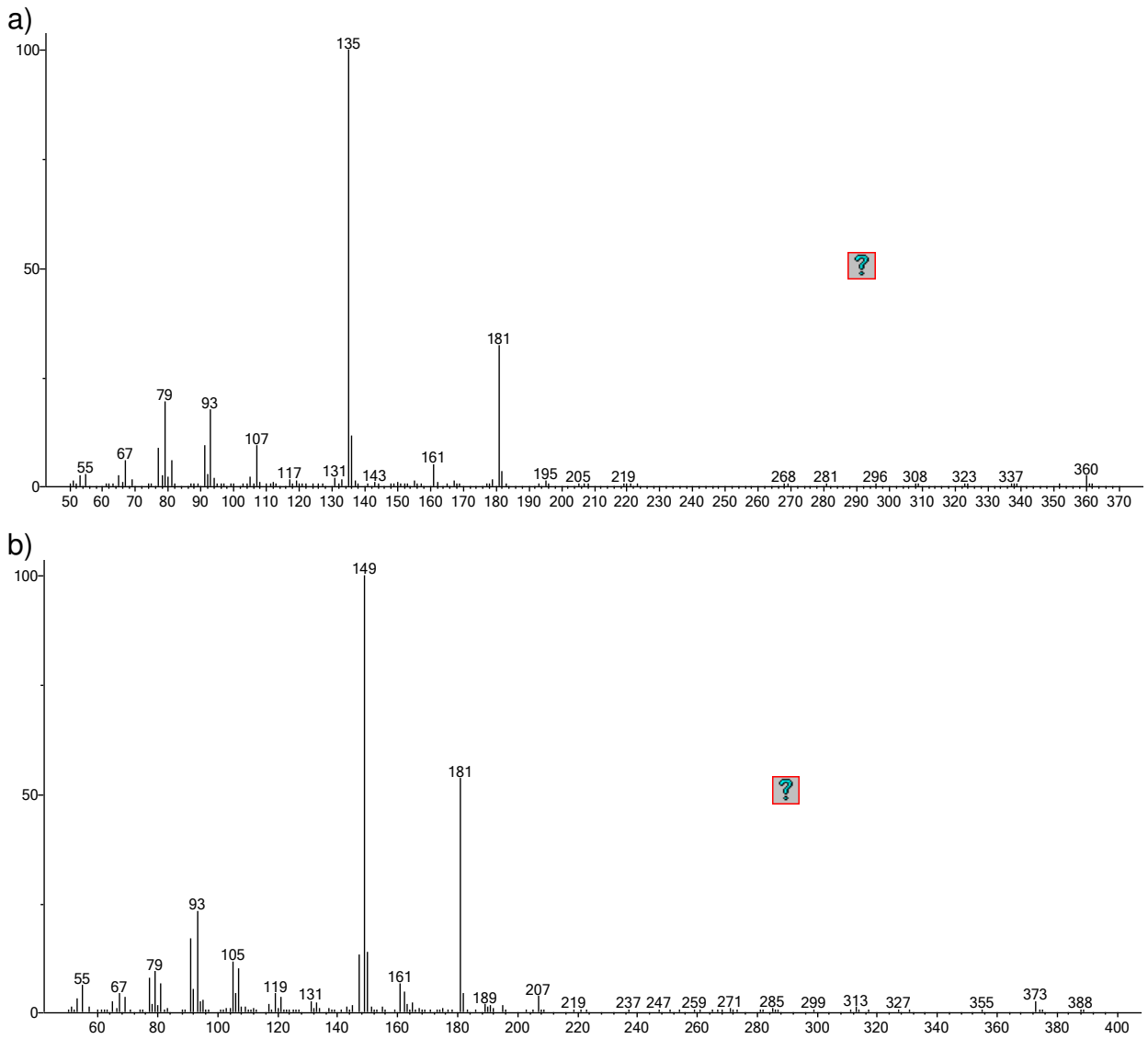


Figure 4.7. Mass spectra of (a) 1-Adamantanecarboxylic acid, pentafluorobenzyl and (b) 3-methyl-1-adamantaneacetic acid, pentafluorobenzyl ester found in the spiked Sigma mixture.

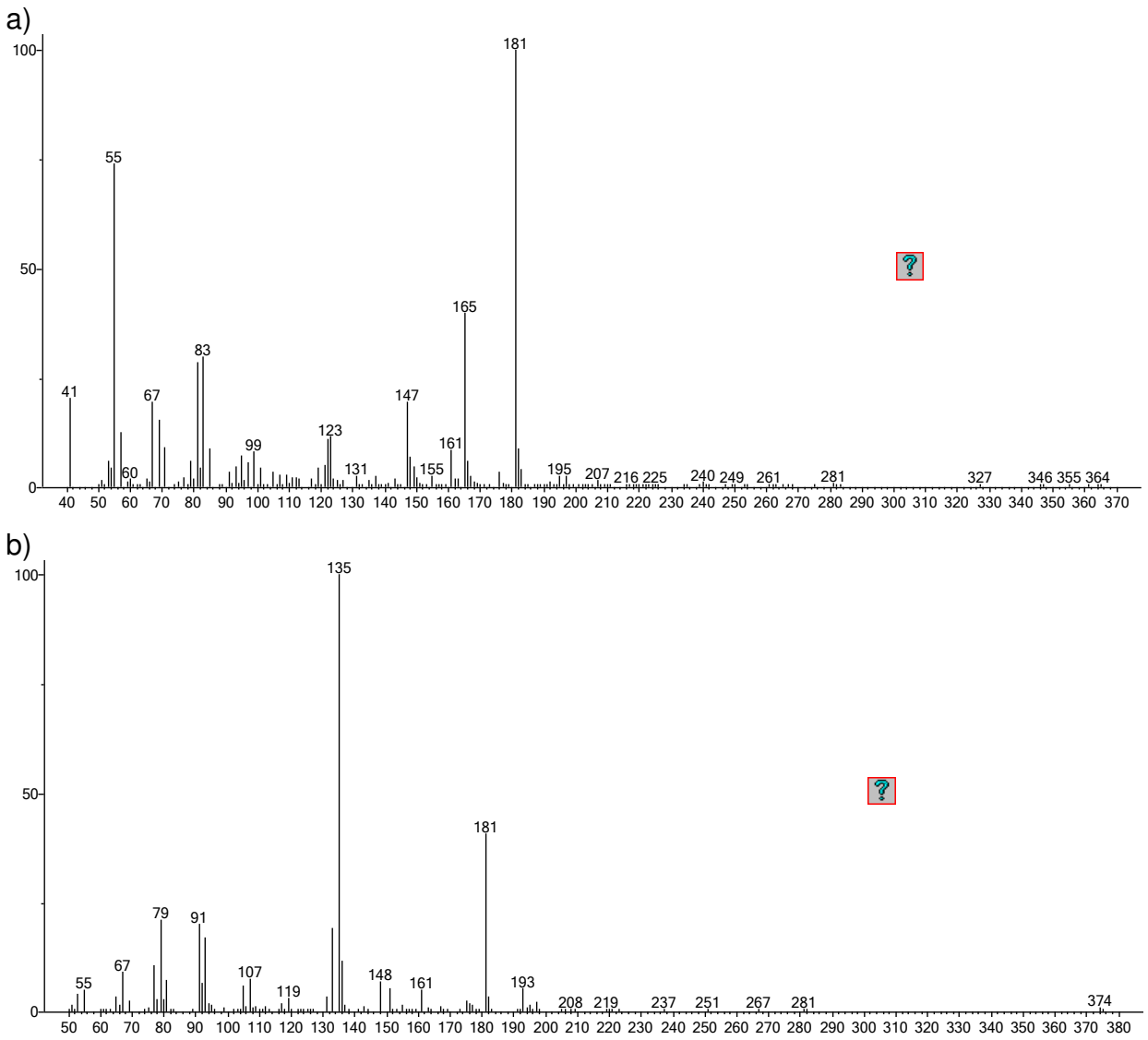


Figure 4.8. Mass spectra of (a) Cyclohexanepentanoic acid, pentafluorobenzyl ester and (b) 1-Adamantaneacetic acid, pentafluorobenzyl ester found in the spiked Merichem mixture.

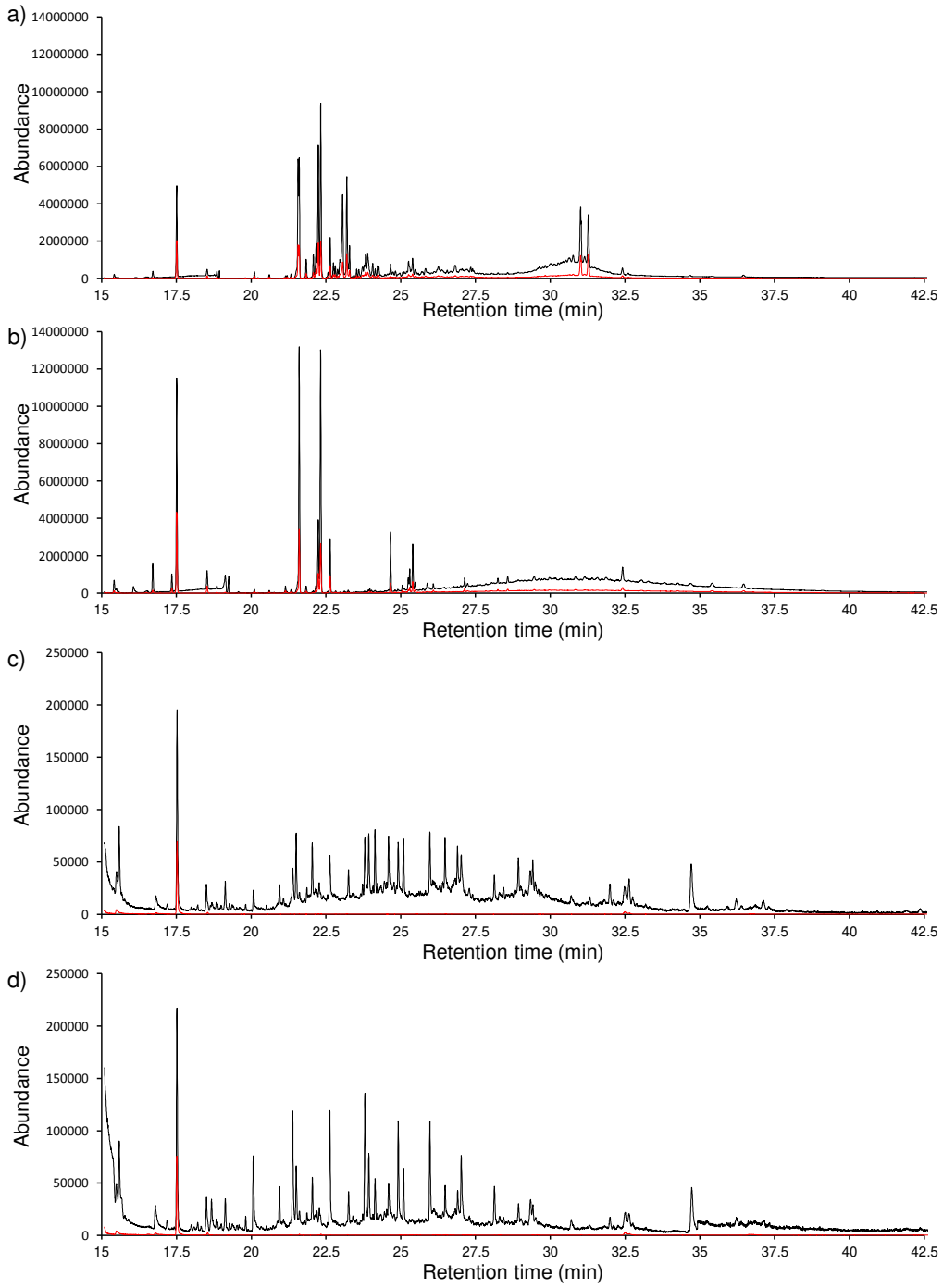


Figure 4.9. GC-EIMS TIC of the polar fraction of a Sigma (a) and a Merichem (b) NA mixture obtained by column chromatography and derivatized with PFBBr, shown in black. GC-EIMS TIC of the non-polar fraction of Sigma (c) and a Merichem (d) NA mixture obtained by column chromatography and derivatized with PFBBr, shown in black. GC-EIMS EIC (181 *m/z*) of a Sigma (a and c) and a Merichem (b and d) NA mixture derivatized with PFBBr, shown in red.

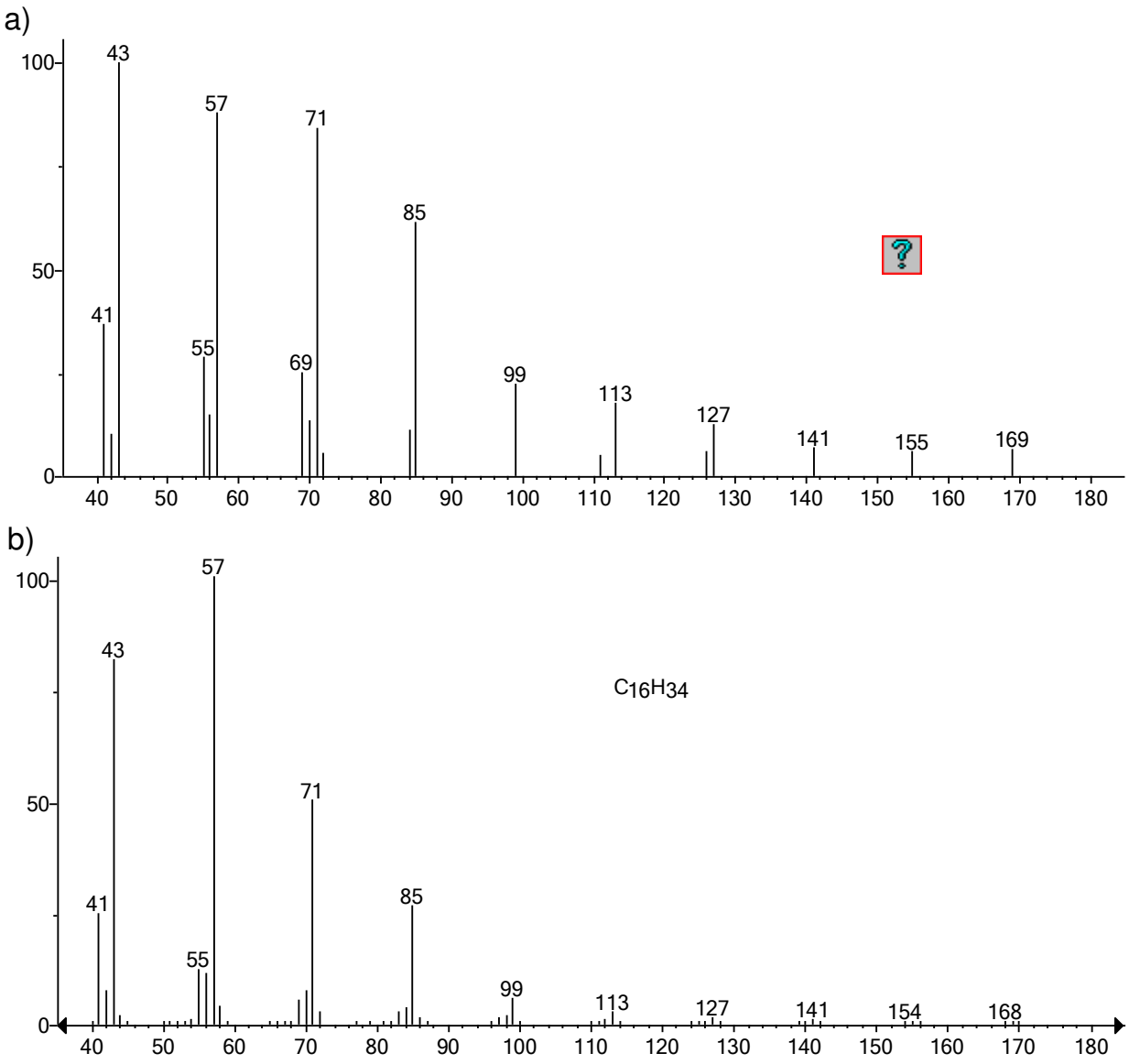


Figure 4.10. Mass spectra of a peak at 26.9 min in the hexane fraction of the Sigma extract (a). Mass spectra of hexadecane by NIST (b).

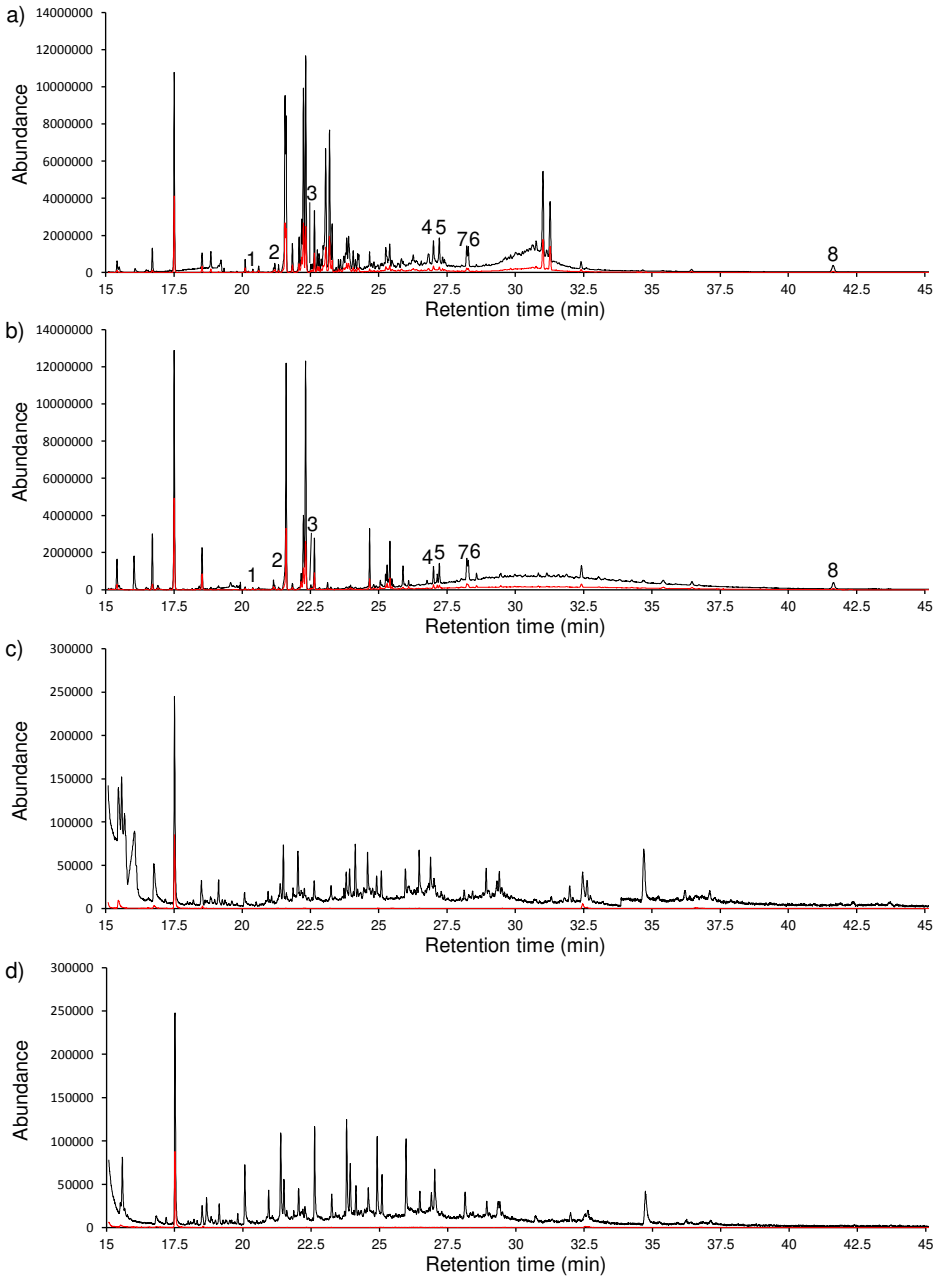


Figure 4.11. GC-EIMS TIC of the polar fraction of Sigma (a) and a Merichem (b) NA mixture spiked with 8 standards prior to column separation and derivatized with PFBBr, shown in black. GC-EIMS TIC of the non-polar fraction of Sigma (c) and a Merichem (d) NA mixture spiked with 8 standards prior to column separation and derivatized with PFBBr, shown in black. GC-EIMS EIC (181 *m/z*) of the polar and non-polar fraction of Sigma and a Merichem NA mixtures spiked with 8 standards prior to column separation and derivatized with PFBBr, shown in red. Numbers correspond to the eight NA standards following the codes of Table 1.

Table 4.1. Individual NA standards for derivatization evaluation.

Compound	Name	Molecular weight	Basic structure description	z family	Purity (%)	Supplier
1	Hexanoic acid	116.16	Aliphatic	0	99	Sigma-Aldrich
2	Cyclopentanecarboxylic acid	114.14	Monocyclic	2	99	Sigma-Aldrich
3	Cyclohexanecarboxylic acid	128.17	Monocyclic	2	≥ 98	Sigma-Aldrich
4	Cyclohexanepentanoic acid	184.28	Monocyclic	2	98	Sigma-Aldrich
5	1-adamantanecarboxylic acid	180.24	Adamantane	6	99	Sigma-Aldrich
6	1-adamantaneacetic acid	194.27	Adamantane	6	98	Sigma-Aldrich
7	3-methyl-1-adamantaneacetic acid	208.3	Adamantane	6	99	Sigma-Aldrich
8	Dehydroabietic acid	300.44	Aromatic, Tricyclic	12	≥ 99	CanSyn Chem. Corp.

Table 4.2. Factors and results for the derivatization of a mixture of NA standards.

Experiment	BF_3/MeOH			MTBSTFA			PFBBr		
	Molar ratio	Time (min)	Signal Yield	Molar ratio	Time (min)	Signal Yield	Molar ratio	Time (min)	Signal Yield
1	250	10	3.2	25	10	11.8	10	10	6.7
2	250	30	5.9	25	30	11.0	10	30	9.8
3	250	50	7.5	25	50	9.2	10	50	10.2
4	750	10	7.1	50	10	9.4	30	10	10.8
5	750	30	10.8	50	30	10.7	30	30	10.3
6	750	30	11.8	50	30	10.6	30	30	10.8
7	750	50	13.4	50	50	9.2	30	50	10.5
8	1250	10	12.0	75	10	9.5	50	10	11.1
9	1250	30	14.4	75	30	9.9	50	30	10.1
10	1250	50	14.0	75	50	8.7	50	50	9.7

Table 4.3. Analysis of variance (ANOVA) results showing the terms of each model. DF: degree of freedom; Adj SS: adjusted sum of squares; Adj MS: adjusted mean squares.

Model	Source	DF	Adj SS	Adj MS	F-Value	*P*-Value
BF3/MeOH	Model	3	124.768	41.5894	32.24	< 0.001
	Linear terms	2	120.867	60.4333	46.84	< 0.001
	Molar ratio	1	94.407	94.4067	73.18	< 0.001
	Time (min)	1	26.46	26.46	20.51	0.004
	Square term					
	Molar ratio*Molar ratio	1	3.901	3.9015	3.02	0.133
	Error- Lack of fit	5	7.241	1.4482	2.90	0.418
MTBSTFA	Model	4	7.55573	1.88893	7.97	0.021
	Linear terms	2	4.70374	2.35187	9.92	0.018
	Molar ratio	1	2.53834	2.53834	10.71	0.022
	Time (min)	1	2.1654	2.1654	9.13	0.029
	Square term					
	Time (min)*Time (min)	1	2.01884	2.01884	8.52	0.033
	2-Way Interaction					
	Molar ratio*Time (min)	1	0.83316	0.83316	3.51	0.12
	Error- Lack of fit	4	1.17968	0.29492	51.31	0.104
PFBBr	Model	4	11.8825	2.9706	7.36	0.025
	Linear terms	2	3.48	1.74	4.31	0.082
	Molar ratio	1	2.94	2.94	7.29	0.043
	Time (min)	1	0.54	0.54	1.34	0.3
	Square term					
	Molar ratio*Molar ratio	1	2.4	2.4	5.95	0.059
	2-Way Interaction					
	Molar ratio*Time (min)	1	6.0025	6.0025	14.88	0.012
	Error- Lack of fit	4	1.8925	0.4731	3.79	0.366

Table 4.4. Comparison of NA signal yields for different chemical derivatizations applied to a mixture of NA standards.

Derivatization reagent	∑ NA signal yield percent	Average NA signal yield percent
BF_3/MeOH	20.1	2.5
MTBSTFA	33.5	4.2
PFBBr	46.3	5.8

Table 4.5. Chromatographic and mass spectral properties of NA derivatives by GC-EIMS using BF_3/MeOH, MTBSTFA, and PFBBr. Relative abundances to maximum are given in parentheses (base 100). ND, non-detected.

Methylated NAs													
Analyte	Rt	Ri	Fragmentation pattern (*m/z*) and relative abundance (%)										M^+
Hexanoic acid, methyl ester	12.13	906	101 (8.8)	99 (19.3)	87 (34.8)	74 (99.9)	71 (9.3)	59 (27.8)	55 (18.6)	43 (52.9)	42 (18.1)	41 (27.1)	130
Cyclopentanecarboxylic acid, methyl ester	12.92	950	128 (6.8)	100 (20.3)	97 (15.1)	87 (99.9)	69 (51.6)	68 (12.8)	67 (16.3)	59 (8.1)	55 (17.3)	41 (46.1)	128
Cyclohexanecarboxylic acid, methyl ester	14.78	1054	142 (22.8)	110 (174)	87 (68.8)	83 (61.2)	82 (20.9)	74 (27.5)	67 (21.5)	59 (17.6)	55 (99.9)	41 (49.6)	142
Cyclohexanepentanoic acid, methyl ester	21.03	1474	148 (17.1)	122 (37.0)	87 (75.5)	83 (21.4)	81 (13.4)	74 (99.9)	67 (20.1)	59 (19.8)	55 (81.1)	41 (49.3)	198
1-adamantanecarboxylic acid, methyl ester	21.16	1483	194 (9.5)	136 (10.9)	135 (99.9)	107 (9.8)	93 (20.4)	91 (13.4)	79 (26.2)	77 (13.3)	67 (7.5)	41 (8.4)	194
1-adamantaneacetic acid, methyl ester	22.31	1575	136 (11.0)	135 (99.9)	107 (6.3)	93 (12.7)	92 (8.0)	91 (17.0)	79 (19.5)	77 (9.6)	67 (6.4)	41 (7.6)	208
3-methyl-1-adamantaneacetic acid, methyl ester	22.35	1579	150 (12.0)	149 (99.9)	133 (7.6)	107 (12.2)	105 (14.9)	93 (28.6)	91 (21.3)	79 (92)	77 (10.2)	41 (9.4)	222
Dehydroabietic acid, methyl ester	31.91	2379	281 (12.6)	240 (20.5)	239 (99.9)	208 (10.7)	141 (10.8)	133 (9.6)	128 (10.0)	59 (11.0)	43 (12.0)	41 (11.0)	314
MTBS-NAs													
Analyte	Rt	Ri	Fragmentation pattern (*m/z*) and relative abundance (%)										M^+
Hexanoic acid, tert-butyldimethylsilyl ester	18.48	1288	174 (11.0)	173 (84.0)	131 (13.9)	76 (7.1)	75 (99.9)	73 (15.4)	47 (6.3)	45 (7.0)	43 (8.9)	41 (13.6)	ND
Cyclopentanecarboxylic acid, tert-butyldimethylsilyl ester	19.23	1341	172 (13.4)	171 (99.9)	76 (72)	75 (98.4)	73 (18.4)	69 (12.8)	57 (8.0)	47 (6.7)	45 (7.7)	41 (24.0)	ND
Cyclohexanecarboxylic acid, tert-butyldimethylsilyl ester	20.69	1448	186 (14.7)	185 (99.9)	83 (7.4)	76 (6.5)	75 (89.3)	73 (16.3)	57 (8.6)	55 (14.8)	45 (6.2)	41 (17.9)	ND
Cyclohexanepentanoic acid, tert-butyldimethylsilyl ester	25.65	1870	242 (19.9)	241 (99.9)	131 (10.5)	129 (30.0)	117 (11.8)	81 (10.6)	75 (74.6)	73 (17.4)	55 (31.0)	41 (17.4)	298
1-adamantanecarboxylic acid, tert-butyldimethylsilyl ester	25.69	1874	238 (19.6)	237 (99.9)	135 (45.6)	93 (10.1)	91 (7.4)	79 (13.7)	77 (7.7)	75 (19.7)	73 (11.4)	41 (10.9)	ND
3-methyl-1-adamantaneacetic acid, tert-butyldimethylsilyl ester	26.75	1975	266 (21.6)	265 (99.9)	150 (11.6)	149 (96.2)	107 (11.5)	93 (29.3)	91 (14.0)	75 (37.0)	73 (18.9)	41 (13.3)	ND

1-adamantaneacetic acid, tert-butyldimethylsilyl ester	26.81	1981	252 (19.8)	251 (96.5)	136 (11.1)	135 (99.9)	93 (14.3)	91 (11.0)	79 (17.7)	75 (36.2)	73 (16.6)	41 (12.5)	ND
Dehydroabietic acid, tert-butyldimethylsilyl ester	38.02	2670	358 (29.2)	357 (99.9)	239 (35.2)	173 (14.3)	171 (22.6)	143 (21.8)	117 (28.8)	115 (14.5)	75 (40.5)	73 (38.8)	414

PFB-NAs

Analyte	Rt	Ri	Fragmentation pattern (*m/z*) and relative abundance (%)										M^+
Hexanoic acid, pentafluorobenzyl ester	20.41	1427	181 (99.9)	161 (10.1)	99 (8.9)	97 (20.4)	73 (9.7)	71 (8.0)	69 (32.2)	55 (11.4)	43 (26.9)	41 (17.3)	296
Cyclopentanecarboxylic acid, pentafluorobenzyl ester	21.21	1487	182 (7.7)	181 (96.4)	161 (14.7)	113 (15.0)	97 (10.0)	95 (31.9)	69 (99.9)	67 (45.9)	55 (8.8)	41 (64.4)	294
Cyclohexanecarboxylic acid, pentafluorobenzyl ester	22.53	1593	181 (99.9)	161 (15.5)	127 (12.9)	109 (40.4)	83 (56.1)	81 (82.7)	67 (8.8)	55 (98.3)	53 (8.6)	41 (42.6)	308
Cyclohexanepentanoic acid, pentafluorobenzyl ester	27.03	2001	181 (99.9)	165 (38.2)	147 (19.0)	123 (10.9)	83 (30.6)	81 (29.3)	69 (14.2)	67 (19.8)	55 (74.7)	41 (35.3)	364
1-adamantanecarboxylic acid, pentafluorobenzyl ester	27.23	2020	181 (33.0)	136 (11.2)	135 (99.9)	107 (9.0)	93 (18.6)	91 (9.8)	79 (21.3)	77 (9.9)	67 (6.8)	41 (6.0)	360
3-methyl-1-adamantaneacetic acid, pentafluorobenzyl ester	28.24	2110	181 (54.5)	150 (13.4)	149 (99.9)	147 (12.9)	107 (10.5)	105 (12.4)	93 (24.4)	91 (17.9)	79 (9.9)	77 (8.4)	388
1-adamantaneacetic acid, pentafluorobenzyl ester	28.30	2115	181 (50.0)	136 (11.3)	135 (99.9)	133 (18.6)	93 (19.5)	91 (22.3)	79 (22.2)	77 (11.5)	67 (9.5)	41 (9.5)	374
Dehydroabietic acid, pentafluorobenzyl ester	41.71	2794	240 (20.1)	239 (99.9)	181 (62.9)	155 (9.7)	143 (9.0)	141 (10.9)	131 (8.6)	129 (10.1)	128 (9.4)	43 (11.7)	480

Table 4.6. Comparison of limits of quantification (LOQ) and limits of detection (LOD) of the eight standards measured by GC-EIMS using BF_3/MeOH, MTBSTFA, and PFBBr as derivatization reagents. Linear equations and coefficients of determination (R^2) are also described.

