Sulforaphane Regulation of LncRNA in a Transplacental Nrf2 Knockout Mouse Model Exposed to Dibenzo[def,p]chrysene

by
Karan Arun Patel

A THESIS
Submitted to
Oregon State University
University Honors College

in partial fulfillment of
the requirements for the
degree of

Honors Baccalaureate of Science in Microbiology
(Honors Scholar)

Presented May 30, 2018
Commencement June 2018
AN ABSTRACT OF THE THESIS OF

Karan Arun Patel for the degree of Honors Baccalaureate of Science in Microbiology presented on May 30, 2018. Title: Sulforaphane Regulation of LncRNA in a Transplacental Nrf2 Knockout Mouse Model Exposed to Dibenzo[def,p]chrysene.

Abstract approved:

_____________________________________________________________

David E. Williams

Dibenzo[def,p]chrysene (DBC) is a highly potent, but less prevalent, environmental carcinogen belonging to a class of compounds known as Polycyclic Aromatic Hydrocarbons (PAHs). They are highly ubiquitous and arise as a byproduct of natural and anthropogenic combustion processes. Previous studies have documented carcinogenic effects upon in utero exposure of PAHs. Nuclear Factor Erythroid-2-Related Factor (Nrf2) is a transcription factor vital to the oxidative stress response. Consequently, Nrf2 deficiency in animal models has shown increased tumor incidence compared to those with Nrf2 when treated with an environmental carcinogen.

Sulforaphane (SFN), a chemopreventive agent from cruciferous vegetables, induces phase II antioxidant genes via interaction with Nrf2. This enhances resistance to carcinogenesis.

The primary objective of this study is to analyze the up or down regulation of various long non-coding RNA (LncRNA >200bp) utilizing a Nrf2 Knockout Mouse Model. LncRNA have been found to possess a myriad of functions as molecular signals and molecular decoys that move to activate and suppress gene expression. They have the ability to guide ribonucleoprotein complexes to assemble at specific chromatin sites. Consequently, cis or trans gene expression is induced.
LncRNA regulation will be evaluated in lung tissue from adult offspring exposed \textit{in utero} to sulforaphane. Sulforaphane was used as a dietary chemopreventive agent in some treatment groups to compare tumor incidence in 10-month old adult offspring. This phytochemical was exposed to the fetus transplacentally and through lactational exposure. The two treatment groups utilized are as follow: control lung from adult offspring; lung from adult offspring born to mothers given a sulforaphane diet. A commercially available LncProfiler RT-qPCR array (System Biosciences, Mountain View, CA. #RA930A-1) was used to examine the lncRNA profiles of the treatment group relative to the control.

This study is being conducted in efforts to document a potential correlation between fetal exposure to PAHs, abundance of cancer associated lncRNA and the impact of the \textit{Nrf2} genotype in the presence of the prototypical \textit{Nrf2} agonist, sulforaphane.

Keywords: Transplacental model, polycyclic aromatic hydrocarbons, dibenzo\textit{[def,p]}chrysene, sulforaphane, Nuclear Factor Erythroid-2-Related Factor (\textit{Nrf2}), long non-coding RNA (LncRNA)

Corresponding e-mail address: patekara@oregonstate.edu
Sulforaphane Regulation of LncRNA in a Transplacental Nrf2 Knockout Mouse Model Exposed to Dibenzo[def,p]chrysene

by
Karan Arun Patel

A THESIS
Submitted to
Oregon State University
University Honors College

in partial fulfillment of
the requirements for the
degree of

Honors Baccalaureate of Science in Microbiology
(Honors Scholar)

Presented May 30, 2018
Commencement June 2018
Honors Baccalaureate of Science in Microbiology project of Karan A. Patel presented on May 30, 2014.

APPROVED:

David E. Williams, Mentor, representing Environmental and Molecular Toxicology/Linus Pauling Institute

Lisbeth K. Siddens, Committee Member, representing Environmental and Molecular Toxicology/Linus Pauling Institute

Susan C. Tilton, Committee Member, representing Environmental and Molecular Toxicology

Toni Doolen, Dean, University Honors College

I understand that my project will become part of the permanent collection of Oregon State University, University Honors College. My signature below authorizes release of my project to any reader upon request.

Karan A. Patel, Author
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Metabolic pathways for bioactivation of PAHs</td>
<td>2</td>
</tr>
<tr>
<td>2.</td>
<td>Chemical structures of dibenzo[def,p]chrysene and benzo[a]pyrene</td>
<td>3</td>
</tr>
<tr>
<td>3.</td>
<td>Aryl hydrocarbon receptor (AhR) pathway for metabolic activation of xenobiotics</td>
<td>5</td>
</tr>
<tr>
<td>4.</td>
<td>Regulation of GSTA2 and NQO1 gene expression through the Nrf2-KEAP1-ARE pathway</td>
<td>6</td>
</tr>
<tr>
<td>5.</td>
<td>Generalized mechanisms of lncRNA involved in cancer initiation and progression</td>
<td>8</td>
</tr>
<tr>
<td>6.</td>
<td>Hydrolysis of glucoraphanin by myrosinase to produce SFN metabolites</td>
<td>9</td>
</tr>
<tr>
<td>7.</td>
<td>ICR-(\text{Nrf}) transplacental study design</td>
<td>14</td>
</tr>
<tr>
<td>8.</td>
<td>Synthesis of LncRNA cDNA for use in LncProfiler qPCR array</td>
<td>18</td>
</tr>
<tr>
<td>9.</td>
<td>Electropherograms of RNA isolated from lung tissue of 10-month old progeny born to SFN treated dams</td>
<td>20</td>
</tr>
<tr>
<td>10.</td>
<td>LncRNA fold regulation in adult \textit{male} progeny</td>
<td>22</td>
</tr>
<tr>
<td>11.</td>
<td>LncRNA fold regulation in adult \textit{female} progeny</td>
<td>22</td>
</tr>
<tr>
<td>12.</td>
<td>LncRNA fold regulation in adult \textit{male} progeny \textit{heterozygous} for Nrf2</td>
<td>23</td>
</tr>
<tr>
<td>13.</td>
<td>LncRNA fold regulation in adult male progeny \textit{null} for Nrf2</td>
<td>23</td>
</tr>
<tr>
<td>14.</td>
<td>LncRNA fold regulation in adult \textit{female progeny} \textit{heterozygous} for Nrf2</td>
<td>24</td>
</tr>
<tr>
<td>15.</td>
<td>LncRNA fold regulation of MSUR1 in adult \textit{female progeny} \textit{null} for Nrf2</td>
<td>24</td>
</tr>
</tbody>
</table>
INTRODUCTION

