A SYSTEMS BIOLOGY-BASED APPROACH TO INVESTIGATE FORMALDEHYDE'S EFFECTS ON MICRORNA EXPRESSION PROFILES

Julia Elizabeth Rager

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Environmental Sciences and Engineering of the Gillings School of Global Public Health

Chapel Hill
2013

Approved by:

Dr. Rebecca C. Fry

Dr. David Diaz-Sanchez

Dr. Ilona Jaspers

Dr. Kenneth G. Sexton

Dr. James A Swenberg

ABSTRACT

JULIA RAGER: A Systems Biology-Based Approach to Investigate Formaldehyde's Effects on MicroRNA Expression Profiles
(Under the Direction of Dr. Rebecca C. Fry)

Formaldehyde is a common indoor and outdoor air pollutant that adversely impacts global public health. Many toxicological studies have shown that formaldehyde causes nasopharyngeal cancer, possibly through tissue damage, increased cell proliferation, and/or DNA damage. However, there is lack of knowledge regarding formaldehyde's effects at the systems biology level and whether epigenetic mechanisms may contribute to cellular responses. Furthermore, whether formaldehyde is capable of altering genomic and epigenomic processes throughout sites distal to the respiratory tract is unknown. This topic is of high interest, as the link between formaldehyde inhalation exposure and leukemia development is currently under heated debate. Epidemiological studies have shown evidence supporting a link between formaldehyde exposure and leukemia development, while toxicological investigations have yet to provide evidence supporting formaldehyde's ability to influence sites distant to the respiratory tract. Before this dispute is resolved, further evaluation of the biological mechanisms linking formaldehyde to disease is clearly necessary. In particular, formaldehyde-induced changes to epigenetic contributors to transcriptional programs are extremely understudied, where microRNA (miRNA) expression profiles have yet to be investigated in relation to formaldehyde.

We set out to test the novel hypothesis that miRNAs have altered expression profiles within the respiratory and hematopoietic systems upon exposure to formaldehyde. Our studies were the first to show that formaldehyde significantly disrupts miRNA expression patterns *in vitro*, within cultured human lung cells, and *in vivo*, within the nasal epithelium of nonhuman primates. Using a rodent model, the impact of formaldehyde exposure on miRNA-related processes in direct contact and distant tissues, including the nasal mucosa, circulating white blood cells, and bone marrow, was evaluated. Formaldehyde was found to significantly alter

miRNA expression profiles within the nose and blood, but not the bone marrow. Evaluating the epigenetic effects of formaldehyde exposure at the systems biology level, putative miRNA-mediated responses were mapped onto interacting networks. Signaling related to inflammation, cell death, and cancer was identified as enriched. Taken together, our research increases the knowledge of under-studied mechanisms linking formaldehyde exposure to disease, acting as an important foundation for future research in public health and toxicology.

ACKNOWLEDGEMENTS

I would first like to thank my advisor, Dr. Rebecca C. Fry, for welcoming me into her research group at UNC. She has made a huge impact on my life, and it is difficult to think where I would be without her mentorship these past four years. Her enthusiasm, leadership, and patience are truly commendable, to say the least. From her, I've learned ways to succeed at work while juggling family and friends. It has been a pleasure seeing Dr. Fry's lab grow to the success that it is today, and being able to play a role in that growth is extremely gratifying. I will forever be grateful to her for making my graduate experience at UNC so incredibly meaningful.

I would also like to thank our Research Specialist and Lab Manager, Lisa Smeester, who dedicated much of her time to helping me plan experiments and learn valuable laboratory techniques.

I would also like to thank my committee members, Dr. David Diaz-Sanchez, Dr. Ilona Jaspers, Dr. Kenneth G. Sexton, and Dr. James A. Swenberg, for their valuable input and advice in preparing my research project and dissertation.

Lastly, I would like to thank my parents, Brent and Diane Rager, who have been extremely supportive of me throughout the years. I appreciate their ongoing help and guidance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Understanding the biological impacts upon exposure to formaldehyde via inhalation exposure is crucial, as formaldehyde in a ubiquitous air pollutant present throughout the environment. The combined studies in this work take several approaches to investigate novel mediators of response to formaldehyde exposure. The ultimate goal for this investigation is to increase the current knowledge of mechanisms and biological pathways underlying formaldehyde-induced effects, acting as an important foundation for future research in public health and toxicology.

Formaldehyde Exposure Sources

Formaldehyde is a common air toxic that is present in both indoor and outdoor atmospheres. In outdoor environments, formaldehyde is present due to direct emissions from anthropogenic and biogenic sources, and is also formed as a secondary chemical product through hydrocarbon atmospheric chemistry (WHO 2001). Anthropogenic sources of formaldehyde include automobile exhaust, power plants, manufacturing facilities, and incinerators (NTP 2011; WHO 2001). Ambient air is estimated to contain formaldehyde at levels between 0.0008 and 0.02 ppm (WHO 2001). High formaldehyde exposures can occur within indoor environments, where formaldehyde is released from household products (e.g. cleaning agents, carpet, furniture) and cigarette smoke (NTP 2011). The highest formaldehyde levels are found in certain occupational environments. For example, high chronic exposures of 2-5 ppm formaldehyde have been measured in garment and textile industries, during the varnishing of furniture and wooden floors, and in some manufacturing jobs related to board mills and foundries (Duhayon et al. 2008). Acute exposures to high levels of formaldehyde ($\geq$ 3 ppm) have been reported for pathologists, embalmers, and paper workers (Duhayon et al. 2008). A range of lower formaldehyde levels have also been measured in industries related to the production of resins, plastics, fibers, abrasives, and rubber (Duhayon et al. 2008).

It is also important to note that formaldehyde is produced endogenously in all cells, as it is an intermediate of serine, glycine, methionine, and choline metabolism (IARC 2006). It is also an essential intermediate formed during the biosynthesis of purines, thymidine, and certain amino acids (IARC 2006). Because of the constant presence of both endogenous and environmental formaldehyde exposure, understanding the exposure response and biological basis of formaldehyde-induced health effects is of utmost importance.

Mechanisms Underlying Formaldehyde-Induced Cancer of the Upper Respiratory Tract

The exact mechanisms linking formaldehyde exposure to cancer remain unknown. Formaldehyde is currently classified as a known human carcinogen (IARC 2006), although the entire mechanism by which formaldehyde induces cancer is not fully understood (Liteplo et al. 2003; NTP 2011). The genotoxic effects of formaldehyde exposure on direct target cells within the respiratory tract have been studied extensively. Specifically, formaldehyde has been shown to cause a variety of types of genetic damage *in vitro* and *in vivo*. These types of damage include DNA-protein crosslinks, which have been detected *in vitro* (Merk et al. 1998), in the nasal mucosa of rats (Casanova et al. 1989; Casanova et al. 1994), and in the nasal turbinates, nasopharynx, larynx, trachea, carina, and proximal bronchi of nonhuman primates (Casanova et al. 1991) exposed to formaldehyde. DNA double strand breaks have also been shown to result from formaldehyde exposure *in vitro* (Noda et al. 2011). Formaldehyde also causes a multitude of DNA adducts to form *in vitro*, including N^2-hydroxymethyl-deoxyguanosine, N^6-hydroxymethyl-deoxyadenosine, and N^4-hydroxymethyl-deoxycytosine (Swenberg et al. 2011). N^2-hydroxymethyl-deoxyguanosine adducts have also been found to be induced exogenously in the nasal mucosa of rats (Lu et al. 2011) and nonhuman primates (Moeller et al. 2011) exposed to formaldehyde.

Formaldehyde-induced genetic damage can lead to cell death. For example, *in vitro* studies show that the amount of formaldehyde-induced DNA-protein crosslinks inversely relates to cell survival (Merk et al. 1998; Ross et al. 1980). Furthermore, a pharmacodynamic model has been created to describe the relationship between formaldehyde exposure, DNA-protein crosslink formation, and DNA replication arrest using in vivo measurements (Heck et al. 1999). During DNA replication arrest, lesions can be repaired. However, if the damage is not repaired properly, DNA damage-response signaling triggers cell death by apoptosis or cellular senescence

(Jackson et al. 2009). Alternatively, carcinogenesis can occur from the activation of an oncogene or the inactivation of a tumor suppressor gene (Jackson et al. 2009). Altogether, the various types of formaldehyde-induced DNA damage may underlie the observed increases in cytotoxicity *in vitro* (Lovschall et al. 2002; Merk et al. 1998; Quievryn et al. 2000) and tissue damage *in vivo* (Chang et al. 1983; Monticello et al. 1991; Monticello et al. 1996).

Cytotoxic damage to target cells is known to cause regenerative cell proliferation to replace and repair damaged tissue (Boobis 2010; Cohen et al. 2008). Increases in the rate of cellular proliferation can then increase the probability of *de novo* mutations in critical oncogenes (Boobis 2010). In the case of formaldehyde, it is postulated that regenerative cell proliferation resulting from formaldehyde-induced cytotoxicity increases the number of DNA replications, and thereby, increases the likelihood of DNA damage causing DNA replication errors, resulting in mutations and eventually carcinogenesis (Liteplo et al. 2003; McGregor et al. 2006). This proposed mode of action is supported by previous investigations showing that formaldehyde causes sustained cell proliferation in direct target tissue *in vivo* (**Figure 1**). For example, many studies show that formaldehyde exposure increases cell proliferation rates in the upper respiratory epithelium of rodents and nonhuman primates (Chang et al. 1983; Monticello et al. 1989; Monticello et al. 1991; Roemer et al. 1993). Furthermore, sites of increased cellular proliferation rates have been shown to correlate with regions of nasal tumor incidence (Monticello et al. 1991; Monticello et al. 1996).

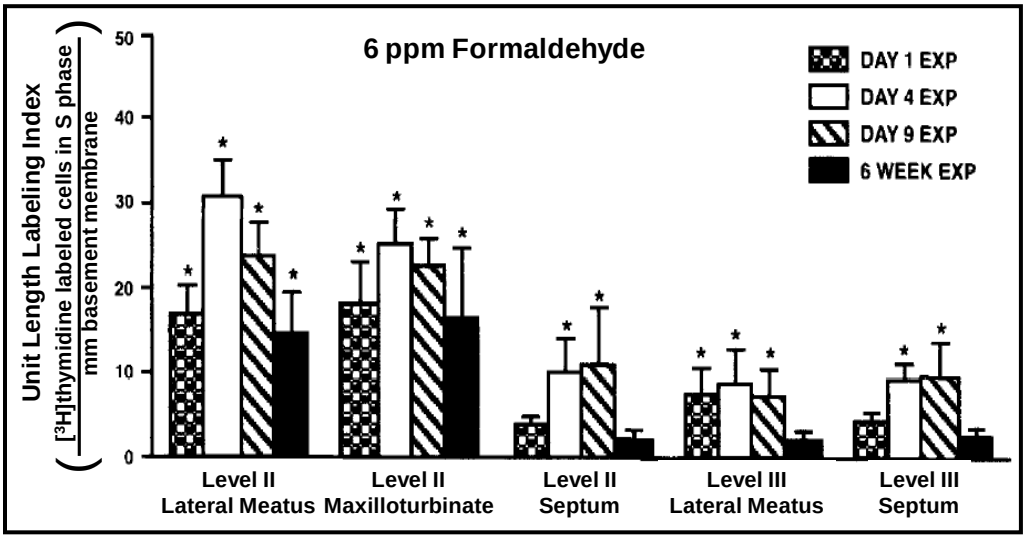

Figure 1: As a part of formaldehyde's carcinogenic mode of action, formaldehyde causes significant increases in cell proliferation in direct target cells (modified from Monticello et al. 1991). Increases in cell proliferation rates, as measured using labeling indices, across several nasal passage levels were found in rats exposed to formaldehyde (Monticello et al. 1991).

While there is a link between cell proliferation and carcinogenesis, other mechanisms influencing formaldehyde-induced cancer are understudied, including signaling pathways that may play a role in altered cell proliferation and carcinogenesis. Our study, therefore, contributes to the current understanding of mechanisms underlying cancer caused by formaldehyde by evaluating epigenetically-regulated processes at the systems-level. Our project is the first to use a systems biology approach to assess the epigenetic effects of formaldehyde on direct target cells within the respiratory tract.

Incongruent Findings Between Epidemiological and Toxicological Studies

There is disparity between epidemiological and toxicological findings regarding possible links between formaldehyde and leukemia. Formaldehyde is a known human carcinogen (IARC 2006), but its association with hematological cancers is currently undergoing debate. In June 2011, the U.S. National Toxicology Program added formaldehyde as a human lymphohematopoietic carcinogen, while stating that the mechanisms by which formaldehyde causes leukemia are unknown (NTP 2011). This classification is largely based on epidemiological evidence supporting formaldehyde-induced myeloid leukemia, as identified in cohorts of embalmers (Hauptmann et al. 2009), garment workers (Pinkerton et al. 2004), and workers in various formaldehyde-related industries (Beane Freeman et al. 2009; Zhang et al. 2009). However, some of these epidemiological findings are controversial (Bachand et al. 2010). These types of occupational exposures are common, where more than two million U.S. workers are exposed to formaldehyde (USDL 2011). Because formaldehyde exposure and leukemia development are both prevalent, understanding the biological basis linking formaldehyde to disease is extremely important.

Despite epidemiological evidence supporting formaldehyde as a leukemogen, the biological plausibility underlying formaldehyde-induced leukemia is still debated. To elaborate, leukemia is a cancer of the blood or bone marrow, where recognized environmental leukemogens typically cause hematopoietic toxicity/genotoxicity and subsequent leukemia development (McHale et al. 2012; Mukherjee et al. 2012). However, formaldehyde is reactive and undergoes metabolism rapidly (IARC 2006), and formaldehyde blood concentrations do not change after inhalation exposure (Casanova et al. 1988; Heck et al. 1985). As a result, some scientists believe it is improbable for such a compound to cause toxicity at sites distant from the respiratory tract

(Cole et al. 2004; Golden et al. 2006; Heck et al. 2004; Pyatt et al. 2008). Nevertheless, a few mechanisms underlying formaldehyde-induced leukemia have been proposed (Goldstein 2011; Zhang et al. 2009). One key event common amongst these proposed mechanisms requires damages or alterations within hematopoietic stem or progenitor cells. Still, researchers have yet to elucidate how formaldehyde, either directly or indirectly, may alter hematopoietic cells. With the goal of filling a component of this research void, our study investigates changes in miRNA signaling throughout the body, representing an area of research that has yet to be evaluated in relation to formaldehyde.

MicroRNAs

MicroRNAs are important epigenetic regulators of gene expression that may play key roles in formaldehyde-induced health effects. These recently discovered molecules, clearly a part of the epigenetic machinery (Iorio et al. 2010), play a large role in the regulation of mRNA abundance and protein production. By base pairing to target mRNAs, miRNAs can cause mRNA degradation and/or translational repression (Filipowicz et al. 2008; Friedman et al. 2009). In some cases, miRNAs can even cleave newly translated proteins (Friedman et al. 2009) (**Figure 2**). To quantify, mammalian miRNAs are estimated to regulate more than 60% of all protein-coding genes (Friedman et al. 2009). Because miRNAs play such pivotal roles in gene regulation, miRNAs have received increasing attention throughout medical and toxicological research fields. Several cancer-related studies have shown that miRNA expression profiles are drastically altered in tumors in comparison to healthy tissue. For example, miRNA expression profiles have been shown to be significantly disrupted in nasopharyngeal carcinoma (Chen et al. 2009) and leukemia (Bousquet et al. 2008; Cammarata et al. 2010; Garzon et al. 2008; Wang et al. 2011). Whether formaldehyde exposure is capable of affecting miRNA expression profiles, which may ultimately influence disease, is currently unknown.

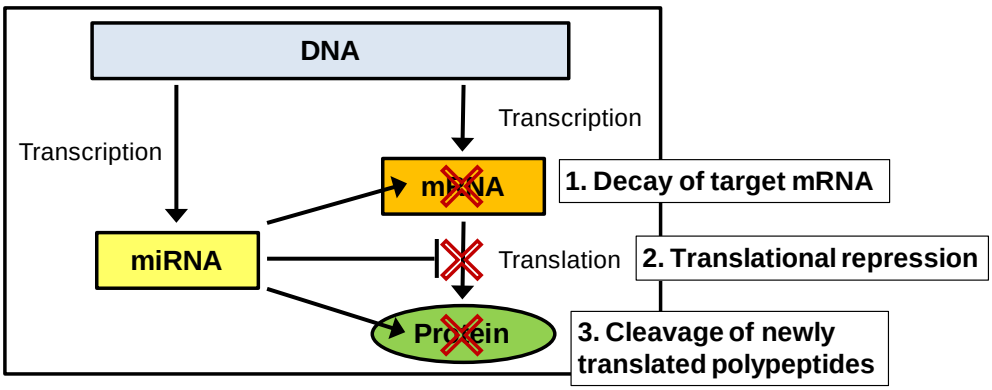

Figure 2: MicroRNAs are important regulators of gene expression and protein production. Three routes are illustrated through which miRNAs can influence gene expression and protein production.

MicroRNAs and Leukemia

Leukemia development is known to be influenced by miRNAs. Hematological function is heavily influenced by miRNAs, since these molecules are important regulators of hematopoietic stem/progenitor cell differentiation (Marcucci et al. 2011; Yendamuri et al. 2009), cell cycle (Han et al. 2010), and apoptosis (Garzon et al. 2009). Distinct miRNA expression profiles also exist in leukemia patients, where miRNAs have been shown to classify various risk groups (Calin et al. 2006; Marcucci et al. 2009; Marcucci et al. 2011). Further linking miRNAs to leukemia, several miRNAs have been implicated as leukemia-related tumor suppressors (e.g. miR-29b) (Garzon et al. 2009) and oncogenes (e.g. miR-125b, miR-155, miR-29a) (Costinean et al. 2006; Han et al. 2010; O'Connell et al. 2008). Because of the current interest regarding formaldehyde's link to leukemia, we find the lack of research on formaldehyde's epigenetic effects surprising. We therefore address this scientific gap by investigating formaldehyde-altered miRNAs throughout multiple target tissues, revealing novel responses that have yet to be investigated.

Project Approach

This research focuses on epigenetic responses to formaldehyde inhalation exposure across direct contact and distant tissues. **The primary hypothesis to be tested is that miRNAs have altered expression profiles within the respiratory and hematopoietic systems upon exposure to formaldehyde.**

The project employs an integrated approach to perform cross-tissue and cross-species comparative analyses of epigenetic responses to formaldehyde exposure (**Figure 3**). We assess miRNA expression profiles across three separate species in multiple tissues and cells, representing possible key events linking formaldehyde inhalation exposure to respiratory and hematopoietic disease. It is notable that our research is the first to: (i) investigate miRNA expression profiles in formaldehyde-exposed cells, (ii) to compare formaldehyde-induced miRNA responses across three species, and (iii) to compare miRNA expression profiles between direct contact and distant targets of formaldehyde inhalation exposure using the rodent model. Using this strategy, we reveal novel mechanisms underlying formaldehyde exposure-induced effects on biological signaling.

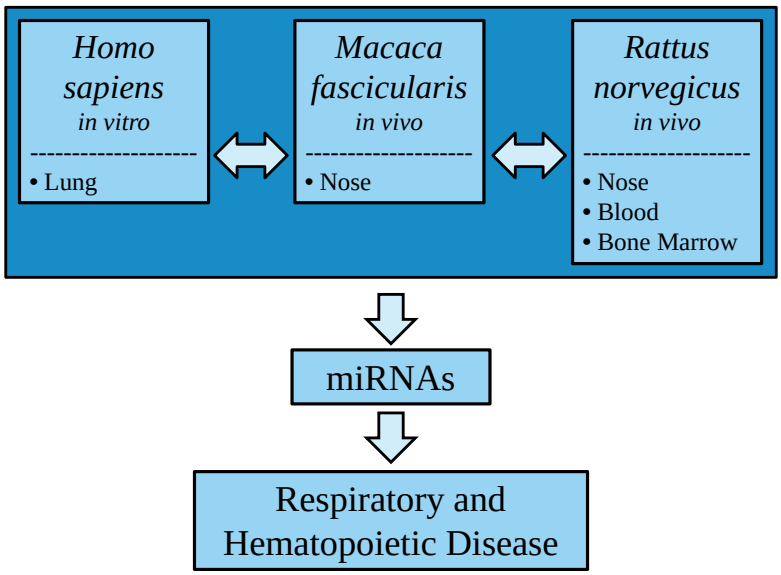

Figure 3: Project overview. The effect of formaldehyde inhalation exposure on miRNA expression profiles is assessed across three species. Using this experimental design, this research tests the novel hypothesis that miRNAs have altered expression profiles within the respiratory and hematopoietic systems upon exposure to formaldehyde.

Dissertation Organization

This dissertation is organized into three chapters. The first chapter describes an *in vitro* study which was the first to show that formaldehyde exposure significantly alters miRNA expression profiles. The study detailed in the second chapter expanded the initial *in vitro* findings using an *in vivo* model. Here, nonhuman primates exposed to formaldehyde were found

to also show significantly altered miRNA expression profiles within the nose, a direct target of formaldehyde inhalation exposure. The last study, described in chapter three, compares miRNA expression profiles altered by formaldehyde exposure across the nose, circulating white blood cells, and bone marrow cells in a rodent model. It is of great interest that miRNA expression patterns were identified as disrupted by formaldehyde exposure in the nose and white blood cells, but not the bone marrow. This study also included a time-series analysis, where formaldehyde was found to disrupt miRNA expression profiles more drastically after shorter periods of exposure. In all three studies, miRNA responses were further evaluated at the mechanistic level by mapping results to the biological networks. Significant pathways involved in cellular regulation were identified as likely disrupted via formaldehyde's influence on miRNA expression profiles. The studies contained within these three chapters provide novel insights into the mechanisms and biological pathways underlying formaldehyde-induced effects.

CHAPTER 1
FORMALDEHYDE EXPOSURE ALTERS MICRORNA EXPRESSION PROFILES IN HUMAN LUNG CELLS

1.1 Overview

Exposure to formaldehyde, a known air toxic, is associated with cancer and respiratory disease. Despite its adverse health effects, the mechanisms underlying formaldehyde-induced disease remain largely unknown. Research investigations have uncovered microRNAs (miRNAs) as key post-transcriptional regulators of gene expression that may influence cellular disease state. While studies have compared different miRNA expression patterns between diseased and healthy tissue, this is the first study to examine perturbations in global miRNA levels resulting from formaldehyde exposure.

We set out to investigate whether cellular miRNA expression profiles are modified by formaldehyde exposure in human lung cells. We hypothesized that formaldehyde exposure disrupts miRNA expression levels within lung cells, representing a novel epigenetic mechanism through which formaldehyde may induce disease.

Human lung epithelial cells were grown at air-liquid interface and exposed to gaseous formaldehyde at 1 ppm for 4 hours. Small RNAs and protein were collected and analyzed for miRNA expression using microarray analysis or IL-8 protein levels by ELISA, respectively.

Gaseous formaldehyde exposure altered the miRNA expression profiles in human lung cells. Specifically, 89 miRNAs were significantly down-regulated in formaldehyde exposed samples versus controls. Functional and molecular network analysis of the predicted miRNA transcript targets revealed that formaldehyde exposure potentially alters signaling pathways associated with cancer and inflammatory response. IL-8 release was increased in cells exposed to formaldehyde, and results were confirmed by real-time PCR.

Formaldehyde alters miRNA patterns which regulate gene expression, potentially leading to the initiation of a variety of diseases.

1.2 Study Objectives

For this study, we set out to test the novel hypothesis that formaldehyde exposure can disrupt miRNA levels within lung cells. We tested this hypothesis by exposing human lung epithelial cells to formaldehyde using a direct air-liquid interface that physically mimics the human respiratory tract. Using microarray analysis, we assessed genome-wide miRNA expression profiles and identified miRNAs altered at the expressed level by formaldehyde exposure. To expand our findings to the systems-level, we predicted transcriptional targets of the formaldehyde-altered miRNAs and mapped them onto molecular interaction networks. Here, critical biological pathways related to putative miRNA-mediated responses to formaldehyde were identified. Taken together, this research suggests a novel epigenetic mechanism by which formaldehyde may induce disease.