Standard	Linear equation	R^2	LOQ (mM)	LOD (mM)
Hexanoic acid, methyl ester	y = 4E+07x - 2E+06	0.996	1.3E-02	7.8E-03
Cyclopentanecarboxylic acid, methyl ester	y = 4E+07x - 331551	0.999	5.5E-03	3.4E-03
Cyclohexanecarboxylic acid, methyl ester	y = 6E+07x - 478784	0.997	2.2E-03	1.0E-04
Cyclohexanepentanoic acid, methyl ester	y = 2E+08x - 2E+07	0.978	1.8E-02	1.4E-02
1-adamantanecarboxylic acid, methyl ester	y = 3E+08x - 2E+06	0.995	1.2E-03	5.0E-04
1-adamantaneacetic acid, methyl ester	y = 3E+08x - 3E+06	0.988	1.5E-03	1.0E-04
3-methyl-1-adamantaneacetic acid, methyl ester	y = 3E+08x - 2E+06	0.992	1.9E-03	5.4E-04
Dehydroabietic acid, methyl ester	y = 6E+06x - 435680	0.983	3.9E-02	1.3E-02
Hexanoic acid, tert-butyldimethylsilyl ester	y = 5E+07x - 418382	0.994	4.4E-03	1.6E-03
Cyclopentanecarboxylic acid, tert-butyldimethylsilyl ester	y = 6E+07x - 472855	0.995	2.1E-03	6.5E-04
Cyclohexanecarboxylic acid, tert-butyldimethylsilyl ester	y = 1E+08x - 1E+06	0.996	1.6E-03	2.0E-04
Cyclohexanepentanoic acid, tert-butyldimethylsilyl ester	y = 3E+08x - 1E+07	0.982	6.5E-03	3.7E-03
1-adamantanecarboxylic acid, tert-butyldimethylsilyl ester	y = 3E+08x - 1E+06	0.998	8.0E-04	1.0E-04
3-methyl-1-adamantaneacetic acid, tert-butyldimethylsilyl ester	y = 3E+08x - 2E+06	0.997	2.0E-03	5.7E-04
1-adamantaneacetic acid, tert-butyldimethylsilyl ester	y = 4E+08x - 5E+06	0.994	1.8E-03	4.0E-04
Dehydroabietic acid, tert-butyldimethylsilyl ester	y = 3E+08x - 3E+06	0.993	5.7E-03	4.3E-03
Hexanoic acid, pentafluorobenzyl ester	y = 1E+08x - 2E+06	0.995	7.0E-03	4.2E-03
Cyclopentanecarboxylic acid, pentafluorobenzyl ester	y = 2E+08x - 1E+06	0.997	2.6E-03	5.0E-04
Cyclohexanecarboxylic acid, pentafluorobenzyl ester	y = 2E+08x - 2E+06	0.997	2.4E-03	1.0E-03
Cyclohexanepentanoic acid, pentafluorobenzyl ester	y = 4E+08x - 9E+06	0.996	4.2E-03	2.1E-03
1-adamantanecarboxylic acid, pentafluorobenzyl ester	y = 4E+08x - 651360	0.998	8.0E-04	3.8E-04
3-methyl-1-adamantaneacetic acid, pentafluorobenzyl ester	y = 4E+08x - 623204	0.999	7.0E-04	7.0E-05
1-adamantaneacetic acid, pentafluorobenzyl ester	y = 5E+08x - 5E+06	0.994	1.7E-03	3.0E-04
Dehydroabietic acid, pentafluorobenzyl ester	y = 5E+08x - 6E+06	0.994	3.0E-03	2.0E-04

Chapter 5

Profiling volatile organic compounds from naphthenic acids, acid extractable organic mixtures, and oil sands process-affected water by solid-phase microextraction-gas chromatography-electron impact mass spectrometry and chemometric analysis

5.1. Abstract

Naphthenic acids (NAs) are complex mixtures of carboxylic acids. They are natural components of oil and can be found in nature after oil spills. They are also produced in large quantities as an undesired waste product from oil sands in Alberta, Canada. There is ongoing research on the toxicity and chemical composition of such mixtures. In this chapter, headspace solid-phase microextraction (SPME) coupled to gas chromatography-electron impact mass spectrometry (GC–EIMS) was used for the qualitative analysis of the volatile organic compound (VOC) profile of three commercial NA blends, and an acid-extractable organic (AEO) mixture from a tailings pond in the Canadian Oil Sands region. In the samples analyzed, 54, 56, 40 and 4 compounds were identified in S1, S2, Merichem NAs and AEOs, respectively. The compounds identified in the mixtures included aliphatic and cyclic hydrocarbons, carboxylic acids, alkylbenzenes, phenols, naphthalene and alkyl-naphthalene, and decalin compounds. A sample of oil sands process-affected water (OSPW) from a tailings pond and aqueous solutions of the NA blends were analyzed to evaluate the matrix effect on the VOC profile. The compound identification was greatly limited by low signal and co-elution in the AEO and OSPW samples. The chromatograms were later subjected to principal

component analysis (PCA) and clustering analysis to determine similarities in the various mixtures. The commercial extracts were found to be closely related to each other and distinct from the AEO and OSPW samples, indicating similarities in the VOC profile of the commercial mixtures. The identified compounds have different levels of reported toxicity, including genotoxicity and carcinogenicity, and represent safety and environmental hazards.

5.2. Introduction

In the Canadian oil sands, oil is present in the form of bitumen. Bitumen is a form of petroleum in solid state, generally found mixed with sand and water. Mining is one of the extraction methods of bitumen used in the oil sands area. In 2013, approximately 2.5 barrels of water were required to produce one barrel of oil from bitumen. In 2016, an increase in industrial efficiency decreased the amount of water expenditure per oil barrel produced to 1.1 [153]. Nevertheless, oil sands process-affected water (OSPW) is continuously produced and held in tailings ponds. As of 2015, these ponds reached a volume of 1,075 Mm^3 [154]. The OSPW contains numerous pollutants such as polycyclic aromatic hydrocarbons (PAH), solvents including benzene, toluene, ethylbenzene, and xylenes, and naphthenic acids (NAs) [14,18,71,155].

Naphthenic acids are carboxylic acids naturally present in oil [22] and are the main source of toxicity in OSPW [20] with concentrations up to 128 mg/L [30]. NAs also result from oil spills, and have industrial applications as ink driers, and wood and fabric preservatives [65–67]. The acid-extractable organics (AEOs) from the OSPW and

commercial mixtures of NAs are toxic to aquatic organisms [88], including to developing tadpoles of *Silurana (Xenopus) tropicalis* (see Chapter 2).

Commercial mixtures of NAs and AEOs have been previously analyzed by gas chromatography-mass spectrometry (GC-MS) as methyl, silyl [129], and pentafluorobenzyl derivatives [83], but only a few compounds have been identified [54]. During derivatization processing, the samples are dried under nitrogen flux and low molecular weight volatiles may be lost. Moreover, the volatile components of such mixtures have not been previously examined. Equally important is the environmental impact of volatile discharge from commercial oil extraction. In 2017, Li et al. reported the emission of up to 70 ± 22 T/d of volatiles from surface mining areas [156]. Furthermore, during mining, the removal of the overburden exposes the underlying bitumen to abiotic factors. In 2009, Kelly et al. reported the migration of oil particles equivalent to 600 T of bitumen within 50 km of oil sands facilities [18], which may release organic compounds. Therefore, the composition of volatile fractions is crucial for understanding the toxicity of NA mixtures, AEOs, and OSPW, evaluating risk, and establishing regulations.

Solid-phase microextraction coupled to GC-MS (SPME-GC-MS) has been previously used for the analysis and identification of volatile organic compounds (VOCs) in water [157], and air [158]. This technique does not need sample evaporation and can therefore offer information on the volatile fraction of complex mixtures. In this chapter, SPME-GC-MS is used to explore the VOC profile of 3 commercial mixtures of NAs, an AEO, and an OSPW sample. Clean water was spiked with the commercial mixtures to observe the changes of the NA profile in the water.

In January 2017, I visited Dr. Jorge Molina Torres' laboratory at CINVESTAV IPN Mexico to conduct the experiments for this chapter. I conducted all these experiments under the supervision of Dr. Juan Vazquez Martinez and M. Sc. Enrique Ramirez Chavez, who helped me refine the experimental design and solving problems.

5.3. Methodology

5.3.1. Analytes

Three commercial mixtures were used in this study, two were purchased from Sigma-Aldrich with lot number BCBC9959V and BCBK0736V, hereafter called Sigma 1 (S1) and Sigma 2 (S2), respectively. The third, a Merichem NA mixture, was provided by Dr. John Headley (Environment and Climate change Canada, Saskatoon, SK). The acid-extractable organics (AEOs) from OSPW was obtained as previously reported [77]. Shortly, after collection, the OSPW was left to settle and the clarified OSPW was acidified to pH 2.5 using sulfuric acid. One L of OSPW was placed in a funnel, 500 mL of dichloromethane were added and the mixture was manually shaken for a couple of minutes. The organic phase was collected and concentrated using a rotary evaporator set to 40 ºC. 150 mL of 0.1 N NaOH was added to the AEO OSPW extract. The pH was reduced to 10 using sulfuric acid. The solution was filtered using 1000 mw cut-off membranes. The AEO extract analyzed here is the result of the re-extraction the AEOs that was dissolved in 0.1 N NaOH using the same methodology. A sample of OSPW was obtained from a pond in the oil sands area (courtesy of P. Knaga and S. Hughes, Shell), that was kept at 4 ºC covered and stored in the dark until analysis.

5.3.2. SPME procedures

To analyze the VOCs from the different samples, solid-phase microextraction was performed with Divinylbenzene/Carboxen/Polydimethylsiloxane (DVB/CAR/PDMS) 50/30 µm fibers, purchased from Supelco, USA. The four mixtures (100 mg) were placed into 2 mL screw-cap amber vials and tightly capped with PTFE/Silicone caps (Phenomenex). To understand the changes of the VOC profile of NA mixtures in water, ultrapure water (Milli-Q) was spiked with the commercial extracts (S1, S2, and Merichem) at a concentration of 50 mg/L. This concentration is within the range previously reported in OSPW samples [17]. Ten grams of OSPW and the aqueous solutions were individually placed in 125 mL screw top Erlenmeyer flask tightly capped with PTFE/Silicone caps. Ten grams of sodium sulfate (Na_2SO_4) anhydrous was used as a drying agent [159,160] and added to each of the solutions and to the OSPW sample. The vials and Erlenmeyer flasks were heated for 15 min at 70 °C in a water bath, the fiber was exposed to the headspace for 40 min at 40 °C in water bath. The fiber head did not contact the mixtures and aqueous solutions. Once sampling was finished, the fiber was withdrawn into the needle and transferred to the injection port of the GC-EIMS. After the extraction by SPME, the volatile compounds were desorbed from the SPME fiber in the GC injection port using manual injection for 10 min. All analyses were performed on three independent replicates. When material was limited, duplicates were analyzed. Blank analysis was run to control for possible contaminants from the fiber, empty vials or empty Erlenmeyer flasks. A C7 to C30 alkane calibration standard (Sigma-Aldrich) was used to calculate the retention indices (Ri) following the methodology of Sun and Stremple [144].

5.3.3. Instrumentation

A gas chromatograph (Agilent Technologies model 7890A GC System) coupled with an electron impact ionization mass spectrometer (Agilent technologies model 5975 Mass Selective Detector) was used for the analysis. The data obtained by the GC-EIMS was collected with the software MassHunter Workstation version B.06.00 (Agilent Technologies, Inc.). The injector temperature was set at 230 °C. The components separation was performed in an HP-5MS ultra inert GC column (30 m x 250 µm x 0.25 µm) and helium was used as carrier gas at a constant flow rate of 1.5 mL/min. The GC oven program began at 40 ºC, held for 4 min, and then was increased at a rate of 15 ºC/min to 300 ºC, held for 5 min. The transfer line temperature was set at 240 ºC. Electron impact mass spectra were obtained at 70 eV of electron energy. Measurements were performed in SCAN mode with *m/z* range set to 25-450. The ion source and quadrupole analyzer temperature were 230 ºC and 150 ºC respectively and operated at 2.9 scans per second.

5.3.4. Identification of components

The software Automated Mass Spectral Deconvolution and Identification System "AMDIS" (http://www.amdis.net/) was used for the determination of the retention time and extracted mass spectrum for each component of the chromatograms of the analytes. The identification of the compounds was performed by comparison of the generated mass spectra with those of the reference compounds of the NIST Mass Spectral Search Program (NIST 02 version mass spectral database, National Institute of Standards and Technology, Washington, DC, USA). As indicated in the manual of the NIST, a match of 900 or greater is considered excellent, 800-900, good, and 700-800 a fair match. Here, the components with a match greater than 800 and after inspection of

the molecular weight, retention index, parent ion and the distribution of the 10 major ions were considered as an identified compound.

5.3.5. Chemometric analysis

Data analysis was performed to identify similarities among the groups of samples. The GC-EIMS chromatograms from retention times of 3-40 min were used for pattern analysis. The software R was used [161] and R Studio as the interface [162] for chemometric analysis. The R package FactoMineR was used for Principal Component Analysis (PCA) [163]. The components with eigenvalues higher than 1 were chosen for the plots. Clustering and bootstrapping analysis were performed using the package pvclust in R [164]. Clustering analysis was performed using Ward's algorithm [165] based on the Euclidean distance. Statistical support for each node in the dendrogram was obtained by the bootstrap method with 1000 replicates. The R package pvclust [164], provides AU (approximately unbiased) *p*-values and BP (bootstrap probability) *p*-values. Nodes with unbiased *p*-values > 95 were considered supported. The R package RcolorBrewer was used to assign colorblind-friendly colors to the PCA plots [166]. The R package dendextend was used to improve the visualization of the dendrograms [167].

5.4. Results and discussion

5.4.1. VOC profile

Using ChemStation to analyze the total ion chromatogram (TIC) of the mixtures, the number of peaks detected was 265, 267, 390, and 267 for S1, S2, Merichem, and AEOs (Fig. 5.1), respectively. However, the number of peaks underestimates the number of components in each of the mixtures, since co-elution was observed during the deconvolution analysis in AMDIS. The signals obtained in the chromatogram of the

volatile components of the AEOs mixture were lower compared to the commercial mixtures (Fig. 5.1), likely due to the processing of the AEO mixture. Acid extraction of OSPW includes the use of DCM that later is removed with a rotavapor and nitrogen flux. The low signal limited identification to four compounds in the AEO sample.

A total of 102 VOCs were identified in the four different mixtures. Out of the compounds that were identified, 86% are hydrocarbons, 5% were carboxylic acid and 9% others, including but not limited to phenol and acetone (Table 5.1, Fig. 5.2). In the different mixtures, 54, 56, 40, and four VOCs were successfully identified in S1, S2, Merichem, and AEOs, respectively (Table 5.1). The identification of the components was restricted by library limitations and co-elution. Therefore, some mass spectra could not be analyzed. For the analysis, NIST similarity matches ≥ 800 with manual revision of each spectrum was considered acceptable.

The highest number of compounds (30) were shared between S1 and S2. Merichem and S1 shared 20 compounds, while S2 and Merichem shared 18. No VOCs were shared between the AEOs and the commercial extracts (Fig. 5.3, Table 5.1). The similarities observed between S1 and S2 could be due to their common supplier and possibly oil source, but cannot be confirmed as Sigma Aldrich does not disclose the source. The similarities of S1 and S2 were also observed in the organic profiles (Chapter 2) determined using electrospray ionization high-resolution mass spectrometry (ESI-HRMS). This showed that S1 and S2 have a similar proportion of O2 species, and slightly different double bond equivalent (DBE) profiles and in carbon number distributions.

In chapter 4, PFBBr was used to derivatize extracts of NAs. The mass spectra analyses showed that there were components that did not react with PFBBr indicating the lack of the carboxylic group and therefore not NAs. The presence of non-NA components in commercial mixtures of NAs has implications in the analysis of concentration assessment of NAs in water. Commercial extracts, including Sigma and Merichem extracts, have been used as standards for the determination of the concentration of measuring NA concentrations in water [30,124]. The concentrations reported in these articles might be inaccurate since commercial NA mixtures contain non-NA components. Also, previous studies may be difficult to replicate, since two extracts from the same supplier have different SPME-GC-EIMS profiles and might also contain a different proportion of NAs and non-NA components.

The compounds identified in the several mixtures included linear alkanes, alkylbenzenes, phenols, naphthalenes, and decalin compounds, among others. Linear alkanes ranging from C10 to C16 were found in the S1 mixture, C9 to C16 for the S2 mixture and C12 to C17 in the Merichem extract (Table 1, Fig. 5.2 a). The C7 to C30 alkane calibration standard was used for the calculation of the Ri and as the standard for confirming the presence of the linear alkanes found in the different mixtures. Alkylbenzenes, including p-xylene, was one of the groups of the compounds that were identified in S1 and S2 (Table 5.1, Fig. 5.2 b). These compounds naturally occur in oil and belong to the benzene, toluene, ethylbenzene, and xylene (BTEX) group. Xylenes can be found at a relatively high concentration in water after oil spills [168]. Other groups of compounds such as phenols were also found in the mixtures (Table 5.1). Naphthalene, the smallest PAH according to the US EPA [169], was identified in S2

(Table 5.1). Alkylated naphthalenes were also identified in the three commercial mixtures (Table 5.1). The EPA considers 16 parent PAHs on the priority list, however, it has been proposed to include alkylated PAHs [170]. Decalin compounds were found in the 3 commercial mixtures (Table 5.1).

D-limonene was found in the VOC profile of the AEOs sample (Table 5.1, Fig. 5.2 d). However, it probably comes from forest disturbances from mining extraction, since D-limonene is a monoterpene naturally found in conifers [171]. In the past, dehydroabietic acid (DHAA), a resin acid toxic to fish [172] and to fish glial cells in vitro [173], was previously identified in an OSPW sample [34]. However, this compound is also believed to be result of forest disturbances.

The compounds found in the mixtures have different levels of known toxicity. For example, alkanes have been reported to have toxic effects in vertebrates, including hepatoxicity [174], damage in the central nervous system [175], and skin irritation [176,177]. Xylene is genotoxic [178] and can induce DNA methylation, thus part of its toxicity is at the epigenetic level [179]. Phenols are components of petroleum [180], they are corrosive to skin [181] and toxic to aquatic life, including trout [182]. Naphthalene can induce anemia in humans [183] and has been classified as a possible human carcinogen [184]. Decalin naturally occurs in oil [185] and is carcinogenic to rats [186]. The extracts S1, S2, and AEOs are toxic, induced morphological abnormalities in a dose-response fashion, and disrupt gene networks related to membrane function, metabolic processes, and gut function in *Silurana (Xenopus) tropicalis* embryos (see Chapter 2 and 3). The identification of toxic non-NA compounds indicates that the lethal

and sub-lethal effects of S1, S2, and AEOs may be partially due to the presence of compounds in addition to NAs.

5.4.2. VOC profile of aqueous samples

From the aqueous solutions of S1 and S2, it was not possible to identify any compound with the criteria herein established. However, in the aqueous solution of Merichem, 8 hydrocarbons were identified including the linear alkanes C13, C14, C15, and C16 (Table 5.2). A lower signal coupled with compound co-elution was observed in all of the mixtures, limiting the identification of the components (Fig. 5.4). The differences in the signal could be due to the concentration or to matrix interferences. Volatility is dependent on intermolecular forces, such as hydrogen bonds. In the different mixtures, several compounds that could form hydrogen bonds with water were identified (Table 5.1). The VOC signal of the OSPW was also low, thus compound identification was not possible (Fig. 5.4 d). This sample was taken in 2016 and analyzed in January 2017. The containers were sealed and protected from light at 4°C. There are logistical limitations, including the on-site OSPW collection, handling, and analysis of these samples, therefore the analysis of fresher samples is recommended.

5.4.3. Chemometric analysis

The relationship between the VOC profiles of the different mixtures was analyzed by chemometric methodologies. Principal component analysis is a non-supervised multivariate method for dimensionality reduction where observations with similar characteristics are closely located in the space [187]. When PCA was performed with the chromatographic data of S1, S2, Merichem and the AEO extracts, PC 1 and PC 2 accounted for 93% of the overall variance (Fig. 5.5 a). In the PCA plot, the AEOs are

separated from the commercial extracts. The PCA also indicates that the VOC profiles of S1 and S2 have major similarities. The replicates of a given sample are tightly grouped together. The cluster analysis shows two main clades (Fig. 5.5 b) that marks the separation of the AEOs and S1, S2, and Merichem extracts. A second clade is composed of the commercial extracts and marks the separation of the Sigma and the Merichem extract. The extracts S1 and S2 are sister groups, suggesting highly similar VOC composition and relative abundances. In the same way, S1 and S2 are sister groups with Merichem. However, the AEOs forms an outgroup to S1, S2, and Merichem, illustrating the differences in their VOC composition. The different replicates are located in their respective cluster and are not mixed between groups indicating good reproducibility. The cluster analysis showed that the closest groups are S1 and S2 having the lowest distance among all clusters. The AU and BP values (*p*-values > 95) of the bootstrap analysis showed that all of the clades have significant statistical support.

The relationship between the VOC profiles of the different mixtures in aqueous solutions at 50 mg/L and a sample of OSPW were determined by PCA and clustering with bootstrap analysis. When PCA was performed with the chromatographic data of the extracts in aqueous solution, PC 1 and PC 2 accounted 88% of the overall variance (Fig. 5.6 a). The PCA plot shows 3 main groups, the Sigma blends were closely related to each other, suggesting that the VOC profile of S1 and S2 in aqueous solutions have similarities. The second and third group are formed by the Merichem mixture in aqueous solution and the OSPW sample. The cluster analysis shows two main clades (Fig. 5.6 b) that marks the separation of the S1 and S2 from AEOs and Merichem extracts, suggesting similarities in the S1 and S2 VOC composition and relative abundance in

aqueous solution. However, the AEOs and the Merichem extract form a second clade. The clustering of the Merichem and the AEOs could be due to matrix interference. The different replicates are located in their respective cluster and there are no replicates mixed between groups, indicating good reproducibility. The AU and BP values of the bootstrap analysis showed that most of the clades have significant statistical support (*p*-values > 0.95). This does not happen in the clade that separates replicate 2 and 3 of the Merichem extract in aqueous solution. This could mean that the differences between these two replicates are not large enough to form a different highly supported clade.

5.5. Conclusion

The analysis of the VOC profile using SPME-GC-EIMS is a relatively fast approach to understand the contribution of oil extraction to volatiles into the environment and hazards for oil workers. In this chapter, the identification of more than 100 volatile components of commercial extracts and an AEO extract are reported. Most of the components were hydrocarbons and a few compounds with a carboxylic acid moiety. The chemometric analysis grouped the different extracts according to their VOC chromatogram. Some of the compounds identified are genotoxic, skin irritants and some such as naphthalene are classified as possible human carcinogens.

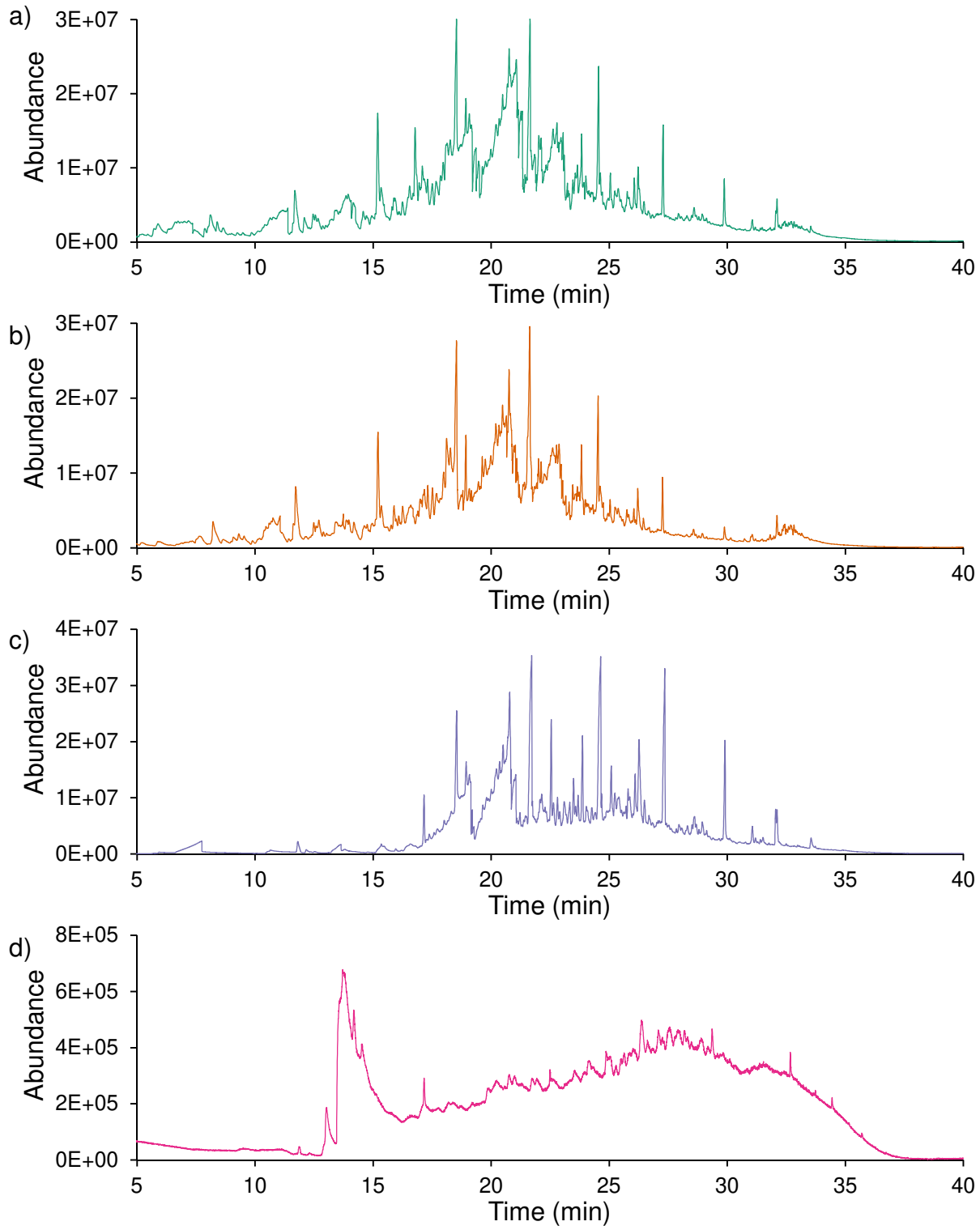


Figure 5.1. Total ion chromatogram (TIC) of the volatile components of S1 (a), S2 (b), Merichem (c), and AEOs (d), using SPME-GC-EIMS.