Polycyclic Aromatic Hydrocarbons

Polycyclic Aromatic Hydrocarbons (PAHs) are ubiquitous environmental pollutants, formed through the incomplete combustion of carbon materials such as coal, oil, diesel, gas, tobacco and coke oven emissions. PAHs are suspected to be lung carcinogens seeing that these byproducts of incomplete combustion can inevitably enter the human respiratory tract (Zhang et al., 2012). PAHs are chemically lipophilic and have a characteristic fused aromatic ring structure that is oftentimes planar. The various configurations that PAHs form allow them to be assigned to one of three classes; non-bay region, bay region, and fjord region (Zhang et al., 2012). Several PAHs are carcinogenic, especially those with 5 or more rings (IARC, 2010; ATSDR, 1995). The general population is susceptible to PAH exposure primarily through diet, especially higher-molecular weight PAHs (Jakszyn et al., 2004; Phillips, 1999). Human exposure to these high molecular weight PAHs is predominantly through diet and GI uptake (Diggs et al. 2011). The half-life of PAHs in animals is on the order of hours due to various metabolic processes that facilitate their elimination.

Bioactivation and Metabolism of PAHs

PAHs are chemically activated and metabolized by the CYP1 family of enzymes, which includes CYP1A1, CYP1A2, and CYP1B1. Three major metabolic pathways are responsible for the bioactivation of PAHs (Figure 1). In DBC metabolism, the peroxidase pathway is known to generate a radical cation that can form depurinating DNA adducts.
The diol epoxide pathway is dependent on cytochrome P450 enzymes, such as cytochrome P450 1B1 (CYP 1B1), and epoxide hydrolase to produce (+) and (-) \textit{trans}-DBC-11,12 dihydrodiol (DHD) products. Further epoxidation of these two DHDs causes the production of (-)\textit{-anti}-DBC-11,12,-\textit{trans}-dihydrodiol-13,14-epoxide (DBCDE) products, leading to the formation of dA DNA adducts. The 2 DHDs produce 4 enantiomers of DBCDE; the (-)-anti, (-)-cis; (+)-anti and (+)-cis. The aldo-keto reductase (AKR) pathway is responsible for producing semi-quinones and quinones (Penning, 2014). The intercalation or covalent reactions between metabolically activated PAHs and DNA induce mutations that lead to carcinogenesis (Katz et al., 1998).

\textbf{Figure 1.} Metabolic pathways for bioactivation of PAHs using the fjord containing dibenzo[def,p]chrysene as an example. The pathway in this figure is intercepted by phase II enzymes, to catalyze conjugation reactions that eliminate PAHs and their metabolites. The three pathways shown are the peroxidase pathway, CYP-dependent epoxygenation pathway, and the aldo-keto reductase (AKR) pathway. Note: PAH diols can undergo redox cycling to produce reactive oxygen species (ROS) (Siddens et al., 2012).
The metabolism of benzo[a]pyrene (B[a]P) has been extensively studied by the research community. B[a]P is a widespread environmental contaminant that is formed through pyrolysis and incomplete combustion processes (IARC, 2010). It has been found to act as both a local and systemic carcinogen, as well as a transplacental carcinogen. The research laboratory of Dr. David E. Williams is studying dibenzo[def,p]-chrysene (DBC, also known as dibenzo[a,l]pyrene). DBC is classified by the IARC as a 2A probable human carcinogen and has a relative potency factor (RPF) value 30-fold higher than B[a]P, as determined in animal cancer models (Madeen et al., 2015). Despite DBC having a greater potency, it is classified as a group 2A carcinogen due to insufficient evidence of carcinogenicity in humans, but sufficient evidence in animal models (IARC, 2010). DBC is a highly potent PAH involved in the initiation of cancers in lung, liver, and the reproductive system (Yu et al., 2006; Castro et al., 2008). It contains configurations of both a bay region and fjord region, compared to B[a]P, which only has a bay region. A distinguishing feature of DBC is its non-planar distorted structure resulting from molecular overcrowding in the fjord region (Katz et al., 1998).

\[\text{(A)}\]

\[\text{(B)}\]

Figure 2. Chemical structures of (A) dibenzo[def,p]chrysene and (B) benzo[a]pyrene highlighting the location of bay regions and fjord regions (Katz et al., 1998).
Aryl Hydrocarbon Receptor