1.3 Materials and Methods

1.3.1 Cell Culture

Human A549 type II lung epithelial cells derived from a human lung adenocarcinoma were cultured according to standard protocol (ATCC). Cells were grown in growth media containing F-12K plus 10% FBS plus 1% penicillin and streptomycin. Cells were plated onto 24 mm diameter collagen-coated membranes with 0.4 μM pores (Trans-CLR; Costar, Cambridge, MA). Upon confluence, cells were cultured in phenol red-free F-12K nutrient mixture without FBS. Immediately prior to exposure, media above each membrane was aspirated in order to create direct air-liquid interface culture conditions. The media beneath each membrane remained to supply nourishment for cells throughout the exposure.

1.3.2 Formaldehyde Treatment

Gaseous formaldehyde was generated by heating 143 mg paraformaldehyde (Aldrich Chemical Company, Inc., Milwaukee Wi, lot no. 05910EI) in an air-flushed “U-tube” until the powder was completely vaporized within a dark un-irradiated 120 m^3 environmental chamber. The walls of the chamber are made of chemically non-reactive film, as detailed previously (Sexton et al. 2004). The chamber was naturally humidified from pre-flushing with HEPA

filtered ambient air during cloudy conditions. This resulted in a formaldehyde concentration of 1 ppm (1.2 mg/m^3) which was then drawn through a cellular exposure chamber (Billups-Rothenberg, Modular Incubator Chamber, Del Mar, CA) at 1.0 L/min. The exposure chamber was positioned within an incubator where CO_2 was added to the formaldehyde exposure source stream at 0.05 L/min and a small water dish provided proper humidification. Prepared lung cells were exposed to 1 ppm formaldehyde for 4 hours, while mock-treated control cells were exposed to humidified air under similar conditions. Experiments were carried out with six technical replicates for each exposure condition, generating a total of 12 samples. After nine hours, cells were scraped and stored at -80°C in TRIzol® Reagent (Invitrogen Life Technologies), and basolateral supernatants were aspirated and stored at -80°C.

1.3.3 Cytotoxicity Analysis

To measure formaldehyde exposure's cytotoxicity, the enzyme lactate dehydrogenase (LDH) was measured within the supernatant of each sample. Measurements were acquired using a coupled enzymatic assay, according to the supplier's instructions (Takara Bio Inc., Japan). LDH fold increase was calculated as $\mu_{LDH,\ FE} / \mu_{LDH\ C}$, where μ represents the mean LDH activity, FE represents formaldehyde exposed samples, and C represents controls.

1.3.4 Microarray Processing

RNA molecules of at least 18 nucleotides in length were isolated using Qiagen's miRNeasy® Kit according to the manufacturer's protocol (Qiagen, Valencia CA). RNA was quantified with the NanoDrop™ 1000 Spectrophotometer (Thermo Scientific, Waltham MA) and its integrity was verified with an Agilent Technologies 2100 Bioanalyzer (Santa Clara, CA). RNA was labeled and hybridized to the human miRNA microarray (version 1) manufactured by Agilent Technologies (Santa Clara, CA). This microarray measures the expression levels of 534 human miRNAs. Three of the six total samples from each exposure condition, three formaldehyde-exposed and three mock-treated samples, were hybridized using 400 ng of input RNA per sample. RNA labeling and hybridization were performed according to the manufacturer's protocol, and microarray results were extracted using Agilent Feature Extraction Software. Data were submitted to NCBI's Gene Expression Omnibus (GEO) database (www.ncbi.nlm.nih.gov/geo/) and are available under accession #GSE22365 (Edgar et al. 2002).

1.3.5 Microarray Analysis

The resulting expression levels for each of the miRNAs measured by the microarrays were calculated and filtered for miRNAs expressed above a background level (background was set at 30, approximating the median signal per array). This resulted in a reduction of probesets from 12033 to 4900 records. Differential miRNA expression was defined as a significant difference in miRNA expression levels between treated samples and untreated samples, where the following three statistical requirements were set: (1) fold change of ≥ 1.5 or ≤ -1.5 (treated versus untreated); (2) p-value < 0.05; and (3) false discovery rate (FDR) < 0.05. P-values and FDRs were generated using the Comparative Marker Selection tool in GenePattern (www.broadinstitute.org/cancer/software/genepattern/) (Reich et al. 2006). Here, 2000 permutation tests were carried out using the signal-to-noise (SNR) ratio analysis and smoothed p-values were determined for each miRNA. SNR is defined by the equation $SNR = (\mu_A - \mu_B) / (\sigma_A + \sigma_B)$, where μ represents average sample intensity and σ represents standard deviation (Golub et al. 1999). SNRs have been shown to provide one of the most accurate classification prediction methods (Cho et al. 2002). False discovery rates (FDRs) were calculated as the expected fraction of false positives among probesets reported as significant using the Benjamini and Hochberg procedure (Benjamini et al. 1995). Targets for the most differentially expressed miRNAs were identified using miRDB (www.mirdb.org) (Wang 2008) where targets with a score of >70 were investigated.

1.3.6 Enriched Biological Functions and Network Analysis

Enriched biological functions and molecular network analyses were performed using the Ingenuity database (Ingenuity® Systems, www.ingenuity.com, Redwood City, CA). The Ingenuity database provides a collection of gene to phenotype associations, molecular interactions, regulatory events, and chemical knowledge accumulated to develop a global molecular network. The lists of putative targets for each miRNA were overlaid onto this global molecular network, where protein networks significantly associated with the targets were algorithmically constructed based on connectivity. Associated enriched canonical pathways within these networks were also identified. Functional analysis was carried out to identify biological functions and disease signatures most significantly associated with the input targets. Statistical significance of each biological function or disease was calculated using a Fischer's

exact test. This test generated a p-value signifying the probability that each function or disease was associated with the miRNA targets by chance alone. Only enriched functions with p-values < 0.005 were assessed.

1.3.7 RT-PCR Verification of miRNA Expression

Expression levels of the five most significantly modified miRNAs were also tested using real-time reverse-transcriptase PCR (RT-PCR). The TaqMan® MicroRNA Primer Assays for hsa-miR-33 (ID 002135), hsa-miR-450 (ID 2303), hsa-miR-330 (ID 000544), hsa-miR-181a (ID 000516), and hsa-miR-10b (ID 002218) were used in conjunction with the TaqMan® Small RNA Assays PCR kit (Applied Biosystems). The Bio-Rad MyCycler Thermal Cyler was used for the reverse transcription step, and the Roche Lightcycler 480 was used for the real-time step. The same three control and three formaldehyde exposed samples from the microarray were used for RT-PCR, which was performed in technical duplicate. Statistical significance was evaluated using a t-test.

1.3.8 Interleukin-8 Measurement

The protein abundance of the cytokine interleukin-8 (IL-8) was measured using the basolateral supernatant from all 12 samples. A BD OptEI^{ATM} human IL-8 enzyme-linked immunosorbent assay (ELISA) was performed and analyzed according to the manufacturers' protocol (BD Biosciences, San Jose, California). Experiments were carried out with 12 technical replicates for each exposure condition. Scanned absorbance reading outliers were identified through the Grubbs' test (www.graphpad.com) where outliers were identified as those with less than a 5% probability of occurring as an outlier by chance alone, as based off a normal distribution (Grubbs 1969). IL-8 fold increase was calculated as $\mu_{IL\text{-}8\ FE} / \mu_{IL\text{-}8\ C}$, where μ represents the mean, FE represents formaldehyde exposed samples, and C represents controls. Statistical significance of the treated versus untreated IL-8 levels was calculated using a t-test with Welch's correction.

1.4 Results

1.4.1 Formaldehyde Exposure Modulates miRNAs in Human Lung Cells

In this study, we set out to identify whether formaldehyde exposure alters the expression levels of miRNAs in lung cells. Human lung epithelial cells (A549) were exposed to gaseous formaldehyde drawn directly from an un-irradiated (dark) environmental chamber into an exposure chamber or were mock-treated. This exposure resulted in a 6.68 fold increase in LDH release. Comparisons to cell viability demonstrate that this fold change in LDH is associated with minimal cell killing. After exposure, small RNAs were collected and their relative abundance measured using microarrays. A total of 343 unique miRNAs were detectable above background in these cells. The 343 miRNAs were further assessed for formaldehyde-induced changes in expression level. A total of 89 miRNAs showed a significant decrease in expression in the formaldehyde exposed lung samples compared to control samples (**Figure 4**, see **Supplementary Table 1**). There were no miRNAs identified with significantly increased expression levels in response to formaldehyde. The five most significantly differentially expressed miRNAs, as determined through microarray analysis, were miR-33 (FC -5.5), miR-450 (FC -3.6), miR-330 (FC -2.4), miR-181a (FC -2.1), and miR-10b (FC -2.1). Here, fold change (FC) represents the ratio of miRNA abundance in exposed relative to the control samples.

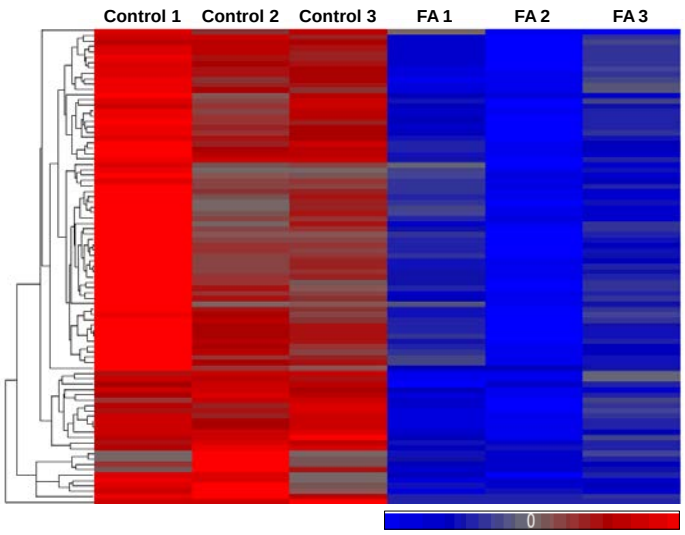

Figure 4: Formaldehyde modulates the expression of 89 miRNAs in human lung cells. A heat map displays the relative expression levels of the 89 miRNAs, where data are mean standardized and hierarchical clustering is performed. Blue indicates relative low expression while red indicates relative high expression. Formaldehyde-treated samples are abbreviated as FA.

1.4.2 miRNA Expression Changes are Validated through RT-PCR

RT-PCR was used to confirm the array-based findings, where the decreased miRNA expression induced by formaldehyde exposure was verified. Specifically, miR-330 showed a formaldehyde-induced FC of -1.3, miR-181a showed a FC of -7.4, miR-33 showed a FC of -1.2, and miR-10b showed a FC of -1.5 (**Figure 5**). miR-450 showed minimal expression changes with a FC of -1.04 (data not shown). As it was not validated with RT-PCR, further analysis on miR-450 was not performed. To assess the similarity of the RT-PCR and array-based expression level quantification, the average relative miRNA abundances were compared against the raw microarray expression levels. This analysis shows high correlations (0.81 for control samples, 0.76 for treated samples) between the average miRNA abundance measured with both RT-PCR and microarray (**Figure 5**). These analyses support that the direction of miRNA differential expression induced by formaldehyde was consistent between the RT-PCR and microarray analyses. It is important to note that there is a difference in the magnitude of expression change with the microarray results generally greater than those obtained with RT-PCR.

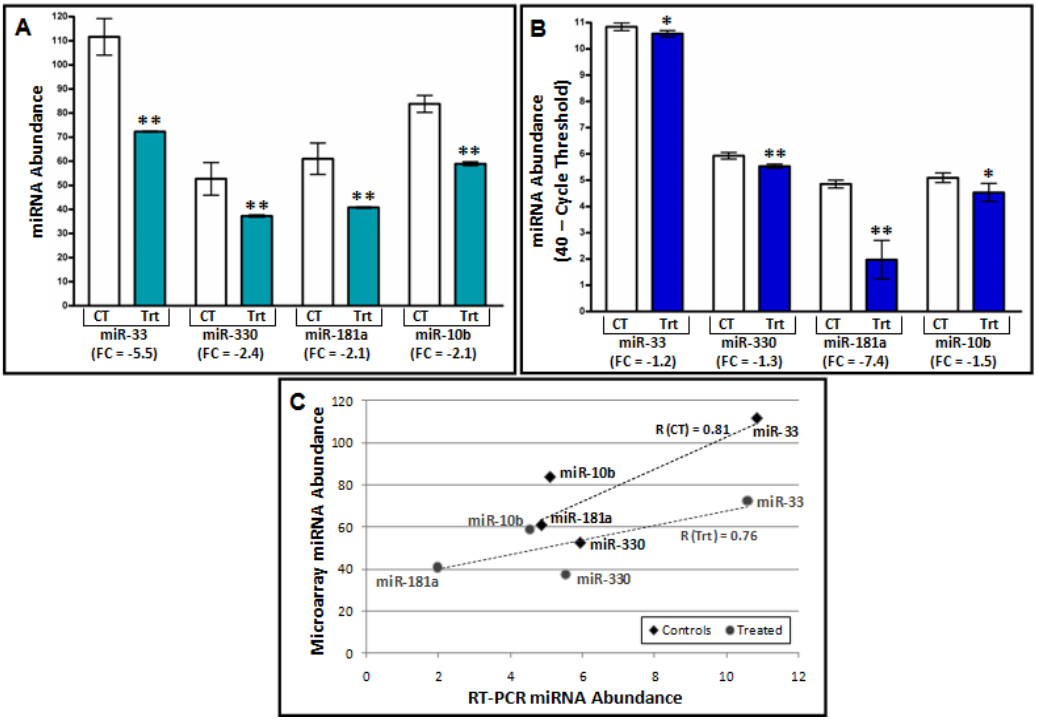

Figure 5: Microarray results align with RT-PCR results. (**A**) miRNA microarray results are displayed as miRNA abundance obtained from raw microarray data. (**B**) miRNA RT-PCR results are displayed in terms of miRNA abundance relative to 40, the maximum cycle threshold. (*) represents p-value < 0.1, and (**) represent p-value < 0.05. Each column represents either control samples (CT) or treated samples (Trt). Average fold changes (FC) are shown, and error bars represent S.E.M. (**C**) Correlation between average miRNA abundance measured by microarrays and RT-PCR is illustrated.

1.4.3 miRNA Targets are Integrated into Biological Networks

In order to identify potential biological pathways affected by formaldehyde exposure, the 89 miRNAs that showed significant changes in expression levels were ranked according to their fold changes in expression, p-values of significance, and RT-PCR results (see **Supplementary Table 1**). Here, the four miRNAs with the most significant formaldehyde-induced changes in expression were further investigated: miR-33, miR-330, miR-181a, and miR-10b. For each of these four miRNAs, we identified their putative mRNA targets. Using a stringent cutoff of a match score between each miRNA and its mRNA targets followed by analysis for unique mRNAs per target list, we identified a total of 67 targets of miR-33, 217 targets of miR-330, 334 targets of miR-181a, and 25 targets of miR-10b (see **Supplementary Table 2**). Among this list of 643 mRNAs, there are 42 that are common to at least two of the modulated miRNAs.

Once the predicted transcriptional targets were identified for the most significant miRNAs, they were overlaid onto molecular pathway maps enabled through the Ingenuity® Systems Knowledge Base. Networks containing miRNA targets were algorithmically constructed based on connectivity and known relationships among proteins. The predicted targets of miR-33, miR-330, miR-181a, and miR-10b resulted in the generation of a total of 40 networks (see **Supplementary Table 3**). For each of the miRNA targets, the most significant (p-values range from 10^{-23} to 10^{-43}) network has been highlighted for further evaluation (**Figure 6**). The proteins identified within these networks were queried for their enrichment for various canonical pathways. A comparison of the canonical pathways highlighted the conservation of a cancer-associated pathway common to all four miRNA-generated networks (see **Supplementary Table 4**). Overlaying the pathway information onto the most significant networks resulted in the identification of enrichment for the nuclear factor kappa beta (NFκB) pathway and the interleukin-8 (IL-8) signaling pathway, among others.

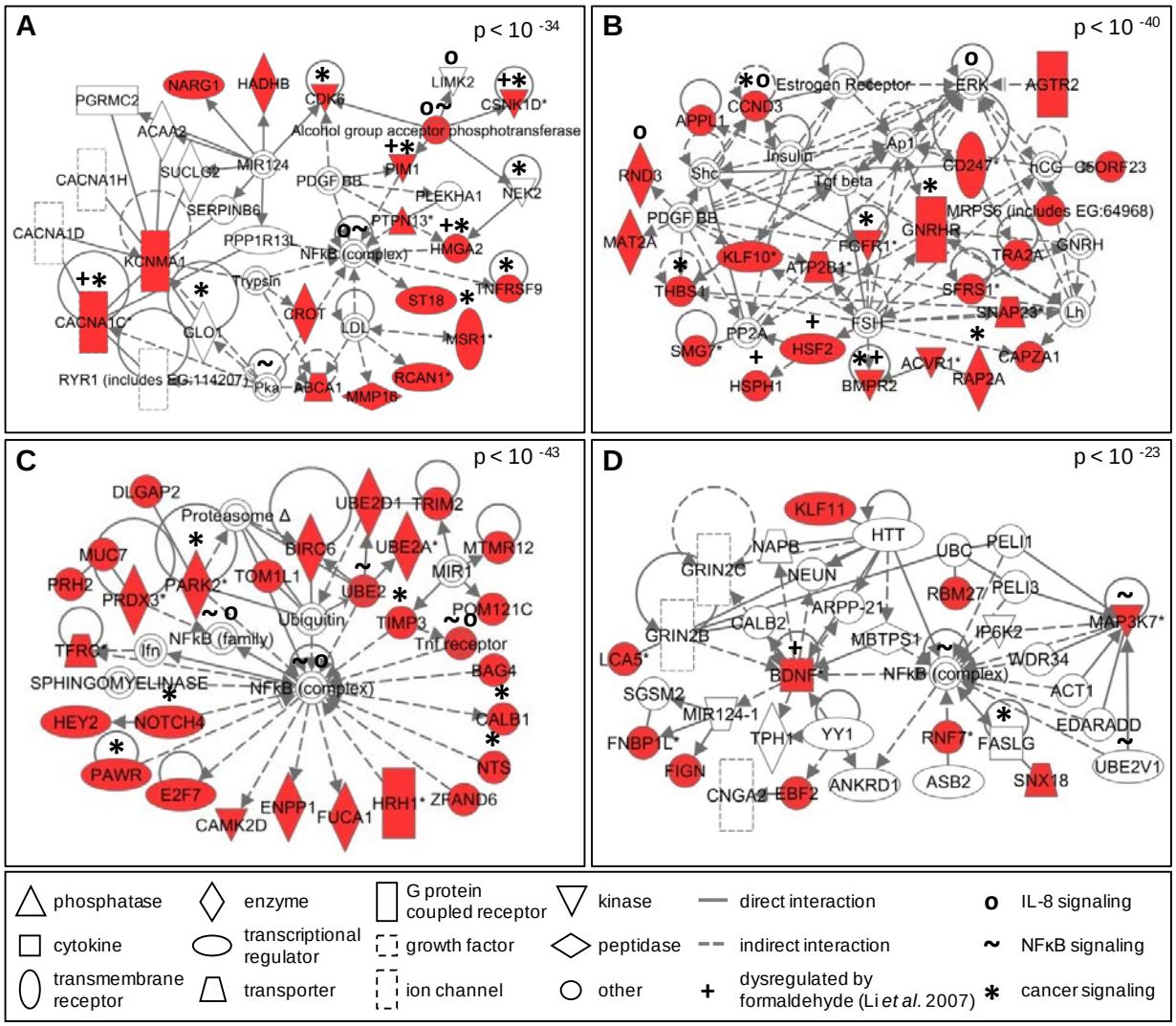

Figure 6: Significant molecular networks of miRNA-mediated signaling likely affected by formaldehyde exposure in human lung cells. Protein networks display interactions using the transcriptional targets of (**A**) miR-33, (**B**) miR-330, (**C**) miR-181a, and (**D**) miR-10b. Networks are displayed with symbols representing predicted miRNA targets (red symbols) or proteins associated with the predicted targets (white symbols).

Using a biological process enrichment analysis, the 40 networks encoded by the transcriptional targets for each miRNA were queried for biological processes that were most significantly modulated by formaldehyde exposure. A total of 71 unique biological processes were found (see **Supplementary Table 5**). Across the mRNA targets, common enrichment was found for 13 different cellular biological processes. These processes included inflammatory response (p-value = 0.0029) and endocrine system development/function (p-value = 0.0018) which were enriched within the targets of all four miRNAs (**Table 1**).

Table 1: Biological functions significantly associated with all predicted target sets of miR-33, miR-330, miR-181a, and miR-10b, the miRNAs with the greatest expression alterations upon exposure to formaldehyde in human lung cells.

Enriched Functions	Average p-value
Cellular Development	0.0011
Small Molecule Biochemistry	0.0015
Nervous System Development and Function	0.0016
Cell-To-Cell Signaling and Interaction	0.0017
Cell Morphology	0.0017
Tissue Development	0.0017
Cellular Function and Maintenance	0.0017
Cellular Movement	0.0017
Endocrine System Development and Function	0.0018
Gene Expression	0.0018
Cellular Growth and Proliferation	0.0021
Inflammatory Response	0.0029
Hematological System Development and Function	0.0033

1.4.4 Conservation of Predicted and Observed mRNA Targets

In our analysis, we used a stringent computational metric to match miRNAs to their predicted mRNA targets to better understand the biological implications of formaldehyde exposure. As these mRNA targets were computationally predicted, we also compared our results with those of an existing genomic database established from a study that analyzed human tracheal fibroblast cells exposed to formaldehyde (Li et al. 2007). In this comparison, we found overlap between the predicted mRNA targets of the formaldehyde-modulated miRNAs and the tested formaldehyde-responsive genes previously identified (Li et al. 2007). Specifically, brain-derived neurotrophic factor (*BDNF*), bone morphogenetic protein receptor, type II (serine/threonine kinase) (*BMPR2*), calcium channel voltage-dependent L type, alpha 1C subunit (*CACNA1C*), casein kinase 1 delta (*CSNK1D*), high mobility group AT-hook 2 (*HMGA2*), heat shock transcription factor 2 (*HSF2*), heat shock 105kDa/110kDa protein 1 (*HSPH1*), and Pim-1 oncogene (*PIM1*), are found within the four most significant networks associated with the identified miRNA targets (**Figure 6**).

We expanded our comparison by performing network analysis on the formaldehyde-associated genes identified by Li *et al*. (2007). Here, networks were constructed and related biological functions were identified, as done with the miRNA predicted target network analysis.

Networks related to cancer (p-value = 1.9 x 10^{-19}), inflammation (p-value = 1.1 x 10^{-8}), and endocrine system disorders (p-value = 3.15 x 10^{-4}) were generated (see **Supplementary Table 6**).

1.4.5 Inflammatory Cytokine IL-8 is Released in Response to Formaldehyde

Based on our findings from the canonical pathway and biological process enrichment analyses that showed the IL-8 pathway as potentially dysregulated by miRNAs associated with formaldehyde exposure, we set out to confirm whether IL-8 protein levels may be influenced by such exposure. After cells were exposed to formaldehyde, IL-8 protein release was assessed. The investigation of the inflammatory response protein IL-8 showed that human lung cells activate an inflammatory response after exposure to formaldehyde. Specifically, an average 16.9 fold increase (p-value < 0.05) in cytokine release was observed in formaldehyde exposed cells relative to control samples (**Figure 7**).

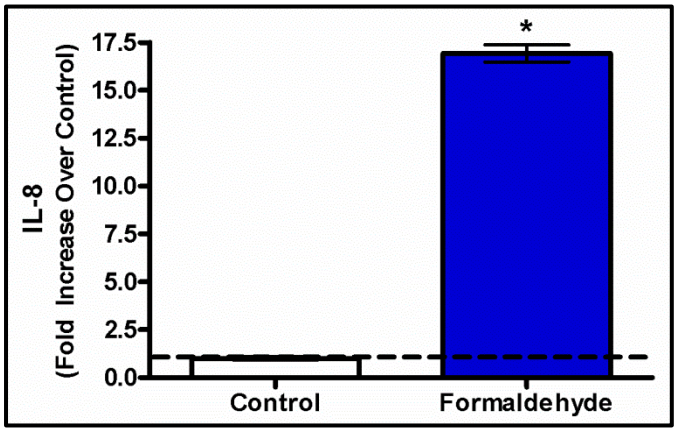

Figure 7: Interleukin-8 levels are significantly elevated in formaldehyde-treated lung cells compared to untreated cells. Results are displayed as fold increase over control +/- S.E.M. (*) indicates statistical significance compared to control samples (p-value < 0.05).