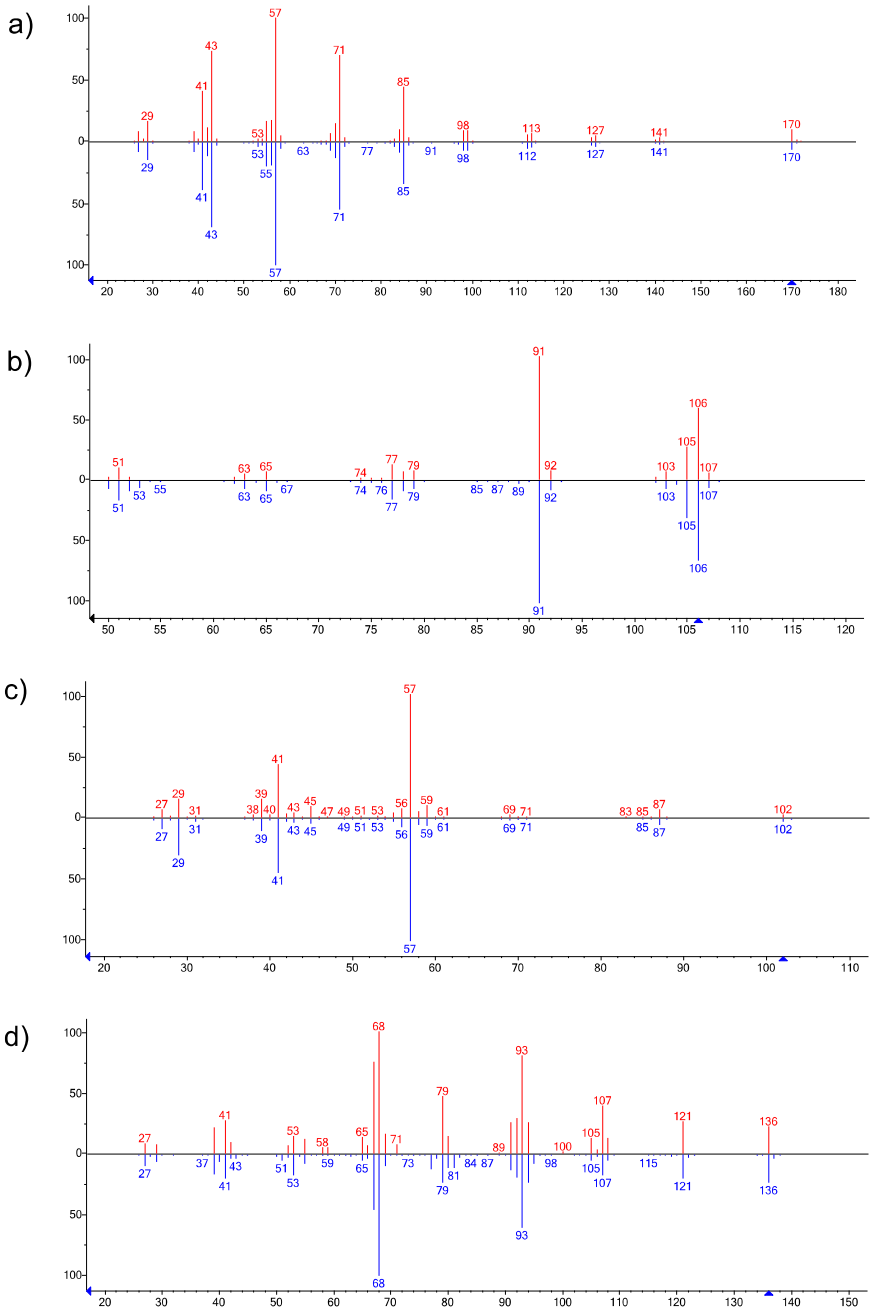


Figure 5.2. Mass spectra of a peak at 18.53 min in S1 identified as dodecane (red spectrum) and NIST standard of dodecane (blue spectrum) (a). Mass spectra of a peak at 7.49 min in S2 identified as p-xylene (red spectrum) and NIST standard of p-xylene (blue spectrum) (b). Mass spectra of a peak at 7.74 min in Merichem identified as 2,2-dimethyl-propanoic acid (red spectrum) and NIST standard of 2,2-dimethyl-propanoic acid (blue spectrum) (c). Mass spectra of a peak at 13.54 min in AEOs identified as D-limonene (red spectrum) and NIST standard of D-limonene (blue spectrum) (d).

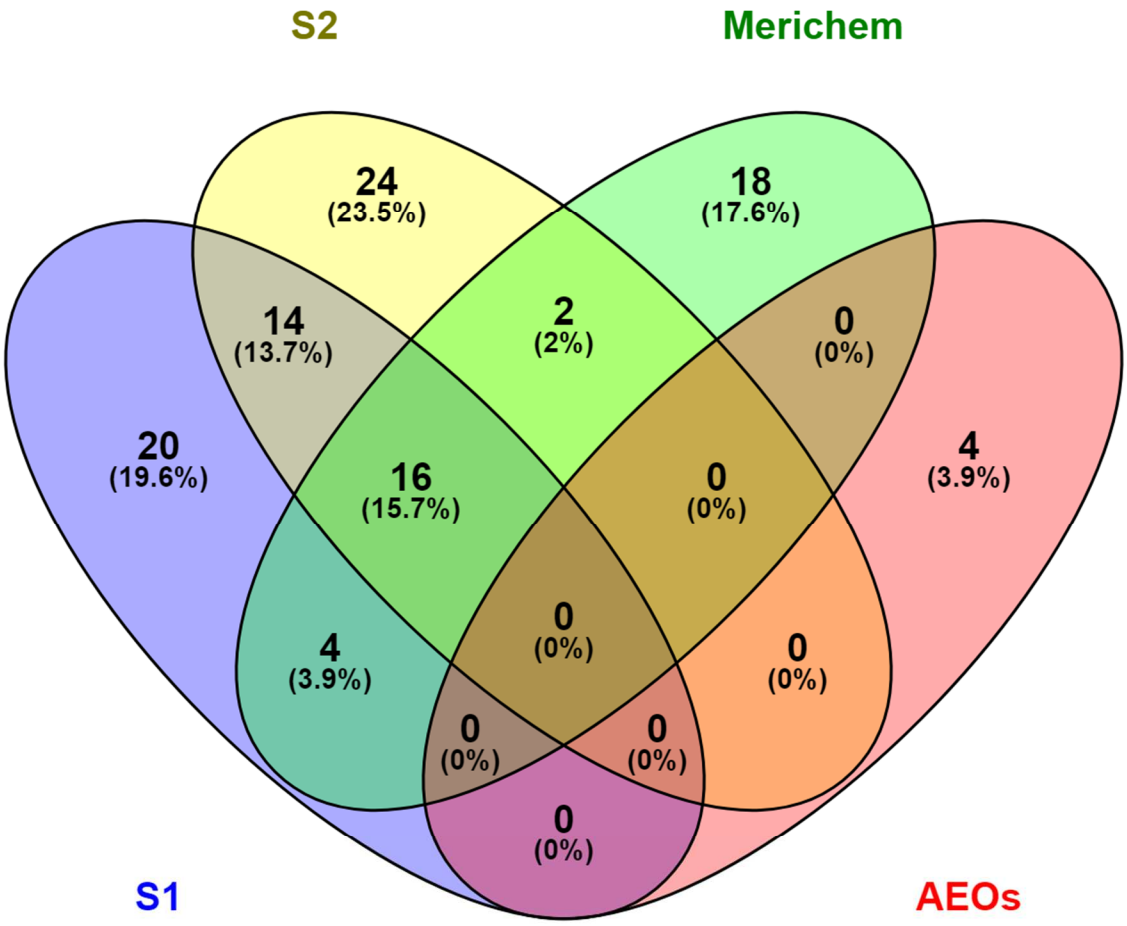


Figure 5.3. Venn diagram of the VOCs identified in S1, S2, Merichem, and AEOs using SPME-GC-EIMS.

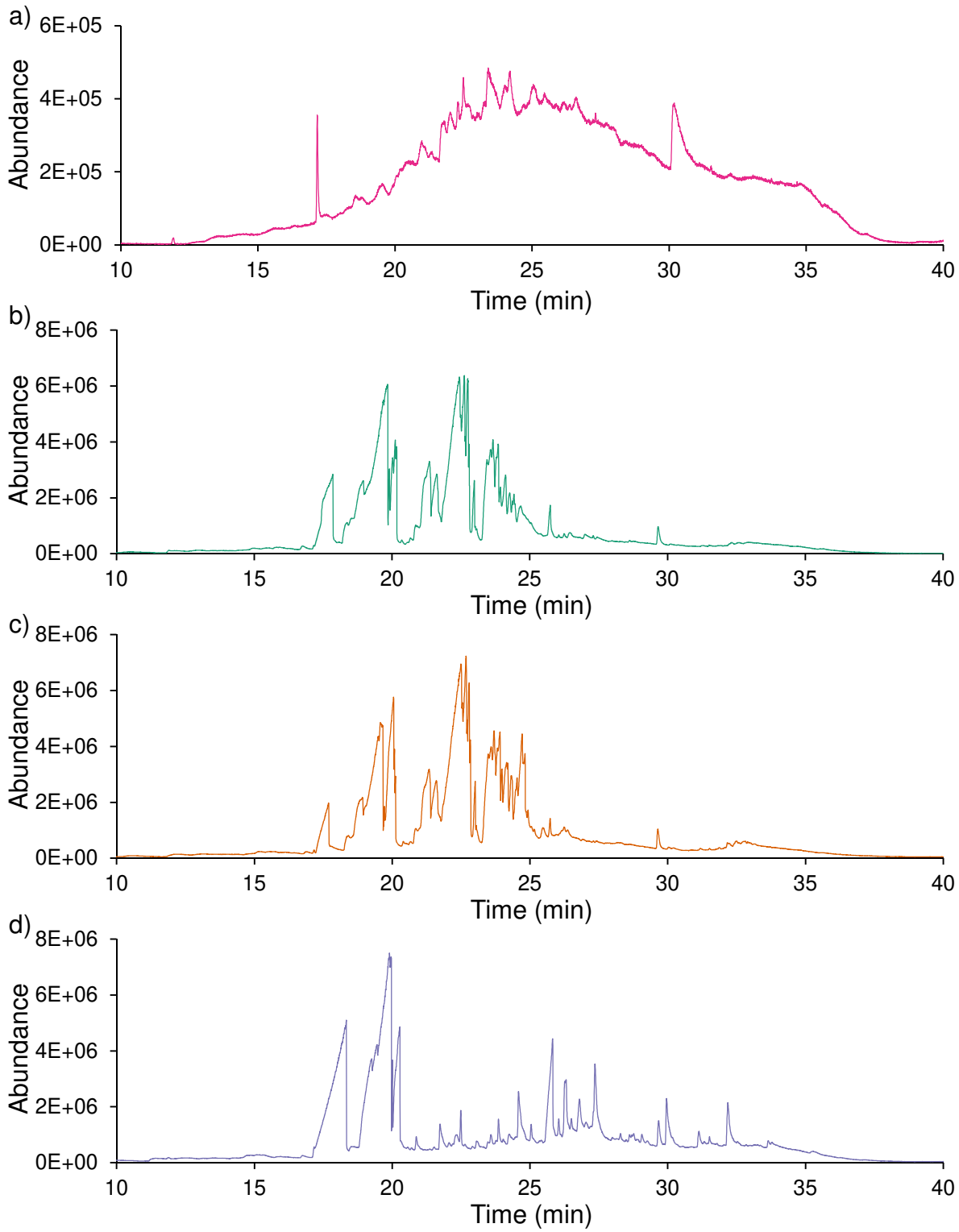


Figure 5.4. SPME-GC-EIMS total ion chromatogram (TIC) of the volatile components of aqueous samples of S1 (a), S2 (b), Merichem (c) at 50 mg/L, and OSPW (d).

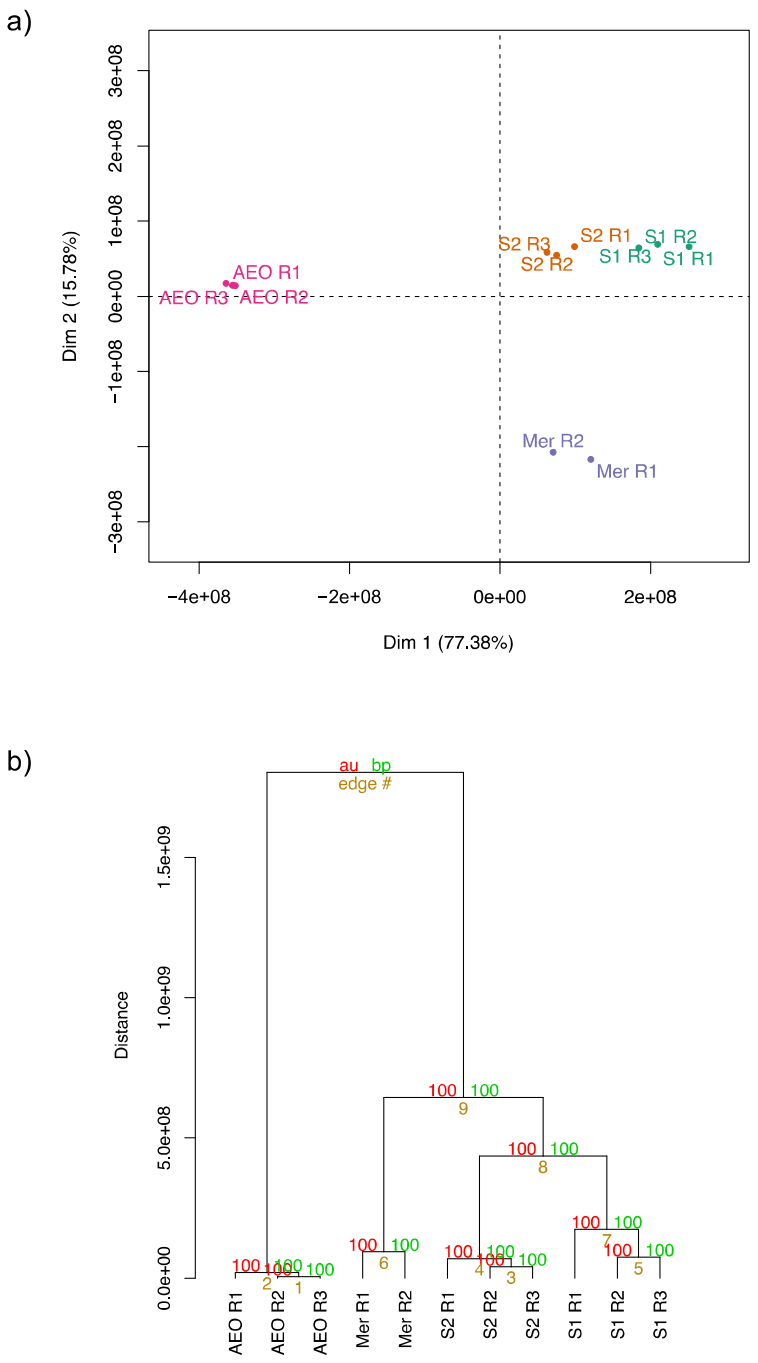


Figure 5.5. Score plot for PCA of the SPME-GC-EIMS chromatograms of S1, S2, Merichem and AEOs in a). The percentages along with Dim 1 and Dim 2 indicate their contribution to the total variance in the PCA. Cluster dendrogram of the SPME-GC-EIMS chromatograms of S1, S2, Merichem, and AEOs. Bootstrap support for cluster provided as AU (Approximately Unbiased) *p*-values in red and BP (Bootstrap Probability) value in green (b).

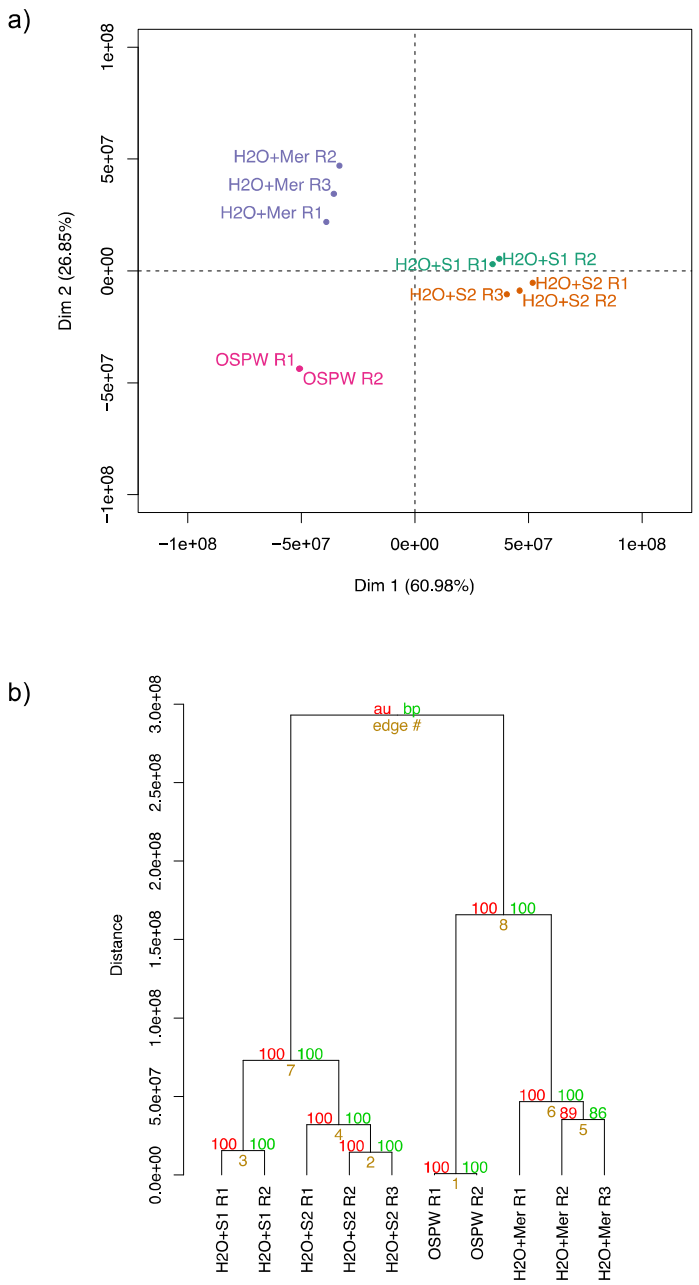
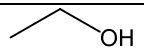
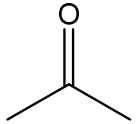
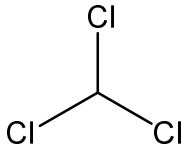
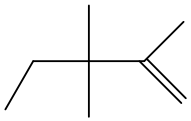
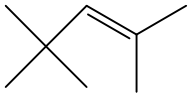
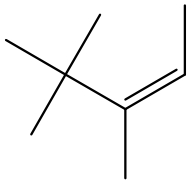
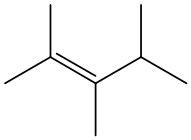
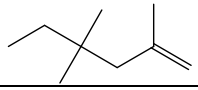
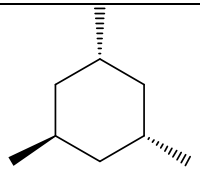
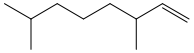
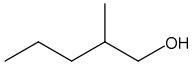
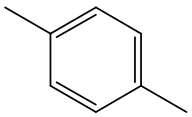
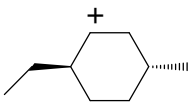
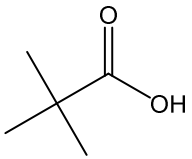
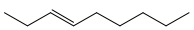
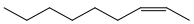
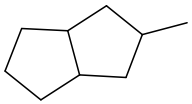
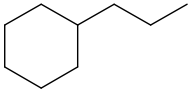
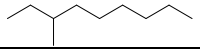
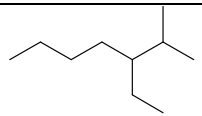
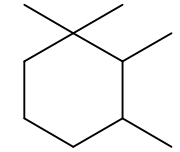
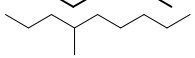
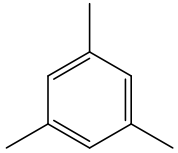
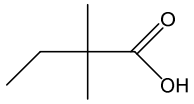
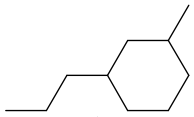
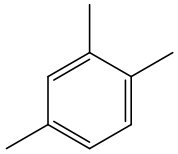
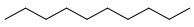
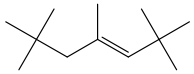
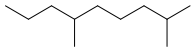
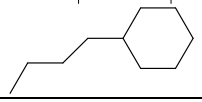


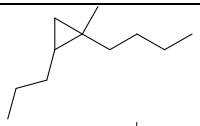
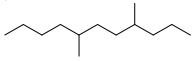
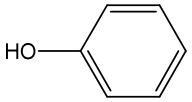
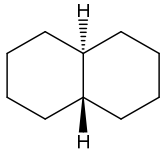
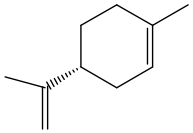
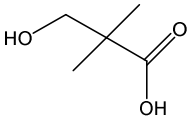
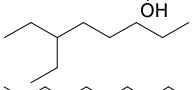
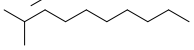
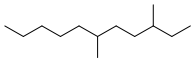
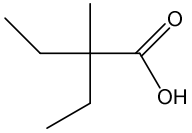
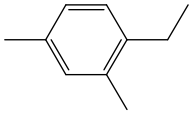
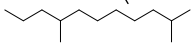
Figure 5.6. Score plot for PCA of the SPME-GC-EIMS chromatograms of aqueous solutions of S1, S2, and Merichem at 50 mg/L, and an OSPW sample in a). The percentages along with Dim 1 and Dim 2 indicate their contribution to the total variance in the PCA. Cluster dendrogram of the SPME-GC-EIMS chromatograms of aqueous solutions of S1, S2, and Merichem at 50 mg/L and an OSPW sample. Bootstrap support for cluster provided as AU (Approximately Unbiased) *p*-values in red and BP (Bootstrap Probability) value in green (b).

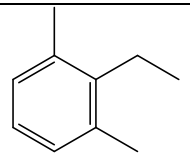
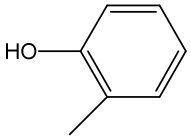
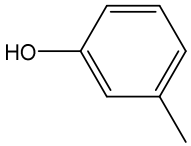
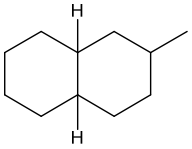
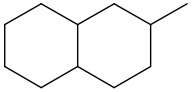
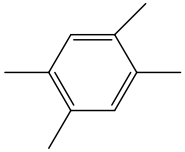
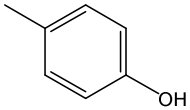
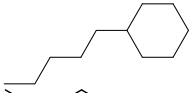
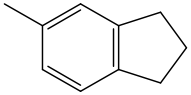
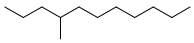
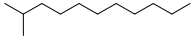
Table 5.1. List of volatile organic compounds identified in mixtures of NAs (S1, S2, and Merichem), and AEOs by SPME coupled to GC-EIMS. Rt, retention time. Ri, retention index.

