Aquatic Toxicology 185 (2017) 48–57

Contents lists available at ScienceDirect

Aquatic Toxicology

journal homepage: www.elsevier.com/locate/aquatox

De novo assembly of the ringed seal (*Pusa hispida*) blubber transcriptome: A tool that enables identification of molecular health indicators associated with PCB exposure

Tanya M. Brown[a,b], S. Austin Hammond[a,c], Bahar Behsaz[c], Nik Veldhoen[a], Inanç Birol[c], Caren C. Helbing[a,*]

[a] Department of Biochemistry and Microbiology, University of Victoria, Victoria, British Columbia V8W 3P6, Canada
[b] Memorial University, St. John's, Newfoundland A1B 3X9, Canada
[c] Canada's Michael Smith Genome Sciences Centre, British Columbia Cancer Agency, Vancouver, BC V5Z 4S6, Canada

ARTICLE INFO

Article history:
Received 5 November 2016
Received in revised form 2 February 2017
Accepted 3 February 2017
Available online 3 February 2017

Keywords:
Transcriptome
de novo assembly
Pinniped
Ringed seal
Polychlorinated biphenyls
RNA-seq

ABSTRACT

The ringed seal, *Pusa hispida*, is a keystone species in the Arctic marine ecosystem, and is proving a useful marine mammal for linking polychlorinated biphenyl (PCB) exposure to toxic injury. We report here the first *de novo* assembled transcriptome for the ringed seal (342,863 transcripts, of which 53% were annotated), which we then applied to a population of ringed seals exposed to a local PCB source in Arctic Labrador, Canada. We found an indication of energy metabolism imbalance in local ringed seals ($n = 4$), and identified five significant gene transcript targets: plasminogen receptor (*Plg-R(KT)*), solute carrier family 25 member 43 receptor (*Slc25a43*), ankyrin repeat domain-containing protein 26-like receptor (*Ankrd26*), HIS30 (not yet annotated) and HIS16 (not yet annotated) that may represent indicators of PCB exposure and effects in marine mammals. The abundance profiles of these five gene targets were validated in blubber samples collected from 43 ringed seals using a qPCR assay. The mRNA transcript levels for all five gene targets, (*Plg-R(KT)*, $r^2 = 0.43$), (*Slc25a43*, $r^2 = 0.51$), (*Ankrd26*, $r^2 = 0.43$), (HIS30, $r^2 = 0.39$) and (HIS16, $r^2 = 0.31$) correlated with increasing levels of blubber PCBs. Results from the present study contribute to our understanding of PCB associated effects in marine mammals, and provide new tools for future molecular and toxicology work in pinnipeds.

1. Introduction

Persistent organic pollutants (POPs) such as polychlorinated biphenyls (PCBs) can bioaccumulate to relatively high concentrations in high-trophic-level arctic marine mammals (Borgå et al., 2004; Muir et al., 2000). Certain health endpoints have been associated with exposure to environmentally relevant concentrations of POPs, and PCBs in particular, have been identified as putative drivers of these deleterious biological effects. For example, changes in hepatic and circulatory vitamin A levels and hepatic *Ahr* and *Cyp1a1* mRNA levels in beluga whales (*Delphinapterus leucas*) from the western Canadian Arctic have been associated with PCBs (Desforges et al., 2013; Noël et al., 2014). Detoxification enzyme activities, including phase 1 enzyme and glutathione-S-transferase (GST) were positively associated with PCBs in ringed seals (*Pusa hispida*) from the Baltic Sea (Routti et al., 2008). Endocrine disruption, reproductive impairment, and immunotoxicity in polar bears (*Ursus maritimus*) have been associated with PCB exposure (Braathen et al., 2004; Sonne et al., 2004; Villanger et al., 2011). However, it remains difficult to ascribe cause-and-effect relationships to any single contaminant class, and questions linger about the nature and significance of observed effects with respect to animal health.