1.5 Discussion

In this study, we exposed human A549 lung epithelial cells to formaldehyde using an *in vitro* exposure system that physically replicates *in vivo* human lung gas exposures (Bakand et al. 2005). It is important to note that A549 cells are carcinoma cells that exhibit differences in certain signaling compared to non-cancerous cells. For instance, A549 cells are enriched for Nrf2

detoxifying pathways and are more resistant to apoptosis in comparison to normal cells (Kweon et al. 2006). While we recognize that A549 cells may not completely mimic normal lung cell response, there are several advantages to using these cells for air toxicant studies. For example, when exposed to gases at an air-liquid interface, A549 cells secrete enough surfactant to mimic airway surface tension (Blank et al. 2006). As a result, A549 cells are routinely used to study the effects of environmental air exposures (Doyle et al. 2004; Doyle et al. 2007; Jaspers et al. 1997; Sexton et al. 2004), including formaldehyde (Quievryn et al. 2000; Speit et al. 2008; Speit et al. 2010). A549 cells have also shown the same sensitivity and removal efficiency towards formaldehyde-induced DNA protein crosslinks as primary human nasal epithelial cells (Speit et al. 2008).

Our microarray analysis revealed that formaldehyde exposure resulted in the down-regulation of 89 miRNAs. It was interesting that all of the modulated miRNAs were down-regulated by formaldehyde exposure. This general trend of miRNA down-regulation has been observed in rat lung cells exposed to cigarette smoke (Izzotti et al. 2009), as well as in multiple tumor cell types, including lung cancer, breast cancer, and leukemia (Lu et al. 2005).

We focused a detailed analysis on the four most significantly down-regulated miRNAs, as determined through microarray analysis and RT-PCR: miR-33, miR-330, miR-181a, and miR-10b. These miRNAs have been studied, to some extent, and knowledge about their regulation and association to disease is growing. For example, miR-33 shows decreased expression levels in tissues from patients with lung carcinomas (Yanaihara et al. 2006). Also, miR-330 expression has been measured at significantly lower levels in human prostate cancer cells when compared against nontumorigenic prostate cells (Lee et al. 2009). Furthermore, miR-330 has been suggested to act as a tumor suppressor by regulating apoptosis of cancer cells (Lee et al. 2009). In addition, miR-10b shows altered expression levels within breast cancer tissue, and is one of the most consistently dysregulated miRNAs able to predict tumor classification (Iorio et al. 2005; Ma et al. 2007). These findings suggest that miR-33, miR-330, and miR-10b may influence cellular disease state, specifically related to cancer.

Formaldehyde exposure also altered the expression level of miR-181a, which has known associations with leukemogenesis (Marcucci et al. 2009). The specific link between formaldehyde exposure and leukemia is currently debated, as numerous epidemiological studies show evidence for possible association to this disease (Hauptmann et al. 2009; Pinkerton et al.

2004; L Zhang et al. 2010), as well as against it (Bachand et al. 2010; Marsh et al. 2004). However, it is important to note that our study evaluates miRNA expression in lung cells, which likely differ from leukemia target cells' responses to formaldehyde exposure, or exposure to formaldehyde's metabolic products. Nevertheless, it is worth highlighting the observation of the dysregulation of miR-181a upon exposure to formaldehyde.

To expand our analysis, we used a systems biology approach to understand the potential biological implications of the miRNA expression changes induced by acute formaldehyde exposure. For this analysis, we used a stringent computational matching approach to identify predicted mRNA targets for miR-33, miR-330, miR-181a, and miR-10b. The identified mRNA targets were used to construct associated molecular networks and were analyzed for their known involvement in signaling pathways and biological functions. The identified networks showed enrichment for various canonical pathways including nuclear factor kappa-B (NFκB) and interleukin-8 (IL-8) signaling. Although very few predicted targets overlapped between the four miRNAs, proteins involved with cancer mechanisms including that of the NFκB pathway were found within the miRNA target networks. Importantly, NFκB has clear links to inflammation and cancer development (Karin et al. 2005). Also related to inflammation, IL-8-related signaling molecules were present in the miRNA target networks. Previous studies have shown IL-8 release in lungs cells representing inflammatory response after exposure to other air pollutants (Jaspers et al. 1997; Sexton et al. 2004). In addition, investigations have shown increased IL-8 levels in lungs of patients with diseases such as acute lung injury (McClintock et al. 2008), adult respiratory distress syndrome (Jorens et al. 1992), and asthma (Bloemen et al. 2007). Inflammation is a recognized formaldehyde-induced response, as formaldehyde is known to irritate the respiratory system (Tuthill 1984) and increase asthmatic response (Rumchev et al. 2002; Wieslander et al. 1997). Our findings suggest that the canonical pathways associated with formaldehyde-induced miRNA alterations may affect the regulation of biological pathways associated with various disease states, including cancer and inflammation.

As a method to further verify our results, we compared the protein levels of cytokine interkeukin-8 (IL-8) in formaldehyde-exposed cells versus mock-treated controls. We found that, indeed, IL-8 showed significantly increased protein expression levels in the formaldehyde-exposed cells. These results support our findings that IL-8 signaling is altered in lung cells exposed to formaldehyde. Interestingly, IL-8 levels are also increased in formaldehyde-exposed

lung cells after pre-sensitization to tumor necrosis factor alpha (TNFα) (Persoz et al. 2010). TNFα is a proinflammatory mediator shown to have increased levels upon exposure to formaldehyde (Bianchi et al. 2004). Our network analyses suggest that cytokine signaling may be altered through changes in miRNA expression levels. Supporting this is a recent study that shows modifications to miRNAs may influence the expression of cytokines, including IL-6 and IL-8 (Jones et al. 2009). Future research will test whether the observed miRNA expression changes are directly associated with IL-8 signaling.

In an effort to gain further understanding of formaldehyde's effects on gene expression, we compared our results with those of an existing genomics database (e.g. mRNA) from a study that evaluated human lung cells exposed to formaldehyde (Li et al. 2007). Using the predicted targets in our most significant miRNA networks, we found the following genes overlap with the existing database: *BDNF*, *BMPR2*, *CACNA1C*, *CSNK1D*, *HMGA2*, *HSF2*, *HSPH1*, and *PIM1*. These genes have been shown to play a role in various diseases. For example, *BDNF*, or brain-derived neurotrophic factor, modulates neurogenesis after injury to the central nervous system (Ming et al. 2005). *CSNK1D*, or casein kinase 1 delta, has been identified as up-regulated in breast cancer tissue (Abba et al. 2007). *HMGA2*, or high mobility group AT-hook 2, is oncogenic in many cells, including lung carcinoma cells, and is regulated by the tumor-suppressive miRNA let-7 (Lee et al. 2007). Lastly, *PIM1*, or Pim-1 oncogene, is found at increased levels within prostate cancer tissue (Dhanasekaran et al. 2001). Network analysis of all formaldehyde-responsive genes identified through the Li et. al. (2007) study revealed significant associations with cancer, inflammation, and endocrine system regulation, which also overlap with our findings. These genes are therefore linked with formaldehyde-induced changes in miRNA abundance as well as mRNA alterations, and they are related to a diverse range of cellular responses including tumorigenesis.

In conclusion, our study provides evidence of a potential mechanism that may underlie the cellular effects induced by formaldehyde, namely the modification of miRNA expression. We identify a set of 89 miRNAs that are dysregulated in human lung cells exposed to formaldehyde. Mapping the most significantly changed miRNAs to their predicted transcriptional targets and their network interactomes within the cell reveals the association of formaldehyde exposure to inflammatory response pathways. We also validate our findings by: (1) performing RT-PCR; (2) integrating our predicted networks with known formaldehyde-

induced mRNA expression changes; and (3) examining protein expression changes of a key inflammatory response mediator, IL-8. Future research will investigate whether the expression levels of these miRNAs may serve as potential biomarkers of formaldehyde exposure in humans. Such biomarkers can be utilized to better monitor human exposure to environmental toxicants and relate them to health effects. Based on our findings, we believe that miRNAs likely play an important role in regulating formaldehyde-induced gene expression and may represent a possible link between exposure and disease.

CHAPTER 2
FORMALDEHYDE AND EPIGENETIC ALTERATIONS: MICRORNA CHANGES IN THE NASAL EPITHELIUM OF NONHUMAN PRIMATES

2.1 Overview

Formaldehyde is an air pollutant present in both indoor and outdoor atmospheres. Because of its ubiquitous nature, it is imperative to understand the mechanisms underlying formaldehyde-induced toxicity and carcinogenicity. MicroRNAs (miRNAs) can influence disease caused by environmental exposures, yet miRNAs are understudied in relation to formaldehyde. Our previous investigation demonstrated that formaldehyde exposure in human lung cells caused disruptions in miRNA expression profiles.

Here, we expand our preliminary *in vitro* findings to an *in vivo* model. We set out to test the hypothesis that formaldehyde inhalation exposure significantly alters miRNA expression profiles within the nasal epithelium of nonhuman primates.

Cynomolgus macaques were exposed by inhalation to approximately 0, 2, or 6 ppm formaldehyde for 6 hours/day for two consecutive days. Small RNAs were extracted from nasal samples and assessed for genome-wide miRNA expression levels. Transcriptional targets of formaldehyde-altered miRNAs were computationally predicted, analyzed at the systems level, and assessed using RT-PCR.

Expression analysis revealed that 3 and 13 miRNAs were dysregulated in response to 2 and 6 ppm formaldehyde, respectively. Transcriptional targets of the miRNA with the greatest increase (miR-125b) and decrease (miR-142-3p) in expression were predicted and analyzed at the systems level. Enrichment was identified for miR-125b targeting genes involved in apoptosis signaling. The apoptosis-related targets were functionally tested using RT-PCR, where all targets showed decreased expression in formaldehyde-exposed samples.

Our study reveals that formaldehyde exposure disrupts miRNA expression profiles within the nasal epithelium, and these alterations likely influence apoptosis signaling.

2.2 Study Objectives

For this study, we set out to test the novel hypothesis that formaldehyde inhalation exposure significantly alters miRNA expression profiles within the nasal epithelium of nonhuman primates. We tested this hypothesis by exposing nonhuman primates (cynomolgus macaques) to ~ 0, 2, or 6 ppm formaldehyde for 6 hours/day across two days. After exposure, nasal epithelial tissue was assessed for formaldehyde-induced changes in miRNA expression profiles across the genome. Transcriptional targets of the miRNAs with the highest increase and highest decrease in expression were predicted *in silico* and mapped onto molecular interaction networks. Important canonical pathways enriched within the constructed networks were identified and further assessed at the gene expression level. Altogether, this study reveals a novel epigenetic mechanism through which formaldehyde may influence critical signaling pathways known to influence disease.

2.3 Materials and Methods

2.3.1 Animals

Cynomolgus macaques were treated humanely and with regard for alleviation of suffering. Animals were exposed, sedated, and euthanized using protocols approved by the Lovelace Research Institute's animal care and use committee (FY10-104A). For this study, eight male cynomolgus macaques (*Macaca fascicularis*) were selected from the Lovelace Respiratory Research Institute colony. Animals were approximately six years of age and weighed between 4.48 and 8.56 kilograms. Animals were conditioned to whole body exposure chambers for 30, 60, 180, and 360 minutes prior to the first day of exposure, as previously described (Moeller et al. 2011).

2.3.2 Formaldehyde Exposures

Animals were exposed to formaldehyde over the course of two days for six hours each day using whole body exposure chambers. Target exposure concentrations were 0, 2, and 6 ppm formaldehyde. Exposure conditions were created by vaporizing [$^{13}CD_2$]-paraformaldehyde.

Formaldehyde was isotope-labeled for the purposes of a previous investigation (Moeller et al. 2011). Chamber concentrations were monitored by collecting samples with a Waters XpoSure Aldehyde Sampler cartridge every five minutes throughout each exposure period. Samples from the cartridges were analyzed using high-performance liquid chromatography with an attached detector monitoring ultraviolet absorbance at 360 nm (Lu et al. 2011; Moeller et al. 2011). Two control animals were placed in whole body exposure chambers containing clean air. Three monkeys were exposed to a target concentration of 2 ppm formaldehyde, where the measured concentration averaged 1.9 ppm across the exposure periods. Three monkeys were exposed to a target concentration of 6 ppm formaldehyde, where the measured concentration averaged 6.1 ppm across the exposure periods. For more detailed methods, see Moeller et al. (Moeller et al. 2011).

2.3.3 Sample Collection

Approximately 15 minutes after the second exposure period, animals were serially sedated with Ketamine (10 mg/kg, intramuscular) and euthanized with Euthasol (>1 ml/4.5 kg, intravenous). Animals underwent necropsy one at a time with each necropsy requiring approximately 45 minutes. All samples were collected within 3 hours of the exposure. Sample collection started immediately after the last exposure in order to parallel sacrifice and sample collection times used in our previous studies (Lu et al. 2011; Moeller et al. 2011). During necropsy, nasal epithelial tissue from the maxilloturbinate regions were collected, placed in RNAlater® (Qiagen, Valencia, CA), and stored at -80°C. Samples were shipped by overnight courier on dry ice to the University of North Carolina at Chapel Hill.

2.3.4 Sample Processing

Small RNAs were isolated from nasal tissue samples. Samples were first disrupted and homogenized using a TissueRuptor (Qiagen) in the presence of TRIzol (Invitrogen Life Technologies, Carlsbad, CA), and RNA was isolated using the miRNeasy® kit (Qiagen). Extracted RNA was quantified with a Nanodrop 1000 spectrophotometer (Thermo Scientific, Waltham, MA) and its integrity verified with a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA). RNA was then labeled and hybridized to the Agilent Human miRNA Microarray (v1.0). This microarray assesses the relative expression levels of 534 miRNAs measured using

11080 probesets. Microarray results were extracted using Agilent Feature Extraction software. Microarray data have been submitted to National Center for Biotechnology Information (NCBI) Gene Expression Omnibus repository (Edgar et al. 2002) and are available under accession number GSE34978 (NCBI 2010).

2.3.5 Microarray Analysis

Microarray data were normalized by quantile normalization. To eliminate background noise, miRNA probes with signal intensities less than the median signal (signal = 40) across all replicates were removed. Differential expression was defined as a significant difference in miRNA levels between exposed versus unexposed samples, where three statistical requirements were set: (i) fold change of ≥ 1.5 or ≤ -1.5 (average exposed versus average unexposed); (ii) p-value < 0.05 (ANOVA); and (iii) a false discovery rate corrected q-value < 0.1. Analysis of variance (ANOVA) p-values were calculated using Partek® Genomics Suite™ software (St. Louis, MO). To control the rate of false positives, q-values were calculated as the minimum "positive false discovery rate" that can occur when identifying significant hypotheses (Storey 2003).

2.3.6 RT-PCR Confirmation of miRNA Expression Changes

To confirm formaldehyde-induced miRNA expression changes, we performed real-time reverse transcriptase polymerase chain reaction (RT-PCR) using two miRNAs identified as the most increased in expression (miR-125b and miR-152) and two miRNAs identified as the most decreased in expression (miR-145 and miR-142-3p) following 6 ppm formaldehyde exposure. TaqMan® MicroRNA Primer Assays for hsa-miR-125b (ID 000449), hsa-miR-152 (ID 000475), hsa-miR-145 (ID 002278) and hsa-miR-142-3p (ID 000464) were used in conjunction with the TaqMan® Small RNA Assays PCR kit (Applied Biosystems, Carlsbad, CA). The same control and formaldehyde-exposed samples from the microarray analysis were used for RT-PCR, and was performed in technical triplicate. The resulting RT-PCR cycle times were normalized against the U6 housekeeping miRNA, and fold changes in expression were calculated using the $\Delta\Delta$Ct method. Statistical significance of the difference in miRNA expression levels between the formaldehyde-exposed and unexposed samples was calculated using an ANOVA (Partek®).

2.3.7 Predicting Targets of miR-125b and miR-142-3p

In order to understand the effects formaldehyde-responsive miRNAs may cause at the gene expression level, computational predictions of the mRNA targets of miR-125b and miR-142-3p were carried out. These two miRNAs were selected as they showed the largest increase (miR-125b) or decrease (miR-142-3p) in expression after 6 ppm formaldehyde exposure. Here, TargetScanHuman (Whitehead 2011) algorithms were employed to identify potential matches between 3' untranslated mRNA regions and miRNA seed sequences (Lewis et al. 2005). The resulting predicted miRNA-mRNA interactions were filtered for the probability of preferentially conserved targeting (PCT) ≥ 0.9. This PCT filter controlled for background conservation across mammals by accounting for mutational biases, dinucleotide conservation rates, and individual untranslated region conservation rates (Friedman et al. 2009).

2.3.8 Pathway Enrichment Analysis of Predicted Targets

Network analysis was performed to understand the systems level response to formaldehyde inhalation exposure possibly mediated via epigenetic (e.g. miRNA) regulation. For this analysis, the predicted mRNA targets of miR-125b and miR-142-3p were overlaid onto a global interaction network. Here, networks were algorithmically constructed based on connectivity, as enabled through Ingenuity Pathway Analysis (Ingenuity Systems®, Redwood City, CA). Canonical pathways within the constructed networks were then identified. Over-represented pathways were defined as pathways than contain more targets than expected by chance, as calculated using the right-tailed Fisher's Exact Test. Here, pathways with enrichment p-values < 0.05 were considered significantly enriched with the predicted targets of miR-125b or miR-142-3p.

2.3.9 Testing miRNA Targets using RT-PCR

All apoptosis-associated genes (n=4) predicted to be regulated by formaldehyde-responsive miR-125b, and all integrin-linked kinase (ILK)-associated genes (n=2) predicted to be regulated by formaldehyde-responsive miR-142-3p, were tested at the gene expression level using RT-PCR. QuantiTect Primer Assays were used with QuantiTect SYBR® Green PCR kits (Qiagen) and the LightCycler® 480 (Roche Applied Science). Specifically, BCL2-antagonist/killer 1 (*BAK1*) (Catalog Number QT00228508), caspase 2, apoptosis-related cysteine

peptidase (*CASP2*) (QT01342509), integrin, beta 8 (*ITGB8*) (QT00038507), mitogen-activated protein kinase kinase 7 (*MAP2K7*) (QT00090545), myeloid cell leukemia sequence 1 (BCL2-related) (*MCL1*) (QT00094122), and rapamycin-insensitive companion of mTOR (*RICTOR*) (QT00065793) were evaluated for potential changes in gene expression levels induced by formaldehyde exposure. Resulting RT-PCR cycle times were normalized against the β-actin housekeeping gene, and fold changes in expression were calculated using the ΔΔCt method. Statistical significance comparing the expression levels between exposed and unexposed samples was calculated using an ANOVA (Partek®).

2.4 Results

2.4.1 Formaldehyde Disrupts miRNA Expression Profiles in Nasal Tissue

To study the effects of formaldehyde inhalation exposure, cynomolgus macaques were exposed to ~ 0, 2, or 6 ppm formaldehyde 6 hr/day for two days. After treatment, nasal epithelial tissue samples were collected and assessed for genome-wide changes in miRNA expression profiles using the Agilent Human miRNA Microarray. A human microarray was used because a miRNA microarray is not currently available for nonhuman primates. This array is suitable for these experimental purposes based on the high degree of similarity in DNA sequences as well as conserved basal gene expression profiles between humans and cynomolgus macaques (Walker 2008). Nevertheless, it is recognized that certain cynomolgus macaque-specific miRNAs may not be accounted for in these analyses, resulting in the potential for underestimation of formaldehyde's true impact on genome-wide miRNA profiles.

Microarray analysis showed three miRNAs with significantly decreased expression levels upon exposure to 2 ppm formaldehyde (**Table 2**). In comparison, exposure to 6 ppm formaldehyde significantly disrupted the expression levels of 13 miRNAs, represented by 15 array probesets (**Table 2**). Of the 13 miRNAs, four were significantly increased and nine were significantly decreased in expression. Interestingly, the three miRNAs that were significantly decreased in response to 2 ppm formaldehyde (e.g. miR-142-3p, miR-145, and miR-203) were also significantly decreased in response to 6 ppm formaldehyde.

Table 2: Formaldehyde inhalation exposure in the nasal epithelium of nonhuman primates significantly disrupts the expression levels of 13 unique miRNAs, represented by 15 array probesets. Significant fold change (FC) comparisons between exposed and unexposed samples are indicated with * (representing p-value < 0.01, q-value < 0.1).

		2 ppm			6 ppm		
miRNA	**Array Feature Number**	**log2FC**	**p-value**	**q-value**	**log2FC**	**p-value**	**q-value**
miR-125b	2637	0.44	6.1E-01	0.666	2.86*	2.2E-04	0.090
miR-152	1548	0.79	3.0E-03	0.297	1.29*	1.3E-04	0.072
miR-219-5p	1180	0.36	8.8E-02	0.451	1.22*	1.7E-04	0.075
miR-532-5p	1259	0.35	3.4E-02	0.390	1.09*	8.1E-05	0.055
miR-520f	14457	-0.61	3.3E-04	0.188	-0.77*	1.4E-04	0.072
miR-26b	12607	-1.13	9.3E-05	0.146	-1.38*	5.2E-05	0.050
miR-140-5p	12026	-0.69	3.6E-04	0.188	-1.56*	2.4E-05	0.036
miR-22	12927	-0.69	4.8E-04	0.203	-1.70*	2.6E-05	0.036
miR-374a	14431	-1.68	1.2E-04	0.148	-1.77*	1.1E-04	0.067
miR-203	12162	-1.98*	4.7E-05	0.098	-2.11*	4.1E-05	0.046
miR-203	11451	-1.75	1.0E-04	0.146	-2.12*	6.7E-05	0.055
miR-142-3p	12366	-4.12*	1.1E-06	0.009	-2.92*	1.6E-06	0.011
miR-29a	13448	-3.24	2.5E-04	0.188	-3.15*	2.6E-04	0.099
miR-145	15649	-3.15*	3.0E-05	0.098	-3.56*	2.6E-05	0.036
miR-142-3p	14658	-2.81	3.1E-04	0.188	-5.01*	1.8E-04	0.075

2.4.2 RT-PCR Confirmed Formaldehyde-Induced miRNA Expression Changes

To confirm that formaldehyde inhalation exposure significantly disrupts the expression of miRNAs, RT-PCR was performed. Specifically, the two miRNAs most increased in expression (miR-125b and miR-152) and the two miRNAs most decreased in expression (miR-145 and miR-142-3p), in response to 6 ppm formaldehyde were validated using this alternative method. Comparing the exposed versus unexposed samples confirmed that miR-125b and miR-152, were, indeed, significantly ($p<0.05$) increased in expression upon exposure to 6 ppm formaldehyde **(Figure 8**). The microarray analysis' stringent multiple test correction filter excluded miR-125b from the list of miRNAs significantly differentially expressed by 2 ppm formaldehyde. However, RT-PCR analysis showed that miR-125b was significantly increased in expression in the 2 ppm formaldehyde-exposed animals. Similar confirmation was observed for miR-145 and miR-142-3p, where expression levels were significantly ($p<0.05$) decreased following 6 ppm formaldehyde exposure (**Figure 8**).

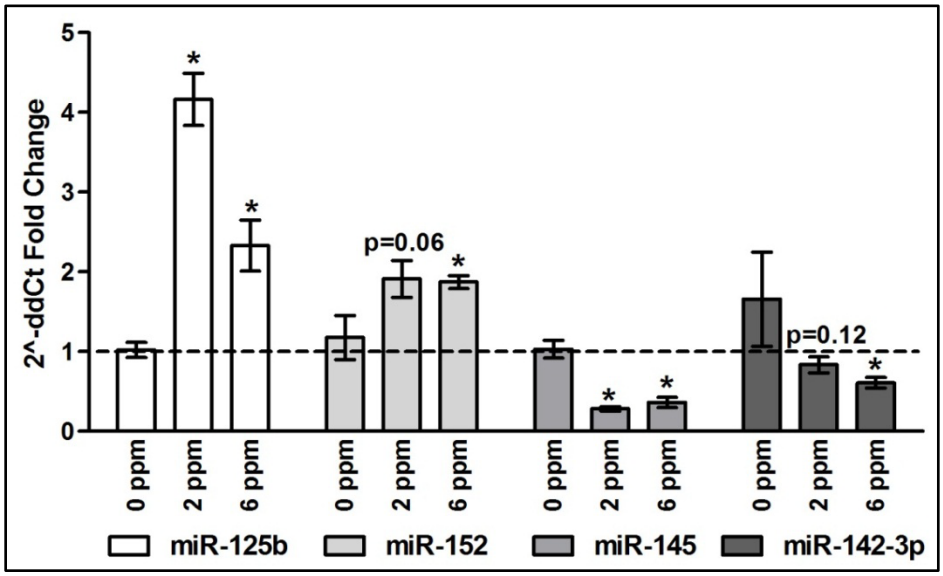

Figure 8: RT-PCR confirms the altered expression of selected miRNAs upon exposure to formaldehyde within the nonhuman primate nasal epithelium. Mean fold changes (exposed / unexposed) in gene expression are displayed (± SE), where (*) represents $p<0.05$.