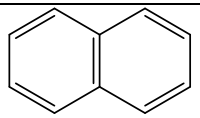
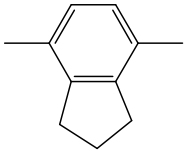
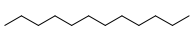
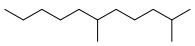
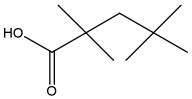
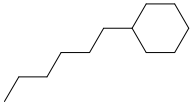
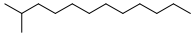
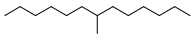
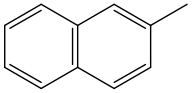
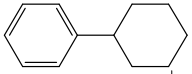
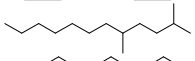
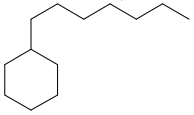
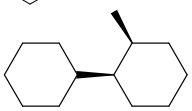
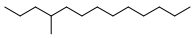
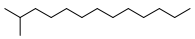
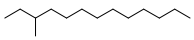
No.	Rt (min)	Ri	Detected compound	Group	Chemical formula	Chemical Structure	CAS number	S1	S2	Mer	AEOs
1	1.78		Ethanol	Other	C2H6O		64-17-5			✓	
2	1.91	-	Acetone	Other	C3H6O		67-64-1	✓			
3	2.54	-	Trichloromethane	Other	CHCl3		67-66-3			✓	
4	3.39	-	1-Pentene, 2,3,3-trimethyl-	Hydrocarbon	C8H16		560-23-6	✓	✓		
5	3.75	-	2-Pentene, 2,4,4-trimethyl-	Hydrocarbon	C8H16		107-40-4	✓			
6	3.85	-	Z-3,4,4-trimethyl-2-pentene	Hydrocarbon	C8H16		39761-64-3		✓		
7	4.44	-	2-Pentene, 2,3,4-trimethyl-	Hydrocarbon	C8H16		565-77-5		✓		
8	5.89	-	2,4,4-Trimethyl-1-hexene	Hydrocarbon	C9H18		51174-12-0	✓			

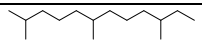
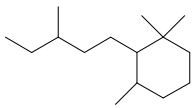
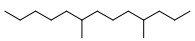
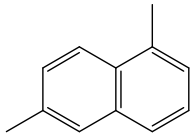
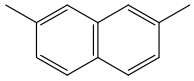
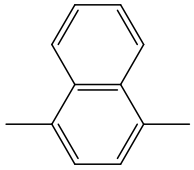
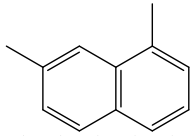
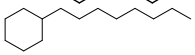
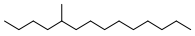
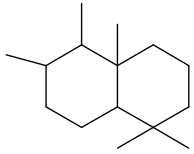
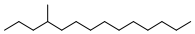
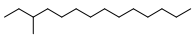
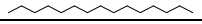
9	6.41	-	Cyclohexane, 1,3,5-trimethyl-, (1α,3α,5β)-	Hydrocarbon	C9H18		1795-26-2		✓	
10	6.63	-	1-Octene, 3,7-dimethyl-	Hydrocarbon	C10H20		4984-01-4	✓		
11	6.74	-	1-Pentanol, 2-methyl-	Other	C6H14O		105-30-6	✓		
12	7.49	-	p-Xylene	Hydrocarbon	C8H10		106-42-3	✓	✓	
13	7.67	-	Cyclohexane, 1-ethyl-4-methyl-, trans-	Hydrocarbon	C9H18		6236-88-0		✓	
14	7.74	-	Propanoic acid, 2,2-dimethyl-	Carboxylic acid	C5H10O2		75-98-9			✓
15	7.86	-	1-Nonene	Hydrocarbon	C9H18		124-11-8	✓		
16	8.13	903	3-Nonene	Hydrocarbon	C9H18		20063-77-8	✓		
17	8.22	906	Nonane	Hydrocarbon	C9H20		111-84-2		✓	
18	8.40	911	cis-2-Nonene	Hydrocarbon	C9H18		6434-77-1	✓		
19	8.64	917	Pentalene, octahydro-2-methyl-	Hydrocarbon	C9H16		3868-64-2		✓	
20	9.10	930	Cyclohexane, propyl-	Hydrocarbon	C9H18		1678-92-8		✓	
21	9.32	935	Nonane, 3-methyl-	Hydrocarbon	C10H22		5911-04-6		✓	

22	9.53	941	Heptane, 3-ethyl-2-methyl-	Hydrocarbon	C10H22		14676-29-0		✓	
23	9.86	950	Cyclohexane, 1,1,2,3-tetramethyl-	Hydrocarbon	C10H20		6783-92-2	✓	✓	
24	10.48	967	Nonane, 4-methyl-	Hydrocarbon	C10H22		17301-94-9		✓	
25	10.77	975	Benzene, 1,3,5-trimethyl-	Hydrocarbon	C9H12		108-67-8		✓	
26	11.22	987	Butanoic acid, 2,2-dimethyl-	Carboxylic acid	C6H12O2		595-37-9	✓	✓	
27	11.43	992	Cyclohexane, 1-methyl-3-propyl-	Hydrocarbon	C10H20		4291-80-9		✓	
28	11.62	997	Benzene, 1,2,4-trimethyl-	Hydrocarbon	C9H12		95-63-6	✓	✓	
29	11.71	1000	Decane	Hydrocarbon	C10H22		124-18-5	✓	✓	
30	12.13	1012	3-Heptene, 2,2,4,6,6-pentamethyl-	Hydrocarbon	C12H24		123-48-8	✓		✓
31	12.47	1021	Nonane, 2,6-dimethyl-	Hydrocarbon	C11H24		17302-28-2	✓	✓	
32	12.69	1028	Cyclohexane, butyl-	Hydrocarbon	C10H20		1678-93-9	✓	✓	

33	13.02	1037	Cyclopropane, 1-butyl-1-methyl-2-propyl-	Hydrocarbon	C11H22		41977-34-8		✓	
34	13.02	1037	Undecane, 4,7-dimethyl-	Hydrocarbon	C13H28		17301-32-5			✓
35	13.16	1041	Phenol	Other	C6H6O		108-95-2	✓		
36	13.34	1046	Naphthalene, decahydro-, trans-	Hydrocarbon	C10H18		493-02-7	✓		
37	13.54	1051	D-Limonene	Hydrocarbon	C10H16		5989-27-5			✓
38	13.74	1057	Hydroxypivalic acid	Carboxylic acid	C5H10O3		4835-90-9		✓	
39	13.83	1060	Octane, 3-ethyl-	Hydrocarbon	C10H22		5881-17-4	✓		
40	13.97	1064	Decane, 2-methyl-	Hydrocarbon	C11H24		6975-98-0	✓	✓	
41	14.19	1070	Undecane, 3,6-dimethyl-	Hydrocarbon	C13H28		17301-28-9			✓
42	14.25	1071	Butanoic acid, 2-ethyl-2-methyl-	Carboxylic acid	C7H14O2		19889-37-3	✓		
43	14.51	1079	Benzene, 1-ethyl-2,4-dimethyl-	Hydrocarbon	C10H14		874-41-9	✓		
44	14.54	1080	Undecane, 2,8-dimethyl-	Hydrocarbon	C13H28		17301-25-6			✓

45	14.78	1086	Benzene, 2-ethyl-1,3-dimethyl-	Hydrocarbon	C10H14		2870-04-4	✓		
46	15.09	1095	Phenol, 2-methyl-	Other	C7H8O		95-48-7	✓		
47	15.13	1096	Phenol, 3-methyl-	Other	C7H8O		108-39-4		✓	
48	15.20	1098	Undecane	Hydrocarbon	C11H24		1120-21-4	✓	✓	
49	15.34	1102	trans-Decalin, 2-methyl-	Hydrocarbon	C11H20		-	✓		
50	15.91	1120	Decalin, 2-Methyl	Hydrocarbon	C11H20		2958-76-1	✓	✓	✓
51	15.99	1122	Benzene, 1,2,4,5-tetramethyl-	Hydrocarbon	C10H14		95-93-2		✓	
52	16.10	1125	Phenol, 4-methyl-	Other	C7H8O		106-44-5	✓		
53	16.25	1130	Cyclohexane, pentyl-	Hydrocarbon	C11H22		4292-92-6	✓	✓	
54	16.90	1149	1H-Indene, 2,3-dihydro-5-methyl-	Hydrocarbon	C10H12		874-35-1		✓	
55	17.17	1158	Undecane, 4-methyl-	Hydrocarbon	C12H26		2980-69-0		✓	
56	17.34	1163	Undecane, 2-methyl-	Hydrocarbon	C12H26		7045-71-8	✓	✓	✓

57	18.11	1186	Naphthalene	Hydrocarbon	C10H8		91-20-3		✓	
58	18.15	1187	1H-Indene, 2,3-dihydro-4,7-dimethyl-	Hydrocarbon	C11H14		6682-71-9	✓		
59	18.53	1199	Dodecane	Hydrocarbon	C12H26		112-40-3	✓	✓	✓
60	18.93	1211	Undecane, 2,6-dimethyl	Hydrocarbon	C13H28		17301-23-4	✓	✓	✓
61	19.17	1219	Pentanoic acid, 2,2,4,4-tetramethyl-	Carboxylic acid	C9H18O2		3302-12-3	✓		✓
62	19.65	1234	Cyclohexane, hexyl-	Hydrocarbon	C12H24		4292-75-5			✓
63	20.49	1262	Dodecane, 2-methyl	Hydrocarbon	C13H28		1560-97-0	✓	✓	✓
64	20.77	1271	Tridecane, 7-methyl	Hydrocarbon	C14H30		26730-14-3	✓	✓	✓
65	21.67	1300	Tridecane	Hydrocarbon	C13H28		629-50-5	✓	✓	✓
66	22.04	1312	Naphthalene, 2-methyl-	Hydrocarbon	C11H10		91-57-6	✓	✓	✓
67	22.12	1315	Benzene, cyclohexyl-	Hydrocarbon	C12H16		827-52-1		✓	
68	22.14	1316	Dodecane, 2,5-dimethyl-	Hydrocarbon	C14H30		56292-65-0			✓
69	22.81	1339	Heptylcyclohexane	Hydrocarbon	C13H26		5617-41-4			✓
70	23.14	1350	1,1'-Bicyclohexyl, 2-methyl-, cis-	Hydrocarbon	C13H24		50991-08-7			✓
71	23.32	1356	Tridecane, 4-methyl-	Hydrocarbon	C14H30		26730-12-1		✓	✓
72	23.47	1361	Tridecane, 2-methyl-	Hydrocarbon	C14H30		1560-96-9	✓	✓	✓
73	23.68	1369	Tridecane, 3-methyl-	Hydrocarbon	C14H30		6418-41-3			✓

74	23.84	1374	Dodecane, 2,6,10-trimethyl-	Hydrocarbon	C15H32		3891-98-3	✓	✓	✓
75	24.03	1381	Cyclohexane, 1,1,3-trimethyl-2-(3-methylpentyl)-	Hydrocarbon	C15H30		54965-05-8			✓
76	24.57	1399	Tetradecane	Hydrocarbon	C14H30		629-59-4	✓	✓	✓
77	24.68	1403	Tridecane, 4,8-dimethyl-	Hydrocarbon	C15H32		55030-62-1			✓
78	24.91	1411	Naphthalene, 1,6-dimethyl-	Hydrocarbon	C12H12		575-43-9			✓
79	24.96	1413	Naphthalene, 2,7-dimethyl-	Hydrocarbon	C12H12		582-16-1		✓	
80	25.25	1424	Naphthalene, 1,4-dimethyl-	Hydrocarbon	C12H12		571-58-4	✓		
81	25.25	1424	Naphthalene, 1,7-dimethyl-	Hydrocarbon	C12H12		575-37-1		✓	✓
82	25.76	1442	Cyclohexane, octyl-	Hydrocarbon	C14H28		1795-15-9	✓	✓	
83	25.96	1450	Tetradecane, 5-methyl-	Hydrocarbon	C15H32		25117-32-2		✓	
84	26.07	1453	Decahydro-4,4,8,9,10-pentamethylnaphthalene	Hydrocarbon	C15H28		80655-44-3	✓	✓	✓
85	26.11	1455	Tetradecane, 4-methyl-	Hydrocarbon	C15H32		25117-24-2		✓	
86	26.47	1468	Tetradecane, 3-methyl-	Hydrocarbon	C15H32		18435-22-8	✓		✓
87	27.30	1498	Pentadecane	Hydrocarbon	C15H32		629-62-9	✓	✓	✓

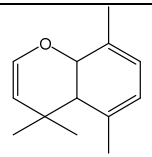
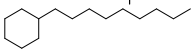
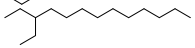
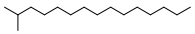
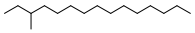
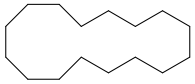
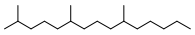
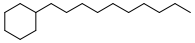
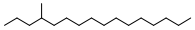
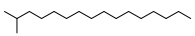
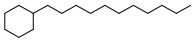
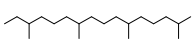
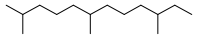
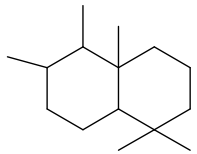
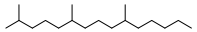
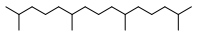
88	27.82	1518	4H-1-Benzopyran, 4,4,5,8-tetramethyl-	Other	C13H16O		82391-06-8	✓		
89	28.56	1547	n-Nonylcyclohexane	Hydrocarbon	C15H30		2883-02-5	✓	✓	
90	28.80	1556	Tridecane, 3-ethyl-	Hydrocarbon	C15H32		13286-73-2			✓
91	28.94	1561	Pentadecane, 2-methyl-	Hydrocarbon	C16H34		1560-93-6		✓	
92	29.12	1568	Pentadecane, 3-methyl-	Hydrocarbon	C16H34		2882-96-4	✓	✓	
93	29.77	1593	Cyclohexadecane	Hydrocarbon	C16H32		295-65-8			✓
94	29.88	1598	Hexadecane	Hydrocarbon	C16H34		544-76-3	✓	✓	✓
95	31.06	1652	Pentadecane, 2,6,10-trimethyl-	Hydrocarbon	C18H38		3892-00-0	✓	✓	✓
96	31.19	1659	Cyclohexane, decyl-	Hydrocarbon	C16H32		1795-16-0			✓
97	31.29	1663	Hexadecane, 4-methyl-	Hydrocarbon	C17H36		25117-26-4			✓
98	31.40	1668	Hexadecane, 2-methyl-	Hydrocarbon	C17H36		1560-92-5			✓
99	32.05	1699	Heptadecane	Hydrocarbon	C17H36		629-78-7			✓
100	32.11	1703	Pentadecane, 2,6,10,14-tetramethyl-	Hydrocarbon	C19H40		1921-70-6	✓		✓
101	32.98	1767	Cyclohexane, undecyl-	Hydrocarbon	C17H34		54105-66-7			✓
102	33.55	1811	Hexadecane, 2,6,10,14-tetramethyl-	Hydrocarbon	C20H42		638-36-8	✓	✓	✓

Table 5.2. List of volatile organic compounds identified in the aqueous solution of Merichem at 50 mg/L by SPME coupled to GC-EIMS. Rt, retention time. Ri, retention index.

No.	Rt (min)	Ri	Detected compound	Group	Chemical formula	Chemical structure	CAS number
1	21.74	1302	Tridecane	Hydrocarbon	C13H28		629-50-5
2	23.86	1373	Dodecane, 2,6,10-trimethyl-	Hydrocarbon	C15H32		3891-98-3
3	24.59	1400	tetradecane	Hydrocarbon	C14H30		629-59-4
4	26.05	1453	Decahydro-4,4,8,9,10-pentamethylnaphthalene	Hydrocarbon	C15H28		80655-44-3
5	27.36	1501	Pentadecane	Hydrocarbon	C15H32		629-62-9
6	29.96	1601	Hexadecane	Hydrocarbon	C16H34		544-76-3
7	31.14	1656	Pentadecane, 2,6,10-trimethyl-	Hydrocarbon	C18H38		3892-00-0
8	32.18	1708	Pentadecane, 2,6,10,14-tetramethyl-	Hydrocarbon	C19H40		1921-70-6

Chapter 6

Fractionation of commercial NA mixtures and AEOs reveals that acute toxicity is mainly due to naphthenic acids

6.1. Abstract

In this chapter, open column chromatography was used to fractionate a commercial naphthenic acid (NA) mixture and an acid extractable organic (AEO) mixture extracted from an oil sands process-affected water (OSPW) sample taken from a tailings pond in the oil sands area in Alberta, Canada. The mixtures and the fractions were analyzed as pentafluorobenzyl derivatives by gas chromatography coupled to electron impact mass spectrometry. The lethal and sublethal toxicity of the fractions was later assessed using *Silurana (Xenopus) tropicalis* embryos. A mixture of classic and non-classic NA standards was used to validate the derivatization with pentafluorobenzyl bromide (PFBBr). PFBBr effectively reacted with classical and non-classical NA standards, including linear, dicarboxylic, sulfur-containing, adamantanes, and cyclic NAs. The mass spectra analysis of the standards showed the fragment ion at 181 *m/z* as one of the 10 most abundant ions and in five cases as the base peak. From the open column chromatography with a Sigma commercial extract (S2), a non-polar and a polar fractions were obtained. The fragment ion at 181 *m/z* was not observed in the non-polar fraction. The recovery rate from the column was 95%. The S2 extract and its polar fraction significantly decreased survival, total length (TL), tail length (TaL), snout-vent length (SVL), and interorbital distance (IOD) ($p < 0.05$). Six fractions from the AEOs were recovered, but only the toxicity of two (F3 and F4) of the fractions was tested due

to limited material. The recovery rate from the column was 93%. The AEO extract significantly reduced survival, TL, TaL, SVL, and IOD ($p < 0.05$). The exposure to F3 did not affect survival up to 48 mg/L ($p > 0.05$), but it decreased the TL, TaL, SVL, and IOD of *S. tropicalis* at 48 mg/L ($p < 0.05$). However, F3 was significantly less toxic than the AEOs ($p < 0.05$). The exposure to F4 did not have any significant effect on *S. tropicalis* ($p > 0.05$). These data indicate that the acute toxicity of commercial NA extracts is mainly due to the presence of NAs. Furthermore, the fractionation of AEOs reduced its toxicity. This suggests loss of toxic compounds during fractionation or that toxicity results from the interaction between fractions.

6.2. Introduction

Naphthenic acids (NAs) are carboxylic acids naturally present in crude oil and account for up to 3% of total weight. The mixtures of NAs are commercialized and their metal salts have multiple uses including wood and textile preservatives [22]. They are generally insoluble in water (< 50 mg/L) [22], however some dissolve during the oil extraction process of the oil sands in Alberta Canada.

The water used in the oil extraction process, or oil sands process-affected water (OSPW), is toxic to aquatic organisms [188]. The OSPW contains high levels of total dissolved solids, residual bitumen, polycyclic aromatic hydrocarbons, and organic acids [189,190]. The toxicity of the OSPW is mainly attributed to its acid fraction, which contains classic and non-classic NAs [20]. Classical NAs are described by the formula $C_nH_{2n-z}O_2$, where *n* indicates the carbon number, *z* specifies the deficiency of hydrogen from the formation of rings (Fig. 1.2) [22]. Non-classical NAs includes dicarboxylic acids,

adamantane, aromatic, and heteroatom NAs [33,78]. As of 2015, a volume of 1,075 Mm^3 of OSPW was kept in tailings ponds to be decontaminated or properly disposed [154]. Therefore, it is essential to identify the most toxic compounds within OSPW to facilitate targeted mitigation.

Commercial mixtures of NAs and AEOs are toxic to aquatic organisms [88] including the frog *Silurana (Xenopus) tropicalis* (see Chapter 2-3). In Chapter 3 and 4, non-NAs components are detected in commercial mixtures of NAs and the composition of the AEOs play an important role in its toxicity [190]. In this chapter, open column chromatography and gas chromatography coupled to electron impact mass spectrometry (GC-EIMS) was used to establish the contribution of NAs to the acute toxicity of commercial mixtures of NAs. The same approach was used to fractionate AEOs with medium and high polarity. The toxicity testing was done on *S. tropicalis*.

In January 2017, I visited Dr. Jorge Molina Torres' laboratory at CINVESTAV IPN Mexico to conduct the experiments for this chapter. I conducted all these experiments under the supervision of Dr. Juan Vazquez Martinez and M. Sc. Enrique Ramirez Chavez, who helped me refine the experimental design and solving problems.

6.3. Methodology

6.3.1. Reagents and chemicals

Seven single NA standards were purchased from two suppliers (Table 6.1). The NA standards were selected to cover a variety of chemical structures, including classical and non-classical NAs, and molecular weights. A commercial mixture was purchased from Sigma-Aldrich Co. Ltd. (St. Louis, United States), with lot number BCBK0736V,

hereafter called S2. All solvents were HPLC-grade (Sigma-Aldrich Co. Ltd, St. Louis, United States). The acid-extractable organics (AEOs) from OSPW was obtained with a methodology previously reported [77]. The OSPW collection and acid extraction were described in Chapter 2. The AEO extract analyzed here is the result of the re-extraction the AEOs that was diluted in 0.1 N NaOH using the same methodology. The derivatization reagent *alpha*-bromo-2,3,4,5,6-pentafluorotoluene (pentafluorobenzyl bromide or PFBBr) was used as purchased (Sigma-Aldrich Chemie, Steinheim, Germany). A C7 to C30 alkane calibration standard (Sigma-Aldrich Chemie, Steinheim, Germany) was used to calculate the retention indices (Ri). Human chorionic gonadotropin (hCG) was purchased from Millipore (Burlington, Massachusetts, United States). The salts for the FETAX solution (NaCl, $NaHCO_3$, KCl, $CaCl_2$, $CaSO_4$, and $MgSO_4$) were purchased from Fisher Scientific (Waltham, Massachusetts, United States) and Sigma-Aldrich (St. Louis, Missouri, United States). L-cysteine was acquired from Sigma-Aldrich.

6.3.2. Open column chromatography

The chromatographic columns were prepared by loading the 44 cm × 2.3 cm glass columns with silica gel 60 (Sigma-Aldrich) suspended in cold hexane so that the column packing was ~28 cm long. Each column was flushed with 20 mL of cold hexane, then loaded with the respective extract (4 g) as a “neat oil” and it was allowed to pass through the column by gravity. Once S2 NA mixture was fully loaded, hexane was added until 150 mL was recovered from the column to collect the hexane fraction. This was followed by 100 mL of ethyl acetate and 120 mL of methanol until the entire fraction passed through the column. For the AEOs column, 125 mL of hexane was added,

followed by 100 mL of ethyl acetate and 300 mL of methanol. The fractions obtained in each of the column experiments were dried under nitrogen flux, weighted, later derivatized with PFBBr and analyzed by GC-EIMS.

6.3.3. Derivatization

The derivatization reactions were performed as we previously reported [83]. For the limits of detection and quantification, an equimolar solution of all of the standards was prepared in methanol. From this solution, a defined volume containing 1 µmol of each standard was transferred to vials and dried under nitrogen flux, to later proceed with the derivatizations. Sigma 2, AEO extract and the fraction stock solutions were prepared in methanol. For these derivatizations, an average molecular weight of 184 amu was assumed based on preliminary experiments. From these solutions, a defined volume containing 5 µmol was transferred to vials and dry under nitrogen flux to later proceed with the derivatizations. The radio analyte derivatization reagent was 1:49. 50 µL of *N*,*N*-Diisopropylethylamine (≥ 99%; Sigma-Aldrich) was added to the reaction and the reaction volume was adjusted to 200 µL with chloroform as organic solvent. The samples were then heated to 60 °C and mixed at 1000 rpm for 10 min. At the end of the reaction the solvent, the weak base, and the PFBBr excess were evaporated to dryness under nitrogen flux, the residue was re-dissolved in a mixture of chloroform/methanol 1:1 to a final concentration of 0.3 mM for each standard and a concentration of 6 µg/µL for S2, AEOs and the fractions.

6.3.4. Instrumentation and software

A gas chromatograph (Agilent Technologies model 7890A GC System) coupled with an electron impact ionization mass spectrometer (Agilent technologies model 5975

Mass Selective Detector) was used for the analysis. The data obtained by the GC-EIMS was collected with the software MassHunter Workstation version B.06.00 (Agilent Technologies, Inc.). The software Automated Mass Spectral Deconvolution and Identification System "AMDIS" (http://www.amdis.net/) was used for the determination of the mass spectrum for each component of the chromatograms of the NA standards. MSD ChemStation Data Analysis version E.02.01.1177 was used to calculate the linearity of the response of the derivatized standards. The conditions and parameters of the chromatography phase were adjusted to obtain the best mixture components resolution with the lowest interference of the noise signal. All the parameters of the MS were optimized to obtain the best signal/noise ratio of the analytes.

Pulsed splitless injection (1 µL) was used. The injector temperature was set at 280 °C. The separation of the components was performed in a capillary column Agilent DB-1ms (60 m × 250 µm × 0.25 µm) and helium used as carrier gas at a constant flow rate of 1 mL/min. The GC oven program began at an initial temperature of 50 °C, held for 2 minutes, and then was increased at a rate of 7.5 °C/min to a final temperature of 300 °C, held for 15 min. The transfer line temperature was set at 280 °C.

Electron impact mass spectra were obtained at 70 eV. Measurements were performed in SCAN mode with *m*/*z* range set to 40-550. The ion source and quadrupole analyzer temperature were 230 °C and 150 °C respectively and operated at 2.9 scans per second. A solvent delay of 12.8 min was selected to avoid damage to the MS filament, without disrupting the measurement of low molecular weight NAs.

6.3.5. GC-EIMS data processing

The Ri of the NA standards derivatives were calculated by alkane linear retention indices (C7 to C) following the methodology of Sun and Stremple [144]. Using a concentration series of mixtures of NA standards ranging from 0.0006 to 0.3 mM were analyzed in order to determine linearity, limits of detection and limits of quantification. Calibration curves were generated to assess the linearity of the GC-EIMS measurements using ChemStation. Limits of detection and quantification were calculated using AMDIS for each of the eight standards on the basis of a signal to noise ratio of 3 and 10, respectively [145].

The compound identification followed methods reported in Chapter 5. Briefly, the software Automated Mass Spectral Deconvolution and Identification System "AMDIS" (http://www.amdis.net/) was used for the determination of the retention time and extracted mass spectrum for the analytes. The generated mass spectra were compared with those of the reference compounds of the NIST Mass Spectral Search Program (NIST 02 version mass spectral database, National Institute of Standards and Technology, Washington, DC, USA). The components with a match greater than 800 and after inspection of the molecular weight, retention index, parent ion and the distribution of the 10 major ions were considered as an identified compound.

6.3.6. Animal husbandry and breeding

All the bioassays were performed at the University of Ottawa in accordance with the guidelines of the Canadian Council of Animal Care and approved by the Animal Care and Veterinary Services. Adult golden strain *S. tropicalis* frogs [79] from our breeding colony were used for bioassays. The frog colony is maintained as stated in Chapter 2. Embryos of *S. tropicalis* are obtained by inducing spawning of adult frogs by

injection of hCG into the posterior lymph sac [62]. Briefly, the frogs were injected with a priming dose of 15 IU hCG and moved to glass tanks with FETAX solution with a temperature of 21 ± 1 ºC. After 1 day, the breeding pairs were injected with a boosting dose of 200 IU hCG and placed in glass tanks with FETAX solution with a temperature of 21 ± 1 ºC. Embryos were collected and the jelly coat was removed using a 2% L-cysteine solution during 2 min exposure [76].

6.3.7. Toxicity testing

Glass Petri dishes were used as exposure vessels and the use of plastics was avoided. Plastic material has been shown to leach chemicals with estrogenic activity [81]. Solutions were prepared using glass cylinders. All glassware was washed as stated in Chapter 2. Solutions of NAs for the different treatments were prepared daily. The extracts were weighed and dissolved in water. The nominal concentrations tested for S2 and its fraction were 0, 0.5, 1, 2, and 4 mg/L. The nominal concentrations tested for AEOs and fractions were 0, 6, 12, 24, and 48 mg/L. Fifteen mL of test solution was added to 6 cm Petri dishes and ten embryos Nieuwkoop-Faber (NF) stage 9-10 were placed in each petri dish. Six glass Petri dishes containing FETAX solution were used as controls. Six replicates with ten embryos were used for each treatment. The tests were conducted in a room with a controlled temperature (27 ± 1 ºC) and 12 h light/dark cycle. Each Petri dish was observed every day, data recorded, and dead embryos were removed. At the same time, exposure solutions were replaced with newly prepared solutions. The doses were selected based on Chapter 2 data. At the end of the exposures, digital images of the embryos were taken under a light stereomicroscope (Nikon SMZ 1500), with a Nikon DS-Fi1 camera and NIS Elements version 3.22.00

software. Images were used for total length (TL), tail length (TaL), snout-vent length (SVL), and interorbital distance (IOD) measurements.

6.3.8. Data processing and statistical analysis

The morphological effects of the different mixtures on the *S. tropicalis* embryos were assessed using ImageJ™ 1.48 (http://imagej.nih.gov/ij/). The survival rate was analyzed as transformed data using the arcsine square root transformation. The Shapiro-Wilk normality and Levene's test for equal variances were performed. Data that failed the normality test or equal variance test were analyzed as transformed data using x^2 transformation. The data were analyzed using two-way ANOVA followed by Tukey's post-hoc test. The morphological effects of the AEO fractions at 48 mg/L were analyzed using one-way ANOVA followed by Tukey's post-hoc test. The alpha (α) level was set to 0.05. The statistical analysis was performed using Sigma Plot 12.0.

6.4. Results and discussion

6.4.1. GC-EIMS analysis of mixtures of standards

Figure 6.1 shows that PFBBr effectively reacted with classical and non-classical NAs. In general, the order of elution of the derivatized standards was based on molecular weight. The derivatives of 2-methylhexanoic acid and heptanoic acid, both with a molecular weight of 310, were successfully resolved. The elution of these two compounds was separated by 1.2 min (Figure 6.1). This indicates that PFB-NA isomers can be resolved by GC-EIMS. The two peaks observed at Rt 34.3 and 35.3 min in a 3:1 proportion correspond to the product of the derivatization of 1,4-Cyclohexanedicarboxylic acid with PFBBr (Fig. 6.1). This result suggests that PFBBr reacts with dicarboxylic acids, producing two isomers.

The mass spectral analysis revealed that the molecular ion $[M]^+$ was detected in all of the standards (Fig. 6.2, Table 6.2). The signature fragment ion of NAs derivatized with PFBBr at 181 *m/z*, which corresponds to the pentafluorobenzyl radical fragment, is one of the 10 most abundant fragments; its relative proportion ranged from 21 to 99%. This fragment was also detected as the base peak in 5 spectra (Fig. 6.2, Table 6.2). This variation is related to the chemical stability of the PFB-NAs. For example, the mass spectrum of PFB-1-Naphthaleneacetic acid has a base peak at 141 *m/z* that corresponds to the bycyclic structure (Fig. 6.2d). The conservation of characteristic ions may help in future compound identification. The GC-EIMS properties of the standards that were used, including Rt, Ri, EIMS spectra, are reported as a contribution to standard libraries, as some of the derivatives are not found in the NIST library. The linearity, limits of quantification, and limits of detection of the derivatization of NAs was assessed using the mixtures of the NA standards (Table 6.3). The linearity ranged between 0.3 and 0.001 mM. The limits of quantification ranged between 0.0002 and 0.0039 mM. The limits of detection ranged between 0.0001 and 0.0025 mM.

6.4.2. Fractionation of NAs

Commercial NA mixtures are partially composed of non-NAs that include aliphatic and cyclic hydrocarbons, alkylbenzenes, phenols, naphthalene and alkyl-naphthalene, and decalin compounds (see Chapters 4 and 5). An open column chromatography approach was used to separate the non-NA components of S2. From the column with S2, two fractions were obtained. The first non-polar fraction corresponded to 9% by weight of S2. The second polar fraction corresponded to 91% by weight of S2. The open column recovery was 95%. A sample of each fraction was

derivatized with PFBBr and analyzed by GC-EIMS. The fragment ion at 181 *m/z* was not detected in the non-polar fraction (Fig. 6.3 b). Linear alkanes ranging from C13 to C24 were identified in the non-polar fraction (Fig. 6.3-6.4; Table 6.4). This is in agreement with the results of Chapter 5, where C12 to C17 linear alkanes were identified in S2. The fragment ion at 181 *m/z* was detected in 97% of the peaks in the methanol fraction (Fig. 6.3 c). This indicates that the methanol fraction is composed mainly of NAs.