The aryl hydrocarbon receptor (AhR) belongs to a family of helix-loop-helix transcription factors. It is a cytosolic receptor that facilitates xenobiotic metabolism through induction of enzymes such as cytochrome P450. AhR forms a complex in the cytosol with heat shock protein 90 (HSP90), X-associated protein 2 (XAP2), and a co-chaperone protein known as p23. This inactive complex will bind to exogenous ligands, such as a PAH, causing nuclear translocation. Binding of the AhR nuclear translocator (ARNT) in the nucleus causes displacement of the HSP90-XAP2-p23 protein complex. This leads to binding of the complex to the dioxin-response element (DRE), also known as the xenobiotic-response elements (XRE), regulating hundreds of genes. Consequently, an increase in expression of cytochrome P450 enzymes is observed, which carry out the biotransformation of PAHs on rough endoplasmic reticulum (Tsay et al., 2013). The AhR repressor (AhHR) effectively competes with ARNT dimerization to inhibit transcriptional activity of XREs. The AhR pathway is prevalent in human bronchial epithelial cells, having many physiological effects in lung tissue, including the initiation and propagation of lung cancer. However, the physiological function of AhR is to metabolize xenobiotics and serve as an adaptive oxidative stress response. The AhR pathway of xenobiotic activation has many underlying consequences. For instance, it can influence cell-cell adhesion and interaction, cytokine signaling, and mucin production (Tsay et al., 2013). Figure 3 illustrates the AhR bioactivation pathway.
Figure 3. Aryl hydrocarbon receptor (AhR) pathway for metabolic activation of xenobiotics (Murray et al., 2014). DRE is in reference to the dioxin-response elements, also referred to as the xenobiotic-response elements (XRE).

Detoxification of PAHs via the Nrf2 Pathway

Nuclear Factor Erythroid-2-Related Factor (Nrf2) is a transcription factor vital to the transcriptional regulation of the oxidative or electrophilic stress response (Li et al., 2013; Nguyen et al., 2009). The Nrf2 factor has been the subject of extensive scientific research and has been correlated to various pathologies such as chronic obstructive pulmonary disease (COPD), inflammatory disorders, and carcinogenesis (Li et al., 2013). The Nrf2-regulatory pathway is responsible for the transcription of mRNA encoding major detoxifying phase II enzymes such as glutathione-S-transferase (GST), NADPH:quinone oxidoreductase (NQO1), and UDP-glucuronosyl transferases (UGT) (Higdon et al., 2007). The expression of these enzymes is mediated by a cis-acting element, termed the antioxidant response element (ARE) that resides in the promoter of these genes.
(Higdon et al., 2007). The ARE is constitutively expressed to counteract the reactive byproducts of aerobic respiration in a non-stressed state, but is upregulated during times of oxidative stress. Furthermore, the Nrf2 factor is continuously tethered to a cytoplasmic protein known as Kelch-like erythroid-cell-derived protein with CNC homology (ECH)-associated protein (KEAP-1), which is responsible for ubiquitination of Nrf2 so that it may be continually degraded in non-stressed cells (Nguyen et al., 2009). Indirect means of activating expression of the ARE have been identified, such as Michael acceptors. Compounds with Michael acceptor groups have the propensity to react with sulfhydryl substituents (i.e. KEAP1-SH), causing dissociation of KEAP-1 from Nrf2 (Magesh et al., 2012).

![Diagram](image)

**Figure 4.** Regulation of GSTA2 and NQO1 gene expression through the Nrf2-KEAP1-ARE pathway. The presence of reactive PAH metabolites generated by CYP1A1/2 bioactivation enhances Nrf2-mediated transcriptional control of AREs, and consequently upregulates the expression of GSTA2 and NQO1 (Nguyen et al., 2009).
**Long non-coding RNA (lncRNA)**

The discovery of non-coding RNA elements has modified the long-held central dogma of molecular biology that mRNA is the only intermediate in the conversion of DNA to protein. A large portion of the untranslated genome, previously considered “junk DNA,” is now known as lncRNA, which have important functions in cellular physiology. Expression of lncRNA is highly organ specific and at relatively lower levels than mRNA expression. A research study examining expression profiles of lncRNA in different organs showed age to be a more significant contributing factor than sex in the variance of lncRNA expression profiles (Wen et al., 2016). Studies have indicated that several cancer risk loci are transcribed into lncRNA, transcripts that play important roles in tumorigenesis (Cheetham et al., 2013; Yang et al., 2014). For example, one such lncRNA, *HOTAIR*, is highly expressed in breast cancer metastases and other tumors predisposed to metastasis (Tsai et al., 2011). There have been ~35,000 lncRNAs identified through the FANTOM3 project, but only a small percentage of these have been functionally characterized (Bhat et al., 2016). The dysregulation of lncRNA levels is often associated with solid tumor formation and leukemia (Tsai et al., 2015).

Functionally, LncRNA can serve as regulatory elements, such as a substrate for DNA-binding proteins that modulate the expression and 3-dimensional configuration of the genome (Cheetham et al., 2013). There are 4 broad categorized mechanisms of lncRNA-mediated regulation; signals for gene activation, guides for chromatin modification, decoys for gene suppression, and scaffolds for recruiting chromatin modifying complexes (Figure 5). *MALAT1* has been implicated in regulating the distribution of serine/arginine
splicing factors (SR), thereby influencing post-transcriptional modification (Cheetham et al., 2013).

Figure 5. Generalized mechanisms of IncRNA involved in cancer initiation and progression. The elucidated functions have generally been at the pre- or post-transcriptional levels, which induces cis or trans gene expression (Adapted from System Biosciences).