2.4.3 Transcriptional Targets of miR-125b and miR-142-3p were Predicted

To understand genomic changes regulated via miRNAs that formaldehyde inhalation exposure may initiate, we computationally predicted mRNA targets of miR-125b and miR-142-3p. These miRNAs were selected for further investigation, as they showed the highest increase or decrease in expression upon exposure to 6 ppm formaldehyde, respectively. In addition, their differential expression was confirmed through RT-PCR analysis. Using seed match-based algorithms, a total of 132 genes were predicted to be targeted by miR-125b (see **Supplementary Table 7**). In comparison, only 13 genes were predicted to be targeted by miR-142-3p (see **Supplementary Table 8**).

2.4.4 Apoptosis Signaling is Associated with miR-125b Predicted Targets

In order to evaluate the potential effects of formaldehyde exposure at the systems level, enriched canonical signaling pathways were evaluated for the 132 predicted targets of miR-125b. Through this network analysis, 11 canonical pathways were identified as significantly over-represented amongst the networks constructed using the predicted targets of miR-125b (see

Supplementary Table 9). The two pathways of highest significance were sphingolipid metabolism (p=0.003) and apoptosis signaling (p=0.003) (**Figure 9**).

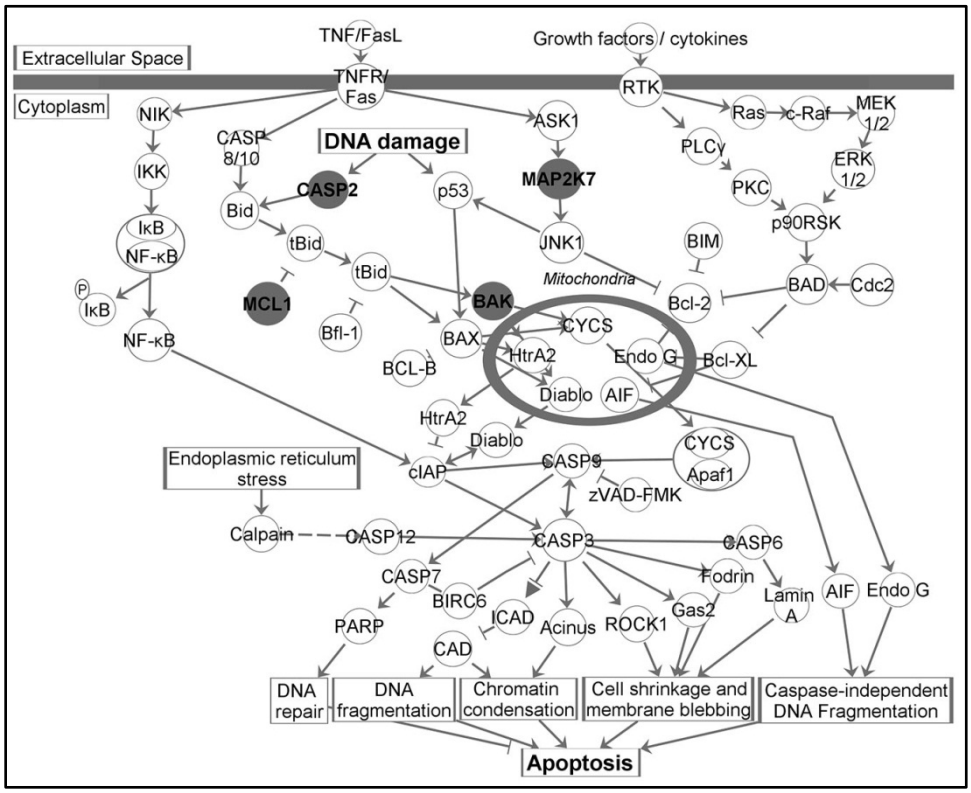

Figure 9: Predicted mRNA targets of formaldehyde-altered miR-125b are involved in apoptosis signaling, suggesting that the regulation of the apoptotic machinery may be modified through formaldehyde's influence on miRNAs. The apoptosis signaling pathway is illustrated, where molecules predicted to be targeted by miR-125b are shaded in dark grey.

2.4.5 Apoptosis-Related miR-125b Targets are Decreased in Expression

All four of the apoptosis-related mRNA molecules predicted to be targeted by miR-125b were tested at the gene expression level using RT-PCR. As miR-125b was increased in expression, it was anticipated that its potential targets would be decreased in expression after formaldehyde exposure. Three of the evaluated targets, *BAK1*, *MAP2K7*, and *MCL1*, showed significantly ($p<0.05$) decreased expression levels in response to both 2 and 6 ppm formaldehyde exposures (**Figure 10**). *CASP2* showed significantly decreased expression in response to 2 ppm formaldehyde. *CASP2* expression was also decreased in response to 6 ppm formaldehyde, but was not statistically significant (p=0.15) (**Figure 10**) Altogether, all four of the apoptosis-related

mRNAs predicted to be regulated by miR-125b showed decreased expression upon exposure to formaldehyde.

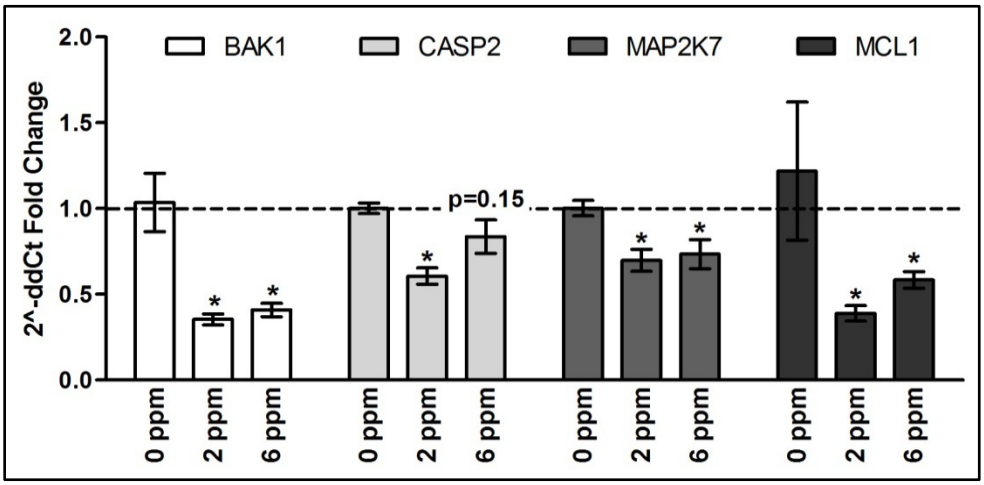

Figure 10: RT-PCR shows the decreased expression of apoptosis signaling-related genes predicted to be targeted by miR-125b, the miRNA with the greatest increased expression resulting from 6 ppm formaldehyde exposure. Mean fold changes (exposed / unexposed) in gene expression are displayed (± SE), where (*) represents $p<0.05$.

2.4.6 ILK Signaling is Associated with miR-142-3p Predicted Targets

To further assess the potential effects of formaldehyde exposure at the systems level, enriched canonical signaling pathways were evaluated for the 13 predicted targets of miR-142-3p. Three canonical pathways were identified as significantly over-represented within the predicted targets of miR-142-3p (see **Supplementary Table 10**). The pathway of highest significance was ILK signaling ($p=0.008$).

2.4.7 ILK-Related miR-142-3p Targets are Altered in Expression

The two ILK signaling-related mRNA molecules predicted to be targeted by miR-142-3p were tested at the gene expression level using RT-PCR. As miR-142-3p was decreased in expression, it was anticipated that its potential targets would have increased expression after formaldehyde exposure. One of the evaluated targets, *ITGB8*, showed significantly increased expression in response to 6 ppm formaldehyde exposure (**Figure 11**). Transcript levels for the other predicted target, *RICTOR*, were significantly decreased in response to 2 and 6 ppm formaldehyde exposure (**Figure 11**).

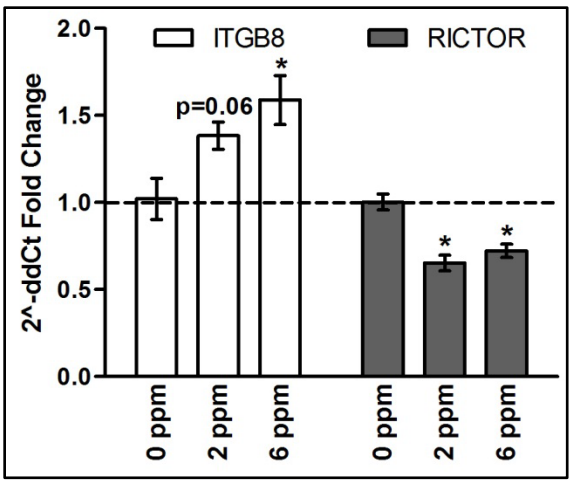

Figure 11: RT-PCR shows the altered expression of ILK signaling-related genes predicted to be targeted by miR-142-3p, the miRNA with the greatest decreased expression resulting from 6 ppm formaldehyde exposure. Mean fold changes (exposed / unexposed) in gene expression are displayed (± SE), where (*) represents $p<0.05$.

2.5 Discussion

This study is the first to evaluate formaldehyde's influence on miRNA expression signatures *in vivo*. In order to study the effects of formaldehyde inhalation exposure, nonhuman primates (cynomolgus macaques) were exposed for 6 hr/day over a course of two days to ~ 0, 2, or 6 ppm formaldehyde. These exposure levels were selected based on previous investigations showing that exposure to 2 ppm and 6 ppm formaldehyde caused DNA-protein crosslinks (Casanova et al. 1991) and DNA adducts (Moeller et al. 2011) within the nasal mucosa of nonhuman primates. The use of nonhuman primates as our animal model is advantageous, as the nasal gross anatomy and pattern of airflow are similar between nonhuman primates and humans (Harkema et al. 2006). Furthermore, there is an extremely high degree of similarity in DNA coding and non-coding sequences between macaques and humans (Walker 2008).

After exposure, animals were euthanized, and nasal epithelial samples from the maxilloturbinate region were collected and assessed for genome-wide changes in miRNA expression profiles. Samples from the maxilloturbinate region were used because inhaled formaldehyde is maximally absorbed within this region (Kepler et al. 1998). In addition, our previous investigation revealed that cynomolgus macaques exposed to isotope labeled [$^{13}CD_2$]-formaldehyde showed detectable amounts of exogenous (i.e. induced by formaldehyde exposure)

and endogenous DNA adducts within nasal samples collected from the maxilloturbinate (Moeller et al. 2011). Specifically, 0.26 ± 0.04 and 0.41 ± 0.05 exogenous N^2-hydroxymethyl-dG / 10^7 dG were present in monkeys exposed to approximately 2 and 6 ppm, respectively, while 2.05 ± 0.53 and 2.49 ± 0.39 endogenous N^2-hydroxymethyl-dG / 10^7 dG adducts were present (Moeller et al. 2011). Furthermore, the respiratory nasal turbinate region of rats exposed to formaldehyde is a site of squamous cell carcinoma formation (Kerns et al. 1983; Monticello et al. 1996).

The expression levels of more than 500 miRNAs were measured across two unexposed, three 2 ppm formaldehyde-exposed, and three 6 ppm formaldehyde-exposed nonhuman primates. Although this sample size was robust enough to detect formaldehyde-responsive miRNAs, we recognize that the size may have limited the power to detect additional changes in miRNA expression. For the genome-wide analysis, a human miRNA microarray was used because a miRNA microarray is not currently available for nonhuman primates. This array is suitable for these experimental purposes based on the high degree of similarity in DNA sequences as well as conserved basal gene expression profiles between humans and cynomolgus macaques (Walker 2008). Baseline human miRNA expression patterns have even been shown to correlate well with cynomolgus macaque miRNA patterns using human miRNA microarrays (Montag et al. 2009). Furthermore, a previous study compared miRNAs identified in the rhesus macaque genome to human homologs and found that 38% of the miRNAs showed 100% homology in precursor sequences (Yue et al. 2008). The remaining 62% of the miRNAs showed between 90 and 100% sequences homology (Yue et al. 2008). Nevertheless, it is recognized that certain cynomolgus macaque-specific miRNAs may not be accounted for in these analyses. Despite these potential limitations, a set of 13 miRNAs with significant differential expression upon exposure to 2 and/or 6 ppm formaldehyde were identified.

Fewer miRNAs were identified as altered by formaldehyde exposure in this *in vivo* study using nonhuman primates than in the *in vitro* study using cultured human airway cells (Chapter 1). This finding was unexpected, as higher formaldehyde levels were used in the nonhuman primates in comparison to the cultured cells. This difference in the number of miRNAs that responded to formaldehyde may be attributable to a variety of factors. As previously discussed, the human miRNA microarray used to assess cynomolgus macaque miRNA responses may have caused a potential underestimation of formaldehyde's true impact on genome-wide miRNA profiles. Also, the cultured airway cells were directly exposed to formaldehyde without the

presence of surrounding tissue. In comparison, the nonhuman primate maxilloturbinate region of the nasal epithelium is surrounded by other tissues that also absorb gaseous formaldehyde, including the anterior lateral meatus, posterior lateral meatus, anterior dorsal septum, anterior mid-septum, and posterior mid-septum, all of which uptake considerable amounts of inhaled formaldehyde (Kimbell et al. 2001). These other tissues may have caused the actual dose of formaldehyde absorbed within the maxilloturbinate region to be lower than a similar *in vitro* exposure. Lastly, there is likely variability between miRNA responses to formaldehyde in the cultured human alveolar epithelial cancer cells versus cynomolgus macaque nasal epithelial non-cancer cells resulting from differences between species, cell type, and/or disease status.

Two of the 13 formaldehyde-responsive miRNAs were among those that we previously showed as altered *in vitro* by formaldehyde, namely miR-26b and miR-140-5p (Rager et al. 2011b). This overlap in response suggests that *in vitro* models may show some responses in common to *in vivo* models at the miRNA level. Many of the formaldehyde-responsive miRNAs in the nonhuman primate have known relationships to disease, where six of the 13 formaldehyde-responsive miRNAs have been identified as differentially expressed in human nasopharyngeal carcinoma. More specifically, miR-142-3p, miR-145, miR-152, miR-203, miR-26b, and miR-29a have all been shown to have altered expression levels in nasopharyngeal cancer tissue in comparison to non-cancerous tissue (Chen et al. 2009; Li et al. 2011; Sengupta et al. 2008; Wong et al. 2012).

In order to evaluate the effects of formaldehyde inhalation exposure at the systems level, molecular targets of miR-125b and miR-142-3p were computationally predicted and analyzed for pathway enrichment. We focused our systems-based analysis on miR-125b and miR-142-3p because these miRNAs showed the highest increase and decrease in expression, respectively, upon exposure to 6 ppm formaldehyde through microarray analysis and were confirmed using RT-PCR analysis. A total of 132 genes were predicted to be targeted by miR-125b, and thereby decreased at the expression level. Far fewer genes were identified for miR-142-3p, where 13 genes were predicted to be targeted by miR-142-3p, and thereby increased at the expression level.

Canonical pathway enrichment analysis revealed a significant association between the predicted targets of miR-125b and apoptosis signaling. To further test this finding, we evaluated the gene expression levels of all four apoptosis signaling-related genes predicted to be targeted

by miR-125b, namely *BAK1*, *CASP2*, *MAP2K7*, and *MCL1*. As predicted, all four genes showed decreased expression levels in the formaldehyde exposed versus unexposed samples. Two of the apoptosis-related genes predicted to be regulated by miR-125b, *MAP2K7* and *MCL1*, have also been shown to have significantly altered expression levels in the nasal epithelium of rats exposed to formaldehyde (Andersen et al. 2010).

The observed decreased expression of genes involved in apoptosis signaling suggests a possible link between formaldehyde exposure and altered regulation of cell death. For example, BAK1 and CASP2 are both pro-apoptotic and have been shown to induce apoptosis *in vitro* and *in vivo* in several cell types (Kumar 2009; Pataer et al. 2000). While the evaluation of proteins encoded by the apoptosis-related genes would further support these findings, such an assessment was not possible here as proteins were not collected. Still, a similar finding has been observed in the nasal epithelium of rats, where nasal instillation of liquid formaldehyde decreased the expression levels of pro-apoptotic genes (Hester et al. 2003). These findings are of high interest, as impaired apoptosis can lead to cellular transformation and cancer development (Hanahan et al. 2011).

Other pathways were also identified as enriched for by the predicted targets of miR-125b, including sphingolipid metabolism. Sphingolipids are an abundant class of lipids present a high levels within eukaryotic membranes (Bartke et al. 2009). Although sphingolipids were first recognized for their structural roles in membrane formation, more recent work shows that sphingolipid metabolites are involved in the regulatory signaling of various biological processes, including apoptosis, cell cycle arrest, inflammation, necrosis, and senescence (Bartke et al. 2009).

Pathway enrichment analysis of the predicted targets of miR-142-3p revealed an enrichment for ILK signaling. It is important to note that this enrichment was not as significant as the enrichment between miR-125b and apoptosis signaling. ILK signaling is involved in a variety of processes within epithelial cells, including cell survival, cell proliferation, and cell adhesion to the extracellular matrix (Gilcrease 2007).

To test our prediction that formaldehyde alters ILK signaling, the expression levels of genes involved in ILK signaling were assessed, including *ITGB8* and *RICTOR*. Because miR-142-3p was decreased in expression by 6 ppm formaldehyde, we anticipated its potential targets to show increased expression. As anticipated, *ITGB8* showed significantly increased expression

resulting from 6 ppm formaldehyde exposure. ITGB8 has been implicated in several biological processes, including airway epithelial cell proliferation (Fjellbirkeland et al. 2003) and airway remodeling involving the extracellular matrix (Kitamura et al. 2011). One of the predicted targets, *RICTOR*, did not show increased transcript levels in formaldehyde-exposed samples. This finding suggests that (i) miR-142-3p may not influence *RICTOR* in the tested conditions, (ii) miR-142-3p may influence RICTOR protein levels by blocking *RICTOR* translation, or (iii) other mechanisms besides miRNA regulation may influence *RICTOR* expression. Some of these scenarios are supported in a recent study where miRNAs were computationally predicted to target hepatic nuclear factor 4α (*HNF4α*) (Ramamoorthy et al. 2012). This study found that a portion of the tested miRNAs successfully targeted *HNF4α*. In addition, some of the miRNAs targeted *HNF4α* by blocking *HNF4α* translation, causing the reduced expression of HNF4α protein while leaving transcript levels unchanged (Ramamoorthy et al. 2012).

It is important to note that our results do not demonstrate that miR-125b directly decreases the expression of *BAK1*, *CASP2*, *MAP2K7*, and *MCL1* upon exposure to formaldehyde, nor that miR-142-3p directly increases the expressed of *ITGB8*. Indeed, this would be difficult to demonstrate *in vivo*. Rather, we show that formaldehyde is associated with the increased expression of the miR-125b and the decreased expression of miR-142-3p, and decreased or increased expression of their respective target genes. However, other studies have confirmed some of these specific miRNA-mRNA interactions. For example, miR-125b has been shown to directly target *BAK1* and down-regulate its expression in prostate cancer cells (Shi et al. 2007) and breast cancer cells (Zhou et al. 2010). Our study thereby employs bioinformatics-based approaches to increase knowledge on the interplay between exposure responses, epigenetics, and signaling pathways.

In conclusion, our study demonstrated that formaldehyde inhalation exposure significantly disrupts miRNA expression profiles within the nasal epithelium *in vivo*. Systems level analysis of the transcriptional targets predicted to be regulated by formaldehyde-responsive miR-125b and miR-142-3p revealed the highest enrichment between genes involved in apoptosis signaling and miR-125b. Apoptosis-related gene targets of miR-125b were functionally validated, as they were shown to have altered transcriptional levels after exposure to formaldehyde in the nasal epithelium. These results provide evidence for a relationship between formaldehyde exposure and altered signaling of the apoptotic machinery, likely regulated via

epigenetic mechanisms. These changes in apoptosis-related signaling are of high importance, as an inappropriate balance between cell death and survival heavily influences cellular disease state. Future research will compare these changes to potential formaldehyde-induced changes occurring in tissues collected from sites distal to the respiratory tract *in vivo*. These comparisons may provide key information related to the pathophysiological mechanisms of action of formaldehyde.

CHAPTER 3
FORMALDEHYDE-INDUCED CHANGES IN MICRORNA SIGNALING IN THE RAT

3.1 Overview

MicroRNAs (miRNAs) are critical regulators of gene expression, yet much remains unknown regarding miRNA changes resulting from environmental exposures and whether they influence pathway signaling across various tissues and time.

To gain knowledge on these novel topics, we set out to investigate *in vivo* miRNA responses to inhaled formaldehyde, an important air pollutant known to disrupt miRNA expression profiles.

Rats were exposed by inhalation to either 0 or 2 ppm formaldehyde (6 hours/day) for 7 days, 28 days, or 28 days followed by a 7 day recovery**.** Genome-wide miRNA expression profiles and associated signaling pathways were assessed within the nasal respiratory mucosa, circulating mononuclear white blood cells (WBC), and bone marrow (BM).

We found that miRNAs were responsive to formaldehyde exposure in the nose and WBC, but not the BM. A transcriptomics-based analysis was performed in the nose and WBC of the rats exposed for 28 days. In the nose, formaldehyde altered the expression of 42 transcripts; of these, 15 (36%) were computationally predicted to be regulated by formaldehyde-responsive miRNAs. Conversely, in the WBC, formaldehyde altered the expression of 130 transcripts; of these, 18 (14%) were predicted to be regulated by miRNAs. Systems-level analyses revealed that the transcripts regulated by miRNAs play diverse roles in cell signaling. Key players include dosage suppressor of mck1 homolog, meiosis-specific homologous recombination (*Dmc1*) and secreted frizzled-related protein 4 (*Sfrp4*) within the nose, involved in cell death signaling. In WBC, key players were v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma) (*Akt3*) and integrin, alpha 2 (*Itga2*), involved in inflammation signaling.

Our study informs critical knowledge towards the biological consequences of inhaled formaldehyde exposure.

3.2 Study Objectives

In this study, we set out to assess the effects of formaldehyde inhalation exposure by examining miRNA endpoints across time and throughout multiple tissues. There is currently a lack of knowledge regarding how miRNAs respond to environmental toxicants over time. In order to fill this knowledge gap, we assessed genome-wide miRNA expression profiles in rats exposed to 0 or 2 ppm formaldehyde for 6 hours/day for 7 days, 28 days, or 28 days after an additional 7 days of recovery. We also performed a novel comparison of formaldehyde-induced miRNA expression changes throughout three tissues: (1) the nasal epithelium, (2) circulating white blood cells (WBC), and (3) the bone marrow (BM). These miRNA responses were integrated with formaldehyde-induced transcriptomic changes, where miRNA-mRNA interactions were predicted and assessed at the systems-level. Taken together, our study contributes critical knowledge towards the understanding of mechanisms underlying formaldehyde-induced health effects, as well as a broader understanding on how miRNAs respond as a function of time and tissue.

3.3 Materials and Methods

3.3.1 Animals

For this study, male Fischer rats (Charles River, Wilmington, MA) were selected from the Lovelace Respiratory Research Institute colony. Animals were exposed, sedated, and euthanized using protocols approved by the Lovelace Research Institute's animal care and use committee (FY10-094). Animals were approximately six to eight weeks of age and weighed between 150 and 250 grams. Animals were housed up to two per cage in approved housing chambers with hardwood chip bedding. Caging and bedding was changed two times per week. The animal room environment was maintained at 18-26°C and 30-70% humidity with a 12 hour light-dark cycle. Animals were provided unlimited access to food (2016C Harlan Global Certified Rodent Chow, Harlan Teklad, Madison, WI) and municipal water except during conditioning and inhalation exposure periods. After a two week quarantine period, animals were conditioned to nose-only exposure tubes for 6.5 hours/day over four conditioning sessions.