Open column chromatography was also used to fractionate an AEO sample by polarity. From the column, 6 fractions were recovered. The first two fractions corresponded to 0.1% of AEOs by weight. The third and fourth fraction corresponded to 84% and 15% by weight to the AEOs, respectively. The last two fractions corresponded to 1.2% by weight to the AEOs. The open column recovery was 93%. The fractions were later derivatized with PFBBr and analyzed by GC-EIMS (Fig. 6.5). The chromatogram of the AEO sample showed a hump with unresolved peaks. The fragment at 181 *m/z* ion was observed in the AEOs and in the fractions 3 and 4. This indicates that the organic compounds of the AEOs and the fraction 3 and 4 reacted with PFFBr and most of the organic compounds of the AEOs and its fractions that entered to the GC-EIMS have a labile hydrogen [132].

6.4.3. Toxicity of fractions

The S2 NA commercial mixture significantly affected survival at 4 mg/L ($p < 0.05$). The AEOs significantly affected survival at 48 mg/L ($p < 0.05$). These results indicate a higher toxicity of commercial mixtures compared to AEOs. The lower toxicity of AEOs could be due to a dilution effect as both oil and non-oil derived organic acids

are extracted from the OSPW. A two-way ANOVA (n = 6) was performed to examine the effect of the S2 fractions on the survival rate and sub-lethal effects in embryos of *S. tropicalis* after 4-day exposure (Fig. 6.6). There was no significant fraction x concentration interaction for survival rate, TL, TaL, SVL, and IOD ($p > 0.05$). There was no significant difference between the fractions on the survival rate, SVL, and IOD on *S. tropicalis* ($p > 0.05$). There was a significant difference between the fractions on TL and TaL on *S. tropicalis* ($p < 0.05$). Increasing concentrations of both the S2 and the polar fraction significantly decreased survival, TL, TaL, SVL, and IOD of *S. tropicalis* ($p < 0.05$). The exposure to S2 and its polar fraction induced abnormalities such as edema, and gut abnormalities (Fig. 6.7). These results indicate that the acute toxicity of S2 is mainly attributed to the presence of NAs. However, more research is needed since genotoxic and possible carcinogenic compounds such as xylene and naphthalene have been identified in S2 (Chapter 5).

The survival rate and sub-lethal effects on embryos of *S. tropicalis* after a 4-day exposure to the AEOs and F3 and F4 were tested (Fig. 6.8). The toxicity of the other fractions was not tested because of the limited material extracted. There was a significant fraction x concentration interaction on survival rate ($p < 0.05$). The AEOs at 48 mg/L were 100% lethal, however, the survival of *S. tropicalis* was not affected by the same concentration of the fractions compared to the control ($p > 0.05$) (Fig. 6.8a). The sub-lethal effects of AEOs and its fractions were tested in two data sets because of the 100% lethality of AEOs at 48 mg/L. The first dataset included the AEOs and its fractions until a concentration of 24 mg/L (Fig. 6.8 b-e). The second set is the sub-lethal effects of F3 and F4 on *S. tropicalis* at 48 mg/L (Fig. 6.8 f-i). There was a significant fraction x

concentration interaction on TL, TaL, and IOD ($p < 0.05$). The exposure to AEOs decreased significantly the TL, TaL, SVL, and IOD of *S. tropicalis* at 24 mg/L ($p < 0.05$) (Fig. 6.8 b-e). The fractions of the AEOs did not affect significantly the TL, TaL, and IOD of *S. tropicalis* compared to the controls up to a concentration of 24 mg/L ($p > 0.05$) (Fig. 6.8 b, c, e). However, the exposure to the F3 at 48 mg/L decreased significantly the TL, TaL, SVL, and IOD of *S. tropicalis* ($p < 0.05$) (Fig. 6.8 f-i). The exposure to F4 at 48 mg/L did not affect significantly the lethal and sub-lethal parameters here measured ($p > 0.05$). The lethal and sublethal toxicity of AEOs was not represented by F3 and F4 was not toxic. Thus, the results may indicate that some toxic compounds were lost in the column or that the toxicity of the AEOs is because of the combination of all of the fractions.

The fractionation of AEOs has been attempted previously by other research groups. In 2008, Frank et al. [191] reported on the fractionation of a methylated OSPW acid extract using vacuum distillation. They separated the compounds by molecular weight and found that the lower molecular weight fraction organic acids were more toxic than the high molecular weight fraction in *Vibrio fischeri* (Microtox assay). In 2015 and 2017, the same methodology was used by Bauer et al. [192,193] to assess the toxicity of the fractions in *Pimephales promelas* (fathead minnow) and *Oryzias latipes* (Japanese medaka) embryos. For these fish, the lower molecular weight fractions were less toxic than the high molecular weight fractions. The differences in the results are not surprising given that one test was with a bacterium and the other a teleost fish. It is clearly necessary to use an array of aquatic species to obtain a complete assessment of the toxicity of the OSPW acids and their fractions. In 2013, Scarlett et al. [38] reported

on the fractionation of the methylated acid extract of a sample of OSPW using Ag+ solid phase extraction. They obtained 10 fractions, but only tested the toxicity of two. The 96 h larval *Danio rerio* (zebrafish) exposure indicated that fraction 6, which contained aromatic compounds, was significantly more toxic than fraction 3, which contained alicyclic compounds. Herein, we report the fraction of an AEO sample by the polarity of its components. Our results may help to design strategies for the mitigation of the OSPW toxicity.

6.5. Conclusion

The derivatization reagent PFBBr reacts with classical and non-classical NAs, including linear, dicarboxylic, sulfur-containing, adamantanes, and cyclic NAs. Isomers of NAs can be resolved by GC-EIMS after derivatization with PFBBr. Open column chromatography approach separated effectively non-NA components from a Sigma commercial extract that represents at least 9% by weight. Overall, there was no significant difference between the toxicity of S2 and its polar fraction. These results indicate that the acute toxicity of the commercial NA mixture is mainly due to the presence of NAs. The AEOs decreased survival and had deleterious sublethal effects on *S. tropicalis*. One of the two fractions tested significantly reduced morphometric parameters in *S. tropicalis*. No observable effects on *S. tropicalis* were found when exposed to the second fraction. However, the toxicity of the AEOS significantly decreased when fractionated.

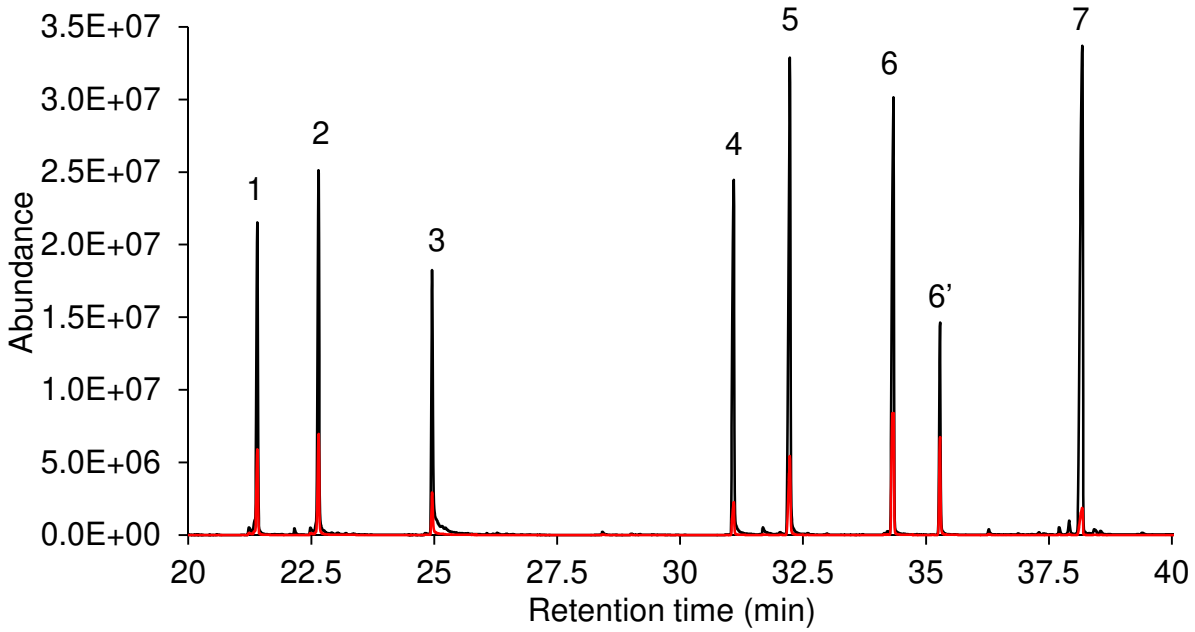


Figure 6.1. GC-EIMS TICs for a derivatized mixture of NA standards with PFBBr shown in black. GC-EIMS EIC (181 *m/z*) for a derivatized mixture of NA standards with PFBBr shown in red. Numbers correspond to NAs derivatives following the codes of Table 6.1.

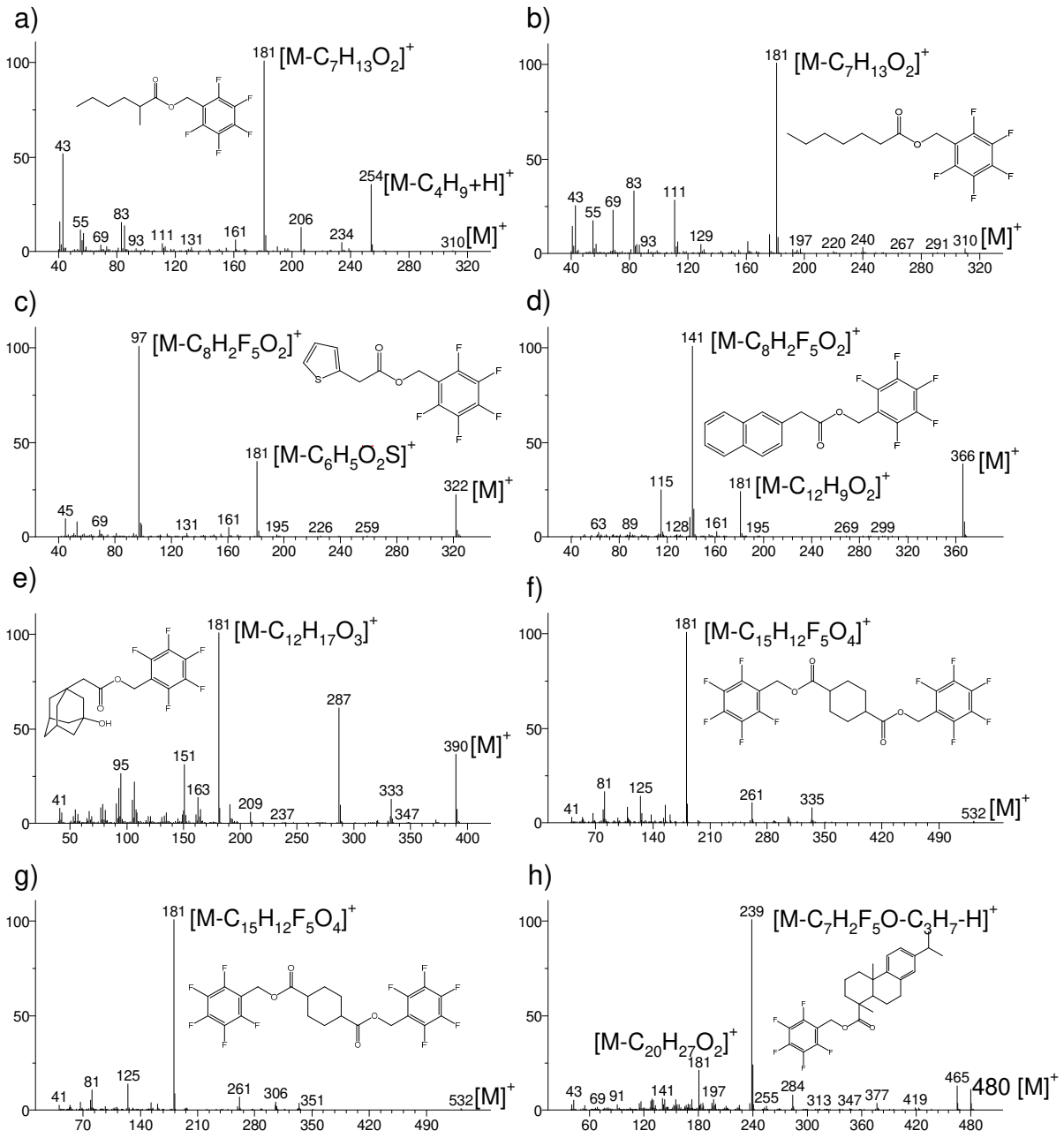


Figure 6.2. Mass spectra of 2-Methylhexanoic acid, pentafluorobenzyl ester (a), Heptanoic acid, pentafluorobenzyl ester (b), 2-Thiopheneacetic acid, pentafluorobenzyl ester (c), 1-Naphthaleneacetic acid, pentafluorobenzyl ester (d), 3-Hydroxyadamantane-1-acetic acid, pentafluorobenzyl ester (e), 1,4-Cyclohexanedicarboxylic acid, pentafluorobenzyl ester (f), Isomer of 1,4-Cyclohexanedicarboxylic acid, pentafluorobenzyl ester (g), and Dehydroabietic acid, pentafluorobenzyl ester (h) and their corresponding structures.

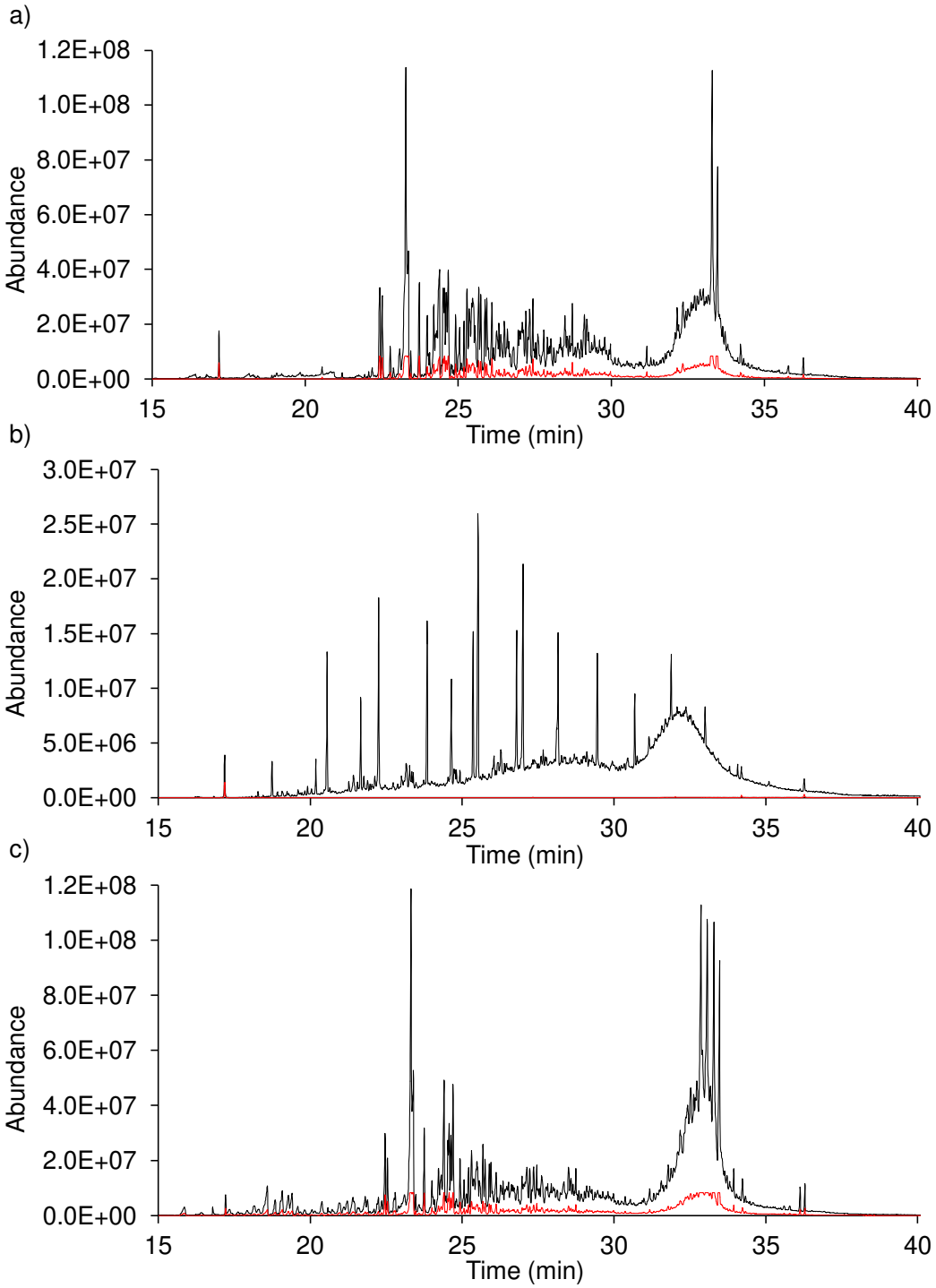


Figure 6.3. GC-EIMS TIC of S2 NA mixture (a), the non-polar fraction of S2 (b), and the polar fraction of S2 (c) obtained by open column chromatography and derivatized with PFBBr, shown in black. GC-EIMS EIC (181 *m/z*) of S2 NA mixture (a), the non-polar fraction of S2 (b), and the polar fraction of S2 obtained by open column chromatography and derivatized with PFBBr, shown in red.

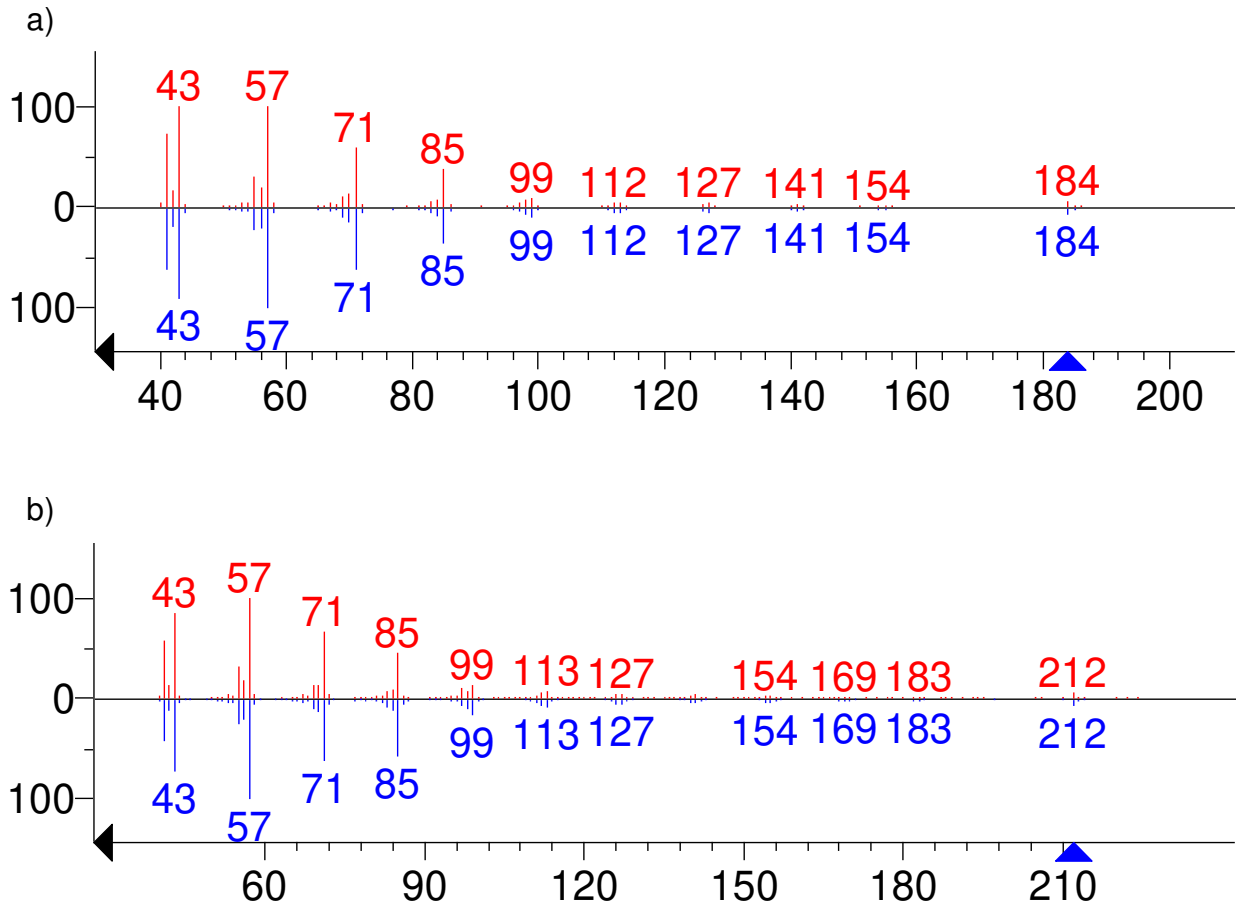


Figure 6.4. Mass spectra of a peak at 18.75 min in the non-polar fraction of S2 identified as tridecane (red spectrum) and the NIST standard of tridecane (blue spectrum) (a). Mass spectra of a peak at 22.26 min in the non-polar fraction of S2 identified as pentadecane (red spectrum) and NIST standard of pentadecane (blue spectrum) (b).

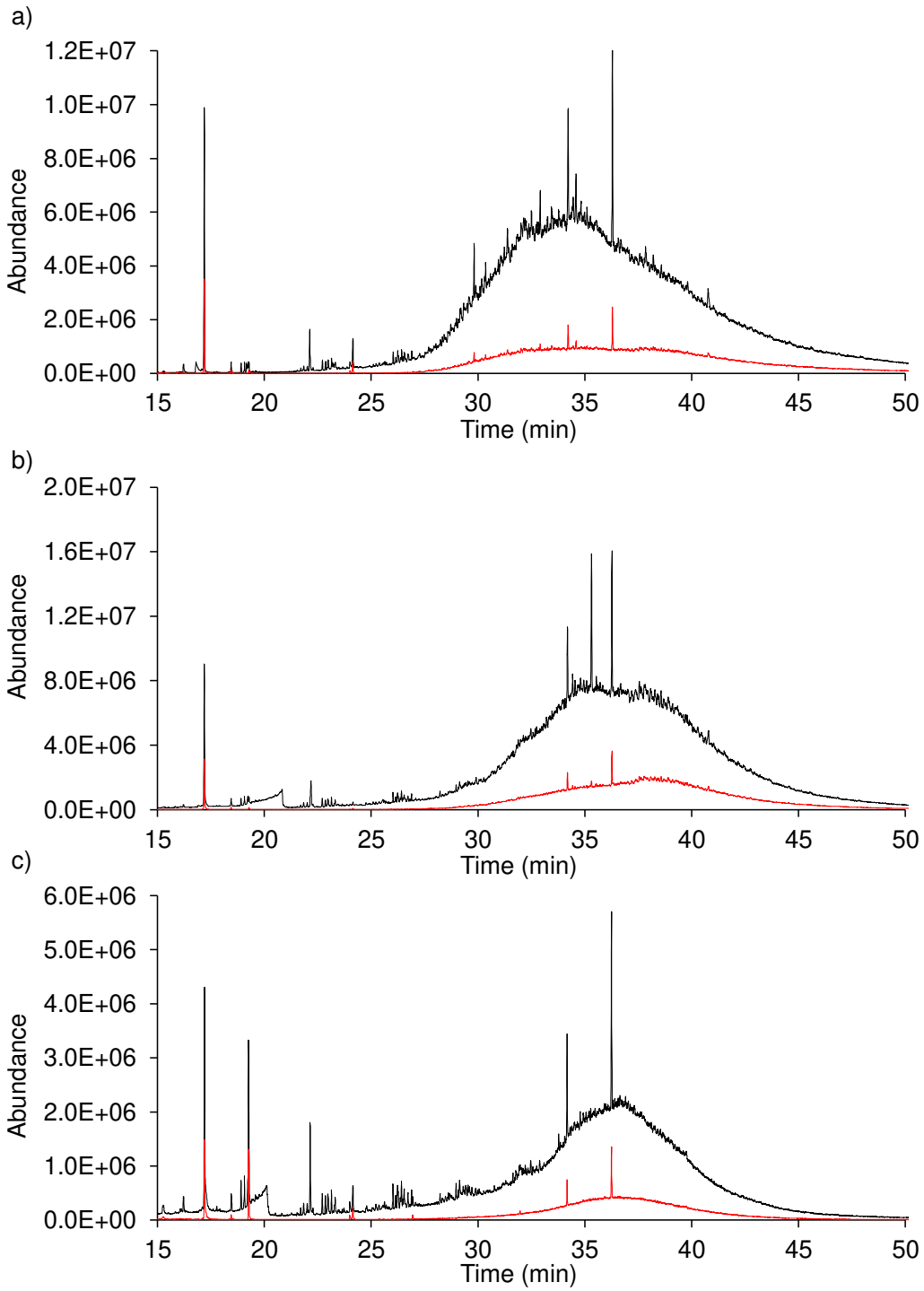


Figure 6.5. GC-EIMS TIC of an AEO mixture (a), fraction 3 (b) and fraction 4 (c) obtained by column chromatography and derivatized with PFBBr, shown in black. GC-EIMS EIC (181 *m/z*) of an AEO mixture (a), fraction 3 (b) and fraction 4 (c) obtained by column chromatography and derivatized with PFBBr, shown in red.

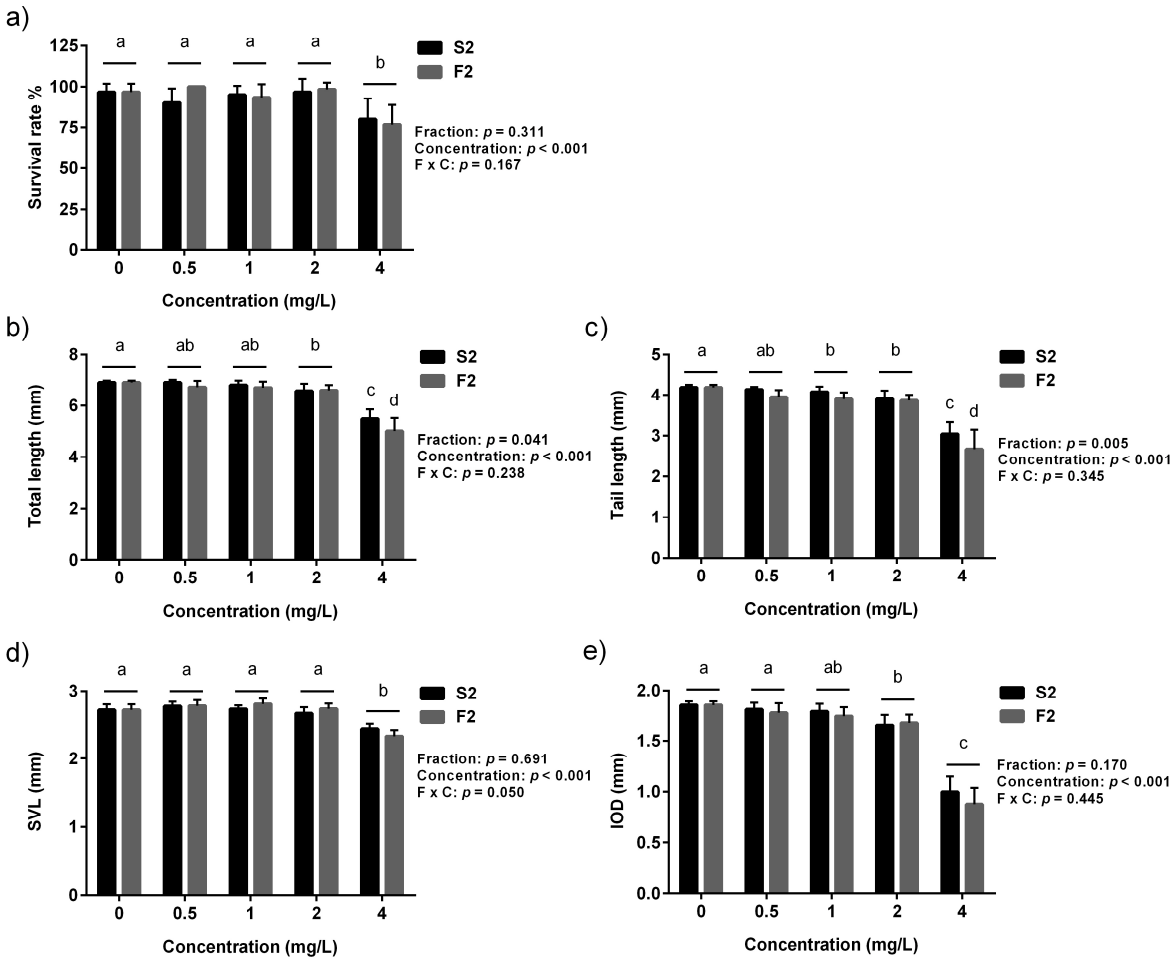


Figure 6.6. Comparison of the survival rate in *S. tropicalis* embryos exposed to S2 and the polar fraction (a). Comparison of teratogenic effects observed in *S. tropicalis* embryos exposed to S1 and the polar fraction. TL (b), SVL (c), TaL (d), and IOD (e) Different letters indicate groups that are significantly different ($p < 0.05$). Error bars represent the standard deviation of the mean, n = 6.