Chemopreventative Properties of Cruciferous Vegetables

Cruciferous vegetables, such as broccoli, brussel sprouts, cabbage, watercress, kale, and cauliflower are an abundant source of glucosinolates. Glucosinolates can hydrolyze to form indoles and isothiocyanates (ITCs). Sulforaphane (SFN), an isothiocyanate, has been implicated as a chemopreventive phytochemical that decreases cancer risk. SFN is present in cruciferous vegetables as the glucosinolate, glucoraphanin, which is cleaved in the presence of myrosinase to form SFN and glucose (Higdon et al., 2007). SFN can be converted into several metabolic products to exert its effect; SFN-cysteine (SFN-cys), SFN-N-acetylcysteine (SFN-NAC), SFN-glutathione conjugate (SFN-GSH) (Figure 6). These metabolites have been linked to chemoprotection against pancreatic, prostate, ovarian, and breast cancers (Chang and Yu, 2016). Extensive studies have identified SFN as a potent inducer of phase II enzymes by increasing Nrf2-
dependent transcription of oxidative stress response genes that contain an ARE (Higdon et al., 2007). Furthermore, previous human-feeding studies involving men with high-grade prostatic intraepithelial neoplasia, revealed changes in TGF-β, insulin, and EGF signaling, which is indicative of SFN’s role in signaling pathways other than the Nrf2 regulatory pathway (Watson et al., 2013).

**Figure 6.** Hydrolysis of glucoraphanin from cruciferous vegetables via the myrosinase enzyme to produce sulforaphane metabolites in the liver. The purpose of this pathway is to generate conjugated sulforaphane metabolites that can be rapidly excreted through urine (Amjad et al., 2015).

ITCs, such as sulforaphane, inhibit carcinogenesis by upregulating detoxification, inflammation, apoptosis, cell cycle arrest and epigenetic regulation of cancer initiation. Additionally, SFN could also be involved in the inhibition of promotion and/or progression of cancer. ITC metabolites, such as SFN, have been shown to be involved in
inhibition of histone-deacetylases (HDACs), causing hyper-acetylation of histone proteins leading to formation of euchromatin (open confirmation) (Myzak et al., 2004; Gerhauser, 2013; Change and Yu, 2016). Consequently, this promotes the transcription of genes. Experiments scrutinizing SFN and its mechanism of action primarily investigate the KEAP1-Nrf2-ARE regulatory pathway. This pathway has been shown to reduce the growth and metastasis of cancers, in the presence of the prototypical agonist, SFN.

**Transplacental Carcinogenesis of Dibenzo[def,p]chrysene (DBC)**

Despite the abundance of clinical studies on extracts of cruciferous vegetables and their potential for disease intervention, there is still much to be unveiled. In particular, the effects of maternal dietary supplementation of phytochemical extracts from vegetables during gestation and their ability to reduce disease risk later in life (Shorey et al., 2013). Exposure to environmental chemicals and xenobiotics during pregnancy can adversely affect the susceptibility of the fetus to disease risk later in life (EPA, 2005). This postulate is a component of the “Developmental Origin of Health and Disease” (DOHaD) phenomenon, also known as the “Barker Hypothesis” (Barker, 2004).

Previously, our laboratory has employed a preclinical model to investigate chemoprevention of DBC induced transplacental carcinogenesis (Yu et al., 2006). These studies have shown that maternal dietary supplementation with indole-3-carbinol, an isothiocyanate from cruciferous vegetables, protected the fetus from DBC-induced T-cell acute lymphoblastic leukemia (T-ALL) mortality and lung tumor incidence (Shorey et al., 2013; Yu et al., 2006). SFN, shown to be a chemopreventive agent in preclinical models, has surprisingly been shown to actually enhance DBC-induced T-ALL mortality (Shorey et al., 2013). As described above, SFN has been shown to exert its chemopreventive
effect through the enhanced expression of the Nrf2 transcription factor in the KEAP1-Nrf2-ARE regulatory pathway. This up-regulates expression of genes that ameliorate carcinogen metabolism, such as phase II enzymes.

To investigate DBC-induced transplacental carcinogenesis and chemoprevention by SFN through Nrf2, an equal number of Nrf2+/− and Nrf2−/− mice, on an ICR genetic background, were bred. SFN failed to provide chemoprotection to pups born to pregnant dams exposed to DBC, despite the abundance and bioavailability of SFN metabolites in neonate plasma (unpublished data). In 10-month progeny expressing Nrf2 (Nrf2+/−) and born to mothers provided with a 400 ppm SFN diet, tumor incidence was not reduced in lung tissue. Surprisingly, lung tumor multiplicity was higher in Nrf2+/− mice in the DBC/SFN treatment group (unpublished data). It has been recognized that the Nrf2 factor may play a dual function as a mechanism of chemoprevention early in carcinogenesis, but also promoting neoplastic growth and survival of the cancer cell (Taguchi and Yamamoto, 2017). It has been proposed that the Nrf2 regulatory pathway potentially exhibits a U-shaped “dose-response” curve, with the degree of Nrf2 activation being associated with enhanced carcinogenesis (Kensler and Wakabayashi, 2010). Thus, the fetus may experience the effects of Nrf2 activation in a similar fashion, although further experimentation is necessary to fully understand this.