3.3.2 Formaldehyde Exposures

In this experiment, male Fischer rats received nose-only inhalation exposures of 2 ppm formaldehyde. Three exposure durations were investigated: (1) 2 ppm formaldehyde exposure, 6 hours/day, for 7 days (7-day group), (2) 2 ppm formaldehyde exposure, 6 hours/day, for 28 days (28-day group), and (3) 2 ppm formaldehyde exposure, 6 hours/day, for 28 days, with a 7 day recovery period following the last exposure (28-day plus recovery group). Control (unexposed) rats were placed in nose-only exposure tubes containing room air for the same duration.

An exposure dose of 2 ppm formaldehyde was used, as this exposure level has been shown to alter gene expression profiles (Andersen et al. 2008) and DNA adduct levels (Lu et al. 2011) in sites of direct contact in rats. Isotope labeled [$^{13}CD_2$]-formaldehyde was used in order to allow the detection of exogenously produced DNA adducts in a parallel experiment, similar to our previous investigations (Lu et al. 2011; Moeller et al. 2011). To generate the exposure conditions, deuterated/^{13}C labeled paraformaldehyde (Cambridge Isotope Laboratories, Inc) was vaporized and directed through a delivery line and into a Tedlar® bag. During each of the 6 hour exposures, one Tedlar® bag was diluted with pre-filtered air and delivered to the inhalation chambers at a target formaldehyde concentration of 2 ppm. Nose-only chamber concentrations were monitored by collecting breathing port samples using Waters XpoSure Aldehyde Sampler cartridges. Cartridges were analyzed by ultraviolet high-performance liquid chromatography. For more detailed exposure protocol descriptions, see our previous publication (Lu et al. 2011).

3.3.3 Sample Collection

After the last exposure period (or the last recovery period for the 28-day plus recovery group), animals were euthanized using an intraperitoneal injection of pentobarbital-based euthanasia solution (Euthasol, Virbac Corp, Fort Worth, TX). Sample collection started immediately after the last exposure in order to parallel sacrifice and sample collection times used in our previous studies (Lu et al. 2011; Moeller et al. 2011). To collect nasal epithelial tissue, the rat skull was split with a slight bias to one side (to preserve septal mucosa) and tissue from the maxilloturbinate was collected and stored in RNAlater® (Qiagen, Valencia, CA, USA). Whole blood was collected by cardiac stick using a heparin-laced syringe, and WBC were isolated using Vacutainer® CPT™ cell preparation tubes (Becton Dickinson, Franklin Lakes, NJ, USA) and stored in TRIzol LS (Life Technologies, Grand Island, NY, USA). To collect BM cells, femurs

were flushed with saline and BM cells stored in RNAlater® (Qiagen). All samples were stored at -80°C and shipped by overnight courier on dry ice to the University of North Carolina at Chapel Hill for further processing.

3.3.4 RNA Isolation

The isolation of total RNA molecules was required to assess miRNA and mRNA expression levels associated with formaldehyde exposure in rats. Nasal epithelial tissue samples were disrupted and homogenized using a TissueRuptor (Qiagen) and RNA isolated using the miRNeasy® kit (Qiagen). For the WBC, samples were homogenized in TRIzol (Life Technologies) and RNA isolated according to the standard TRIzol protocol. The BM samples were filtered through 70-μm nylon mesh filters (Fischer Scientific, Waltham, MA, USA) to remove any bone fragments. Filtered bone marrow samples were then homogenized in TRIzol (Life Technologies) and RNA isolated according to the standard TRIzol protocol. Extracted RNA was quantified with a Nanodrop 1000 spectrophotometer (Thermo Scientific, Waltham, MA) and its integrity verified with a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA), using both the Nano and Small RNA kits. We employed quality control to incorporate RNA integrity number assessment and electrophoretogram peak assessment of small RNA molecules.

3.3.5 MiRNA Microarray Analysis

To assess whether formaldehyde inhalation exposure modifies the expression levels of miRNAs within the nose, WBC, and BM, RNA samples were labeled and hybridized to the Agilent Rat miRNA Microarray, as based off miRBase v16.0. For the nose and BM samples, exposed and unexposed samples were assessed in biological triplicate. For the WBC samples, exposed samples were assessed in biological quadruplicate and unexposed samples were assessed in biological triplicate. Microarray results were extracted using Agilent Feature Extraction software. Microarray data have been submitted to National Center for Biotechnology Information (NCBI) Gene Expression Omnibus repository (Edgar et al. 2002) and are available under accession number GSE42393 (www.ncbi.nlm.nih.gov/geo). To analyze the miRNA microarray results, data were first normalized by quantile normalization. In the case where multiple microarray chips were used, batch effect was identified and removed using Partek® Genomics SuiteTM software (St. Louis, MO). To eliminate background noise, miRNA probes

with signal intensities less than the median signal across all replicates were removed. Differential expression was then defined as a significant difference in miRNA levels between exposed versus unexposed samples, where three statistical requirements were set: (i) fold change of ≥ 1.5 or ≤ -1.5 (average exposed versus average unexposed); (ii) p-value < 0.05 (ANOVA); and (iii) a false discovery rate corrected q-value < 0.1. Analysis of variance (ANOVA) p-values and q-values were calculated using Partek® Genomics Suite™ software (St. Louis, MO). To control the rate of false positives, q-values were calculated as the minimum "positive false discovery rate" that can occur when identifying significant hypotheses (Storey 2003). The miRNAs that met these statistical requirements were identified as significantly altered at the expression level after exposure to formaldehyde, or formaldehyde-responsive.

3.3.6 Transcript (mRNA) microarray analysis

To assess the influence of formaldehyde exposure on gene expression levels, nose and WBC RNA samples from the 28-day group were assessed. Here, exposed and unexposed samples were assessed in biological triplicate. RNA was labeled and hybridized to the Affymetrix GeneChip® Rat Gene 1.0 ST Array. Microarray data have been submitted to National Center for Biotechnology Information (NCBI) Gene Expression Omnibus repository (Edgar et al. 2002) and are available under accession number GSE42394 (www.ncbi.nlm.nih.gov/geo). To analyze the mRNA microarray results, data were first normalized by robust multi-chip average (Irizarry et al. 2003), and background noise was eliminated by removing mRNA probes with signal intensities less than the median signal across all replicates. Because the resulting mRNA microarray data were far more robust than the miRNA microarray data, a slightly different set of statistical parameters was used in this microarray analysis. Specifically, differential expression was defined as a significant difference in mRNA levels between exposed versus unexposed samples, where two statistical requirements were set: (i) fold change of ≥ 1.5 or ≤ -1.5 (average exposed versus average unexposed); and (ii) p-value < 0.01 (ANOVA). The q-value filter was not applied in this analysis; however, a stricter p-value requirement of $p < 0.01$ was required to ensure statistical significance. The genes that met these statistical requirements were identified as differentially expressed upon exposure to formaldehyde.

3.3.7 Computationally Predicting miRNA Transcriptional Targets

In order to gain further understanding of potential interactions between formaldehyde-responsive miRNAs and mRNAs, mRNA targets of the formaldehyde-responsive miRNAs were predicted *in silico*. Here, TargetScanHuman algorithms (release 6.2) were employed to identify potential matches between 3' untranslated mRNA regions and miRNA seed sequences (Whitehead 2012). These algorithms are frequently updated to ensure prediction of targets above the background of false-positive predictions (Grimson et al. 2007). The resulting predicted miRNA-mRNA interactions were filtered for total context plus scores < - 0.1. The total context plus score controlled for factors influencing miRNA targeting, including miRNA binding site type and location, local adenine and uracil content, supplementary pairing, target site abundance, and seed-pairing stability (Garcia et al. 2011). The list of predicted mRNA targets of formaldehyde-responsive miRNAs was then compared to the mRNAs identified as differentially expressed by formaldehyde exposure. mRNAs that were both measured as differentially expressed by formaldehyde and predicted to be targeted by formaldehyde-responsive miRNAs were identified and referred to as the predicted miRNA-mRNA interactions resulting from formaldehyde exposure.

3.3.8 Systems-Level Analysis of Predicted miRNA-mRNA Interactions

Network analysis was performed to understand the systems-level response to formaldehyde inhalation exposure possibly mediated via miRNAs. For this analysis, the mRNAs differentially expressed by formaldehyde and also predicted to be regulated by formaldehyde-responsive miRNAs were overlaid onto a global interaction network. Here, networks were algorithmically constructed based on connectivity, as enabled through Ingenuity Pathway Analysis (Ingenuity Systems®, Redwood City, CA). Significance for each constructed network was evaluated based off a modified Fisher's exact test (Calvano et al. 2005), where networks containing known interactions between formaldehyde-associated proteins and associated signals were mapped. Functional enrichment analysis was performed by querying the networks for over-represented pathways involved in diseases and disorders or molecular and cellular functions. Over-represented pathways were defined as pathways than contain more targets than expected by random chance, as calculated using the right-tailed Fisher's Exact Test. Here, functions with

enrichment p-values < 0.005 were considered significantly enriched within the constructed networks.

3.3.9 Confirming miRNA Microarray Results using RT-PCR

To confirm the miRNA microarray results, formaldehyde-induced changes in miRNA expression were validated using real-time reverse transcriptase polymerase chain reaction (RT-PCR). The same sample replicates used for the microarray analysis were used for RT-PCR, plated in technical triplicate. TaqMan® MicroRNA Primer Assays for hsa-miR-31 (ID 000185) and the U6 housekeeping miRNA (ID 001973) were used in conjunction with the TaqMan® Small RNA Assays PCR kit (Applied Biosystems, Carlsbad, CA). The MyCyler Thermal Cycler (Bio-Rad, Hercules, CA) was used for the reverse transcription step, and the Lightcycler 480 (Roche, Indianapolis, IN) was used for the real-time step. The resulting RT-PCR cycle times were normalized against the housekeeping miRNA, and fold changes in expression were calculated based off delta delta cycle time values. Statistical significance of the difference in miRNA expression levels between the formaldehyde-exposed and unexposed samples were calculated using ANOVA (Partek®).

3.3.10 Confirming mRNA Microarray Results using RT-PCR

To verify the mRNA microarray analysis results for the nose of the 28-day rat group, RT-PCR was performed at the gene expression level. The same sample replicates used for the microarray analysis were used for RT-PCR, plated in technical triplicate. QuantiTect Primer Assays were used in conjunction with QuantiTect SYBR® Green PCR kits (Qiagen) and the Stratagene Mx3005P QPCR System (Agilent Technologies). Specifically, chloride channel calcium activated 3 (*Clca3*) (Cat. No. QT01570870), C-type lectin domain family 11, member a (*Clec11a*) (Cat. No. QT00407043), dosage suppressor of mck1 homolog, meiosis-specific homologous recombination (*Dmc1*) (Cat. No. QT01627241), and secreted frizzled-related protein 4 (*Sfrp4*) (Cat. No. QT00179830) were evaluated for changes in gene expression levels induced by formaldehyde exposure. Resulting RT-PCR cycle times were normalized against the β-actin housekeeping gene, and fold changes in expression were calculated based off delta delta cycle time values. Statistical significance comparing the expression levels between exposed and unexposed samples was calculated using ANOVA (Partek®).

3.4 Results

3.4.1 Study Design

Our study set out to investigate formaldehyde-induced changes in miRNA expression profiles at three different exposure times and three different tissues. Specifically, male Fischer rats received nose-only inhalation exposures of 2 ppm formaldehyde (6 hours/day) for either (1) 7 days, (2) 28 days, or (3) 28 days followed by a 7 day recovery (**Figure 12**). Control (unexposed) rats were placed in nose-only exposure tubes containing room air for the same duration. In order to examine the effects of formaldehyde across multiple tissues, samples were collected from (1) the nasal epithelium, (2) circulating mononuclear WBC, and (3) BM cells immediately after the last treatment.

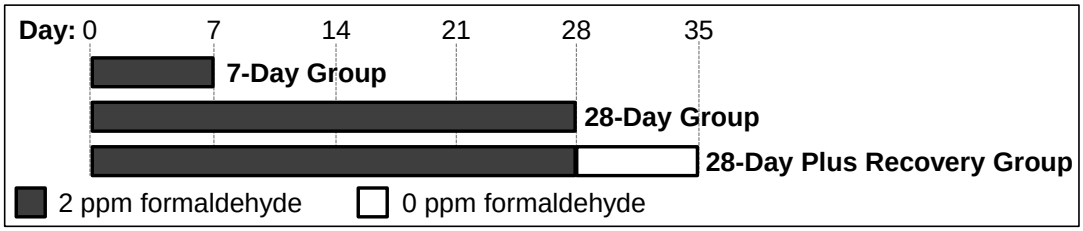

Figure 12: Study design. Responses to formaldehyde were evaluated across three exposure durations, where rats were exposed to 2 ppm formaldehyde (6 hours/day) for either: (1) 7 days, (2) 28 days, or (3) 28 days followed by a 7 day recovery. Control (unexposed) rats were treated using matched conditions.

3.4.2 Formaldehyde Alters miRNA Expression in the Nose and WBC

In order to determine whether formaldehyde inhalation exposure modifies the expression levels of miRNAs within the rat nose, WBC, and BM, small RNAs from these samples were assessed using the Agilent Rat miRNA Microarray, developed using miRBase v16.0. This recently updated array measures the expression levels of 695 rat miRNAs. Using stringent statistical requirements (fold change ≥ 1.5 or ≤ -1.5, p-value < 0.05, q-value < 0.10), microarray analysis revealed that formaldehyde exposure altered the expression of 84, 59, and 0 miRNAs in the nose in the 7-day, 28-day, and 28-day plus recovery groups, respectively. Together, these represent 108 formaldehyde-responsive miRNAs in the nose. In the WBC, formaldehyde exposure altered the expression of 31, 8, and 3 miRNAs in the 7-day, 28-day, and 28-day plus recovery groups, respectively. These represent 40 total formaldehyde-responsive miRNAs in the

WBC. In the BM, no miRNAs were identified as significantly altered at the expressed level using our statistical filters (**Figure 13**, **Supplementary Table 11**).

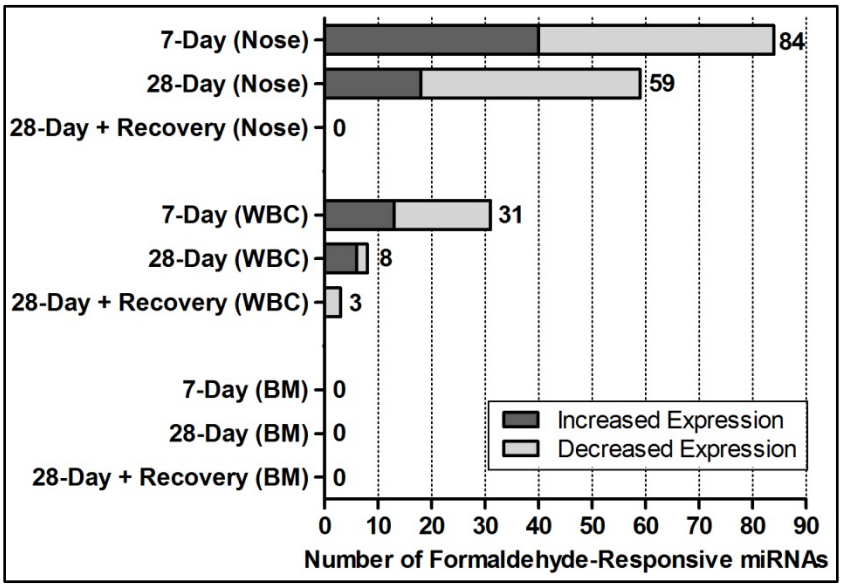

Figure 13: Distribution of formaldehyde-responsive miRNAs across three exposure conditions and three tissues in the rat. Formaldehyde was found to significantly disrupt miRNA expression profiles in the nose and WBC, but not the BM. More miRNAs showed altered expression after 7 days in comparison to 28 days of exposure.

3.4.3 Formaldehyde-Responsive miRNAs are Tissue-Specific

Comparing the formaldehyde-responsive miRNAs across the different tissues reveals largely distinct miRNA responses. To specify, out of a total 108 formaldehyde-responsive miRNAs within the nose, only 10 were also responsive within the WBC (**Figure 14A,B**). Furthermore, most of these overlapping miRNAs displayed different directions of altered expression.

3.4.4 Formaldehyde-Responsive miRNA Expression is Largely Sustained in the Nose

Comparing the formaldehyde-responsive miRNAs across the various exposure conditions reveals that many miRNAs show similar expression patterns in the nose, but not in the WBC (**Figure 14A,C**). To specify, 35 miRNAs were altered at the expression level in the nose in both the 7-day and 28-day exposure groups. Furthermore, 34 of these 35 miRNAs showed altered expression in the same direction. Seven of the 34 miRNAs with sustained decreased expression

over time, let-7a, let-7c, let-7f, miR-10b, miR-126, miR-21, and miR-23a, have previously been shown to be significantly down-regulated in expression in cultured lung cells exposed to 1 ppm formaldehyde (Rager et al. 2011b). One of the 34 miRNAs with sustained decreased expression over time, miR-203, has also previously been shown to be significantly decreased in expression within the nasal epithelium of nonhuman primates exposed to 2 and 6 ppm formaldehyde across two days. It is important to note that these miRNA expression changes did not persist in the nose after 7 days of recovery.

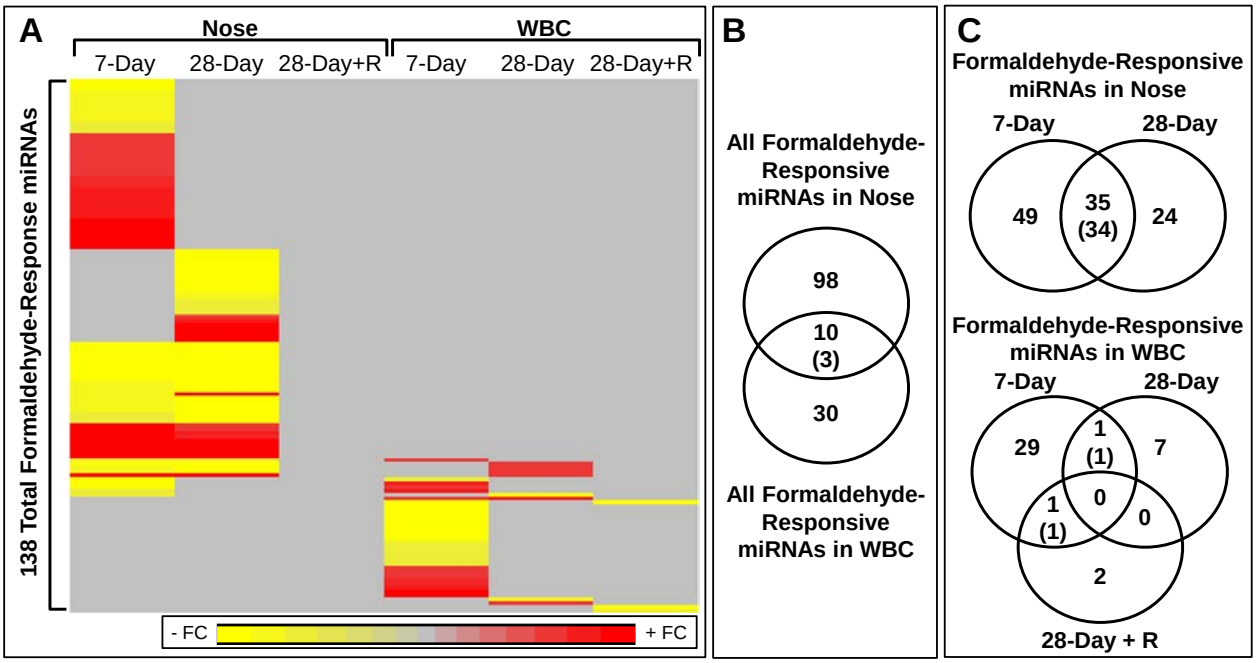

Figure 14: Formaldehyde-responsive miRNAs throughout the nose and WBC. (A) Formaldehyde-responsive miRNAs are largely tissue-specific. Significantly responsive miRNAs are shaded according to fold change value (FC = Exposed / Unexposed). **(B) Few formaldehyde-responsive miRNAs overlap between the nose and WBC. (C) Formaldehyde-responsive miRNAs are largely sustained over time in the nose but not the WBC.** Overlapping miRNAs are enumerated in the centers of the venn diagrams and the number of overlapping miRNAs with the same direction of altered expression is shown within parentheses.

Formaldehyde-responsive miRNAs were less sustained over time within the WBC. To detail, only two miRNAs were significantly altered at the expression level when comparing between the 7-day, 28-day, and 28-day plus recovery groups (**Figure 14C**). The first miRNA, miR-326, showed increased expression in both the 7-day and 28-day groups. The second miRNA, miR-212, showed decreased expression in both the 7-day and 28-day plus recovery

groups. It is important to note that miR-212 also showed decreased expression (FC = -2.34, p-value = 0.008, q-value = 0.155) in formaldehyde-exposed WBC samples from the 28-day group, but it did not pass the stringent multiple test correction requirement.

3.4.5 Formaldehyde Causes Tissue-Specific Gene Expression Changes

In order to assess the influence of formaldehyde inhalation exposure at the gene expression level, a transcriptomics-based analysis was performed using the nose and WBC samples from the 28-day rat group. This additional analysis was performed on the 28-day group in order to understand the genomic effects resulting from longer exposure conditions. Only the nose and WBC samples were assessed at the gene expression level, as no changes in miRNA expression were detected in the BM, and our ultimate goal was to gain insight into possible miRNA-mRNA interactions resulting from formaldehyde exposure.

Here, transcript levels were measured using the Affymetrix GeneChip® Rat Gene 1.0 ST Array, which assesses the expression levels of 27,342 genes. Comparing the gene expression levels within exposed versus unexposed samples revealed that formaldehyde inhalation exposure caused the differential expression of 42 genes in the nose, 3 of which were up-regulated and 39 were down-regulated in expression (**Figure 15A**, **Supplementary Table 12**). In the WBC, 130 genes were differentially expressed, where 123 were up-regulated and 7 were down-regulated in expression (**Figure 15B**, **Supplementary Table 13**). Of the genes differentially expressed by formaldehyde in the nose and WBC, only two overlapped between the two tissues, and these genes showed different directions of differential expression associated with formaldehyde (**Figure 15C**). This analysis, therefore, shows that the transcriptomic changes induced by formaldehyde inhalation exposure are tissue-specific.

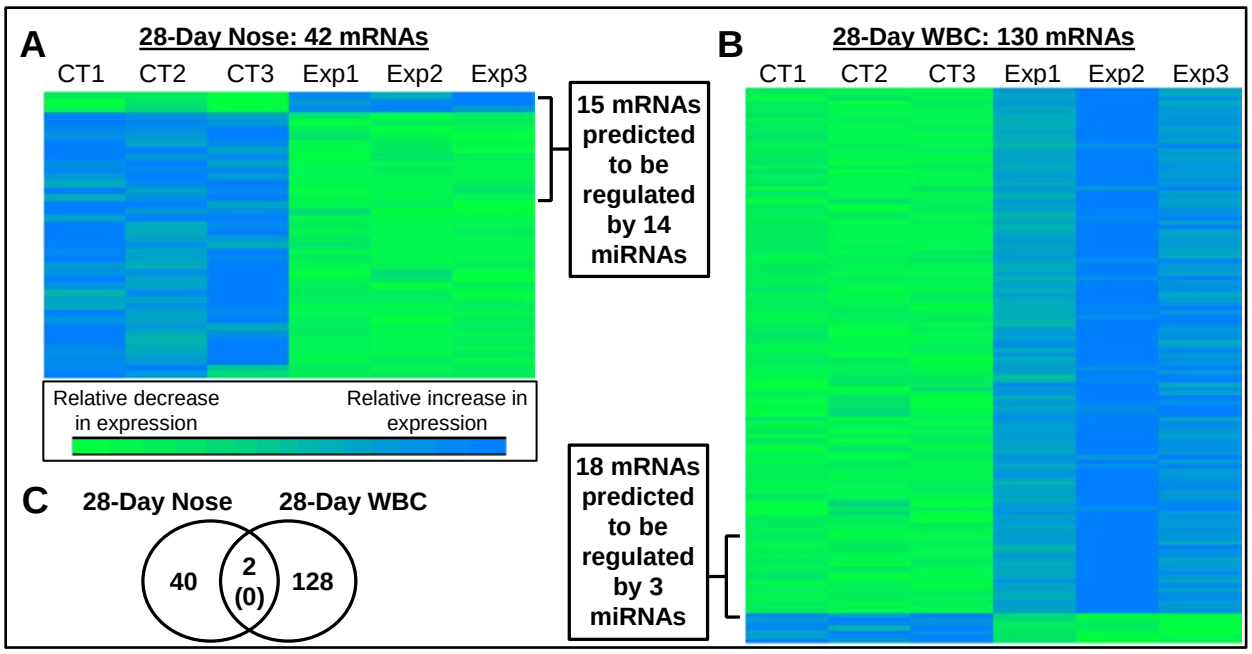

Figure 15: Gene expression changes induced by formaldehyde exposure in the rat nose and WBC of the 28-day group. (A) Formaldehyde changes the expression levels of 42 transcripts in the nose. (B) Formaldehyde changes the expression levels of 130 transcripts in the WBC. The relative, z-score normalized expression levels of messenger RNAs (mRNAs) are displayed. CT refers to unexposed control samples, and Exp refers to formaldehyde-exposed samples. **(C) Transcriptional changes induced by formaldehyde exposure are tissue-specific.** The number of differentially expressed mRNAs for each tissue is shown, where overlapping mRNAs are enumerated in the center of the venn diagram and the number of overlapping mRNAs differentially expressed in the same direction is shown within parentheses.