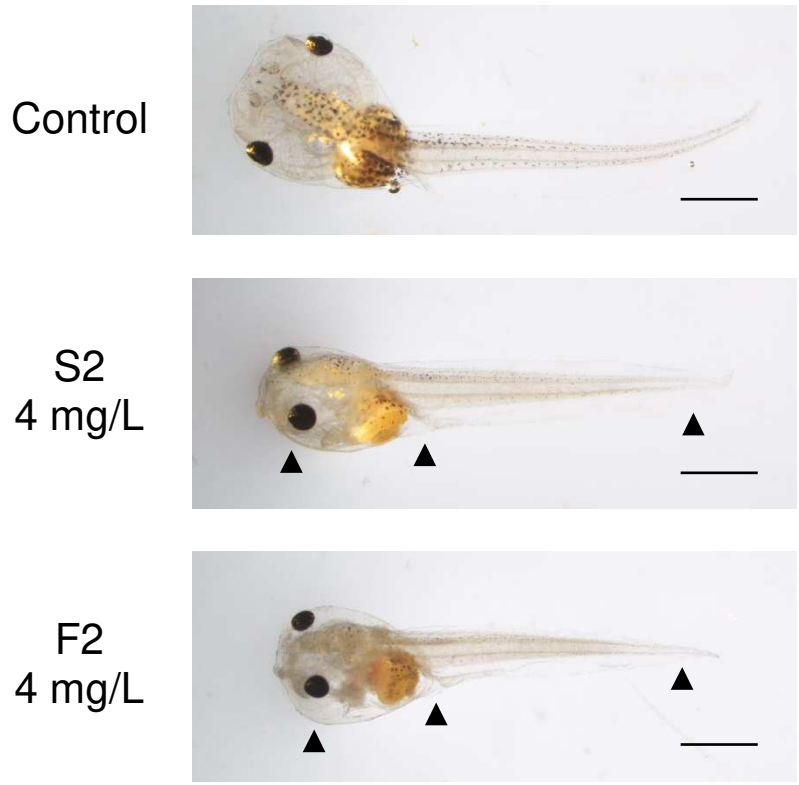


Figure 6.7. Photographs of *S. tropicalis* indicating the main teratogenic effects induced by exposure to S2 and the polar fraction of S2 (F2). Some of the observed abnormalities are indicated with arrows that include edema, size abnormalities, and gut abnormalities. Scale bar is equal to 1 mm.

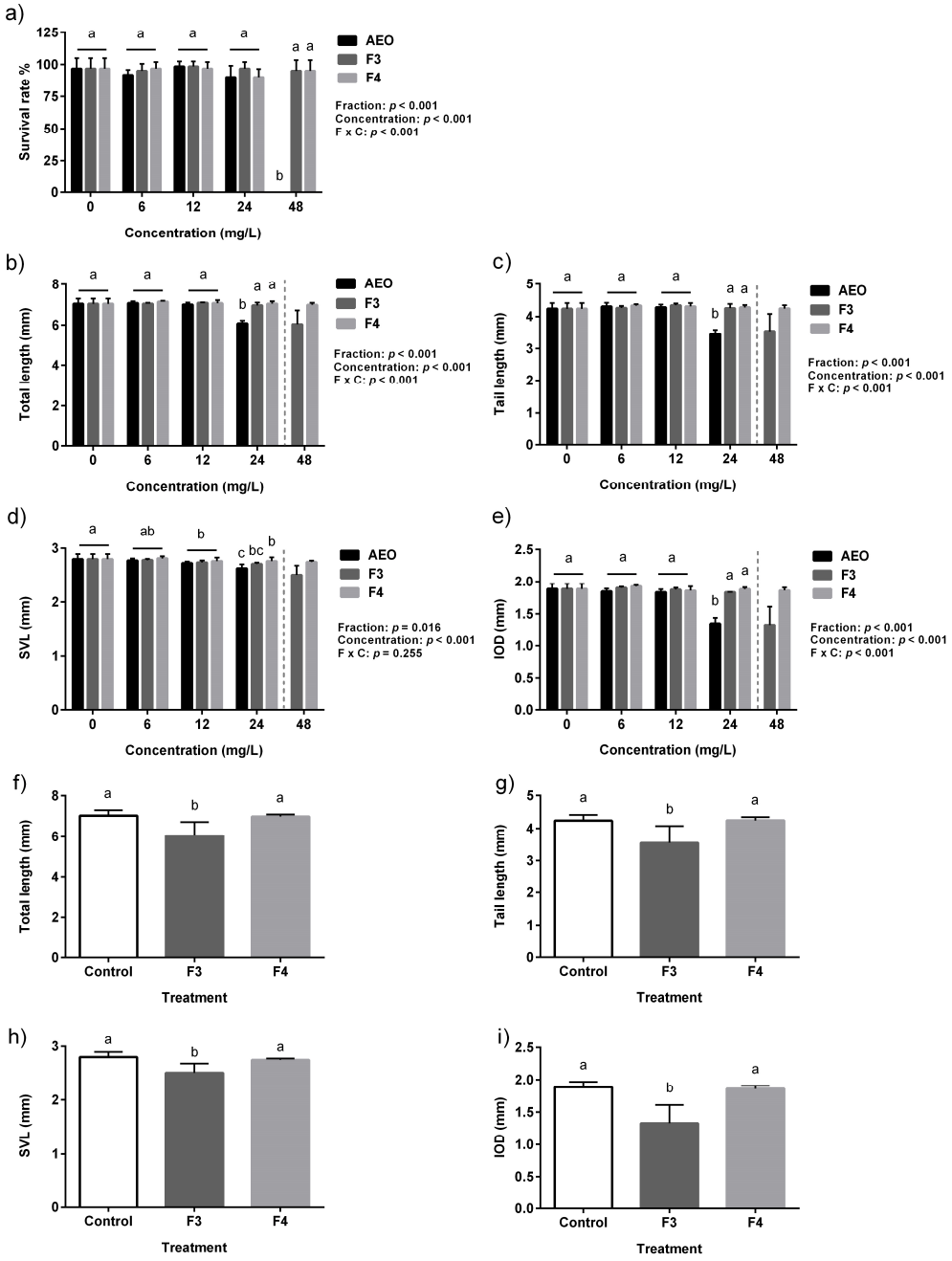


Figure 6.8. Comparison of the survival rate in *S. tropicalis* embryos exposed to AEO, F3, and F4 (a). Comparison of teratogenic effects observed in *S. tropicalis* embryos exposed to S1 and the polar fraction. TL (b), TaL (c), SVL (d), and IOD (e). Comparison of TL (f), TaL (g), SVL (h), and IOD (i) after 4 days exposure to F3 and F4 at 48 mg/L. Different letters indicate groups that are significantly different ($p < 0.05$). Error bars represent the standard deviation of the mean, n = 6.

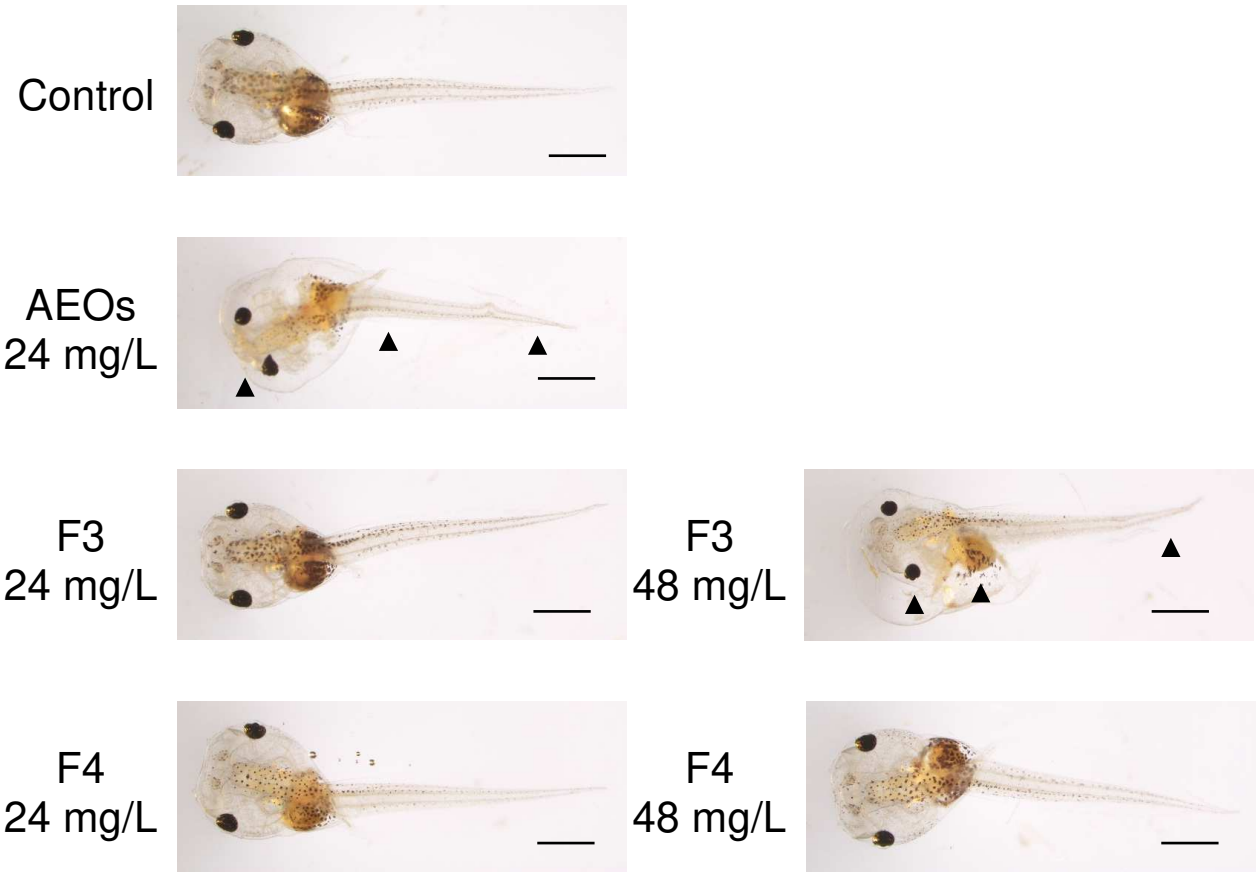


Figure 6.9. Photographs of *S. tropicalis* indicating the main teratogenic effects induced by exposure to AEOs, F3, and, F4. Scale bar is equal to 1 mm. Some of the observed abnormalities are indicated with arrows that include edema, size abnormalities, and eye abnormalities.

Table 6.1. Individual NA standards used to assess limit of quantification and detection.

Compound	Name	Molecular weight	Purity (%)	Supplier
1	2-methylhexanoic acid	130.18	99	Sigma-Aldrich
2	Heptanoic acid	130.18	≥99	Sigma-Aldrich
3	2-thiopheneacetic acid	142.18	98	Sigma-Aldrich
4	1-Naphthaleneacetic acid	186.21	≥95	Sigma-Aldrich
5	3-Hydroxyadamantane-1-acetic acid	210.27	97	Sigma-Aldrich
6	1,4-Cyclohexanedicarboxylic acid	172.18	99	Sigma-Aldrich
7	Dehydroabietic acid	300.44	≥99	CanSyn Chem. Corp.

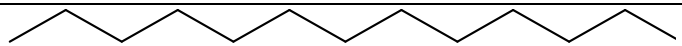
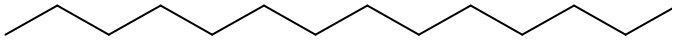
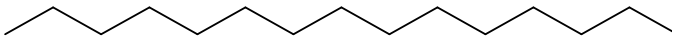
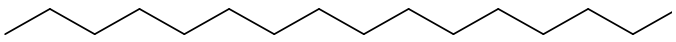
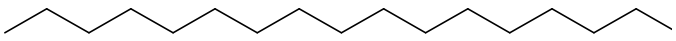
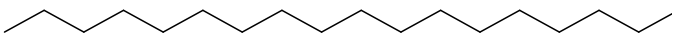
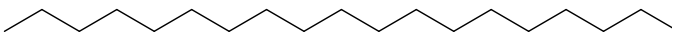
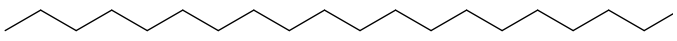
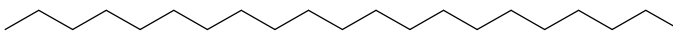
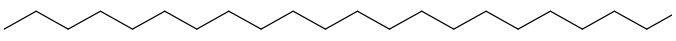
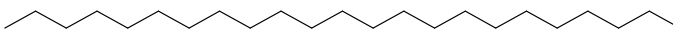
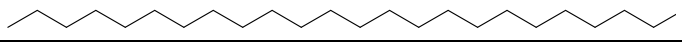
Table 6.2. Chromatographic and mass spectral properties of NA derivatives by GC-EIMS using PFBBr. Relative abundances to maximum are given in parentheses (base 100).

Analyte	Rt	Ri	Fragmentation pattern (m/z) and relative abundance (%)										M+
2-Methylhexanoic acid, pentafluorobenzyl ester	21.40	1448	254 (35.1)	206 (12.5)	182 (8.3)	181 (99.9)	85 (13.4)	83 (15.3)	57 (9.3)	55 (11.2)	43 (51.3)	41 (15.4)	310
Heptanoic acid, pentafluorobenzyl ester	22.65	1522	181 (99.9)	182 (8.3)	176 (9.9)	161 (6.3)	111 (27.9)	83 (32.7)	69 (22.5)	55 (17.0)	43 (24.9)	41 (14.2)	310
2-Thiopheneacetic acid, pentafluorobenzyl ester	24.97	1672	323 (3.3)	322 (22.2)	181 (39.5)	161 (4.9)	99 (6.3)	98 (7.2)	97 (99.9)	69 (3.7)	53 (7.7)	45 (9.4)	322
1-Naphthaleneacetic acid, pentafluorobenzyl ester	31.08	2131	367 (7.9)	366 (38.2)	181 (23.7)	161 (2.8)	142 (14.5)	141 (99.9)	139 (10.1)	116 (2.4)	115 (24.6)	63 (2.3)	366
3-Hydroxyadamantane-1-acetic acid, pentafluorobenzyl ester	32.23	2229	390 (36.0)	333 (12.5)	287 (60.4)	181 (99.9)	163 (13.3)	151 (30.7)	107 (21.5)	105 (12.10	95 (26.1)	93 918.5)	390
1,4-Cyclohexanedicarboxylic acid, pentafluorobenzyl ester	34.33	2421	335 (7.9)	261 (10.5)	182 (9.7)	181 (99.9)	155 (9.0)	125 (14.0)	109 (8.0)	81 (16.3)	79 (6.8)	67 (5.0)	532
Isomer of 1,4-Cyclohexanedicarboxylic acid, pentafluorobenzyl ester	35.29	2513	334 (3.6)	306 (4.0)	261 (6.7)	182 (8.7)	181 (99.9)	153 (3.9)	125 (13.5)	81 (10.4)	79 (5.1)	67 (4.1)	532
Dehydroabietic acid, pentafluorobenzyl ester	38.16	2774	480 (10.8)	465 (12.5)	284 (7.8)	240 (23.6)	239 (99.9)	181 (20.8)	173 (5.5)	143 (5.4)	141 (5.9)	129 (5.7)	480

Table 6.3. Limits of quantification (LOQ) and limits of detection (LOD) of the 7 standards measured by GC/EIMS using PFBBr as derivatization reagents. Linear equations and coefficients of determination (R^2) are also described.

	Analyte	Linear equation	R^2	LOQ (mM)	LOD (mM)
1	2-Methylhexanoic acid, pentafluorobenzyl ester	y = 2E+09x + 2E+06	0.992	0.0020	0.0013
2	Heptanoic acid, pentafluorobenzyl ester	y = 2E+09x - 5E+06	0.999	0.0039	0.0025
3	2-Thiopheneacetic acid, pentafluorobenzyl ester	y = 2E+09x - 1E+07	0.998	0.0038	0.0017
4	1-Naphthaleneacetic acid, pentafluorobenzyl ester	y = 2E+09x - 3E+06	0.997	0.0015	0.0009
5	3-Hydroxyadamantane-1-acetic acid, pentafluorobenzyl ester	y = 3E+09x - 683507	0.997	0.0004	0.0002
6	1,4-Cyclohexanedicarboxylic acid, pentafluorobenzyl ester	y = 3E+09x + 9E+06	0.995	0.0003	0.0001
6’	Isomer of 1,4-Cyclohexanedicarboxylic acid, pentafluorobenzyl ester	y = 1E+09x - 9E+06	0.997	0.0020	0.0006
7	Dehydroabietic acid, pentafluorobenzyl ester	y = 4E+09x + 3E+07	0.992	0.0002	0.0001

Table 6.4. List of organic compounds identified in the non-polar fraction of S2 by GC-EIMS. Rt, retention time. Ri, retention index.

No.	Rt	Ri	Detected compound	Chemical formula	Chemical structure	CAS number
1	18.75	1297	Tridecane	C13H28		629-50-5
2	20.56	1398	Tetradecane	C14H30		629-59-4
3	22.26	1498	Pentadecane	C15H32		629-62-9
4	23.85	1598	Hexadecane	C16H34		544-76-3
5	25.37	1698	Heptadecane	C17H36		629-78-7
6	26.80	1798	Octadecane	C18H38		593-45-3
7	28.16	1898	Nonadecane	C19H40		629-92-5
8	29.46	1998	Eicosane	C20H42		112-95-8
9	30.69	2098	Heneicosane	C21H44		629-94-7
10	31.89	2199	Docosane	C22H46		629-97-0
11	33.01	2298	Tricosane	C23H48		638-67-5
12	34.08	2397	Tetracosane	C24H50		646-31-1

Chapter 7

General discussion and conclusion

7.1 Thesis summary and discussion

Naphthenic acids (NAs) are oil-derived carboxylic acids that are produced in large amounts as a waste product of the oil sands extraction activities in Alberta, Canada. The NAs are found in relatively high concentrations in sediments after oil spills and their salts have several commercial applications. I have explored the toxicity of commercial NA mixtures and an acid-extractable organics (AEOs) mixture derived from an oil sands process-affected water (OSPW) sample from a tailings pond in Canada's oil sands region. The analytical methods that were developed and employed to examine NA chemical profiles are illustrated in Figure 7.1.

The exposure of *Silurana (Xenopus) tropicalis* at early stages of development to the mixtures increased mortality the incidence of teratogenic effects including reduced body size, cranial abnormalities, heart abnormalities, gut abnormality, edema, and ocular abnormalities. The rank order potency based on LC_{50} and EC_{50} with and without normalization for the quantity of O_2 species is S1 > S2 > AEO as shown in Chapter 2. As the AEOs were extracted from an OSPW sample exposed to environmental conditions in Northern Alberta, a proportion of its components are not oil derived organic acids. This may generate a dilution effect, resulting in lower toxicity compared to commercial mixtures. The mixtures of NAs have several mechanisms of toxicity including the disruption of metabolism and the cell membrane function as shown in the transcriptomic profiling results in Chapter 3. Other toxic mechanisms related to the

several abnormalities observed in the *S. tropicalis* larvae after exposure to NAs include gut function, cartilage differentiation, actin cytoskeleton, cytoskeleton organization, and edema.

The AEOs are chemically more complex compared to commercial NAs, as shown by the electrospray ionization high-resolution mass spectrometry (ESI-HRMS) analysis in Chapter 2. The commercial blends considered here are 99% composed of O_2 species, mainly of aliphatic NAs. However, in the AEOs the O_2 content is estimated to be ~ 60%, composed mainly of 3.5 and 4.5 double bond equivalences (DBE), presumably double and triple ring NAs. A derivatization method for NAs using pentafluorobenzyl bromide (PFBBr) was developed, optimized, and validated in Chapter 4. The PFB derivatives have a characteristic fragment from the ion at 181 *m*/*z* which can be used to differentiate compounds without labile hydrogens, greatly increasing the accuracy of analyzing complex NA mixtures. The GC-EIMS analysis of NA mixtures derivatized with PFBBr indicated that non-NA components are present in the mixtures identified by the lack of the fragment ion at 181 *m*/*z*. The derivatization of NAs with PFBBr offers several advantages, including increased derivatization signal yields, resolution, identification accuracy and sensitivity compared to boron trifluoride-methanol (BF_3/MeOH) or *N*-*tert*-Butyldimethylsilyl-*N*-methyltrifluoroacetamide (MTBSTFA), as shown in the results in Chapter 4. The GC-EIMS analysis corroborates that commercial extracts are mainly composed of aliphatic carboxylic acids. There is no single approach that can resolve all the organic components in a mixture. Therefore, a combination of methodologies is more suitable for the analysis of complex mixtures.

Sample processing prior GC-EIMS may lead to loss of key low molecular weight compounds. These compounds were targeted in chapter 5, where I identified more than 100 volatile components of commercial extracts and an AEO mixture using headspace solid phase microextraction coupled to GC-EIMS (HS-SPME-GC-EIMS). Most of the identified components were hydrocarbons and a few compounds with a carboxylic acid moiety. The chemometric analysis grouped the different extracts according to their VOC chromatogram. Some of the compounds identified are genotoxic or skin irritants, and some such as naphthalene are classified as possible human carcinogens. In Chapter 6, I used open column chromatography to fractionate a commercial NA blend and an AEO mixture. The non-NA components from a Sigma commercial extract represent at least 9% by weight. Overall, there was no significant difference between the toxicity of S2 and the polar fraction. Therefore, the acute toxicity of NA commercial mixtures is mainly due to the presence of NAs. The AEOs decreased survival and had deleterious effects on *S. tropicalis*. Six fractions of the AEOs were obtained, and only fraction 3 and 4 were tested due to material limitations. The fraction 3 significantly reduced morphometric parameters in developing embryos of *S. tropicalis,* but was less toxic than AEOs. No observable effects on *S. tropicalis* were found when exposed to the fraction 4.

7.2. Significance of the research

The toxicological dose descriptors and morphometric analysis herein presented may contribute to the development of environmental guidelines for NAs and AEOs. Furthermore, only a few scientific articles have reported the possible mechanism of toxicity of NAs, including narcosis [49] and oxidative stress [88]. Using transcriptional

profiling by microarray, I report several possible mechanisms of toxicity and present evidence on the disruption of metabolism and the cell membrane by AEOs and commercial extracts. Other pathways also significantly enriched were related to the edema and gut abnormalities observed in the *S. tropicalis* larvae after exposure to NAs. The chemical characterization of commercial NA mixtures, AEOs and OSPW may also contribute to mitigation strategies for OSPW. The chemical composition of mixtures of NAs have now been characterized by numerous research groups [122–127] however, very few compounds have been successfully identified [54]. I reported the use of PFBBr as a derivatization reagent for NAs and identified over 100 different compounds.

7.3. Future perspectives

Future research should focus on specific tissues as gut abnormalities were one of the most relevant phenotypic abnormalities observed during exposures and in the transcriptomic analysis. The link between gene expression and/or protein changes and such phenotype abnormalities should be addressed. The effects of NA mixtures and AEOs on adult frogs remain unknown and should be explored, particularly the potential impact on reproductive fitness. Concentrations of up to 128 mg/L are reported in the tailings ponds [30], however, much lower concentrations are recorded in the water of the Athabasca River [28]. Therefore, the effects of chronic, low dose exposures to OSPW, AEOs, and commercial extracts should be explored in *S. tropicalis*. Considering that about 40% of frog species are in decline, in part related to habitat loss, climate change and pollution [194], the effects of the oil industry in the boreal forests of Alberta should

also be investigated using local species such as wood frogs (*Lithobates sylvaticus*) in both larval and adult stages.

The complete elucidation of the chemical composition of OSPW, AEOs and NA mixtures remains a challenge. The ESI-HRMS and GC-EIMS analyses revealed hundreds of components, making AEOs particularly complex mixtures. One important step towards the analysis of AEOs is the differentiation of oil derived components from the biotic contribution to the OSPW. Research should focus on the chemical profile of OSPW at different points during oil sands and bitumen processing. Furthermore, in Chapter 5, I used a bipolar fiber coated with divinylbenzene/carboxen/poly-dimethylsiloxane (DVB/CAR/PDMS) that targets C3-C20 compounds. The use of fibers coated with polar materials such as polyacrylate and carbowax (CW) could be used to examine polar volatiles in the mixtures. Deforestation and removal of the top soil layer allow air and snow to carry oil particles [18], which may release contaminants into the environment. Therefore, the volatile and semivolatile component of contaminated snow collected around the oil sands extraction area should also be carefully analyzed.

Complementary experiments with the fractions should be considered in future research. Would the original toxicity of the AEOs or other environmental samples be recovered by mixing the fractions? Further fractionation and toxicity testing of the commercial extracts should be explored, with the specific goal of bioassay-guided identification of individual chemicals. The experiments in Chapter 6 show that the acute toxicity of the commercial blends is caused by NAs, however, this conclusion might change under a chronic exposure scenario. Therefore, lifetime exposures should be conducted to reveal the full toxic potential of NAs and OSPW-derived contaminants.

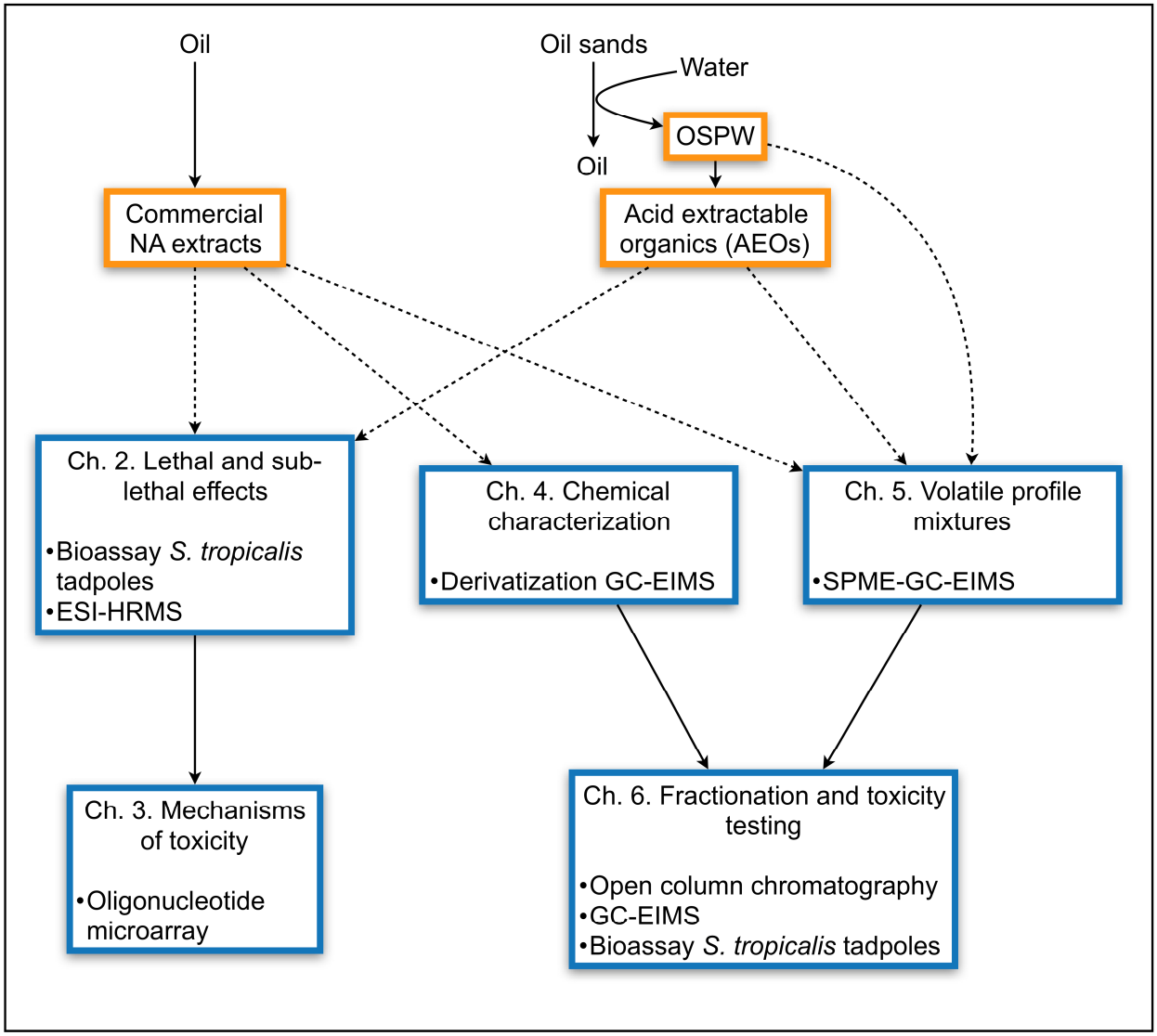


Figure 7.1. Diagram showing the research conducted on this Ph.D. project. *S. tropicalis*; *Silurana (Xenopus) tropicalis*. ESI-HRMS; electrospray ionization coupled to high-resolution mass spectrometry. GC-EIMS; gas chromatography coupled to electron impact mass spectrometry. SPME; solid phase microextraction.