The primary objective of this study was to analyze the up or down regulation of various long non-coding RNA (lncRNA >200bp) using tissues from the Nrf2 knockout mouse model previously described. LncRNA have the ability to guide ribonucleoprotein complexes to assemble at specific chromatin sites, causing cis or trans gene expression. As such, the role of lncRNA in transplacental carcinogenesis and chemoprevention
potentially reaffirms the postulates of the Barker Hypothesis. The focus of this specific project was to examine regulation of lncRNA in 10-month progeny born to mothers given a SFN diet, but not DBC. To examine this, archived lung tissue from adult progeny was used. It was hypothesized that SFN, through Nrf2, would be associated with a significant change in expression levels of various cancer related lncRNAs.
Materials and Methods

Vehicle and SFN Transplacental Diet

Custom diets and semi-purified control diets, AIN93G and AIN93M, were purchased from Research Diets, Inc. (New Brunswick, NJ). SFN, Cat #S699115, was purchased from Toronto Research Chemicals (North York, Ontario). These prepared diets were stored and protected from light at -20°C for the duration of the experiment.

Animal Treatment Protocol

The following describes the protocol employed in the handling and care of animals used in this transplacental study. Protocols for the handling and treatment of mice were reviewed by the Oregon State University Institutional Animal Care and Use Committee (IACUC) and conducted in an AAALAC accredited facility. Nrf2 knockout mice, originally from RIKEN BioResource Center (Ibaraki, Japan), were maintained as an ICR strain colony for the Linus Pauling Institute’s Cancer Prevention and Intervention Program. Male Nrf2\(^{-/-}\) were bred to female Nrf2\(^{+/-}\) mice to produce litters containing 50% heterozygotes and 50% knockouts. All animals were housed in micro-isolator cages (Life Products, Inc., Seaford, DE) with CareFRESH bedding at 20 ± 1°C and 50 ± 10% humidity with a light/dark cycle of 12h The appearance of a vaginal plug marked day 0 of gestation, at which time the mice were segregated by sex and females weighed.

Mice were given a diet of AIN93G ad libitum throughout the breeding process and up to gestation day 9 (GD 9). On GD 9, the maternal diet of AIN93G was maintained in the control diet group, but replaced with a diet containing 400 ppm of a SFN in the dietary intervention group. Diets were provided as pellets and changed weekly to
maintain the stability of active ingredients. On GD 17, pregnant mice were administered an oral gavage of 5 ml/kg body weight of a corn oil vehicle or 15 mg/kg of DBC in corn oil. Maternal dietary supplementation of SFN was continued until weaning on post-natal day 21 (PND21). Pups were kept with the mother until PND21 when the progeny were weaned, given an AIN93G diet, and segregated by sex into different micro-isolator cages. Animals were monitored daily for 10 months. If signs of morbidity appeared, the animal was immediately euthanized with CO₂ asphyxiation and cervical dislocation followed by necropsy. Ten-month old animals were euthanized and necropsied with portions of lung and liver snap frozen in liquid N₂ for future analysis. This study focuses on lung tissue harvested from 10-month-old tumor study animals evaluated for lncRNA expression.

Males and females (n=16) from the corn oil/control diet group were compared to the corn oil/SFN diet group.

Figure 7. ICR-Nrf2 transplacental study design. This transplacental model scrutinized in utero exposure to DBC and its effect on tumor incidence in progeny. GD9 marks the start of a SFN dietary intervention (400 ppm), which was continued until weaning on PND21. Pregnant dams were given an oral gavage of 15 mg/kg DBC on GD17. The focus of this lncRNA study was on pups from the SFN/CO treatment, and not the DBC/CO treatment.
Genotyping for Nrf2

Ear snips collected from 10 month old progeny were incubated at 55°C in 100 μL DirectPCR® lysis reagent containing 0.06 mg proteinase K for 6-12 hours to extract DNA from tissues. Lysates were heated to 85°C for 45 minutes to inactivate the proteinase enzyme. Samples were centrifuged at 16,000 x g for 1 minute to pellet cell debris. The supernatant (0.5-1 μL) was used directly in a 10 μL five primer PCR reaction containing 10X Biolase™ polymerase buffer, 1 unit Biolase™ DNA polymerase, 2 mM MgCl2, 0.2 mM (each) dNTPs, 0.4 μM of each primer; muNrf2 f-tggacgggaccttggaagctg, muNrf2 r-ttgttgacctccaggggc, Nrf2/LacZ complex r-gcggttgacctgtgtaatgggatagg, and a technical positive control muFMO4 f-cagtgtgaacaaagcgccgcag, muFMO4 r-caggaaaggtctccagaggt. PCR products were separated on 2% agarose gels and imaged with a ChemiDoc™ MP system BioRad Laboratories, Inc. (Hercules, CA). Animals heterozygous for Nrf2 resulted in both a 232 bp (wild type allele) and 400 bp (null allele) while knockouts contained only the 400 bp product. PCR amplifications were validated by using mouse FMO4 primers which produced a 150 bp product in all successful reactions.

Total RNA Isolation

TRIzol Reagent was used for total RNA isolation (Life Technologies, Carlsbad, CA). For each sample extracted, roughly 30-50 mg of lung tissue was homogenized using TRIzol and an electric homogenizer. The homogenate was transferred to a 1.5 mL microcentrifuge tube and incubated at room temperature for 5 minutes to permit complete dissociation of the nucleoprotein complex. Then 0.2 mL of chloroform per 1 mL of TRIzol reagent used for homogenization was added, followed by vigorous shaking by
hand for 30 seconds. After incubation at room temperature for 2-3 minutes, samples were centrifuged at 12,000xg at 4°C for 15 minutes. After collection of the supernatant, 0.5 mL of 100% isopropanol was added to this aqueous fraction to precipitate the RNA. Samples were incubated at room temperature for 10 minutes and then centrifuged at 12,000xg for 10 minutes at 4°C. A solvent degasser was used to remove the supernatant, leaving only an RNA pellet. The RNA pellet was washed with 1 mL of 75% ethanol and then centrifuged at 7500xg for 5 minutes. After careful removal of the ethanol, the RNA pellet was air dried and then re-suspended in 0.02 mL of nuclease/RNase-free water. Finally, dissolution of the RNA was facilitated by momentarily placing the solution in a heat block at 55°C.