3.4.6 Formaldehyde-Responsive miRNAs May Mediate a Fraction of Transcriptional Effects

To gain insight into the potential impact miRNAs may have on gene expression, interactions between formaldehyde-responsive miRNAs and differentially expressed transcripts were computationally predicted based on sequence matches between 3' untranslated mRNA regions and miRNA seed sequences. Here, 15 of the 42 (36%) differentially expressed transcripts in the nose were predicted to be regulated by 14 miRNAs in the 28-day group (**Figure 15A**). In the WBC, on the other hand, 18 of the 130 (14%) differentially expressed transcripts were predicted to be regulated by 3 miRNAs (**Figure 15B**). It is important to note that the computational miRNA database did not recognize all of the formaldehyde-responsive miRNAs, leaving some potential interactions unexamined. Still, these predictions suggest that the magnitude of miRNA-mediated gene expression regulation may differ between tissues. The

predicted miRNA-mRNA interactions resulting from formaldehyde exposure are illustrated in more detail within **Figure 16**.

3.4.7 Systems-Level Analysis Reveals miRNA-Mediated Signaling of Critical Pathways

A systems biology-based analysis was performed to understand formaldehyde-associated pathways and cellular functions that are possibly mediated by miRNAs. For this network analysis, the proteins encoded by transcripts differentially expressed by formaldehyde and also predicted to be regulated by formaldehyde-responsive miRNAs (n=15 for the nose, n=18 for the WBC) were mapped onto molecular networks (**Supplementary Table 14**). Within these significant networks ($p < 0.001$), signaling of proteins involved in certain molecular/cellular functions and diseases was enriched (**Table 3**). Key players involved in the enriched functions and diseases include DMC1 and SFRP4 within the nose, involved in signaling related to cancer, cell death and survival, cell development, and cell growth. In WBC, key players include AKT3 and ITGA2, involved in signaling related to cell growth, cell movement, and inflammation. While the identified gene sets are small, p-values are adjusted for size and represent enriched functions. All the constructed networks were combined for each tissue, and only molecules required to maintain network structure are shown illustrating important miRNA-mediated signaling within the nose and WBC (**Figure 16**).

In the nose, key players involved in multiple enriched signaling patterns include proteins encoded by dosage suppressor of mck1 homolog, meiosis-specific homologous recombination (*Dmc1*) and secreted frizzled-related protein 4 (*Sfrp4*). These two transcripts, predicted to be regulated by formaldehyde-responsive miRNAs, are both involved in signaling related to cell growth and proliferation and cell death and survival, among others (**Table 3**). Also within the miRNA-mediated signaling in the nose are signaling interactions involving the transcription factor, hepatocyte nuclear factor 4, alpha (HNF4A).

In WBC, key players involved in multiple enriched signaling patterns include proteins encoded by v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma) (*Akt3*) and integrin, alpha 2 (*Itga2*). These two transcripts, predicted to be regulated by formaldehyde-responsive miRNAs, are both involved in signaling related to inflammatory response, among others (**Table 3**).

Table 3: Functions and diseases that are enriched within the miRNA-mediated signaling networks associated with formaldehyde exposure in the (A) nose and (B) WBC of the rat 28-day group.

Function or Disease	p-value	Proteins Encoded by Formaldehyde-Associated Transcripts
(A) Nose		
Cellular Development	0.0005	CLEC7A, CLEC11A, DMC1, EHD4, NOV, OSMR, SFRP4
Cellular Growth and Proliferation	0.0006	CLEC7A, CLEC11A, DMC1, EHD4, GBP2, NOV, OSMR, SFRP4
Cell Cycle	0.0008	DMC1, GBP2
Cellular Movement	0.0008	CLEC11A, NOV, SFRP4
Dermatological Diseases and Conditions	0.0008	CLEC7A,CLEC11A,GBP2, OSMR
Developmental Disorder	0.0008	HS3ST1, MID1, OSMR
DNA Replication, Recombination, and Repair	0.0008	DMC1
Hereditary Disorder	0.0008	CLEC7A, MID1, OSMR
Hypersensitivity Response	0.0008	CLEC11A
Infectious Disease	0.0008	CLEC7A, CLEC11A
Inflammatory Response	0.0008	CLEC7A, CLEC11A
Lipid Metabolism	0.0008	CLEC7A, CLEC11A, FAR1
Nucleic Acid Metabolism	0.0008	GBP2, FAR1
Small Molecule Biochemistry	0.0008	CLEC7A, CLEC11A, FAR1, GBP2, SFRP4
Cell Morphology	0.0016	CLEC11A, DMC1, EHD4, MID1, NOV, OSMR
Metabolic Disease	0.0016	NOV, OSMR
Molecular Transport	0.0023	CLEC7A, CLEC11A, SFRP4
Cancer	0.0024	CLEC11A, DMC1, EHD4, GBP2, NOV, SFRP4
Cell-To-Cell Signaling and Interaction	0.0024	CLEC7A,CLEC11A,GBP2, NOV
Hematological Disease	0.0024	CLEC11A
Cell Death and Survival	0.0032	CLEC11A, DMC1, EHD4, NOV, SFRP4
Cell Signaling	0.0040	GBP2
Cardiovascular Disease	0.0048	NOV
(B) WBC		
Cellular Development	0.0010	AKT3, ITGA2
Developmental Disorder	0.0010	ITGA2
Respiratory Disease	0.0010	ITGA2
Cellular Growth and Proliferation	0.0020	AKT3, ITGA2, SKAP2, TES
Neurological Disease	0.0020	AKT3
Cell Morphology	0.0030	AKT3, DNAJC3, ITGA2
Small Molecule Biochemistry	0.0030	AKT3, DNAJC3, TPK1
Inflammatory Response	0.0030	AKT3, ITGA2
Nucleic Acid Metabolism	0.0030	AKT3, TPK1
Cell Signaling	0.0030	MUC20
Cell-To-Cell Signaling and Interaction	0.0030	ITGA2
Vitamin and Mineral Metabolism	0.0030	TPK1

Function or Disease	p-value	Proteins Encoded by Formaldehyde-Associated Transcripts
Cellular Movement	0.0037	AKT3, ITGA2
Gene Expression	0.0040	ITGA2

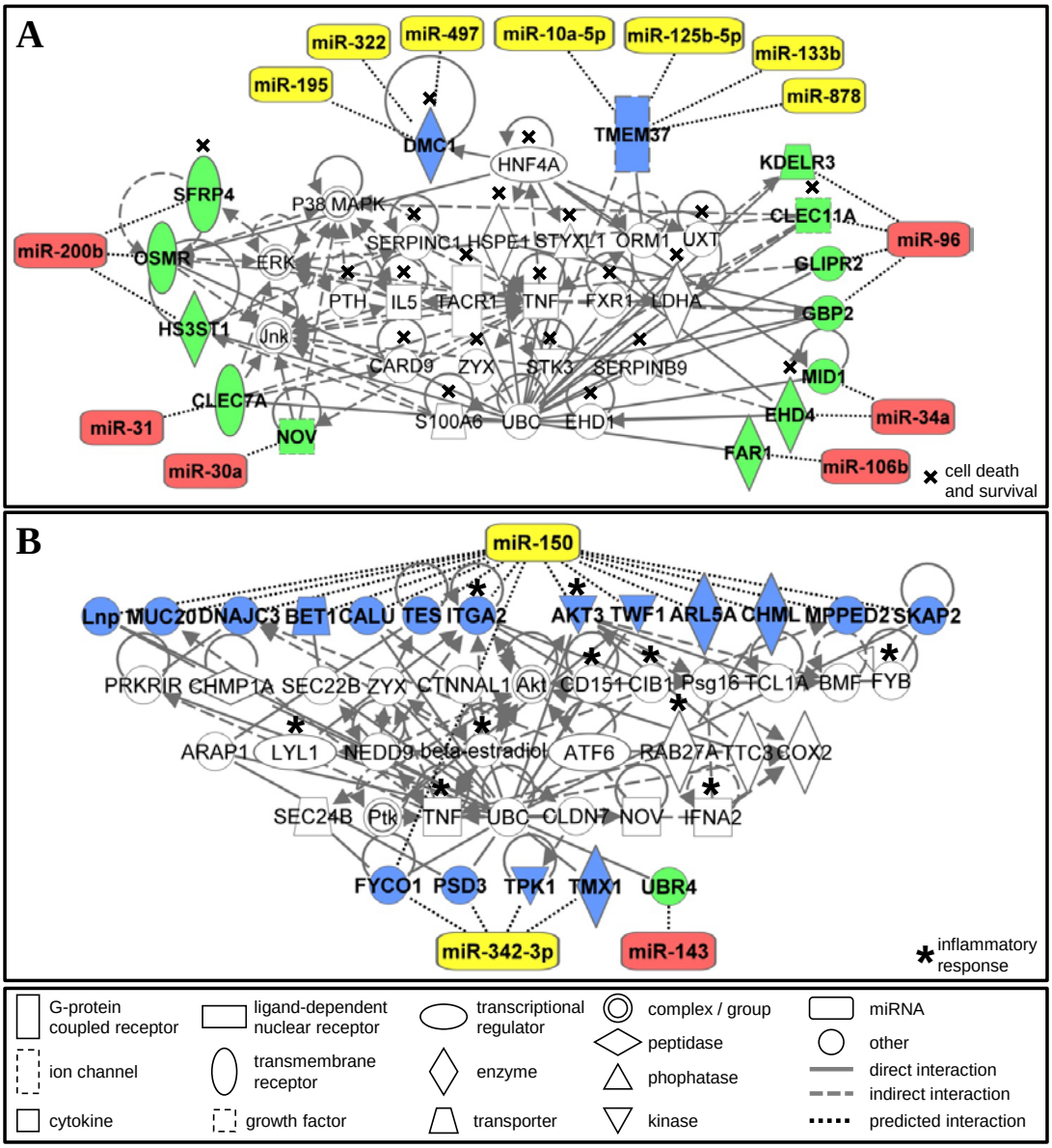

Figure 16: Network interactomes associated with formaldehyde exposure in the rat 28-day group. (A) miRNA-mediated signaling in the nose involves cell death and survival. (B) miRNA-mediated signaling in the WBC involves inflammatory response. Genes that have formaldehyde-induced increased expression are shown in blue, and genes that have decreased expression are shown in green. MiRNAs that have formaldehyde-induced increased expression are shown in red, and miRNAs that have decreased expression are shown in yellow. Proteins associated with the network signaling are white.

3.4.8 Microarray Results were Confirmed using RT-PCR

To confirm the microarray analyses, formaldehyde-induced changes in miRNA and transcript expression levels were validated using RT-PCR. To specify, RT-PCR was performed on miR-31, the miRNA showing the greatest increase in expression in the nose of the 28-day group and also an increase in expression in the nose of the 7-day group. Supporting the microarray findings, miR-31 was measured as significantly increased in expression in the formaldehyde-exposed nose samples of the 7-day and 28-day groups through RT-PCR (**Figure 17**). A subset of genes differentially expressed by formaldehyde in the nose of the 28-day group was also tested using RT-PCR. Specifically, *Dmc1* and *Clca3* were tested, as they showed the greatest increase and decrease in expression resulting from exposure, respectively. Two other genes, *Clec11a* and *Sfrp4*, were evaluated using RT-PCR in order to test genes that were not as drastically altered at the expression level, but still identified as differentially expressed. All of the tested genes showed significant formaldehyde-induced changes in expression that aligned with the microarray findings (**Figure 17**).

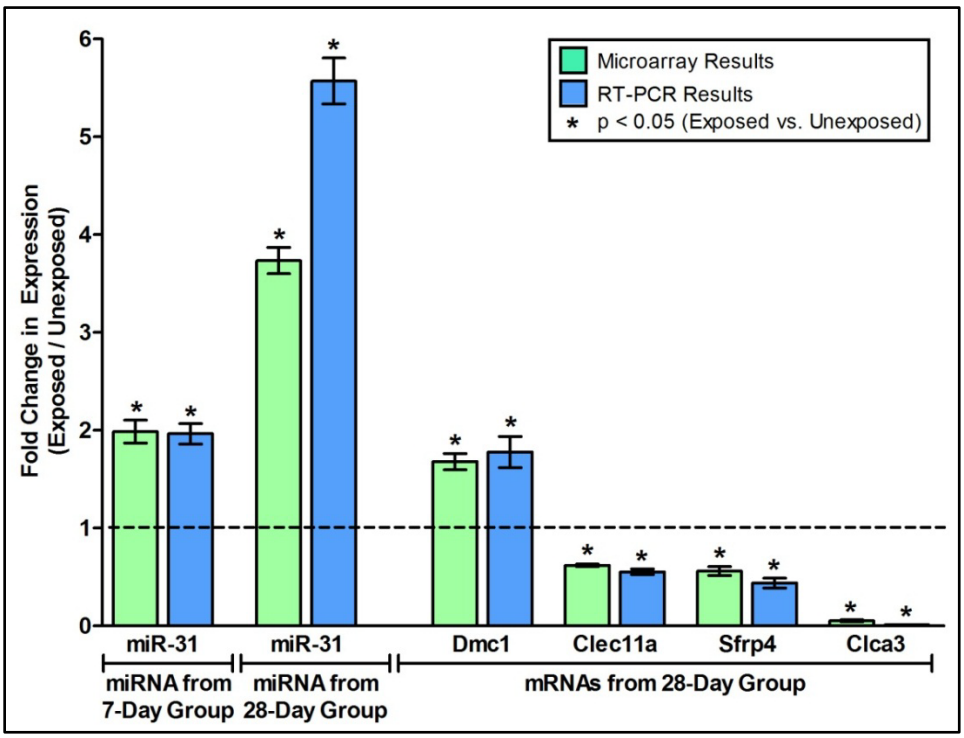

Figure 17: Microarray results align with RT-PCR results in the rat nose.

3.5 Discussion

This study employed a multi-tiered approach to enable an understanding of the genome-wide miRNA responses to an important environmental toxicant and relate these to alterations in transcription-based signaling pathways. Changes in miRNA expression were investigated over a time series through the evaluation of various exposure durations. Changes in miRNA expression were also evaluated over a spatial distribution through the evaluation of various sites across the body. The environmental air pollutant formaldehyde was selected as the toxicant for this investigation as it is currently ranked as a high priority contaminant of concern and our previous study showed that formaldehyde significantly disrupts miRNA expression profiles *in vitro* (Rager et al. 2011b). For the current study, formaldehyde was further evaluated *in vivo*, where rats received nose-only inhalation exposures of 2 ppm formaldehyde 6 hours/day for either 7 days (7-day group), 28 days (28-day group), or 28 days followed by a 7 day recovery (28-day plus recovery group). Control (unexposed) rats were treated using matched conditions.

This investigation is the first to simultaneously examine exposure-induced perturbations in genome-wide miRNA expression throughout the nose, circulating WBC, and BM. The nose was the site of greatest change in miRNA expression profiles (n=108 miRNAs) resulting from formaldehyde inhalation exposure followed by the WBC (n=40 miRNAs) after 7 and 28 days of exposure. There were no miRNAs that showed significantly altered expression in the BM using the statistical parameters of FC $\geq \pm$ 1.5, p-value < 0.05, and q-value < 0.10 (exposed versus unexposed). These findings suggest that cells in direct contact with formaldehyde exposure display more drastic responses at the miRNA-level than cells distant from the exposure contact sites immediately after exposure.

It is noteworthy that both the nose and WBC showed a higher number of formaldehyde-responsive miRNAs after 7 days of exposure in comparison to 28 days of exposure, suggesting that both these tissues may display a degree of adaptive or compensatory responses to 2 ppm formaldehyde exposure over time. This adaptive trend coincides with a previous genomics study showing that 2 ppm formaldehyde inhalation exposure significantly disrupts the expression of more genes after 5 days of exposure than 15 days of exposure in the rat nasal epithelium (Andersen et al. 2008). It is important to note that this adaptive trend in miRNA expression may be dose-dependent and not apparent at higher concentrations. For example, 6 ppm formaldehyde exposure has been shown to cause sustained alterations in gene expression (Andersen et al. 2008)

and increases in cell proliferation across time within the rat nasal epithelium (Monticello et al. 1991). Future studies will evaluate how miRNA expression levels change throughout time in higher exposure conditions.

The degree to which formaldehyde-responsive miRNAs are sustained across the exposure time points within the rat nose is substantial. More specifically, 34 of the 59 (58%) formaldehyde-responsive miRNAs in the nose of the 28-day group were also altered in the same direction in the 7-day group. Many of these miRNAs with sustained expression alterations have also been shown to be altered in cultured human lung cells exposed to formaldehyde (Rager et al. 2011b). Specifically, let-7a, let-7c, let-7f, miR-10b, miR-126, miR-21, and miR-23a were all significantly decreased in expression in cultured lung cells exposed to formaldehyde (Rager et al. 2011b) and in the nose of rats exposed to formaldehyde for 7 and 28 days. Of these miRNAs, let-7a, let-7c, let-7f, and miR-10b are also decreased in expression in nasopharyngeal carcinoma tissue in comparison to healthy tissue (Li et al. 2011). These common patterns of miRNA expression associated with both exposure and disease could represent etiologic relationships which warrant further investigation.

In contrast to the nose, the WBC showed few miRNAs with shared expression patterns across the time points. The only miRNA common to the 7 day and 28-day plus recovery groups was miR-212, which showed significantly decreased expression in WBC. This miRNA was also decreased in expression in the 28-day group but was not statistically significant. Although the function of miR-212 within circulating WBC is unknown, in lung cancer cells miR-212 plays an important role in cell death, where inhibition of miR-212 has been shown to decrease apoptosis involving the tumor necrosis factor-related apoptosis-inducing ligand (Incoronato et al. 2010).

One of the goals of this study was to compare formaldehyde-altered miRNA expression to transcript levels in order to determine the relationship between these two genomic events. To achieve this goal, we next performed a genome-wide mRNA expression analysis in the nose and WBC of the 28-day group. Gene expression analysis was not performed in the BM, as miRNA expression levels within the BM were not significantly altered by formaldehyde. Within the nose, we found that formaldehyde exposure caused the differential expression of 42 genes. It is not surprising that many of these genes have previously been identified as differentially expressed by formaldehyde exposure. For example, C-type lectin domain family 11, member a (*Clec11a*), midline 1 (*Mid1*), phospholipase A2, group IVA (cytosolic, calcium-dependent)

(*Pla2g4a*), and schlafen 2 (*Slfn2*) expression levels were identified here as well as in a previous study evaluating formaldehyde-induced changes in the rat nasal epithelium (Thomas et al. 2007).

Within the WBC of the rats exposed for 28 days, 130 genes were differentially expressed by formaldehyde exposure. This transcriptomic response was unexpected, as this trend was reversed at the miRNA-level, where fewer miRNAs were altered by formaldehyde in the WBC in comparison to the nose. These data suggest that other transcriptional regulators, for example, DNA methylation and histone modifications (Cedar et al. 2009), likely play a role in formaldehyde-induced genomic response.

There are very few studies that have evaluated the effects of formaldehyde inhalation exposure on gene expression within blood. It has been shown that the expression levels of six genes in human whole blood samples correlate with formaldehyde exposure dose, as evaluated using urinary concentrations of a formaldehyde adduct, thiazolidine-4-carboxylate (Li et al. 2007). Contrasting this finding, another study found that formaldehyde inhalation exposure of up to 0.7 ppm for 5 days (4 hours/day) did not cause significant changes in gene expression patterns within the blood of humans (Zeller et al. 2011). Our study is the first to evaluate the transcriptomic effects of formaldehyde inhalation exposure within the blood of rodents.

A previous rodent study evaluated formaldehyde's effect on blood at the proteomics-level, where rats were exposed to 0, 5, or 10 ppm formaldehyde for two weeks at 6 hours/day and 32 plasma proteins were identified as potential biomarkers of exposure (Im et al. 2006). This study also identified two cytokines with modified levels associated with formaldehyde inhalation exposure, suggesting that formaldehyde induces an inflammatory effect within the blood (Im et al. 2006). None of these proteins were encoded by genes we identified as differentially expressed in the WBC. However, we did identify other genes with altered expression in the WBC that play a role in inflammation. To date, there is a lack of studies with consistent findings regarding formaldehyde's influence on gene expression and protein levels within circulating blood, making our study's findings critical to elucidating the systemic effects of formaldehyde inhalation exposure.

Realizing that miRNAs are not the only regulators of gene expression in the cell, we set out to predict which of the formaldehyde-induced transcriptional changes may be attributable to the miRNAs themselves. The aim of this analysis was to determine potential miRNA-mRNA interactions that occur within the nose and WBC upon exposure to formaldehyde. For this

analysis, interactions between formaldehyde-responsive miRNAs and differentially expressed transcripts were computationally predicted based on seed match-based algorithms in the nose and WBC of the 28-day group. It is important to note that some rat miRNAs and transcripts were not recognized by the computational prediction database, and as databases become more populated with more species-specific information, our ability to perform such analyses will improve. Taking this limitation into account, we predicted that 15 of the 42 (36%) differentially expressed transcripts were predicted to be regulated by formaldehyde-responsive miRNAs in the nose. In the WBC, 18 of the 130 (14%) differentially expressed transcripts were predicted to be regulated by formaldehyde-responsive miRNAs. These findings suggest that the extent of miRNA-mediated control may differ between tissues, and that other mechanisms likely contribute to the formaldehyde-induced genomic response.

A systems biology-based analysis was performed to identify formaldehyde-associated pathways and cellular functions that are likely influenced by miRNAs within the rat nose and WBC. For this systems-level analysis, molecular networks were constructed of proteins encoded by the differentially expressed transcripts predicted to be regulated by formaldehyde-responsive miRNAs. Specifically, the 15 transcripts predicted to be regulated by miRNAs in the nose, and the 18 transcripts predicted to be regulated by miRNAs in the WBC, were used to construct molecular signaling networks.

Within the resulting networks, key players involved in multiple enriched functional signaling patterns were identified in the miRNA-mediated signaling in the nose and WBC. Within the nose, enriched cell proliferation and cell death-related signaling involving the key players, DMC1 and SFRP4, was present. For example, DMC1 is essential for meiotic recombination, where targeted Dmc1 disruption causes apoptosis (Yoshida et al. 1998). SFRP4 is a secreted-type WNT signaling inhibitor which plays an important role in the regulation of cell proliferation and cell death (Katoh et al. 2007). The enrichment for cell proliferation and cell death-related signaling in the nose coincides with previous studies showing that formaldehyde inhalation exposure causes cell death and cytotoxicity-induced cell proliferation (Monticello et al. 1991) and altered genomic signaling related to cell proliferation and apoptosis (Andersen et al. 2010; Hester et al. 2003) in the rat nasal epithelium. Results of the network analysis highlight the role of the transcription factor, HNF4A, within miRNA-mediated signaling in the nose.

Interestingly, this transcription factor has previously been implicated in the response of lung cells exposed to air pollutant mixtures containing formaldehyde (Rager et al. 2011a).