POPs in the Arctic have been largely attributed to long-range transport through atmospheric processes (Macdonald et al., 2000), although local sources (e.g. military sites) also exist and can contaminate food webs (Bright et al., 1995; Kuzyk et al., 2005b). Saglek Bay, Labrador, Canada, has been the site of a military radar station since the 1950s; however, it was not until 1996 that the PCB contamination in excess of 50 parts per million (ppm) – the maximum allowable amount specified in the Canadian Environmental Protection Act for PCB material storage regulations – was discovered at the site, along with evidence that PCB had entered the marine ecosystem (Kuzyk et al., 2005b). Previous studies from the

* Corresponding author.
E-mail address: chelbing@uvic.ca (C.C. Helbing).

http://dx.doi.org/10.1016/j.aquatox.2017.02.004
0166-445X/

area indicated that shorthorn sculpin (*Myoxocephalus scorpius*) and black guillemot (*Cepphus grylle*) nestlings were at risk of impaired reproduction or death (Kuzyk et al., 2005a). More recently, changes in hepatic *Ahr*, *Il1b*, *Esr1*, *Igf1*, and *Nr3c1* mRNA levels in ringed seals from the affected area were associated with exposure to PCBs (Brown et al., 2014b).

The primary tools used in molecular profiling (i.e. quantitative real-time polymerase chain reaction; qPCR and DNA microarray) of arctic marine mammals to assess toxicological risk have relied heavily on the evaluation of a select number of known gene targets expressed within the transcriptome and determination of mRNA transcript levels (Brown et al., 2014b; Noël et al., 2014; Routti et al., 2010). While qPCR demonstrates great sensitivity, it is limited in its breadth across the transcriptome for each experimental run (Veldhoen et al., 2012). In contrast, massive parallel sequencing methods, such as RNA sequencing (RNA-seq), are a powerful means to obtain large amounts of transcriptome data allowing for molecular profiling across a greater representation of the transcriptome and augmentation in the development of targeted qPCR-based tools. Unlike DNA microarrays and qPCR, RNA-seq can be implemented as a discovery platform to increase our understanding of constituents of tissue-specific transcriptomes (Wang et al., 2009). RNA-seq data are typically generated as tens to hundreds of millions of paired-end 75-mer to 150-mer reads, which are then mapped onto an annotated reference genome for identification and quantitation (Trapnell et al., 2010). In the absence of a reference genome, *de novo* assembly of RNA-seq reads has enabled the study of transcriptomes originating from non-model species (Birol et al., 2015; Frischkorn et al., 2014; Gao et al., 2013; Meyer et al., 2011). Thus, adoption of a combinatorial approach that harnesses the power of RNA-seq along with the targeted approach of qPCR can aid in the development of sensitive and informed molecular endpoints for application towards ecotoxicogenomic assessment of wildlife species.

The objectives of the present study were to increase the available molecular biological information related to ringed seals and identify novel gene transcripts associated with PCB exposure. The combination of a relatively low long-range POP "background" level and a local PCB "hotspot" on the Labrador coast provides an invaluable opportunity to identify gene expression endpoints that serve as potential targets for monitoring PCB-associated health effects in marine mammals. We *de novo* assembled and annotated the blubber transcriptome of ringed seal and used this new resource to inform qPCR-based assessment of PCB-associated health effects in an affected population. To our knowledge this work represents the first reported transcriptome specific for ringed seals, and will serve as an important resource for continued development of species-specific molecular tools.

2. Materials and methods

2.1. Sample collection

Tissue samples from adult (≥6 years old) and subadult (<6 years old) ringed seals ($n = 43$) were obtained from Inuit hunters in four marine inlets (Nachvak Fjord, Saglek Fjord, Okak Bay, and Anaktalak Bay) along the northern Labrador coast during the 2009, 2010 and 2011 fall (September and October) seasons. Ages of the seals were determined at Matson's Laboratory, U.S.A., by longitudinal thin sectioning a lower canine tooth and counting annual growth layers in the cementum using a compound microscope and transmitted light. Morphometric measurements including length, weight, axial girth, and blubber thickness (at the sternum) were collected for each animal. We chose to evaluate blubber which is routinely used in marine mammal studies since it can be obtained using a minimally invasive biopsy procedure. Inner blubber was selected for analysis as previous studies reported this biological material to be more appropriate than outer blubber to investigate the association of POP concentrations and changes in mRNA transcript abundance, likely due to increased vascularisation and metabolic activity in this tissue layer (Strandberg et al., 2008; Tabuchi et al., 2006; Wilson et al., 2007). Blubber samples (~1 g) were preserved directly in the field in RNA*later* tissue preservation solution as per the manufacturer's instructions (Applied Biosystems, Foster City, CA) and subsequently stored at −20 °C until isolation of total RNA. All tissue samples were obtained within 1 h of death.