RNA was quantified using a ND-1000 spectrophotometer (Thermo Fischer Scientific, Wilmington, DE) and checked for quality using an Agilent 2100 bioanalyzer (Agilent Technologies, Santa Clara, CA). The NanoDrop spectrophotometer provides an absorbance spectrum and a A260/A280 ratio used as a metric of quality. Pure RNA is marked by an A260/A280 ratio of 2.0. Ratios lower than 2.0 are indicative of contamination by proteins, phenols, salts, and other molecules. Generally, an A260/A280 ratio of 1.8-2.0 was accepted to be RNA of good quality and was used for cDNA synthesis.

The RNA concentration was normalized to 50-500 ng/µL and assessed for quality using an Agilent 2100 bioanalyzer (Agilent Technologies, Santa Clara, CA). The electropherogram obtained using Agilent software (2100 Expert Software) provided information of the RNA electrophoretic trace, which is used to assign an RNA Integrity Number (RIN) from 0-10 as a metric for RNA quality. RNA samples with RIN values of
7 or higher are suitable for cDNA synthesis and real time quantitative PCR (RT-qPCR). After quality assessment, the isolated RNA was diluted to ~200-400 ng/µL in preparation for first strand synthesis.

**cDNA Synthesis**

A 1-2 µg aliquot of isolated RNA was reverse transcribed to quantifiable cDNA using reagents provided in the Human and Mouse LncProfilers™ qPCR Array Kits (System Biosciences, Mountain View, CA. #RA930A-1). A schematic of the LncProfiler cDNA reaction setup is included in Appendix B.Outlined below in Figure 8 is the detailed mechanism by which LncRNAs are reverse transcribed to cDNA for use in LncProfiler qPCR arrays. LncRNAs are first polyadenylated using a polyA polymerase, which enhances the synthesis of cDNA from lncRNA 100-fold. Polyadenylation further promoted the use of small RNA elements as endogenous reference RNA controls for each array. After incubation at 37˚C for 30 minutes in a GeneAmp PCR System 9700 thermal cycler (Applied Biosystems), oligo dT adaptors are annealed to the site of polyadenylation at 60˚C for 5 minutes, followed by a cooling period for 2 minutes at room temperature. Finally, a random primer mix and reverse transcriptase is added to the reaction cocktail at 42˚C for 60 minutes to convert the lncRNA to cDNA. Nucleotides are recruited to polymerize the complementary strand in the 5’ to 3’ direction. The newly synthesized cDNA was stored at -20˚C until used with the LncProfiler qPCR array.
Figure 8. Synthesis of LncRNA cDNA for use in LncProfiler qPCR array. A three-step protocol required the polyadenylation of LncRNA, annealing of adaptors, and final conversion to cDNA mediated by the reverse transcriptase enzyme (Adapted from System Biosciences).
Real-Time Quantitative Polymerase Chain Reaction (RT-qPCR)

RT-qPCR reactions were prepared by aliquoting 4 µL of primers into 96-well plates. A qPCR reaction master mix was prepared by aliquoting 1750 µL of 2X SYBR Green master mix (Qiagen), 20 µL of LncRNA cDNA, and 1480 µl of nuclease free water into a reservoir. A volume of 26 µL of the RT-qPCR mastermix was transferred into each well of the 96-well LncProfiler qPCR array that contained IncRNA specific primers, with five endogenous reference RNA controls. LncProfiler array reactions were carried out using a Bio-Rad IQ5 thermal cycler (Hercules, CA) with the thermal cycling protocol shown below in Table 1.

Table 1. Thermal cycling protocol used for RT-qPCR using Bio-Rad IQ5.

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Repeats</th>
<th>Step</th>
<th>Time (min)</th>
<th>Temperature (˚C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2:00</td>
<td>50.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>10:00</td>
<td>95.0</td>
</tr>
<tr>
<td>2</td>
<td>40</td>
<td>1</td>
<td>0:15</td>
<td>95.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>1:00</td>
<td>60.0 *</td>
</tr>
</tbody>
</table>

*PCR/Melt Data acquisition

Data Analysis

LncRNA expression was evaluated in data segregated by Nrf2 genotype and/or sex. Raw LncRNA expression data was processed using RT² Profiler PCR Array Data Analysis software, version 3.5 (SABiosciences, Valencia, CA). This open source software utilizes CT values of qPCR raw data to calculate $2^{-\Delta\Delta CT}$, giving net fold regulation of each IncRNA. Beta actin was consistently used as the endogenous reference RNA control. Statistical significance was determined using a p-value <0.05. Select lncRNA displaying significant log-fold regulation will be scrutinized for discussion.
RESULTS

RNA Isolation

RNA was isolated using a standard liquid-liquid extraction protocol with Trizol reagent, which contains guanidinium thiocyanate-phenol-chloroform. After quantification of samples on an ND1000 spectrophotometer and evaluation of extracted RNA, an Agilent 2100 Bioanalyzer for quality, RNA quality was assessed based on the RNA Integrity Number (RIN). Obtained bioanalyzer results are in the form of electropherograms, showing fluorescence units (FU) plotted against time. An electropherogram generated for eukaryotic total RNA shows two strong peaks, corresponding to 18S and 28S ribosomal bands. A decrease in signal intensity of the 28S band is indicative of RNA degradation. RIN values are calculated based on the ratios of signal intensities and areas observed through the electrophoretic trace. Bioanalyzer results showed that RNA was extracted successfully with acceptable quantity and quality (Figure 9), with an average RIN of 9.58 and standard deviation of 0.22.