References

[1] World Energy Council, World Energy Resources: 2013 Survey, (2013). http://www.worldenergy.org/wp-content/uploads/2013/09/Complete_WER_2013_Survey.pdf.

[2] United States Energy Information Administration, Canada, Internet, 2015. https://www.eia.gov/beta/international/analysis.cfm?iso=CAN.

[3] E.D. Attanasi, R.F. Meyer, Natural Bitumen and Extra-Heavy Oil, in: I. Iancu (Ed.), 2010 Surv. Energy Resour., 22nd ed., World Energy Council, 2010: pp. 123–150. http://www.worldenergy.org/documents/ser_2010_report_1.pdf.

[4] R.D. Bott, Canada's Oil Sands, Candian Cent. Energy Inf. (2011). doi:http://www.centreforenergy.com/Shopping/Product.asp?ProductKey=12.

[5] National Energy Board, Canada's Oil Sands, Opportunities and Challenges to 2015, (2004). http://www.neb.gc.ca/clf-nsi/rnrgynfmtn/nrgyrprt/lsnd/pprtntsndchllngs20152004/pprtntsndchllngs20152004-eng.pdf.

[6] U.S. Energy Information Agency, Total Petroleum and Other Liquids Production 2015, (2015). https://www.eia.gov/beta/international/rankings/#?prodact=53-1&cy=2015&pid=53&aid=1&tl_id=1-A&tl_type=a.

[7] United States Energy Information Administration, International Energy Outlook 2011, (2011). http://www.eia.gov/ieo/pdf/0484(2011).pdf.

[8] J.-C. Carrillo, E. Di Caprio, S. Barber, A.G. Sanchez, A. Hedelin, C. Kotsiki, M.L. Paumen, V. Marinov, C. Mertl, S.R. Jiménez, M. Thomas, M. Vaissière, E. Wasilewska, Hazard classification and labelling of petroleum substances in the European Economic Area - 2015, Brussels, 2015. https://www.concawe.eu/wp-

content/uploads/2017/01/rpt_15-9.pdf.

[9] Canadian Association of Petroleum Producers, Crude Oil, Forecast, Markets & Transportation, 2017. http://www.capp.ca/publications-and-statistics/publications/303440.

[10] Government of Alberta Canada, Environmental Management of Alberta's Oil Sands. Resourceful. Responsible., 2009. http://environment.gov.ab.ca/info/library/8042.pdf.

[11] Natural Resources Canada, Multiphase Systems, 2014 (2013). https://www.nrcan.gc.ca/energy/oil-sands/5871.

[12] Natural Resources Canada, Oil sands processes, 2014 (2013). https://www.nrcan.gc.ca/energy/oil-sands/5853.

[13] Syncrude, Sustainability 2012, 2014. http://syncrudesustainability.com/2012/assets/Uploads/Syncrude-Sustainability-2012.pdf.

[14] WWF Canada, WWF, Tailings, A lasting Oil Sands Legacy, WWF, Online, 2010. http://awsassets.wwf.ca/downloads/wwf_tailings_report_october_2010_final.pdf.

[15] Government of Alberta, Oil Sands: Water, (2013). http://oilsands.alberta.ca/water.html.

[16] Canadian Association of Petroleum Producers, Water use in Canada's Oil Sands, 2012. http://www.capp.ca/getdoc.aspx?Docid=193756.

[17] R.A. Frank, J.W. Roy, G. Bickerton, S.J. Rowland, J. V Headley, A.G. Scarlett, C.E. West, K.M. Peru, J.L. Parrott, F.M. Conly, L.M. Hewitt, Profiling Oil Sands Mixtures from Industrial Developments and Natural Groundwaters for Source

Identification, Environ. Sci. Technol. 48 (2014) 2660–2670. doi:10.1021/es500131k.

[18] E.N. Kelly, J.W. Short, D.W. Schindler, P. V Hodson, M. Ma, A.K. Kwan, B.L. Fortin, Oil sands development contributes polycyclic aromatic compounds to the Athabasca River and its tributaries, Proc. Natl. Acad. Sci. 106 (2009) 22346–22351. doi:10.1073/pnas.0912050106.

[19] L.L. Schramm, R.G. Smith, J.A. Stone, Interfacial chemistry of the hot water process for recovering bitumen from the Athabasca Oil Sands, AOSTRA. (1984) 5–14.

[20] M.D. MacKinnon, H. Boerger, Description of two treatment methods for detoxifying oil sands tailing pond water, Water Pollut. Reseach J. Canada. 21 (1986) 496–512.

[21] W.K. Seifert, R.M. Teeter, W.G. Howells, M.J.R. Cantow, Analysis of crude oil carboxylic acids after conversion to their corresponding hydrocarbons, Anal. Chem. 41 (1969) 1638–1647. doi:10.1021/ac60281a006.

[22] J.A. Brient, P.J. Wessner, M.N. Doyle, Naphthenic Acids, in: Kirk-Othmer Encycl. Chem. Technol., John Wiley & Sons, Inc., Hoboken, New Jersey, USA, 2000. doi:10.1002/0471238961.1401160802180905.a01.

[23] W.E. Rudzinski, L. Oehlers, Y. Zhang, B. Najera, Tandem Mass Spectrometric Characterization of Commercial Naphthenic Acids and a Maya Crude Oil, Energy & Fuels. 16 (2002) 1178–1185. doi:10.1021/ef020013t.

[24] F.M. Holowenko, M.D. MacKinnon, P.M. Fedorak, Characterization of naphthenic acids in oil sands wastewaters by gas chromatography-mass spectrometry, Water

Res. 36 (2002) 2843–2855.

[25] A.D. McNaught, A. Wilkinson, IUPAC. Compendium of Chemical Terminology (the "Gold Book"), (1997). doi:doi:10.1351/goldbook.N04085.

[26] P.Y. Bruice, Organich Chemistry, 5th ed., Prentice Hall, 2006.

[27] T.E. Havre, J. Sjöblom, J.E. Vindstad, Oil/Water-Partitioning and Interfacial Behavior of Naphthenic Acids, J. Dispers. Sci. Technol. 24 (2003) 789–801. doi:10.1081/DIS-120025547.

[28] D.M. Grewer, R.F. Young, R.M. Whittal, P.M. Fedorak, Naphthenic acids and other acid-extractables in water samples from Alberta: What is being measured?, Sci. Total Environ. 408 (2010) 5997–6010. doi:10.1016/j.scitotenv.2010.08.013.

[29] F.M. Holowenko, M.D. MacKinnon, P.M. Fedorak, Methanogens and sulfate-reducing bacteria in oil sands fine tailings waste, Can J Microbiol. 46 (2000) 927–937.

[30] A.D. McQueen, C.M. Kinley, M. Hendrikse, D.P. Gaspari, A.J. Calomeni, K.J. Iwinski, J.W. Castle, M.C. Haakensen, K.M. Peru, J. V. Headley, J.H. Rodgers, A risk-based approach for identifying constituents of concern in oil sands process-affected water from the Athabasca Oil Sands region, Chemosphere. 173 (2017) 340–350. doi:10.1016/j.chemosphere.2017.01.072.

[31] J.W.S. Lai, L.J. Pinto, L.I. Bendell-Young, M.M. Moore, E. Kiehlmann, Factors that affect the degradation of naphthenic acids in oil sands wastewater by indigenous microbial communities, Environ. Toxicol. Chem. 15 (1996) 1482–1491. doi:10.1002/etc.5620150909.

[32] D.W. McMartin, J. V Headley, D.A. Friesen, K.M. Peru, J.A. Gillies, Photolysis of

naphthenic acids in natural surface water, J Env. Sci Heal. A Tox Hazard Subst Env. Eng. 39 (2004) 1361–1383.

[33] S.J. Rowland, A.G. Scarlett, D. Jones, C.E. West, R.A. Frank, Diamonds in the rough: identification of individual naphthenic acids in oil sands process water, Env. Sci Technol. 45 (2011) 3154–3159. doi:10.1021/es103721b.

[34] D. Jones, C.E. West, A.G. Scarlett, R.A. Frank, S.J. Rowland, Isolation and estimation of the “aromatic” naphthenic acid content of an oil sands process-affected water extract, J Chromatogr A. 1247 (2012) 171–175. doi:10.1016/j.chroma.2012.05.073.

[35] R. Orrego, J. Guchardi, V. Hernandez, R. Krause, L. Roti, J. Armour, M. Ganeshakumar, D. Holdway, Pulp and paper mill effluent treatments have differential endocrine-disrupting effects on rainbow trout, Env. Toxicol Chem. 28 (2009) 181–188. doi:10.1897/08-191.1.

[36] S.J. Rowland, C.E. West, A.G. Scarlett, D. Jones, R.A. Frank, Identification of individual tetra- and pentacyclic naphthenic acids in oil sands process water by comprehensive two-dimensional gas chromatography/mass spectrometry, Rapid Commun Mass Spectrom. 25 (2011) 1198–1204. doi:10.1002/rcm.4977.

[37] J. V Headley, K.M. Peru, D.W. McMartin, M. Winkler, Determination of dissolved naphthenic acids in natural waters by using negative-ion electrospray mass spectrometry, J AOAC Int. 85 (2002) 182–187.

[38] A.G. Scarlett, H.C. Reinardy, T.B. Henry, C.E. West, R.A. Frank, L.M. Hewitt, S.J. Rowland, Acute toxicity of aromatic and non-aromatic fractions of naphthenic acids extracted from oil sands process-affected water to larval zebrafish,

Chemosphere. 93 (2013) 415–420. doi:10.1016/j.chemosphere.2013.05.020.

[39] C.S. Ho, C.W. Lam, M.H. Chan, R.C. Cheung, L.K. Law, L.C. Lit, K.F. Ng, M.W. Suen, H.L. Tai, Electrospray ionisation mass spectrometry: principles and clinical applications, Clin Biochem Rev. 24 (2003) 3–12.

[40] F.A. Settle, Handbook of instrumental techniques for analytical chemistry, Prentice Hall PTR, 1997. http://books.google.ca/books?id=iDVRAAAAMAAJ.

[41] Q. Hu, R.J. Noll, H. Li, A. Makarov, M. Hardman, R. Graham Cooks, The Orbitrap: a new mass spectrometer, J Mass Spectrom. 40 (2005) 430–443. doi:10.1002/jms.856.

[42] J. V Headley, K.M. Peru, A. Janfada, B. Fahlman, C. Gu, S. Hassan, Characterization of oil sands acids in plant tissue using Orbitrap ultra-high resolution mass spectrometry with electrospray ionization, Rapid Commun Mass Spectrom. 25 (2011) 459–462. doi:10.1002/rcm.4877.

[43] M.C. McMaster, Getting Started in GC/MS, in: GC/MS, John Wiley & Sons, Inc., 2007: pp. 47–55. doi:10.1002/9780470228357.ch5.

[44] D.M. Jones, J.S. Watson, W. Meredith, M. Chen, B. Bennett, Determination of naphthenic acids in crude oils using nonaqueous ion exchange solid-phase extraction, Anal Chem. 73 (2001) 703–707.

[45] M. Merlin, S.E. Guigard, P.M. Fedorak, Detecting naphthenic acids in waters by gas chromatography-mass spectrometry, J Chromatogr A. 1140 (2007) 225–229. doi:10.1016/j.chroma.2006.11.089.

[46] R.F. Young, D.L. Coy, P.M. Fedorak, Evaluating MTBSTFA derivatization reagents for measuring naphthenic acids by gas chromatography-mass

spectrometry, Anal. Methods. 2 (2010) 765–770. doi:10.1039/C0AY00014K.

[47] R.A. Heintz, S.D. Rice, A.C. Wertheimer, R.F. Bradshaw, F.P. Thrower, J.E. Joyce, J.W. Short, Delayed effects on growth and marine survival of pink salmon Oncorhynchus gorbuscha after exposure to crude oil during embryonic development, Mar. Ecol. Prog. Ser. 208 (2000) 205–216. doi:10.3354/meps208205.

[48] G.D. Marty, D.E. Hinton, J.W. Short, R.A. Heintz, S.D. Rice, D.M. Dambach, N.H. Willits, J.J. Stegeman, Ascites, premature emergence, increased gonadal cell apoptosis, and cytochrome P4501A induction in pink salmon larvae continuously exposed to oil-contaminated gravel during development, Can. J. Zool. 75 (1997) 989–1007. doi:10.1139/z97-120.

[49] R.A. Frank, K. Fischer, R. Kavanagh, B.K. Burnison, G. Arsenault, J. V Headley, K.M. Peru, G. Van Der Kraak, K.R. Solomon, Effect of Carboxylic Acid Content on the Acute Toxicity of Oil Sands Naphthenic Acids, Env. Sci Technol. 43 (2008) 266–271. doi:10.1021/es8021057.

[50] U.M. Cowgill, L.R. Williams, A.C.E.-47 on B. Effects, E. Fate, A.C.E.-47 on B. Effects, E.F.S.E. 0. on A. Toxicology, Aquatic Toxicology and Hazard Assessment: 12th volume, ASTM, 1989. doi:https://doi.org/10.1520/STP1027-EB.

[51] L.N. Britton, Toxicity and Persistence of Surfactants Used in the Petroleum IndustrySurfactants, Cambridge University Press, 2000.

[52] S.A. McNeill, C.J. Arens, N.S. Hogan, B. Kollner, M.R. van den Heuvel, Immunological impacts of oil sands-affected waters on rainbow trout evaluated using an in situ exposure, Ecotoxicol Env. Saf. 84 (2012) 254–261.

doi:10.1016/j.ecoenv.2012.07.016.

[53] L.A. Leclair, G.Z. MacDonald, L.J. Phalen, B. Kollner, N.S. Hogan, M.R. van den Heuvel, The immunological effects of oil sands surface waters and naphthenic acids on rainbow trout (Oncorhynchus mykiss), Aquat Toxicol. 142–143 (2013) 185–194. doi:10.1016/j.aquatox.2013.08.009.

[54] S.J. Rowland, C.E. West, D. Jones, A.G. Scarlett, R.A. Frank, L.M. Hewitt, Steroidal aromatic “naphthenic acids” in oil sands process-affected water: structural comparisons with environmental estrogens, Env. Sci Technol. 45 (2011) 9806–9815. doi:10.1021/es202606d.

[55] R.J. Kavanagh, R.A. Frank, K.R. Solomon, G. Van Der Kraak, Reproductive and health assessment of fathead minnows (Pimephales promelas) inhabiting a pond containing oil sands process-affected water, Aquat Toxicol. 130–131 (2013) 201–209. doi:10.1016/j.aquatox.2013.01.007.

[56] B.D. Hersikorn, J.E.G. Smits, Compromised metamorphosis and thyroid hormone changes in wood frogs (Lithobates sylvaticus) raised on reclaimed wetlands on the Athabasca oil sands, Environ. Pollut. 159 (2011) 596–601. doi:10.1016/j.envpol.2010.10.005.

[57] R.H. McKee, C.M. North, P. Podhasky, J.H. Charlap, A. Kuhl, Toxicological assessment of refined naphthenic acids in a repeated dose/developmental toxicity screening test, Int J Toxicol. 33 (2014) 168s–180s. doi:10.1177/1091581813504229.

[58] F.H. Pough, Amphibian biology and husbandry., ILAR J. 48 (2007) 203–213.

[59] D.R. Frost, Amphibian Species of the World: an Online Reference. Version 6.0,

(2017). http://research.amnh.org/herpetology/amphibia/index.html (accessed November 6, 2017).

[60] W.A. Hopkins, Amphibians as models for studying environmental change., ILAR J. 48 (2007) 270–277.

[61] M. Schmid, B.J. Evans, J.P. Bogart, Polyploidy in Amphibia., Cytogenet. Genome Res. 145 (2015) 315–330. doi:10.1159/000431388.

[62] V.S. Langlois, P. Duarte-Guterman, S. Ing, B.D. Pauli, G.M. Cooke, V.L. Trudeau, Fadrozole and finasteride exposures modulate sex steroid- and thyroid hormone-related gene expression in Silurana (Xenopus) tropicalis early larval development, Gen. Comp. Endocrinol. 166 (2010) 417–427. doi:http://dx.doi.org/10.1016/j.ygcen.2009.11.004.

[63] U. Hellsten, R.M. Harland, M.J. Gilchrist, D. Hendrix, J. Jurka, V. Kapitonov, I. Ovcharenko, N.H. Putnam, S. Shu, L. Taher, I.L. Blitz, B. Blumberg, D.S. Dichmann, I. Dubchak, E. Amaya, J.C. Detter, R. Fletcher, D.S. Gerhard, D. Goodstein, T. Graves, I. V Grigoriev, J. Grimwood, T. Kawashima, E. Lindquist, S.M. Lucas, P.E. Mead, T. Mitros, H. Ogino, Y. Ohta, A. V Poliakov, N. Pollet, J. Robert, A. Salamov, A.K. Sater, J. Schmutz, A. Terry, P.D. Vize, W.C. Warren, D. Wells, A. Wills, R.K. Wilson, L.B. Zimmerman, A.M. Zorn, R. Grainger, T. Grammer, M.K. Khokha, P.M. Richardson, D.S. Rokhsar, The genome of the Western clawed frog Xenopus tropicalis., Science (80-.). 328 (2010) 633–6. doi:10.1126/science.1183670.

[64] S.D. Richardson, T.A. Ternes, Water Analysis: Emerging Contaminants and Current Issues., Anal. Chem. 90 (2018) 398–428.

doi:10.1021/acs.analchem.7b04577.

[65] M.K. McNutt, R. Camilli, T.J. Crone, G.D. Guthrie, P.A. Hsieh, T.B. Ryerson, O. Savas, F. Shaffer, Review of flow rate estimates of the Deepwater Horizon oil spill, Proc Natl Acad Sci U S A. 109 (2012) 20260–20267. doi:10.1073/pnas.1112139108.

[66] C. Aeppli, C.A. Carmichael, R.K. Nelson, K.L. Lemkau, W.M. Graham, M.C. Redmond, D.L. Valentine, C.M. Reddy, Oil weathering after the Deepwater Horizon disaster led to the formation of oxygenated residues, Env. Sci Technol. 46 (2012) 8799–8807. doi:10.1021/es3015138.

[67] Y. Wan, B. Wang, J.S. Khim, S. Hong, W.J. Shim, J. Hu, Naphthenic acids in coastal sediments after the Hebei Spirit oil spill: a potential indicator for oil contamination, Env. Sci Technol. 48 (2014) 4153–4162. doi:10.1021/es405034y.

[68] Government of Alberta, Alberta Oil Sands Industry, Quarterly Update, Spring 2016, (2016).

[69] CAPP, What are the Oil Sands?, 2016 (2015). http://www.capp.ca/canadian-oil-and-natural-gas/oil-sands/what-are-oil-sands.

[70] Government of Alberta, Oil Sands: Tailings, (2013). http://oilsands.alberta.ca/tailings.html.

[71] J. Kurek, J.L. Kirk, D.C.G. Muir, X. Wang, M.S. Evans, J.P. Smol, Legacy of a half century of Athabasca oil sands development recorded by lake ecosystems, Proc. Natl. Acad. Sci. 110 (2013) 1761–1766. doi:10.1073/pnas.1217675110.

[72] C.S. Rosenfeld, N.D. Denslow, E.F. Orlando, J.M. Gutierrez-Villagomez, V.L. Trudeau, Neuroendocrine disruption of organizational and activational hormone

programming in poikilothermic vertebrates, J. Toxicol. Environ. Heal. Part B. 20 (2017) 276–304. doi:10.1080/10937404.2017.1370083.

[73] E.J. Crespi, R.J. Denver, Ancient origins of human developmental plasticity., Am. J. Hum. Biol. 17 (2005) 44–54. doi:10.1002/ajhb.20098.

[74] M. Vu, L. Navarro-Martín, J.M. Gutierrez-Villagomez, V.L. Trudeau, Development of an in vitro Ovary Culture System to Evaluate Endocrine Disruption in Wood Frog Tadpoles, J. Toxicol. Environ. Heal. Part A. 78 (2015) 1137–1141. doi:10.1080/15287394.2015.1074970.

[75] J.E.G. Smits, B.D. Hersikorn, R.F. Young, P.M. Fedorak, Physiological effects and tissue residues from exposure of leopard frogs to commercial naphthenic acids, 2012. doi:10.1016/j.scitotenv.2012.07.043.

[76] Standard Guide for Conducting the Frog Embryo Teratogenesis Assay-Xenopus (FETAX), (2012).

[77] V. V. Rogers, K. Liber, M.D. MacKinnon, Isolation and characterization of naphthenic acids from Athabasca oil sands tailings pond water, Chemosphere. 48 (2002) 519–527. doi:10.1016/S0045-6535(02)00133-9.

[78] M.P. Barrow, M. Witt, J. V Headley, K.M. Peru, Athabasca Oil Sands Process Water: Characterization by Atmospheric Pressure Photoionization and Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry, Anal Chem. 82 (2010) 3727–3735. doi:10.1021/ac100103y.

[79] A.W. Olmstead, A. Lindberg-Livingston, S.J. Degitz, Genotyping sex in the amphibian, Xenopus (Silurana) tropicalis, for endocrine disruptor bioassays, Aquat. Toxicol. 98 (2010) 60–66. doi:10.1016/j.aquatox.2010.01.012.

[80] S.D. Melvin, V.L. Trudeau, Growth, development and incidence of deformities in amphibian larvae exposed as embryos to naphthenic acid concentrations detected in the Canadian oil sands region, Environ. Pollut. 167 (2012) 178–183. doi:http://dx.doi.org/10.1016/j.envpol.2012.04.002.

[81] C.Z. Yang, S.I. Yaniger, V.C. Jordan, D.J. Klein, G.D. Bittner, Most Plastic Products Release Estrogenic Chemicals: A Potential Health Problem That Can Be Solved, Environ. Health Perspect. 119 (2011) 989–996. doi:10.1289/ehp.1003220.

[82] J.A. Bantle, Atlas of Abnormalities: A Guide for the Performance of FETAX, Printing Services, Oklahoma State University, 1991. https://books.google.ca/books?id=Lgj_QwAACAAJ.

[83] J.M. Gutierrez-Villagomez, J. Vázquez-Martínez, E. Ramírez-Chávez, J. Molina-Torres, V.L. Trudeau, Analysis of naphthenic acid mixtures as pentafluorobenzyl derivatives by gas chromatography-electron impact mass spectrometry, Talanta. 162 (2017) 440–452. doi:http://dx.doi.org/10.1016/j.talanta.2016.10.057.

[84] J.R. Marentette, R.A. Frank, A.J. Bartlett, P.L. Gillis, L.M. Hewitt, K.M. Peru, J. V. Headley, P. Brunswick, D. Shang, J.L. Parrott, Toxicity of naphthenic acid fraction components extracted from fresh and aged oil sands process-affected waters, and commercial naphthenic acid mixtures, to fathead minnow (Pimephales promelas) embryos, Aquat. Toxicol. 164 (2015) 108–117. doi:10.1016/j.aquatox.2015.04.024.

[85] A.G. Marshall, R.P. Rodgers, Petroleomics: Chemistry of the underworld, Proc. Natl. Acad. Sci. . 105 (2008) 18090–18095. doi:10.1073/pnas.0805069105.

[86] E. Bae, I.J. Yeo, B. Jeong, Y. Shin, K.-H. Shin, S. Kim, Study of Double Bond Equivalents and the Numbers of Carbon and Oxygen Atom Distribution of Dissolved Organic Matter with Negative-Mode FT-ICR MS, Anal. Chem. 83 (2011) 4193–4199. doi:10.1021/ac200464q.

[87] J. Wang, X. Cao, Y. Huang, X. Tang, Developmental toxicity and endocrine disruption of naphthenic acids on the early life stage of zebrafish (*Danio rerio*), J. Appl. Toxicol. 35 (2015) 1493–1501. doi:10.1002/jat.3166.

[88] J.R. Marentette, K. Sarty, A.M. Cowie, R.A. Frank, L.M. Hewitt, J.L. Parrott, C.J. Martyniuk, Molecular responses of Walleye (Sander vitreus) embryos to naphthenic acid fraction components extracted from fresh oil sands process-affected water., Aquat. Toxicol. 182 (2017) 11–19. doi:10.1016/j.aquatox.2016.11.003.

[89] P. Brunswick, L.M. Hewitt, R.A. Frank, M. Kim, G. van Aggelen, D. Shang, A traceable reference for direct comparative assessment of total naphthenic acid concentrations in commercial and acid extractable organic mixtures derived from oil sands process water, J. Environ. Sci. Heal. Part A. 52 (2017) 274–280. doi:10.1080/10934529.2016.1253399.

[90] M.B. Woudneh, M. Coreen Hamilton, J.P. Benskin, G. Wang, P. McEachern, J.R. Cosgrove, A novel derivatization-based liquid chromatography tandem mass spectrometry method for quantitative characterization of naphthenic acid isomer profiles in environmental waters, J. Chromatogr. A. 1293 (2013) 36–43. doi:http://dx.doi.org/10.1016/j.chroma.2013.03.040.

[91] G.D. Morandi, S.B. Wiseman, A. Pereira, R. Mankidy, I.G.M. Gault, J.W. Martin,

J.P. Giesy, Effects-Directed Analysis of Dissolved Organic Compounds in Oil Sands Process-Affected Water, Environ. Sci. Technol. 49 (2015) 12395–12404. doi:10.1021/acs.est.5b02586.

[92] R. Huang, Y. Chen, M.N.A. Meshref, P. Chelme-Ayala, S. Dong, M.D. Ibrahim, C. Wang, N. Klamerth, S.A. Hughes, J. V Headley, K.M. Peru, C. Brown, A. Mahaffey, M. Gamal El-Din, Characterization and determination of naphthenic acids species in oil sands process-affected water and groundwater from oil sands development area of Alberta, Canada., Water Res. 128 (2017) 129–137. doi:10.1016/j.watres.2017.10.003.

[93] Government of Alberta, Oil Sands Facts and Stats, 2017. https://open.alberta.ca/dataset/b6f2d99e-30f8-4194-b7eb-76039e9be4d2/resource/063e27cc-b6d1-4dae-8356-44e27304ef78/download/FSOilSands.pdf.

[94] Government of Canada, Oil Sands: A strategic resource for Canada, North America and the global market, 2013. http://www.nrcan.gc.ca/sites/www.nrcan.gc.ca/files/energy/pdf/eneene/pubpub/pdf/OS-brochure-eng.pdf.

[95] Natural Resources Canada, Oil sands: Land use and reclamation, 2016. http://www.nrcan.gc.ca/energy/publications/18740.

[96] P.D. Nieuwkoop, J. Faber, Normal Table of Xenopus Laevis (Daudin): A Systematical and Chronological Survey of the Development from the Fertilized Egg Till the End of Metamorphosis, Garland Pub., 1994. https://books.google.ca/books?id=lBthQgAACAAJ.

[97] V.S. Langlois, C.J. Martyniuk, Genome wide analysis of Silurana (Xenopus) tropicalis development reveals dynamic expression using network enrichment analysis, Mech. Dev. 130 (2013) 304–322. doi:https://doi.org/10.1016/j.mod.2012.12.002.

[98] S.-Y. Kim, D.J. Volsky, PAGE: Parametric Analysis of Gene Set Enrichment, BMC Bioinformatics. 6 (2005) 144. doi:10.1186/1471-2105-6-144.

[99] P.S. Aranda, D.M. LaJoie, C.L. Jorcyk, Bleach gel: a simple agarose gel for analyzing RNA quality., Electrophoresis. 33 (2012) 366–369. doi:10.1002/elps.201100335.

[100] S. Dhorne-Pollet, A. Thelie, N. Pollet, Validation of novel reference genes for RT-qPCR studies of gene expression in Xenopus tropicalis during embryonic and post-embryonic development., Dev. Dyn. 242 (2013) 709–717. doi:10.1002/dvdy.23972.

[101] H.J. Motulsky, R.E. Brown, Detecting outliers when fitting data with nonlinear regression – a new method based on robust nonlinear regression and the false discovery rate, BMC Bioinformatics. 7 (2006) 123. doi:10.1186/1471-2105-7-123.