The miRNA-mediated signaling in the WBC involved the key players AKT3 and ITGA2, which are both related to inflammation signaling. To elaborate, AKT3 is involved in inflammation and cytokine production (Wang et al. 2012), and ITGA2 is involved in inflammation and platelet aggregation (Jirouskova et al. 2007). Validating our computational predictions, miR-150 has been shown to suppress Akt3 expression in the mouse lung (Wang et al. 2012). In addition to miRNAs mediating inflammation-related gene expression, the altered signaling of miRNAs in the WBC may be a secondary effect associated with formaldehyde-induced inflammation. The expression levels of miRNAs have been shown to be altered under conditions of inflammation, including in the WBC of human subjects exposed to lipopolysaccharides (Schmidt et al. 2009), as well as in response to activation of antiviral signaling proteins (Goncharova et al. 2010). As previously discussed, formaldehyde's ability to influence inflammatory-related proteins within the blood has been shown in one study (Im et al. 2006), but this response is not well characterized. Future research should investigate the relationship of inflammatory response pathways as mediators of formaldehyde-induced miRNA change.

Taken together this study advances the growing body of knowledge on miRNAs by evaluating environmental exposure-induced changes in miRNA expression and relating these changes to gene expression regulation and cell signaling. We find that formaldehyde inhalation exposure significantly disrupts miRNA expression profiles within the rat nose and WBC, but not within the BM. While we recognize that miRNAs are but one regulator of gene expression, this finding could inform our understanding of diseases associated with formaldehyde. We also find that formaldehyde-induced changes in miRNA and transcript expression are largely tissue-specific, where there is minimal overlap between formaldehyde-responsive miRNAs or transcripts in the nose compared to WBC. In addition, a systems biology-based analysis of the miRNA-mediated transcriptional changes reveals that formaldehyde-responsive miRNAs may mediate important pathways involved in critical cell functions, including cell death in the nose and inflammation in the WBC. These results increase the understanding of mechanisms and biological pathways underlying formaldehyde-induced effects, and also broaden our knowledge on how miRNAs respond across time and in different tissues.

DISCUSSION AND CONCLUSIONS

Understanding the biological impacts upon exposure to formaldehyde via inhalation exposure is crucial in order to effectively promote global public health. Formaldehyde is a ubiquitous air pollutant present throughout indoor and outdoor environments, thus it is imperative to understand how it influences human health and identify which target tissues are impacted upon exposure. The studies described within the three chapters contribute meaningful findings to the overall field of toxicology, as they are the first to: (1) investigate formaldehyde's effects on miRNA expression profiles, (2) relate formaldehyde-altered miRNAs to critical biological signaling pathways, and (3) compare and contrast epigenomic responses to an air toxicant throughout multiple regions of the body across multiple exposure durations. Together, these findings contribute novel science to the knowledge of mechanisms and biological pathways underlying formaldehyde-induced effects.

MiRNAs are Key Responders to Formaldehyde in Direct Target Tissues

A common theme throughout this dissertation is the replicated finding that gaseous formaldehyde exposure significantly disrupts miRNA expression profiles in direct target tissues, regardless of the species, exposure level, and exposure duration investigated. In the first chapter, we identified 89 miRNAs with significantly decreased expression upon exposure to 1 ppm formaldehyde in human lung epithelial cells. The research detailed in the second chapter revealed 3 and 13 miRNAs significantly differentially expressed by 2 and 6 ppm formaldehyde exposure, respectively, in nonhuman primate nasal epithelial tissue. Within the third chapter, we identified 84 and 59 miRNAs with significant changes in expression resulting from 7 and 28 days of 2 ppm formaldehyde exposure, respectively, in rodent nasal epithelial tissue. Comparing these formaldehyde-responsive miRNAs across species reveals that 28 miRNAs are altered at the expression level in direct target tissues in at least two of the species investigated (**Figure 18**).

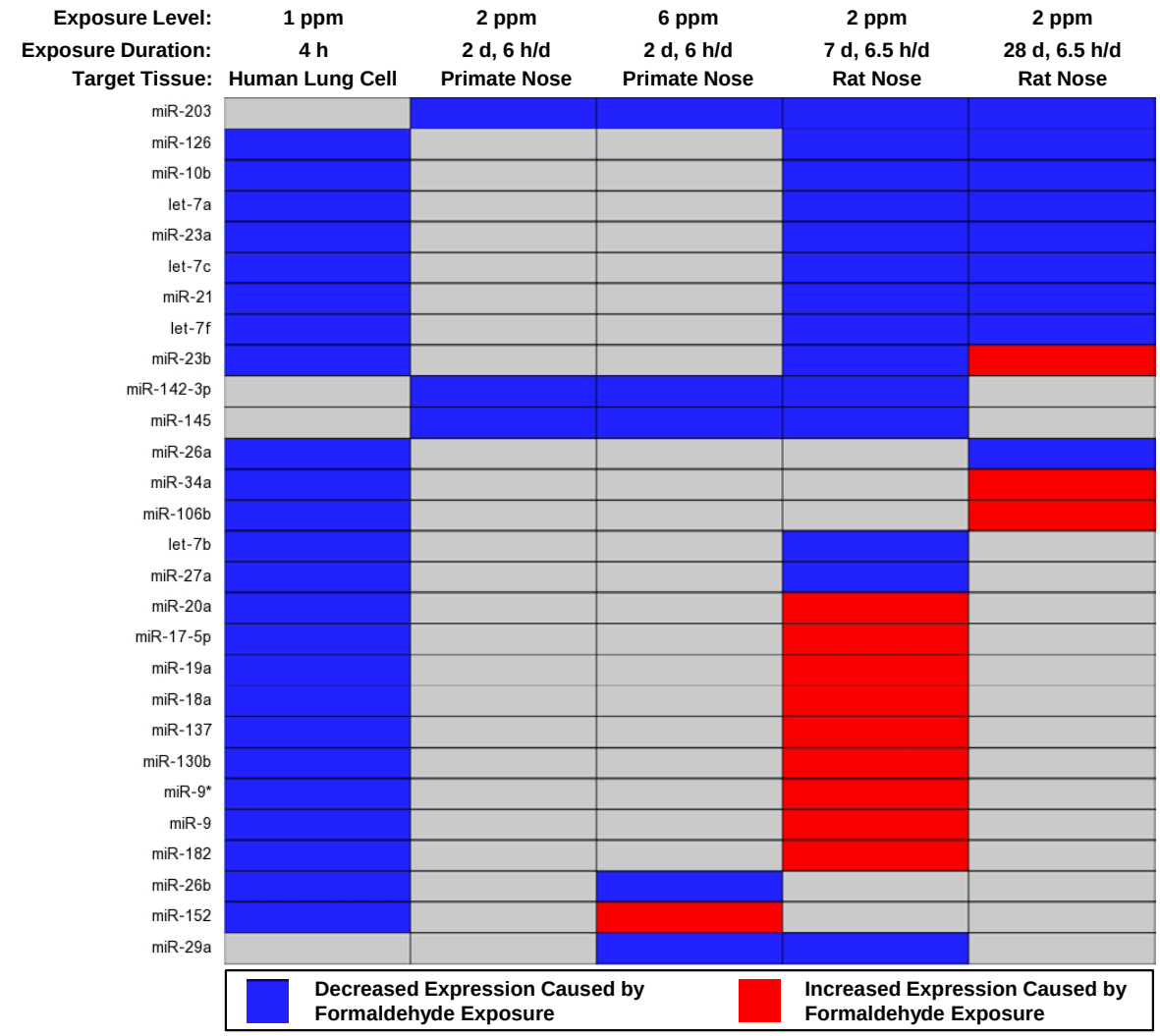

Figure 18: 28 miRNAs that were significantly altered at the expression level within direct target tissues by formaldehyde exposure in at least two of the three models tested. These miRNAs represent critical epigenetic responders to formaldehyde exposure that persist across all three evaluated model systems.

Many of the 28 commonly altered miRNAs have known relationships to disease. For instance, nine of the 28 formaldehyde-responsive miRNAs have also been identified to be dysregulated in nasopharyngeal carcinoma. To specify, let-7a, let-7b, let-7c, let-7f, miR-10b, and miR-203 are all decreased in expression in nasopharyngeal carcinoma tissue (Li et al. 2011; Wong et al. 2012) and also show decreased expression after formaldehyde exposure. Additionally, miR-106b, miR-182, and miR-9* are all increased in expression in nasopharyngeal carcinoma tissue (Wong et al. 2012) and also show increased expression upon exposure to formaldehyde. Although the exact functions of these formaldehyde-altered miRNAs that are

dysregulated in nasal cancer are currently unknown, some miRNAs' functions are beginning to be elucidated. For instance, the disruption of let-7 family members is known to a cause a less differentiated cellular state and contribute to the development of cancer (Roush et al. 2008). The increased expression of miR-106b has also been shown to contribute to tumor cell development by promoting cell cycle progression and modulating the function of cell cycle checkpoints (Ivanovska et al. 2008).

The 28 consistently responsive miRNAs are of high interest, as these molecules could represent future targets for therapeutic strategies. For example, the induction of miR-126 expression has been implicated as a treatment strategy for small cell lung cancer (Miko et al. 2011). This miRNA is commonly found at decreased expression levels in small cell lung cancer cells, but when overexpressed, miR-126 inhibits cancer cell proliferation (Miko et al. 2011). In our studies, we found that miR-126 was significantly decreased in expression by formaldehyde in human lung cells and rat nasal tissue, paralleling the expression profile evident in small cell lung cancer. Our studies thereby support miRNAs as key responders of formaldehyde exposure that may play a role in disease progression, regardless of species.

MiRNAs Respond in a Tissue Type-Specific Manner

The *in vivo* comparison of miRNA expression profiles across tissues revealed that miRNA responses to formaldehyde exposure are largely tissue-specific. As detailed in chapter three, miRNA expression profiles were drastically altered within the rodent nasal epithelium, a region in direct contact with inhaled formaldehyde. Circulating mononuclear white blood cells also showed altered miRNA expression patterns, but to a lesser extent than in the nose. After 7 days of exposure, 84 miRNAs were modified at the expression level in the nose, while 31 were modified in the WBC. Additionally, 28 days of formaldehyde exposure disrupted the expression levels of 59 miRNAs in the nose and 8 miRNAs in the WBC. Comparing the formaldehyde-responsive miRNAs in the nose to the white blood cells revealed only three miRNAs with differential expression in the same direction across both tissues, providing further evidence that formaldehyde alters miRNA expression levels in a highly tissue-dependent manner.

Of high interest, no miRNAs were identified as significantly responsive to formaldehyde exposure within the bone marrow, a region that is distal to tissues directly contacting inhaled

formaldehyde. These results are the first to reveal that, under the tested conditions, the bone marrow does not respond to inhaled formaldehyde at the miRNA level within the rodent model.

Formaldehyde-Responsive miRNAs Regulate Genes Involved in Critical Pathways

To gain insights into the mechanistic consequences of formaldehyde-induced changes in miRNA expression profiles, all three studies employed a systems biology strategy to evaluate putative miRNA-mediated signaling disruptions. All studies performed functional enrichment analyses by querying the miRNA-mediated network signaling likely modified by formaldehyde for known associations to cellular function and disease, and in chapters 2 and 3 these were functionally validated. It is notable that many of the enriched functions overlapped between each species/tissue analysis (**Figure 19**). Strikingly, four of the same functions were identified as enriched across all the species/tissues, namely, cellular development, cellular growth and proliferation, cellular movement, and small molecule biochemistry. Also of note was the enrichment for inflammatory response and cancer, present in most of the putative miRNA-mediated signaling responses. These results suggest that despite potential differences between species or tissues, formaldehyde-associated changes in biological pathway signaling related to similar biological functions remains a common theme.

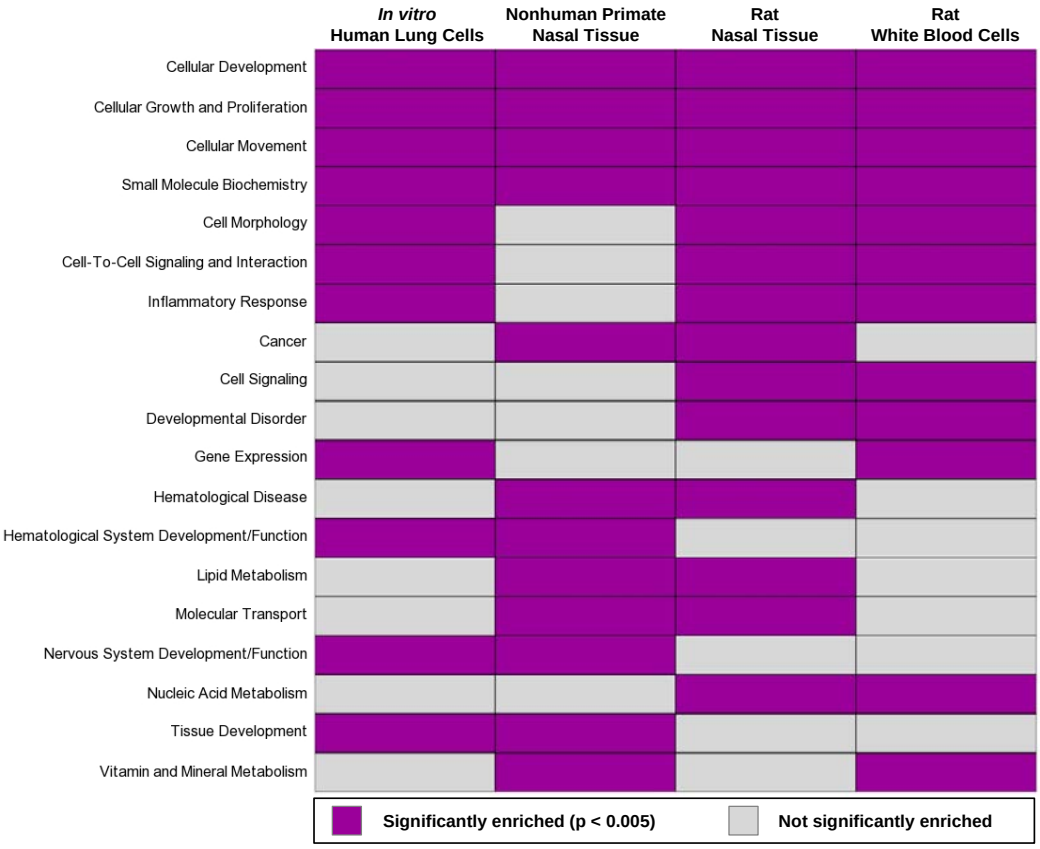

Figure 19: Common biological functions and disease signatures are enriched for by miRNA-mediated signaling responses to formaldehyde exposure across multiple species/tissues.

Proposed Mechanism for Changes in Distal Signaling Induced by Formaldehyde

An important finding from the research detailed in chapter three was that inhaled formaldehyde significantly disrupted miRNA and gene expression profiles within the nasal epithelium and circulating mononuclear white blood cells *in vivo*. **A mode of action linking formaldehyde inhalation exposure to changes in sites distal to the respiratory tract, including the circulating blood and bone marrow, is currently unknown.** Because formaldehyde is highly reactive and undergoes rapid metabolism (IARC 2006), some scientists believe it is unlikely for such a compound to cause toxicity at sites distant from the respiratory tract upon inhalation (Cole et al. 2004; Golden et al. 2006; Heck et al. 2004; Pyatt et al. 2008). However, this line of evidence does not take into account possible mechanisms involving secondary toxicity or downstream events that may influence sites distant from direct targets of exposure. In fact, many scientists articulate that alternative mechanisms underlying

formaldehyde-induced effects distal to the respiratory tract, including leukemia development, should be tested (Goldstein 2011; NTP 2011; Thompson et al. 2011).

In an effort to understand possible mechanisms for the observed changes in miRNA expression distal to the respiratory tract, we propose a novel mode of action (MOA) that may underlie formaldehyde-induced changes in the hematopoietic system. This MOA not only includes mechanisms by which inhaled formaldehyde may influence signaling within the circulating blood, but it also includes mechanisms related to leukemogenesis. This possible disease endpoint was included in order to comprehensively examine possible links between formaldehyde exposure and changes in sites distal to the respiratory tract at the theoretical level. While we did not identify changes in miRNA expression profiles within the bone marrow, our stringent statistical criteria of fold change $\geq \pm 1.5$, p-value < 0.05, and false discovery rate q-value < 0.10 may have excluded more subtle changes in miRNA expression induced by formaldehyde exposure. It may be the case that modest miRNA changes may influence cell signaling. For this reason, our proposed MOA includes possible formaldehyde-induced disruptions in signaling within both the blood and bone marrow.

The proposed MOA integrates published research findings that link formaldehyde exposure, airway responses, epigenetics, hematopoietic regulation, and the possible transfer of signals between organs. Specifically, we postulate that formaldehyde inhalation exposure initiates changes in the airway epithelium, which may lead to dysregulated miRNA expression profiles in distant hematopoietic stem/progenitor cells through three possible pathways (**Figure 20**, **A-C**). While each pathway is drawn as a separate series of events, there is also possibility of cross-talk between the pathways. As this area of research is in its infancy, we recognize that more evidence is needed for the proposed MOA to be considered plausible. Still, we view it is necessary to propose events that may link formaldehyde inhalation exposure to effects in regions distal to the respiratory tract, highlighting the importance of studying possible epigenetic effects of formaldehyde exposure throughout the body.

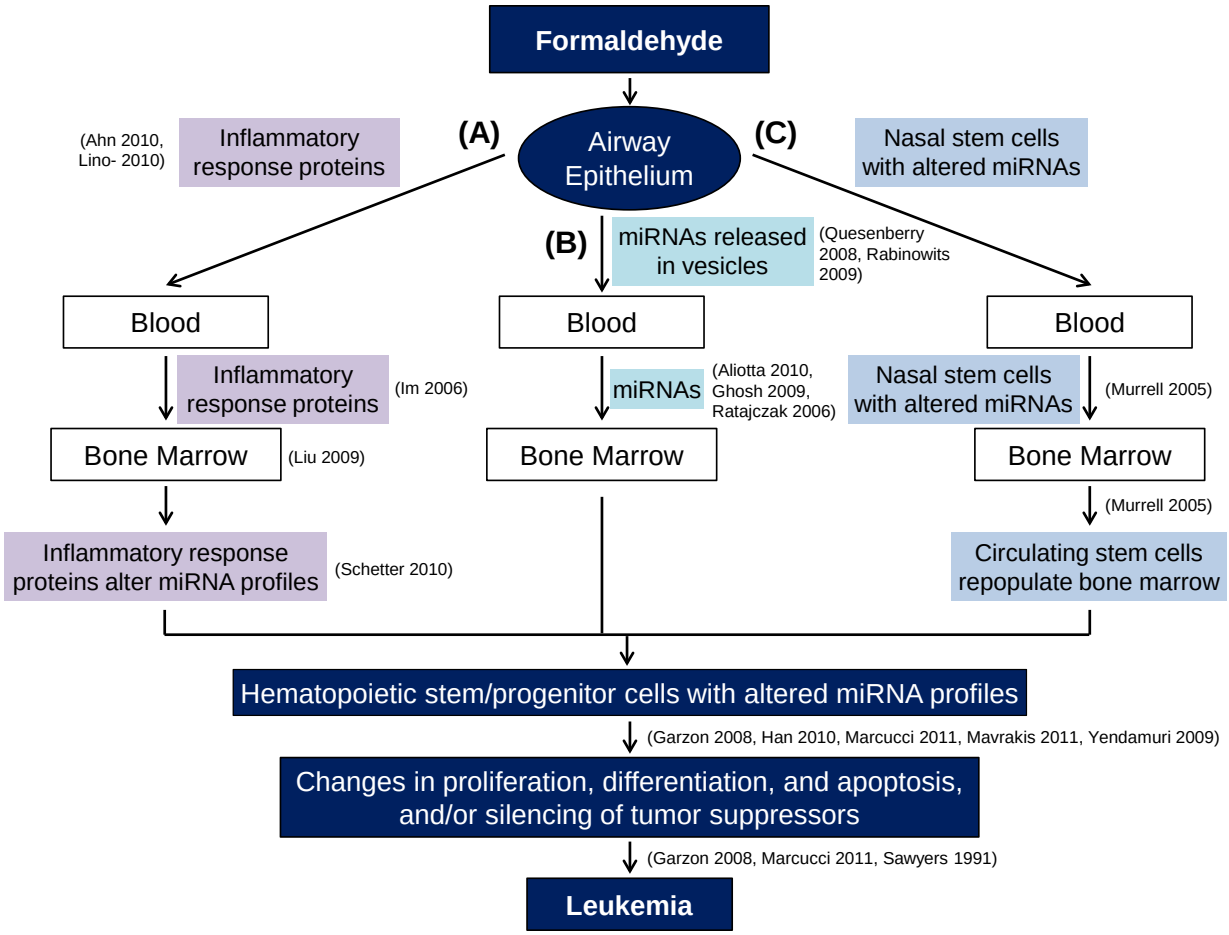

Figure 20: Novel mode of action (MOA) that may link formaldehyde inhalation exposure to hematopoietic changes. We propose three possible pathways (A, B, and C) that join together with a common key event involving the dysregulation of miRNA expression profiles in bone marrow cells. While each pathway is illustrated as a separate series of proposed events, there is possibility of cross-talk between the pathways.

It is important to note that our research did not aim to test this proposed MOA comprehensively. The purpose of outlining a possible MOA is to provide a potential basis for the biological plausibility underlying the observed changes in miRNA expression profiles in tissues distal to the respiratory tract. We did not test each potential interaction mentioned in the MOA. Rather, we focused on the critical steps in the proposed MOA: altered miRNA expression profiles within airway, circulating blood, and bone marrow cells. The pathways involved in the proposed MOA are discussed here.

Pathway A. Inflammatory Response Signaling: One possible pathway linking formaldehyde inhalation exposure to dysregulated miRNA expression profiles in the tissues distal to the respiratory tract may involve downstream inflammation signaling (**Figure 20A**). Formaldehyde exposure may cause local inflammatory responses within airway-related cells, which may progress to more systemic consequences through inflammatory protein release and circulation throughout the body. To detail, gaseous formaldehyde exposure has been shown to induce inflammatory-related protein (e.g. IL-1β, IL-6, VEGF, serpin A3N, serpin B9, and serpin B1A) release in airway cells *in vivo* (Ahn et al. 2010; Lino-dos-Santos-Franco et al. 2010). Inflammatory response proteins can then circulate in the blood, as supported by a previous study showing elevated levels of IL4 and IFNγ formaldehyde-exposed rats (Im et al. 2006). Then, inflammatory-related proteins may enter the bone marrow via the blood, as circulating inflammatory cytokines (e.g. IL-23) have been shown to regulate blood cell production and marrow-derived cell migration (Liu et al. 2009). Once inside bone marrow cells, cytokines and other inflammatory-mediators can heavily impact miRNA expression profiles (Schetter et al. 2010). These proposed events may represent how inhaled formaldehyde possibly induces epigenetic changes within the hematopoietic system through inflammatory response signals.

Pathway B. Circulating miRNAs: The second proposed pathway connecting formaldehyde inhalation exposure to epigenetic changes in tissues distal to the respiratory tract involves circulating miRNAs (**Figure 20B**). Circulating miRNAs that are free of cells, present within secreted membrane vesicles (e.g. exosomes), were recently discovered in 2007 (Valadi et al. 2007). This breakthrough study revealed that mRNA and miRNA molecules are present in cell-derived exosomes present in blood (Valadi et al. 2007). Exosomes are actively secreted membrane vesicles released from many cell types which can travel to distant tissues to influence cell behavior and physiology (Théry 2011). Increasing evidence indicates that exosomes play pivotal roles in cell-to-cell communication via bidirectional exchange of proteins, lipids, and genetic material (i.e. nucleic acids) (Camussi et al. 2010; Théry 2011). This transferrable material has clearly been demonstrated as functional in target cells. For example, transferred mRNAs can be translated into protein in target cells (Ratajczak et al. 2006; Valadi et al. 2007). In addition, transferred miRNAs can cause changes in mRNA levels in target cells (Aliotta et al. 2010; Y Zhang et al. 2010). Because circulating plasma miRNAs have been identified as biomarkers of cancer, including lung cancer (Rabinowits et al. 2009; Silva et al. 2011) and

leukemia (Tanaka et al. 2009), it is possible that vesicle-modulated intercellular communication may occur between the respiratory tract and hematopoietic systems after formaldehyde exposure. This new connection may link formaldehyde inhalation exposure to effects in organs distal to direct target regions.