A sub-set ($n = 4$) of adult (≥14 years old) male ringed seals were chosen for RNA-seq analysis. Two of the seals were from the reference Okak Bay and had a PCB signature that is lighter, less chlorinated and consistent with a long-range background signature, while the remaining two seals were from the PCB-contaminated Saglek Bay, with a heavier, more chlorinated "local" PCB signature [see Brown et al. (2014a) for the established approach used to identify a "local" versus long-range signature seal]. The Total PCB concentrations (average ± SD) in the affected seals (2376 ± 31 ng/g lw) were 4.4-fold greater than the reference seals (543 ± 45 ng/g lw). $\sum$PCB concentrations originate from Brown et al. (2014b). The detailed methodology for extraction, cleanup and quantification of target analytes has been reported elsewhere (Brown et al., 2014a; Drouillard et al., 2004; Lazar et al., 1992). Samples were analyzed for individual PCB congeners by gas chromatography electron capture detection (GC-ECD). Percent lipid was determined using gravimetric lipid determination by weight of extract method with dichloromethane. For each batch of six samples, an in-house reference homogenate tissue, method blank, and the external PCB-34 recovery standard were analyzed for 62 PCB congeners: 19, 18/17, 24/27, 16/32, 26, 28/31, 33/20, 22, 45, 46, 52, 49, 47/48, 44, 42, 64/41/71, 40, 74, 70/76, 66/95, 91, 60/56, 92/84/101, 99, 97, 87, 85, 136, 110, 151/82, 144/135, 149, 118, 134, 146, 153, 105/132, 141, 179, 137, 130/176, 138/163, 158, 178, 187/182, 183, 128/167, 185, 174, 177, 156/171/202, 157/173/200/204, 172, 180, 201, 170/190, 199, 203/196, 207, 194, and 206. All PCB congeners that were detected in 90% of the samples were included in the data analysis, in samples where an individual congener was not detected it was replaced with a random number between the detection limit (0.011–0.150 ng/g) and zero. Average PCB concentrations in the affected seals were above the endocrine and immune function effects thresholds for seals [1300–1370 ng/g lw, (Brown et al., 2014b; Mos et al., 2010)].

Detailed procedures of total RNA extraction are described elsewhere (Buckman et al., 2011; Noël et al., 2014; Tabuchi et al., 2006). Briefly, each sample was homogenized in a 1.5 mL microcentrifuge tube using a Retsch MM301 mixer mill (Thermo Fisher Scientific, Ottawa, ON, Canada) following the addition of 700 μL TRIzol reagent (Thermo Fisher Scientific) and a 3 mm diameter tungsten-carbide bead. Total RNA from the two Okak Bay reference site seals were combined. Similarly, total RNA from the two PCB contaminated seals from Saglek Bay were pooled. For each animal, total RNA concentration was assessed by UV spectrophotometry followed by an evaluation of RNA integrity using an Agilent Bioanalyzer 2100. RIN scores of 7.2 (reference site) and 7.9 (affected site) were obtained and strand-specific RNA-seq libraries were generated using the MultiMACS mRNA Isolation kit (Miltenyi Biotec, Auburn, CA, USA), SuperScript® Double Strand cDNA Synthesis Kit (Thermo Fisher Scientific) with random hexamer primers, and the NEBNext DNA Library Prep Reagent Set for Illumina® (New England Biolabs, Whitby, ON, Canada). A total of 537 million 75 base pair (bp) paired-end reads were generated on an Illumina HiSeq 2000 platform (San Diego, CA, USA) using the TruSeq PE Cluster Kit v3-cBot-HS (Illumina Canada, Victoria, BC, Canada) and TruSeq SBS Kit v3-HS (Illumina Canada) following manufacturer's instruc-