[102] L.-H. Heckmann, P.B. Sørensen, P.H. Krogh, J.G. Sørensen, NORMA-Gene: A simple and robust method for qPCR normalization based on target gene data, BMC Bioinformatics. 12 (2011) 250. doi:10.1186/1471-2105-12-250.

[103] M.C. Newman, W.H. Clements, Ecotoxicology: A Comprehensive Treatment, CRC Press, 2007. https://books.google.ca/books?id=y11sdkzQLKkC.

[104] B.R. Baer, A.E. Rettie, CYP4B1: an enigmatic P450 at the interface between xenobiotic and endobiotic metabolism., Drug Metab. Rev. 38 (2006) 451–476.

doi:10.1080/03602530600688503.

[105] R. Uauy, D.R. Hoffman, P. Peirano, D.G. Birch, E.E. Birch, Essential fatty acids in visual and brain development., Lipids. 36 (2001) 885–895.

[106] R. Allikmets, L.M. Schriml, A. Hutchinson, V. Romano-Spica, M. Dean, A human placenta-specific ATP-binding cassette gene (ABCP) on chromosome 4q22 that is involved in multidrug resistance., Cancer Res. 58 (1998) 5337–5339.

[107] M.N. Chernova, L. Jiang, D.J. Friedman, R.B. Darman, H. Lohi, J. Kere, D.H. Vandorpe, S.L. Alper, Functional comparison of mouse slc26a6 anion exchanger with human SLC26A6 polypeptide variants: differences in anion selectivity, regulation, and electrogenicity., J. Biol. Chem. 280 (2005) 8564–8580. doi:10.1074/jbc.M411703200.

[108] R.W. Freel, M. Morozumi, M. Hatch, Parsing apical oxalate exchange in Caco-2BBe1 monolayers: siRNA knockdown of SLC26A6 reveals the role and properties of PAT-1., Am. J. Physiol. Gastrointest. Liver Physiol. 297 (2009) G918-29. doi:10.1152/ajpgi.00251.2009.

[109] Z. Jiang, J.R. Asplin, A.P. Evan, V.M. Rajendran, H. Velazquez, T.P. Nottoli, H.J. Binder, P.S. Aronson, Calcium oxalate urolithiasis in mice lacking anion transporter Slc26a6., Nat. Genet. 38 (2006) 474–478. doi:10.1038/ng1762.

[110] S.K. Young, T.D. Baird, R.C. Wek, Translation Regulation of the Glutamyl-prolyl-tRNA Synthetase Gene EPRS through Bypass of Upstream Open Reading Frames with Noncanonical Initiation Codons., J. Biol. Chem. 291 (2016) 10824–10835. doi:10.1074/jbc.M116.722256.

[111] E.M. Wright, D.D. Loo, M. Panayotova-Heiermann, M.P. Lostao, B.H. Hirayama,

B. Mackenzie, K. Boorer, G. Zampighi, “Active” sugar transport in eukaryotes., J. Exp. Biol. 196 (1994) 197–212.

[112] W.S. Lee, Y. Kanai, R.G. Wells, M.A. Hediger, The high affinity Na+/glucose cotransporter. Re-evaluation of function and distribution of expression., J. Biol. Chem. 269 (1994) 12032–12039.

[113] E. Turk, B. Zabel, S. Mundlos, J. Dyer, E.M. Wright, Glucose/galactose malabsorption caused by a defect in the Na+/glucose cotransporter., Nature. 350 (1991) 354–356. doi:10.1038/350354a0.

[114] J.T. Lam, M.G. Martin, E. Turk, B.A. Hirayama, N.U. Bosshard, B. Steinmann, E.M. Wright, Missense mutations in SGLT1 cause glucose-galactose malabsorption by trafficking defects., Biochim. Biophys. Acta. 1453 (1999) 297–303.

[115] H. Weintraub, R. Davis, S. Tapscott, M. Thayer, M. Krause, R. Benezra, T.K. Blackwell, D. Turner, R. Rupp, S. Hollenberg, al. et, The myoD gene family: nodal point during specification of the muscle cell lineage, Science (80-.). 251 (1991) 761 LP-766. http://science.sciencemag.org/content/251/4995/761.abstract.

[116] X.Q. Wu, H.M. Jing, Y.G. Zheng, Z.M. Yao, W. Ke, Resistance of Mo-bearing stainless steels and Mo-bearing stainless-steel coating to naphthenic acid corrosion and erosion–corrosion, Corros. Sci. 46 (2004) 1013–1032. doi:http://dx.doi.org/10.1016/S0010-938X(03)00192-6.

[117] X. Wu, H. Jing, Y. Zheng, Z. Yao, W. Ke, Erosion–corrosion of various oil-refining materials in naphthenic acid, Wear. 256 (2004) 133–144. doi:http://dx.doi.org/10.1016/S0043-1648(03)00370-3.

[118] R.D. Kane, M.S. Cayard, A comprehensive study of naphthenic acid corrosion, CORROSION2002. (2002).

[119] S.D. Richardson, T.A. Ternes, Water analysis: emerging contaminants and current issues, Anal Chem. 86 (2014) 2813–2848. doi:10.1021/ac500508t.

[120] M.P. Barrow, J. V Headley, K.M. Peru, P.J. Derrick, Data Visualization for the Characterization of Naphthenic Acids within Petroleum Samples, Energy & Fuels. 23 (2009) 2592–2599. doi:10.1021/ef800985z.

[121] X. Ortiz, K.J. Jobst, E.J. Reiner, S.M. Backus, K.M. Peru, D.W. McMartin, G. O'Sullivan, V.Y. Taguchi, J. V Headley, Characterization of Naphthenic Acids by Gas Chromatography-Fourier Transform Ion Cyclotron Resonance Mass Spectrometry, Anal Chem. 86 (2014) 7666–7673. doi:10.1021/ac501549p.

[122] X. Wang, K.L. Kasperski, Analysis of naphthenic acids in aqueous solution using HPLC-MS/MS, Anal. Methods. 2 (2010) 1715–1722.

[123] R. Hindle, M. Noestheden, K. Peru, J. Headley, Quantitative analysis of naphthenic acids in water by liquid chromatography–accurate mass time-of-flight mass spectrometry, J. Chromatogr. A. 1286 (2013) 166–174. doi:http://dx.doi.org/10.1016/j.chroma.2013.02.082.

[124] P. Brunswick, D. Shang, G. van Aggelen, R. Hindle, L.M. Hewitt, R.A. Frank, M. Haberl, M. Kim, Trace analysis of total naphthenic acids in aqueous environmental matrices by liquid chromatography/mass spectrometry-quadrupole time of flight mass spectrometry direct injection, J. Chromatogr. A. 1405 (2015) 49–71. doi:http://dx.doi.org/10.1016/j.chroma.2015.05.048.

[125] K. Duncan, D. Volmer, C. Gill, E. Krogh, Rapid Screening of Carboxylic Acids

from Waste and Surface Waters by ESI-MS/MS Using Barium Ion Chemistry and On-Line Membrane Sampling, J. Am. Soc. Mass Spectrom. (2015) 1–8. doi:10.1007/s13361-015-1311-y.

[126] R. Huang, K.N. McPhedran, M. Gamal El-Din, Ultra Performance Liquid Chromatography Ion Mobility Time-of-Flight Mass Spectrometry Characterization of Naphthenic Acids Species from Oil Sands Process-Affected Water, Env. Sci Technol. 49 (2015) 11737–11745. doi:10.1021/acs.est.5b03178.

[127] A.S. Pereira, J.W. Martin, Exploring the complexity of oil sands process-affected water by high efficiency supercritical fluid chromatography/orbitrap mass spectrometry, Rapid Commun Mass Spectrom. 29 (2015) 735–744. doi:10.1002/rcm.7156.

[128] W.W. Christie, Why I dislike boron trifluoride-methanol, Lipid Technol. (1994) 66–68.

[129] A.G. Shepherd, V. van Mispelaar, J. Nowlin, W. Genuit, M. Grutters, Analysis of Naphthenic Acids and Derivatization Agents Using Two-Dimensional Gas Chromatography and Mass Spectrometry: Impact on Flow Assurance Predictions, Energy & Fuels. 24 (2010) 2300–2311. doi:10.1021/ef900949m.

[130] C.E. West, A.G. Scarlett, A. Tonkin, D. O'Carroll-Fitzpatrick, J. Pureveen, E. Tegelaar, R. Gieleciak, D. Hager, K. Petersen, K.E. Tollefsen, S.J. Rowland, Diaromatic sulphur-containing "naphthenic" acids in process waters, Water Res. 51 (2014) 206–215. doi:10.1016/j.watres.2013.10.058.

[131] M.J. Wilde, C.E. West, A.G. Scarlett, D. Jones, R.A. Frank, L.M. Hewitt, S.J. Rowland, Bicyclic naphthenic acids in oil sands process water: Identification by

comprehensive multidimensional gas chromatography–mass spectrometry, J. Chromatogr. A. 1378 (2015) 74–87. doi:http://dx.doi.org/10.1016/j.chroma.2014.12.008.

[132] D.R. Parkinson, 2.26 - Analytical Derivatization Techniques, in: J. Pawliszyn (Ed.), Compr. Sampl. Sample Prep., Academic Press, Oxford, 2012: pp. 559–595. doi:http://dx.doi.org/10.1016/B978-0-12-381373-2.00060-0.

[133] W.P. St. John, J. Rughani, S.A. Green, G.D. McGinnis, Analysis and characterization of naphthenic acids by gas chromatography–electron impact mass spectrometry of tert.-butyldimethylsilyl derivatives, J. Chromatogr. A. 807 (1998) 241–251. doi:http://dx.doi.org/10.1016/S0021-9673(98)00085-5.

[134] A.C. Scott, R.F. Young, P.M. Fedorak, Comparison of GC-MS and FTIR methods for quantifying naphthenic acids in water samples, Chemosphere. 73 (2008) 1258–1264. doi:10.1016/j.chemosphere.2008.07.024.

[135] F.C. Damasceno, L.D.A. Gruber, A.M. Geller, M.C. Vaz de Campos, A.O. Gomes, R.C.L. Guimaraes, V.F. Peres, R.A. Jacques, E.B. Caramao, Characterization of naphthenic acids using mass spectroscopy and chromatographic techniques: study of technical mixtures, Anal. Methods. 6 (2014) 807–816. doi:10.1039/C3AY40851E.

[136] F.K. Kawahara, Microdetermination of derivatives of phenols and mercaptans by means of electron capture gas chromatography, Anal Chem. 40 (1968) 1009–1010. doi:10.1021/ac60262a044.

[137] F. Orata, Derivatization reactions and reagents for gas chromatography analysis, in: M. Ali Mohd (Ed.), Adv. Gas Chromatogr. - Prog. Agric. Biomed. Ind. Appl.,

INTECH Open Access Publisher, Rijeka, Croatia, 2012.

[138] Supelco, ed., Pentafluorobenzyl Bromide, Hexaoxacyclooctadecane, Product Specification, (n.d.). http://www.sigmaaldrich.com/content/dam/sigma-aldrich/docs/Supelco/Product_Information_Sheet/4796.pdf.

[139] B. Davis, Crown ether catalyzed derivatization of carboxylic acids and phenols with pentafluorobenzyl bromide for electron capture gas chromatography, Anal Chem. 49 (1977) 832–834. doi:10.1021/ac50014a041.

[140] C.-J. Chien, M.J. Charles, K.G. Sexton, H.E. Jeffries, Analysis of Airborne Carboxylic Acids and Phenols as Their Pentafluorobenzyl Derivatives: Gas Chromatography/Ion Trap Mass Spectrometry with a Novel Chemical Ionization Reagent, PFBOH, Env. Sci Technol. 32 (1998) 299–309. doi:10.1021/es970526s.

[141] E. Gracia-Moreno, R. Lopez, V. Ferreira, Quantitative determination of five hydroxy acids, precursors of relevant wine aroma compounds in wine and other alcoholic beverages, Anal Bioanal Chem. 407 (2015) 7925–7934. doi:10.1007/s00216-015-8959-9.

[142] O. Quehenberger, A.M. Armando, E.A. Dennis, High sensitivity quantitative lipidomics analysis of fatty acids in biological samples by gas chromatography-mass spectrometry, Biochim Biophys Acta. 1811 (2011) 648–656. doi:10.1016/j.bbalip.2011.07.006.

[143] M.A. Bezerra, R.E. Santelli, E.P. Oliveira, L.S. Villar, L.A. Escaleira, Response surface methodology (RSM) as a tool for optimization in analytical chemistry, Talanta. 76 (2008) 965–977. doi:http://dx.doi.org/10.1016/j.talanta.2008.05.019.

[144] G. Sun, P. Stremple, Retention index, Chracterization of Flavor, Fragance and

Many other compounds on DB1 and DB-XLB, J & W Scientific , n.d. http://www.chem.agilent.com/cag/cabu/pdf/b-0279.pdf.

[145] J.I. Cacho, N. Campillo, P. Viñas, M. Hernández-Córdoba, Evaluation of three headspace sorptive extraction coatings for the determination of volatile terpenes in honey using gas chromatography–mass spectrometry, J. Chromatogr. A. 1399 (2015) 18–24. doi:http://dx.doi.org/10.1016/j.chroma.2015.04.041.

[146] J.S. Clemente, N.G. Prasad, M.D. MacKinnon, P.M. Fedorak, A statistical comparison of naphthenic acids characterized by gas chromatography-mass spectrometry, Chemosphere. 50 (2003) 1265–1274.

[147] D.R. Knapp, Handbook of Analytical Derivatization Reactions, Wiley, Hoboken, New Jersey, USA., 1979. https://books.google.ca/books?id=sKZeE-REDswC.

[148] J. Wu, R. Hu, J. Yue, Z. Yang, L. Zhang, Study on the Derivatization Process Using NO-bis-(trimethylsilyl)-trifluoroacetamide, N-(tert-butyldimethylsilyl)-N-methyltrifluoroacetamide, Trimethylsilydiazomethane for the Determination of Fecal Sterols by Gas Chromatography-Mass Spectrometry, World Acad. Sci. Eng. Technol. Int. J. Chem. Mol. Nucl. Mater. Metall. Eng. 2 (2008) 175–178.

[149] A. Shareef, C.J. Parnis, M.J. Angove, J.D. Wells, B.B. Johnson, Suitability of N, O-bis (trimethylsilyl) trifluoroacetamide and N-(tert-butyldimethylsilyl)-N-methyltrifluoroacetamide as derivatization reagents for the determination of the estrogens estrone and 17α-ethinylestradiol by gas chromatography–mass spectromet, J. Chromatogr. A. 1026 (2004) 295–300.

[150] A. Shareef, M.J. Angove, J.D. Wells, Optimization of silylation using N-methyl-N-(trimethylsilyl)-trifluoroacetamide, N, O-bis-(trimethylsilyl)-trifluoroacetamide and

N-(tert-butyldimethylsilyl)-N-methyltrifluoroacetamide for the determination of the estrogens estrone and 17α-ethinylestradio, J. Chromatogr. A. 1108 (2006) 121–128.

[151] J.T. Watson, O.D. Sparkman, Electron Ionization, in: Introd. to Mass Spectrom., John Wiley & Sons, Ltd, Chichester, West Sussex, England, 2008: pp. 315–448. doi:10.1002/9780470516898.ch6.

[152] W. Lu, A. Ewanchuk, L. Perez-Estrada, D. Sego, A. Ulrich, Limitation of fluorescence spectrophotometry in the measurement of naphthenic acids in oil sands process water, J Env. Sci Heal. A Tox Hazard Subst Env. Eng. 48 (2013) 429–436. doi:10.1080/10934529.2013.729802.

[153] Government of Alberta; Oil Sands Information Portal, Oil Sands Water use, (2017). http://osip.alberta.ca/map/ (accessed October 13, 2017).

[154] Government of Alberta; Oil Sands Information Portal, Total Area and Volume of Oil Sands Tailings Ponds over Time, (2017). http://osip.alberta.ca/map/ (accessed October 13, 2017).

[155] T. Siddique, P.M. Fedorak, M.D. MacKinnon, J.M. Foght, Metabolism of BTEX and Naphtha Compounds to Methane in Oil Sands Tailings, Environ. Sci. Technol. 41 (2007) 2350–2356. doi:10.1021/es062852q.

[156] S.-M. Li, A. Leithead, S.G. Moussa, J. Liggio, M.D. Moran, D. Wang, K. Hayden, A. Darlington, M. Gordon, R. Staebler, P.A. Makar, C.A. Stroud, R. McLaren, P.S.K. Liu, J. O'Brien, R.L. Mittermeier, J. Zhang, G. Marson, S.G. Cober, M. Wolde, J.J.B. Wentzell, Differences between measured and reported volatile organic compound emissions from oil sands facilities in Alberta, Canada., Proc.

Natl. Acad. Sci. U. S. A. 114 (2017) E3756–E3765. doi:10.1073/pnas.1617862114.

[157] A. Kuczyńska, L. Wolska, J. Namieśnik, An Attempt to Identify Volatile and Semi-Volatile Organic Compounds Present in the Odra River Waters, Chromatographia. 60 (2004) S279--S289. doi:10.1365/s10337-004-0206-z.

[158] T. Cecchi, Identification of representative pollutants in multiple locations of an Italian school using solid phase micro extraction technique, Build. Environ. 82 (2014) 655–665. doi:https://doi.org/10.1016/j.buildenv.2014.10.009.

[159] H. Rosales-Bravo, H.C. Morales-Torres, J. Vazquez-Martinez, J. Molina-Torres, V. Olalde-Portugal, L.P. Partida-Martinez, Novel consortium of Klebsiella variicola and Lactobacillus species enhances the functional potential of fermented dairy products by increasing the availability of branched-chain amino acids and the amount of distinctive volatiles., J. Appl. Microbiol. 123 (2017) 1237–1250. doi:10.1111/jam.13565.

[160] L. Alonso, M.J. Fraga, Simple and rapid analysis for quantitation of the most important volatile flavor compounds in yogurt by headspace gas chromatography-mass spectrometry., J. Chromatogr. Sci. 39 (2001) 297–300.

[161] R Core Team, R: A Language and Environment for Statistical Computing, (2013). http://www.r-project.org/.

[162] RStudio Team, RStudio: Integrated Development Environment for R, (2016). http://www.rstudio.com/.

[163] S. Lê, J. Josse, F. Husson, FactoMineR: An R Package for Multivariate Analysis, J. Stat. Software; Vol 1, Issue 1 . (2008). https://www.jstatsoft.org/v025/i01.

[164] R. Suzuki, H. Shimodaira, Pvclust: an R package for assessing the uncertainty in hierarchical clustering, Bioinformatics. 22 (2006) 1540–1542. http://dx.doi.org/10.1093/bioinformatics/btl117.

[165] P. Legendre, L.F.J. Legendre, Numerical Ecology, Elsevier Science, 1998.

[166] E. Neuwirth, RColorBrewer: ColorBrewer Palettes, (2014). https://cran.r-project.org/package=RColorBrewer.

[167] T. Galili, dendextend: an R package for visualizing, adjusting, and comparing trees of hierarchical clustering, Bioinformatics. (2015). doi:10.1093/bioinformatics/btv428.

[168] C.M. Reddy, J.S. Arey, J.S. Seewald, S.P. Sylva, K.L. Lemkau, R.K. Nelson, C.A. Carmichael, C.P. McIntyre, J. Fenwick, G.T. Ventura, B.A.S. Van Mooy, R. Camilli, Composition and fate of gas and oil released to the water column during the Deepwater Horizon oil spill., Proc. Natl. Acad. Sci. U. S. A. 109 (2012) 20229–20234. doi:10.1073/pnas.1101242108.

[169] U.S. Department of Health and Human Services, Polycyclic Aromatic Hydrocarbons (PAHs) Fact Sheet, 2009. https://www.epa.gov/north-birmingham-project/polycyclic-aromatic-hydrocarbons-pahs-fact-sheet.

[170] J.T. Andersson, C. Achten, Time to Say Goodbye to the 16 EPA PAHs? Toward an Up-to-Date Use of PACs for Environmental Purposes., Polycycl. Aromat. Compd. 35 (2015) 330–354. doi:10.1080/10406638.2014.991042.

[171] G. Garcia, A. Garcia, M. Gibernau, A. Bighelli, F. Tomi, Chemical compositions of essential oils of five introduced conifers in Corsica., Nat. Prod. Res. 31 (2017) 1697–1703. doi:10.1080/14786419.2017.1285299.

[172] Z. Pandelides, J. Guchardi, D. Holdway, Dehydroabietic acid (DHAA) alters metabolic enzyme activity and the effects of 17β-estradiol in rainbow trout (Oncorhynchus mykiss), Ecotoxicol. Environ. Saf. 101 (2014) 168–176. doi:https://doi.org/10.1016/j.ecoenv.2013.11.027.

[173] L. Xing, J.M. Gutierrez-Villagomez, D.F. Da Fonte, M.J. Venables, V.L. Trudeau, Dehydroabietic acid cytotoxicity in goldfish radial glial cells in vitro, Aquat. Toxicol. 180 (2016) 78–83. doi:https://doi.org/10.1016/j.aquatox.2016.09.009.

[174] S. Khan, K.P. Pandya, Hepatotoxicity in albino rats exposed to n-octane and n-nonane., J. Appl. Toxicol. 5 (1985) 64–68.

[175] O.G. Nilsen, O.A. Haugen, K. Zahlsen, J. Halgunset, A. Helseth, H. Aarset, I. Eide, Toxicity of n-C9 to n-C13 alkanes in the rat on short term inhalation., Pharmacol. Toxicol. 62 (1988) 259–266.

[176] R.J. Babu, A. Chatterjee, E. Ahaghotu, M. Singh, Percutaneous absorption and skin irritation upon low-level prolonged dermal exposure to nonane, dodecane and tetradecane in hairless rats., Toxicol. Ind. Health. 20 (2004) 109–118. doi:10.1191/0748233704th197oa.

[177] F. Muhammad, N.A. Monteiro-Riviere, J.E. Riviere, Comparative in vivo toxicity of topical JP-8 jet fuel and its individual hydrocarbon components: identification of tridecane and tetradecane as key constituents responsible for dermal irritation., Toxicol. Pathol. 33 (2005) 258–266. doi:10.1080/01926230590908222.

[178] L.C.C. Medeiros, F.A.C. Delunardo, L.N. Simoes, M.G. Paulino, T.S. Vargas, M.N. Fernandes, R. Scherer, A.R. Chippari-Gomes, Water-soluble fraction of petroleum induces genotoxicity and morphological effects in fat snook

(Centropomus parallelus)., Ecotoxicol. Environ. Saf. 144 (2017) 275–282. doi:10.1016/j.ecoenv.2017.06.031.

[179] O. Jimenez-Garza, L. Guo, H.-M. Byun, M. Carrieri, G.B. Bartolucci, B.S. Barron-Vivanco, A.A. Baccarelli, Aberrant promoter methylation in genes related to hematopoietic malignancy in workers exposed to a VOC mixture., Toxicol. Appl. Pharmacol. 339 (2017) 65–72. doi:10.1016/j.taap.2017.12.002.

[180] E. Field, F.H. Dempster, G.E. Tilson, Phenolic Compounds from Petroleum Sources, Ind. Eng. Chem. 32 (1940) 489–496. doi:10.1021/ie50364a010.

[181] T.-M. Lin, S.-S. Lee, C.-S. Lai, S.-D. Lin, Phenol burn., Burns. 32 (2006) 517–521. doi:10.1016/j.burns.2005.12.016.

[182] T. Tišler, J. Zagorc-Končan, Comparative assessment of toxicity of phenol, formaldehyde, and industrial wastewater to aquatic organisms, Water. Air. Soil Pollut. 97 (1997) 315–322. doi:10.1007/BF02407469.

[183] K. Santucci, B. Shah, Association of naphthalene with acute hemolytic anemia., Acad. Emerg. Med. 7 (2000) 42–47.

[184] Toxicology and Carcinogenesis Studies of Naphthalene (CAS No. 91-20-3) in B6C3F1 Mice (Inhalation Studies)., Natl. Toxicol. Program Tech. Rep. Ser. 410 (1992) 1–172.

[185] P.S. Makovetskii, D.F. Serdyuk, T.B. Voronina, Naphthene hydrocarbons of the benzine and kerosene-gasoil fraction of petroleum, Chem. Technol. Fuels Oils. 2 (1966) 468–472. doi:10.1007/BF00725973.

[186] NTP toxicology and carcinogenesis studies of decalin (CAS No. 91-17-8) in F344/N rats and B6C3F(1) mice and a toxicology study of decalin in male NBR

rats (inhalation studies)., Natl. Toxicol. Program Tech. Rep. Ser. (2005) 1–316.

[187] C.B.Y. Cordella, PCA: The Basic Building Block of Chemometrics, 2012. doi:10.5772/51429.

[188] C. Li, L. Fu, J. Stafford, M. Belosevic, M. Gamal El-Din, The toxicity of oil sands process-affected water (OSPW): A critical review., Sci. Total Environ. 601–602 (2017) 1785–1802. doi:10.1016/j.scitotenv.2017.06.024.

[189] B.G. Brownlee, P.D. Josephy, N.J. Bunce, Comparison of the Ames Salmonella Assay and Mutatox Genotoxicity Assay for Assessing the Mutagenicity of Polycyclic Aromatic Compounds in Porewater from Athabasca Oil Sands Mature Fine Tailings, Environ. Sci. Technol. 33 (1999) 2510–2516. doi:10.1021/es981343o.

[190] S.A. Hughes, A. Mahaffey, B. Shore, J. Baker, B. Kilgour, C. Brown, K.M. Peru, J. V Headley, H.C. Bailey, Using ultrahigh-resolution mass spectrometry and toxicity identification techniques to characterize the toxicity of oil sands process-affected water: The case for classical naphthenic acids., Environ. Toxicol. Chem. 36 (2017) 3148–3157. doi:10.1002/etc.3892.

[191] R.A. Frank, R. Kavanagh, B. Kent Burnison, G. Arsenault, J. V Headley, K.M. Peru, G. Van Der Kraak, K.R. Solomon, Toxicity assessment of collected fractions from an extracted naphthenic acid mixture, Chemosphere. 72 (2008) 1309–1314. doi:https://doi.org/10.1016/j.chemosphere.2008.04.078.

[192] A.E. Bauer, R. Frank, J. Headley, M. Peru Kerry, M. Hewitt L, G. Dixon D, Enhanced characterization of oil sands acid-extractable organics fractions using electrospray ionization–high-resolution mass spectrometry and synchronous

fluorescence spectroscopy, Environ. Toxicol. Chem. 34 (2015) 1001–1008. doi:10.1002/etc.2896.

[193] A.E. Bauer, R.A. Frank, J. V Headley, K.M. Peru, A.J. Farwell, D.G. Dixon, Toxicity of oil sands acid-extractable organic fractions to freshwater fish: Pimephales promelas (fathead minnow) and Oryzias latipes (Japanese medaka), Chemosphere. 171 (2017) 168–176. doi:https://doi.org/10.1016/j.chemosphere.2016.12.059.

[194] P.J. Bishop, A. Angulo, J.P. Lewis, R.D. Moore, G.B. Rabb, J.G. Moreno, The amphibian extinction crisis-what will it take to put the action into the amphibian conservation action plan?, SAPI EN. S. Surv. Perspect. Integr. Environ. Soc. (2012).

Annexe 1. Other contributions to research during the Ph.D.

Neuroendocrine disruption of organizational and activational hormone programming in poikilothermic vertebrates. Rosenfeld CS, Denslow ND, Orlando EF, **Gutierrez-Villagomez JM**, Trudeau VL. J Toxicol Environ Health B Crit Rev. 2017;20(5):276-304. doi: 10.1080/10937404.2017.1370083. Review. PMID:28895797

Dehydroabietic acid cytotoxicity in goldfish radial glial cells in vitro. Xing L, **Gutierrez-Villagomez JM**, Da Fonte DF, Venables MJ, Trudeau VL. Aquat Toxicol. 2016 Nov;180:78-83. doi: 10.1016/j.aquatox.2016.09.009. Epub 2016 Sep 12. PMID: 27658224

Development of an in vitro Ovary Culture System to Evaluate Endocrine Disruption in Wood Frog Tadpoles. Vu M, Navarro-Martín L, **Gutierrez-Villagomez JM**, Trudeau VL. J Toxicol Environ Health A. 2015;78(18):1137-41. doi:10.1080/15287394.2015.1074970. Epub 2015 Sep 18. PMID: 26383587

Dopamine D1 receptor activation regulates the expression of the estrogen synthesis gene aromatase B in radial glial cells. Xing L, McDonald H, Da Fonte DF, **Gutierrez-Villagomez JM**, Trudeau VL. Front Neurosci. 2015 Sep 2;9:310. doi: 10.3389/fnins.2015.00310. eCollection 2015. PMID: 26388722