![Electropherograms](image)

**Figure 9.** Electropherograms of RNA isolated from lung tissue of 10-month old progeny born to SFN treated dams. The first peak (18S), second peak (28S), and RIN value are metrics of RNA quality.
Fold Regulation of LncRNA

This study was specifically focused on potential changes in lncRNA expression from *in utero* and lactational exposure to SFN, and not DBC. Changes in LncRNA abundance in 10-month-old progeny (lung tissue) born to pregnant dams given a SFN diet were analyzed using a commercially available LncProfiler qPCR array with an assay for 90 different cancer associated lncRNAs. Significant fold regulation (p<0.05) was seen in various lncRNAs when results were separated by gender and *Nrf2* genotype. Fold regulation results were normalized using a beta actin control included in the assay. All lncRNA changes observed in the prepared figures were significant as determined using a p-value <0.05.

SFN treatment was associated with up-regulation of rnascRNA, MSUR1 and Y-RNAs in 10-month-old progeny (Figure 10 and 11). Furthermore, significant down-regulation in LincENC1, LincRNA cox2, and H19 antisense were observed (Figure 10 and 11). When segregating data based on *Nrf2* genotype, rnascRNA and MSUR1 were up-regulated in mice with genotype *Nrf2*+/− (het) or *Nrf2*−/− (null). This is with the exception that MSUR1 was not significantly changed in male progeny with *Nrf2*+/− genotype (Figure 12). The greatest change in MSUR1 was observed in female *Nrf2*+/− progeny, with a fold regulation >10 (Figure 15). Segregating fold regulation data based on *Nrf2* genotype led to the detection of significantly changed lncRNA that were otherwise not detected in the analysis when pooling animals together based on gender. For instance, up-regulation of zfas1 was only seen in female progeny with *Nrf2*+/− genotype (Figure 14). A consideration of the functional roles and consequences of a subset of these lncRNAs is considered in the discussion.
Figure 10. LncRNA fold regulation in 10-month old male progeny exposed to SFN transplacentally during gestation and through lactation. Fold regulation is compared to a beta-actin control. (p<0.05)

Figure 11. LncRNA fold regulation in 10-month old female progeny exposed to SFN transplacentally during gestation and through lactation. Fold regulation is compared to a beta-actin control. (p<0.05)
**Figure 12.** LncRNA fold regulation in 10-month old male progeny *heterozygous* for *Nrf2* (*Nrf2<sup>+/−</sup>*) exposed to SFN transplacentally during gestation and through lactation. Fold regulation is compared to a beta-actin control. (p<0.05)

**Figure 13.** LncRNA fold regulation in 10-month old male progeny *null* for *Nrf2* (*Nrf2<sup>−/−</sup>*) exposed to SFN transplacentally during gestation and through lactation. Fold regulation is compared to a beta-actin control. (p<0.05)
**Figure 14.** LncRNA fold regulation in 10-month old female progeny heterozygous for Nrf2 (Nrf2<sup>+/−</sup>) exposed to SFN transplacentally during gestation and through lactation. Fold regulation is compared to a beta-actin control. (p<0.05)

**Figure 15.** LncRNA fold regulation of MSUR1 in 10-month old female progeny null for Nrf2 (Nrf2<sup>−/−</sup>) exposed to SFN transplacentally during gestation and through lactation. Fold regulation is compared to a beta-actin control. (p<0.05)
DISCUSSION

This study specifically focused on potential changes in lncRNA expression from *in utero* and lactation exposure to SFN, and not DBC. The archived lung tissue used for this analysis was collected from animals used in a larger study investigating transplacental carcinogenesis of DBC and the potential for SFN to provide chemoprotection to the fetus. A similar lncRNA analysis using archived neonate lung tissue was done for all treatment groups. Thus, the results of this study contribute to the objective of comparing lncRNA expression profiles in neonate and adult progeny, determining if such changes initiated in early development are sustained through adulthood.

LncRNA and its connection to SFN is seldom researched. However, a recent study by Beaver et al. (2017) indicated that certain lncRNAs are a chemopreventive target of SFN. Thus, SFN plays an important role in genetic regulatory mechanisms to exert its effect. In a human prostate cancer cell line, the majority of lncRNAs significantly changed were primarily decreased in expression. SFN treatment dysregulated the expression of eight lncRNA including Mir22HG, LINC01351, LINC00883, and LINC01116. Further investigation of LINC01116 by suppressing it with siRNA revealed its role in promoting growth of the prostate cancer cell line PC-3 (Beaver et al., 2017).

*MSUR1*

MSUR1 (mutant SOD1-up-regulated RNA 1) protects cells in a dose dependent manner by decreasing levels of hydroxyl radicals and oxidation of proteins. Analysis of the MSUR1 sequence has been shown to exhibit homology with 18S rRNA. Due to its
rRNA-like sequence, MSUR1 may be functioning as a \textit{cis}-regulatory element that interacts with rRNA or ribosomal proteins to regulate translational efficiency (Chang et al., 2008). The process by which such transcripts originate has yet to be elucidated. A study done by Chang et al. at Ohio State University showed that the protective effects of MSUR1 were amplified in cell lines mutated for superoxide dismutase 1 (SOD1 G93A) (Chang et al., 2008). Free radical levels were significantly reduced in SOD1-mutant cells with up-regulated MSUR1 expression.

In our study, MSUR1 levels were up-regulated in both male and female mice. The largest fold regulation was seen in female mice heterozygous for \textit{Nrf2}, with almost a 15-fold difference compared to the control. Female mice with a \textit{Nrf2} knockout also exhibited MSUR1 upregulation, but not of the same magnitude as female heterozygous mice. Thus, the increased expression of MSUR1 may be \textit{Nrf2}-mediated, when in the presence of the SFN agonist. This result appears to be amplified to a greater extent in females than in males.

\textit{Foxn2-as, Gli2-as, H19-as}

Antisense (\textit{as}) transcripts are a sub-class of non-coding RNA elements involved in the \textit{cis} or \textit{trans} regulation of genes. Antisense IncRNAs are derived from the transcription of the strand opposite to a protein-coding gene. \textit{Cis}-acting antisense IncRNA modulate gene expression by interacting with a gene in proximity to the same DNA region. \textit{Trans}-acting antisense IncRNA exert their effect on different and oftentimes on other chromosomes. Studies have shown that 63\% of transcripts are accompanied by antisense non-coding RNA elements (Villegas and Zaphiropoulos, 2015). It was once believed that antisense RNA was transcriptional noise. However, such
a notion is becoming less accepted as we elucidate their role in biologically mobilizing RNA at various points in the eukaryotic life cycle. Foxn2-as, Gtl2-as, and H19-as have all been determined to function as tumor suppressors (Jiang and Bikle, 2013). Specifically, Gtl2-as is a maternally imprinted lncRNA tumor suppressor. The observed down-regulation of these tumor suppressors by SFN could potentially be a contributing factor to the higher multiplicity of tumors seen in the DBC/SFN treatment. Specifically, SFN enhanced tumorigenicity in Nrf2+/− mice born to pregnant dams given the phytochemical through their diet. Although further experimentation is required, expression differences seen in the SFN/CO treatment when compared to the control could be Nrf2-dependent and correlated with enhanced expression of the pathway by SFN. H19 antisense lncRNA, primarily residing on the maternal chromosome, is also functionally trans-acting and activates the expression of insulin-like growth factor 2 (IGF2) (Berteaux et al., 2008).

### Zfas1

Many lncRNAs have been discovered to have an associated role with post-pubertal development of mammary glands. Recent microarray analysis has revealed the Zfas1 lncRNA to be in great abundance within developing mammary gland tissue. Zfas1 is classified as an antisense RNA that promotes epithelial proliferation and differentiation in mammary glands of mice (Hansji et al. 2016). The Nrf2 pathway has been revealed to play a critical role in the self-renewal of mammary stem cells (Zhang et al. 2014). Specifically, researchers utilized a Nrf2 knockout mouse model to implicate lncRNA ROR as an intermediary in this role. In our transplacental study, we observed significant up-regulation of Zfas1 in female mice heterozygous for Nrf2 and exposed to SFN in
uterus. We did not see this expression difference in Nrf2 knock-out female mice. Although this must be scrutinized further, up-regulation of the Nrf2 pathway by SFN may play an important role in the increase of Zfas1, a lncRNA important to the maintenance of mammary stem cell pluripotency.

The Nrf2 factor is most familiar to researchers as the master regulator that controls the mammalian defense mechanism against chemical carcinogenesis. Such exquisite complexity has many wondering if such a regulatory pathway functions under the influence of non-coding RNA elements of the transcriptome. Although lncRNA are transiently expressed, our understanding of their role in the Nrf2 response will enhance our knowledge surrounding the initiation and propagation of various cancers. Lastly, our potential ability to modulate lncRNA levels in vivo could provide a new avenue for cancer therapies and chemoprevention.
BIBLIOGRAPHY


APPENDICES
Appendix A - LncProfiler cDNA reaction setup (System Biosciences)

**Start:**
In a thin-walled PCR tube or PCR-compatible plate well combine:

**STEP 1:** PolyA Tail
5 µl Total RNA (~1-2 µg)  
2 µl 5X PolyA Buffer  
+ 1 µl 25mM MnCl₂  
1.5 µl 5mM ATP  
0.5 µl PolyA Polymerase  
10 µl  
Incubate for 30 min. at 37°C

**STEP 2:** Anneal Anchor dT Adapter
Add: + 0.5 µl Oligo dT Adapter  
Heat for 5 min. at 60°C  
Let cool to room temp for 2 min.

**STEP 3:** Synthesize cDNAs
Add: 4 µl 5X RT Buffer  
2 µl dNTP mix  
+ 1.5 µl 0.1M DTT  
1.5 µl Random Primer mix  
1 µl Reverse Transcriptase  
20.5 µl Total in tube  
Incubate for 60 min. at 42°C  
Heat for 10 min. at 95°C
Appendix B – 96-well plate LncProfiler RT-qPCR array (System Biosciences)

Endogenous controls (green):

**18S rRNA** – RNA structural component of small subunit in eukaryotic cytoplasmic ribosomes.

**RNU43 (snoRNA)** – small nucleolar RNA important to the modification and processing of ribosomal RNA.

**GAPDH** – Glyceraldehyde-3-phosphate dehydrogenase; important in the aerobic catabolism of glucose to pyruvate.

**Beta Actin** – One of six isoforms of actin. Beta-actin is cytoskeletal actin not found in the sarcoplasm of muscle.

**U6 snRNA** – small nuclear RNA, which plays a functional role in pre-mRNA splicing.