This proposed pathway first requires that direct target cells exhibit altered miRNA expression profiles upon exposure to formaldehyde. With the research included in this dissertation, we provide strong evidence showing that formaldehyde exposure significantly disrupts miRNA expression profiles within directly exposed airway epithelial cells *in vitro* and in the nasal epithelium of rodents and nonhuman primates. These affected cells may then release miRNAs as shedding vesicles through direct budding from the cell plasma membrane, or miRNAs may be released as exosomes through endosomal membrane compartments fusing with the plasma membrane (Camussi et al. 2010). This phenomenon is found *in vivo*, where, for example, miRNAs derived from lung cancer cells have been identified in circulating plasma samples (Rabinowits et al. 2009). Furthermore, lung cells have been shown to release exosomes containing genetic material capable of modifying nearby marrow stem cell differentiation in an *in vitro* model (Quesenberry et al. 2008).

Once the circulating exosomes travel from the respiratory tract to hematopoietic regions, it may be possible for exosomes to transfer material into target bone marrow cells. Exosomes are thought to recognize target cells through the binding of specific ligands on cell surfaces (Théry 2011). After identifying specific target cells, exosomes can enter cells either through endocytic pathways or they can transfer material by fusing to the cell's membrane and releasing contents into the cytoplasm (Théry 2011). General vesicle-mediated communication is thought to occur in four ways (**Figure 21**) (Camussi et al. 2010). It is possible that one of the communication methods involving epigenetic reprogramming of distant target cells may occur after formaldehyde exposure.

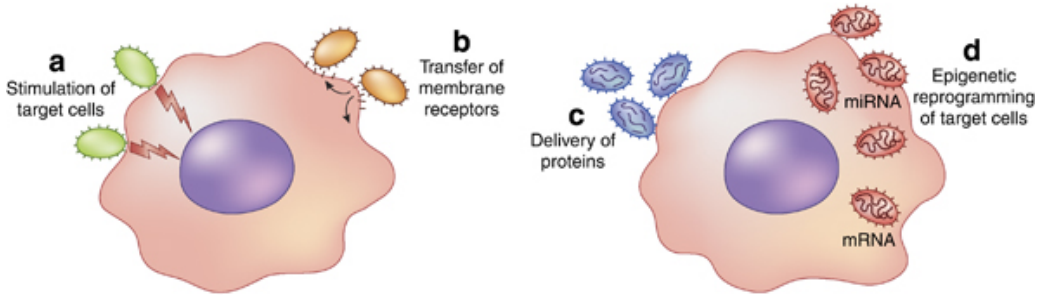

Figure 21: Methods of vesicle-mediated intercellular communication (from Camussi et al. 2010). Vesicles may communicate with target cells through: **(a)** ligands that directly stimulate target cells, **(b)** the transfer of membrane receptors, **(c)** the delivery of functional proteins, or **(d)** the transfer of genetic material, including mRNA and miRNA. This type of communication between tissues may act as a potential link between formaldehyde inhalation exposure and signal disruptions in sites distal to the respiratory epithelium.

Although circulating miRNAs have been implicated as probable messengers during tumor development (Théry 2011), studies have yet to show that miRNAs within circulating vesicles can specifically travel from the respiratory system to the bone marrow and cause leukemia. Still, vesicles have been shown to regulate the expansion of hematopoietic stem cells through the horizontal transfer of protein and mRNA (Ratajczak et al. 2006). In addition, circulating vesicles have been shown to interact and modulate leukemic bone marrow stromal cells (Ghosh et al. 2009). Furthermore, an *in vitro* model showed that mRNAs and miRNAs can transfer via vesicles from lung cells to bone marrow cells separated by a cell-impermeable membrane, causing altered transcriptional induction in marrow cells (Aliotta et al. 2010). It may, therefore, be possible for altered miRNAs originating from the respiratory tract to enter circulating blood and enter hematopoietic stem/progenitor cells, modifying hematopoietic function.

Pathway C. Nasal Stem Cells: The third proposed pathway linking formaldehyde exposure to hematopoietic effects involves the epigenetic alteration of stem cells located within the nose (**Figure 20C**). Pluripotent stem cells are present within the nasal epithelium (Murrell et al. 2005; Roisen et al. 2001), a direct target region for formaldehyde inhalation exposure (Overton et al. 2001). These stem cells may experience altered miRNA expression profiles upon exposure to formaldehyde, as other types of stem cells can show modulated miRNA expression (Chen et al. 2005; Marcucci et al. 2011; Yendamuri et al. 2009). It has been suggested that

formaldehyde-altered nasal stem cells can migrate to the bone marrow during normal cell trafficking or trafficking enhanced by surrounding cytotoxicity (Zhang et al. 2009); although, there is no clear experimental evidence for this step. Next, the circulating stem cells with altered miRNA expression profiles may enter and repopulate the bone marrow. Evidence for this step is provided by a study using enzymatically dissociated olfactory epithelial cells from rats (Murrell et al. 2005). Here, nasal stem cells were injected into the tail vein of irradiated host rats, and the donated stem cells were found to repopulate the bone marrow, generating myeloid and lymphoid cells (Murrell et al. 2005). These nasal stem cells were also transplanted into a chicken embryo model, where numerous cells of hematopoietic lineage were also generated from the donated stem cells (Murrell et al. 2005). Although this proposed pathway requires further experimental evidence, it may represent a set of events linking formaldehyde inhalation exposure to epigenetic alterations within the hematopoietic system.

Dysregulated miRNAs May Cause Hematopoietic Effects: The three proposed pathways, as detailed above, all connect formaldehyde inhalation exposure to altered miRNA expression profiles within the hematopoietic system. As pointed out previously, our research did not detect changes in miRNA expression profiles within rodent bone marrow cells using stringent statistical criteria. Still, more subtle changes in miRNA expression may influence cell signaling, and species variability may exist between rodent and human responses to formaldehyde within the bone marrow. For these reasons, this endpoint is still included in the proposed MOA in order to comprehensively examine all potential effects of altered signaling within tissues distal to the respiratory tract.

Once hematopoietic stem/progenitor cells exhibit altered miRNA expression profiles, a few events may occur to ultimately tie formaldehyde exposure to leukemogenesis. First, dysregulated miRNAs can decrease or increase the transcription and/or translation of important genes (Bartel 2004; Filipowicz et al. 2008; Iorio et al. 2010). Within hematopoietic stem cells, miRNAs are known to regulate genes encoding proteins involved in several steps of hematopoiesis, including differentiation, proliferation, and apoptosis (Garzon et al. 2008; Han et al. 2010; Marcucci et al. 2011; Yendamuri et al. 2009). Because altered hematopoiesis is a key event known to induce leukemia (Sawyers et al. 1991), leukemogenesis may result through this mechanism. In addition, miRNAs within bone marrow cells regulate important leukemia-associated tumor suppressors (Garzon et al. 2008; Mavrakis et al. 2011). When these miRNAs

are dysregulated, they can act as oncogenes and cause leukemia, likely through the inactivation of tumor suppressors (Garzon et al. 2008; Mavrakis et al. 2011). Through this MOA, inhaled formaldehyde may influence the hematopoietic system at the epigenetic level, ultimately contributing to disease.

MOA Summary: Our proposed MOA illustrates three possible pathways linking formaldehyde inhalation exposure to miRNA expression changes in the blood and bone marrow. The relationships within these pathways are substantiated by research findings published in peer-reviewed journals. However, the exact mechanisms linking inhalation exposure to blood and bone marrow responses are unknown, making this proposed MOA of high interest. Our research did not aim to fully test this MOA, as this was not feasible. Instead, we investigated parts of the postulated MOA by establishing whether miRNA expression profiles are altered by formaldehyde inhalation exposure in direct contact and distant sites. These investigations contributed critical knowledge to the MOA linking formaldehyde exposure to cancer.

Public Health Relevance

Our research investigating formaldehyde's influence on miRNA expression profiles reveals novel mechanisms underlying formaldehyde-induced changes in biological signaling, and also lays the foundation for future research on disease prevention. The research has highlighted that formaldehyde is a potent disruptor of miRNAs through the evaluation of three model systems. With this new recognition of miRNAs as key regulators of air toxicant response, there is the opportunity to investigate methods to prevent miRNA expression disruption. The miRNA-mediated biological pathways modified by formaldehyde exposure can also act as potential targets for therapeutic strategies. Our findings also provide sound evidence supporting the alterations of miRNA expression patterns, gene expression patterns, and associated signaling pathways caused by exposure to formaldehyde within circulating blood. The capability of inhaled air pollutants to impact sites distal to the airway epithelium is of high interest, and may link inhaled pollutant exposures to other non-respiratory diseases. Our findings provide the basis for future hypothesis generation for disease prevention and treatment, which can be used to decrease the overall impact of environmental exposure-induced disease.

Relationship of Current Research to Future Hypotheses

Our research revealed novel responses to inhaled formaldehyde which no previous studies have investigated. Because of the originality of the research, some mechanisms underlying the observed responses remain unknown. Our findings provide the basis for future hypothesis generation and research questions, including some questions that are detailed in this section.

1. What triggers altered signaling of formaldehyde-associated miRNAs? Results from our research were the first to indicate that formaldehyde exposure disrupts miRNA expression profiles. With these initial findings, much remains unknown regarding the exact mechanism though which formaldehyde exposure impacts these small RNAs. Other mediators of gene and/or miRNA expression may influence the observed miRNA expression modification, including transcription factors and epigenetic regulators. For instance, exposure to formaldehyde has been shown to cause histone modifications *in vitro* (Lu et al. 2008), which may ultimately affect gene and/or miRNA expression profiles. Another epigenetic mechanism that has yet to be investigated in relation to formaldehyde involves DNA methylation alterations in promoter-associated CpG (cytosine-phosphate-guanine) sites. Increased methylation (i.e. hypermethylation) of promoter CpG islands commonly causes transcriptional silencing (Shames et al. 2007). Conversely, decreased methylation (i.e. hypomethylation) of promoter CpG islands is commonly associated with transcriptional activation (Shames et al. 2007). Because DNA methylation profiles have been identified as heavily influenced by various environmental exposures (Bollati et al. 2010), it is likely that formaldehyde impacts the methylation status of important genes and/or miRNAs. To gain increased knowledge on potential mechanisms linking formaldehyde exposure to changes in miRNA expression levels, the following hypothesis could be tested: Formaldehyde exposure modifies the methylation status of DNA encoding formaldehyde-responsive miRNAs.

This hypothesis could be tested through the use of a methylation inhibitor, such as 5-aza-2′-deoxycytidine. Nasal epithelial cells could be exposed to formaldehyde *in vitro*, and the resulting modifications in miRNA expression profiles could be compared in cells treated with the methylation inhibitor versus cells not treated with the methylation inhibitor. If differences in miRNA responses occur between the two treatments, then the miRNAs that respond differently

are likely to exhibit differential methylation resulting from formaldehyde exposure. Such epigenetic influence has been shown to occur for certain formaldehyde-responsive miRNAs. For example, the expression levels of let-7 family members, many of which were consistently dysregulated by formaldehyde in our studies, are known to be heavily influenced by DNA methylation status (Roush et al. 2008). Results from this future study could therefore provide important information on the mechanisms connecting formaldehyde exposure to miRNA expression modifications.

2. How are miRNA expression patterns altered in the WBC? Our research revealed that inhalation exposure to formaldehyde disrupted the expression levels of miRNAs within circulating WBC in rodents. This finding is of high interest, as the events linking inhaled formaldehyde exposure to altered signaling in tissues distal to the respiratory tract are unknown. A previous toxicological investigation found increased levels of inflammatory mediators IL4 and IFNγ within the circulating blood of rodents exposed to formaldehyde (Im et al. 2006). With this finding, future investigations could increase understanding on how formaldehyde influences tissues that are distant to the respiratory tract by testing the following hypothesis: Formaldehyde exposure changes the levels of inflammatory mediators within the blood which influence signaling within circulating WBC.

In order to support this hypothesis, protein expression analysis could be performed using proteins collected from plasma and WBC samples of formaldehyde-exposed animals. If changes in the levels of inflammatory mediators are identified in formaldehyde-exposed samples, this would represent the first step in supporting this hypothesis. A next step would be to demonstrate that these proteins may link formaldehyde inhalation exposure to changes in miRNA expression and cell signaling in circulating WBC.

To test whether formaldehyde-altered miRNAs are directly modulated by circulating inflammatory mediators, this hypothesis could be further tested using an *in vitro* model. For example, cultured WBC could be treated with the formaldehyde-modulated inflammatory mediators, and changes in miRNA expression could be assessed. If these changes overlap with those observed in response to formaldehyde, then it could support that inflammatory proteins indeed influence miRNA differential expression, acting as a link between formaldehyde exposure and miRNA expression alterations.

3. What is the relationship between formaldehyde-induced DNA damage and miRNA expression modification? Our research established that formaldehyde disrupts miRNA expression profiles *in vitro* and *in vivo*. Research performed by our collaborators also clearly showed that formaldehyde exposure causes DNA damage (Lu et al. 2009; Lu et al. 2011; Moeller et al. 2011). An understudied event underlying exposure-induced toxicity is the influence of DNA damage on miRNA signaling. Because miRNAs have been shown to play an important role in the response to UV-induced DNA damage (Pothof et al. 2009), it is possible that miRNAs are involved in responses to formaldehyde-induced DNA damage. In order to investigate this potential relationship, the following hypothesis could be tested: The expression levels of an important subset of formaldehyde-responsive miRNAs are associated with formaldehyde-induced DNA damage.

This hypothesis could be tested by exposing nasal epithelial cells to various formaldehyde doses *in vitro*. Increases in DNA damage as well as changes in miRNA expression profiles would be observed, where some of the formaldehyde-altered miRNAs' expression levels may correlate with the amount of formaldehyde-induced DNA damage. This subset of miRNAs would likely play important roles in DNA damage response associated with formaldehyde exposure. Findings may be similar to a previous study, where UV-inducible miR-16 was found to regulate cell cycle checkpoint genes and cell proliferation, which are heavily involved in responses to DNA damage (Pothof et al. 2009). Results from this proposed experiment would increase the understanding of relationships between DNA damage and novel mediators of damage response in order to accurately interpret biological events linking environmental exposure responses to disease.

Dissertation Conclusion

To summarize our investigations, we have highlighted formaldehyde as an air toxicant that dramatically alters miRNA expression patterns. These epigenomic responses were established *in vitro*, in cultured human lung epithelial cells, and *in vivo*, in nonhuman primates and rodents. The formaldehyde-associated miRNA expression alterations were evaluated at the systems biology level, where critical pathways likely altered via formaldehyde's influence on miRNA expression were revealed. Taken together, our research increases the knowledge of under-studied

mechanisms linking formaldehyde exposure to disease, acting as an important foundation for future research in public health and toxicology.

SUPPLEMENTARY TABLES

Supplementary Table 1: miRNAs significantly (p-value < 0.05, FDR < 0.05, fold change ≥ +/- 1.5) differentially expressed upon exposure to 1 ppm formaldehyde in human lung epithelial cells.

miRNA	Formaldehyde/Control Ratio
miR-33	-5.48
miR-450	-3.57
miR-330	-2.43
miR-181a	-2.11
miR-10b	-2.11
miR-422b	-2.02
miR-532	-1.84
miR-501	-1.82
miR-487b	-1.80
miR-20a	-1.80
miR-34a	-1.73
miR-93	-1.72
miR-106b	-1.71
miR-137	-1.71
miR-103	-1.70
miR-301	-1.70
miR-10a	-1.70
miR-126	-1.70
miR-17-5p	-1.69
miR-107	-1.69
miR-454-3p	-1.69
miR-140	-1.68
miR-101	-1.68
miR-130a	-1.68
miR-19a	-1.67
miR-26a	-1.67
miR-19b	-1.67
miR-106a	-1.66
miR-99a	-1.66
miR-18a	-1.66
miR-424	-1.65
let-7a	-1.65
miR-20b	-1.65
miR-25	-1.64
miR-590	-1.64
miR-15b	-1.64
let-7b	-1.63
miR-660	-1.63
miR-27b	-1.63
miR-194	-1.62
miR-361	-1.62
miR-192	-1.62
miR-215	-1.62

miRNA	Formaldehyde/Control Ratio
miR-374	-1.62
miR-15a	-1.62
let-7c	-1.61
miR-148b	-1.60
miR-181b	-1.60
miR-425-5p	-1.60
miR-23b	-1.60
let-7d	-1.59
miR-28	-1.58
miR-125a	-1.58
miR-181d	-1.58
miR-130b	-1.58
miR-185	-1.58
miR-324-5p	-1.58
miR-9*	-1.57
miR-452	-1.57
miR-565	-1.57
miR-26b	-1.57
miR-152	-1.57
miR-16	-1.57
miR-650	-1.56
miR-21	-1.56
miR-9	-1.56
miR-186	-1.56
miR-151	-1.56
miR-582	-1.55
let-7e	-1.55
let-7g	-1.55
miR-98	-1.55
miR-224	-1.55
miR-23a	-1.54
miR-27a	-1.54
miR-362	-1.54
let-7f	-1.54
miR-17-3p	-1.53
miR-550	-1.53
miR-29b	-1.53
miR-182	-1.53
miR-100	-1.51
miR-509	-1.51
miR-652	-1.51
miR-331	-1.51
miR-34b	-1.51
miR-189	-1.51
let-7i	-1.51
miR-24	-1.50

Supplementary Table 2: Predicted transcriptional targets for miR-33, miR-330, miR-181a, and miR-10b.

Symbol	GenBank	Entrez Gene ID	miRNAs with Overlapping Targets
1. miR-33			
ABCA1	NM_005502	19	miR-330
ABCE1	NM_001040876	6059	
APPBP2	NM_006380	10513	
ARID5B	NM_032199	84159	
ASAP1	NM_018482	50807	
B3GALT2	NM_003783	8707	
C11ORF41	NM_012194	25758	
CACNA1C	NM_001129834	775	
CDC42BPA	NM_003607	8476	
CDK6	NM_001259	1021	
CLSPN	NM_022111	63967	
CROT	NM_021151	54677	
CSNK1D	NM_001893	1453	
DPY19L1	NM_015283	23333	
DSC3	NM_001941	1825	miR-181a
DYRK3	NM_001004023	8444	
EBF1	NM_024007	1879	
EEA1	NM_003566	8411	
EN2	NM_001427	2020	
ESCO1	NM_052911	114799	
FAM46C	NM_017709	54855	
FGA	NM_000508	2243	
FUT9	NM_006581	10690	miR-181a
GLCCI1	NM_138426	113263	
GOPC	NM_020399	57120	
GRIA3	NM_007325	2892	miR-330
HADHB	NM_000183	3032	
HBS1L	NM_006620	10767	
HIPK2	NM_001113239	28996	
HMGA2	NM_003483	8091	
ING3	NM_019071	54556	
KCNMA1	NM_001014797	3778	
KIAA2018	NM_001009899	205717	
KIF3C	NM_002254	3797	
LCA5	NM_181714	167691	miR-10b
LIPI	NM_198996	149998	
LOC152742	XM_001128848	152742	
LPP	NM_005578	4026	
MMP16	NM_005941	4325	
MSR1	NM_138715	4481	
NAP1L2	NM_021963	4674	
NARG1	NM_057175	80155	
NAT8	NM_003960	9027	
NAT12	NM_001011713	122830	miR-330
PIM1	NM_002648	5292	
PNMA1	NM_006029	9240	
PRDM2	NM_001007257	7799	

Symbol	GenBank	Entrez Gene ID	miRNAs with Overlapping Targets
PTPN13	NM_080685	5783	
RCAN1	NM_203417	1827	miR-330
RIMBP2	NM_015347	23504	
SATB2	NM_015265	23314	
SCN8A	NM_014191	6334	
SEC24C	NM_004922	9632	miR-181a
SIX4	NM_017420	51804	
SLC25A25	NM_001006642	114789	
SLC26A7	NM_052832	115111	
SLC39A14	NM_015359	23516	
SLITRK3	NM_014926	22865	
SLU7	NM_006425	10569	
SPAST	NM_199436	6683	
ST18	NM_014682	9705	
TNFRSF9	NM_001561	3604	
UBE2V2	NM_003350	7336	
ZNF140	NM_003440	7699	
ZNF148	NM_021964	7707	miR-330
ZNF281	NM_012482	23528	
ZNF300	NM_052860	91975	
2. miR-330			
ABCA1	NM_005502	19	miR-33
ACVR1	NM_001105	90	
ADAMTS5	NM_007038	11096	
AFF2	NM_002025	2334	
AFF4	NM_014423	27125	
AGTR2	NM_000686	186	
AK7	NM_152327	122481	
ANGEL2	NM_144567	90806	
ANKH	NM_054027	56172	
AP2M1	NM_004068	1173	
API5	NM_006595	8539	
APPL1	NM_012096	26060	
ARFGEF2	NM_006420	10564	
ARHGAP12	NM_018287	94134	
ARHGAP20	NM_020809	57569	
ARL17P1	NM_001113738	51326	
ATL2	NM_022374	64225	
ATP2B1	NM_001001323	490	miR-181a
ATP2C1	NM_014382	27032	
AZIN1	NM_148174	51582	
BCL9	NM_004326	607	
BCL11B	NM_022898	64919	
BFSP1	NM_001195	631	
BMPR2	NM_001204	659	
BTRC	NM_033637	8945	
C10ORF10	NM_007021	11067	
C14ORF129	NM_016472	51527	miR-181a
C2CD2	NM_015500	25966	
C5ORF15	NM_020199	56951	

Symbol	GenBank	Entrez Gene ID	miRNAs with Overlapping Targets
C5ORF23	NM_024563	79614	
C9ORF5	NM_032012	23731	
C9ORF64	NM_032307	84267	
CALCR	NM_001742	799	
CAPZA1	NM_006135	829	
CBX5	NM_012117	23468	
CCND3	NM_001760	896	
CD247	NM_000734	919	
CDR2	NM_001802	1039	
CHP	NM_007236	11261	
CLCN5	NM_001127899	1184	
CLDN8	NM_199328	9073	miR-181a
CLDN18	NM_016369	51208	
CMPK1	NM_016308	51727	
CNBP	NM_001127195	7555	
CRLS1	NM_019095	54675	
CSNK1G3	NM_001044722	1456	
CYP7A1	NM_000780	1581	
D4S234E	NM_014392	27065	
DAG1	NM_004393	1605	
DCAF7	NM_005828	10238	
DICER1	NM_030621	23405	
DLX1	NM_001038493	1745	
DNM3	NM_015569	26052	
DNM1L	NM_005690	10059	
DOCK5	NM_024940	80005	
DPP10	NM_001004360	57628	
EDEM1	NM_014674	9695	
EEF1A1	NM_001402	1915	
EFHC1	NM_018100	114327	
EIF5	NM_183004	1983	
EIF4E	NM_001968	1977	
EPM2A	NM_005670	7957	
ERAP1	NM_016442	51752	
ERBB4	NM_005235	2066	
ERC1	NM_178039	23085	
ERLIN2	NM_001003791	11160	
EXOC8	NM_175876	149371	
FAM107B	NM_031453	83641	
FAM72D	NM_207418	728833	
FGFR1	NM_023107	2260	
FMO2	NM_001460	2327	
FOXK1	NM_001037165	221937	miR-181a
FRK	NM_002031	2444	
GJC1	NM_005497	10052	
GNRHR	NM_001012763	2798	
GPRASP1	NM_014710	9737	
GRB10	NM_001001550	2887	miR-181a
GRIA3	NM_007325	2892	miR-33
HELZ	NM_014877	9931	

Symbol	GenBank	Entrez Gene ID	miRNAs with Overlapping Targets
HIVEP2	NM_006734	3097	
HNRNPU	NM_031844	3192	
HSF2	NM_004506	3298	
HSPH1	NM_006644	10808	
ID2	NM_002166	3398	
IMPACT	NM_018439	55364	
INO80D	NM_017759	54891	miR-181a
INSL5	NM_005478	10022	
ITM2C	NM_030926	81618	
JPH1	NM_020647	56704	
KANK2	NM_015493	25959	
KAT2B	NM_003884	8850	
KDM4C	NM_015061	23081	
KDSR	NM_002035	2531	
KIAA1012	NM_014939	22878	
KLF10	NM_005655	7071	
KLHL24	NM_017644	54800	
LAPTM4B	NM_018407	55353	
LNX2	NM_153371	222484	
LRPPRC	NM_133259	10128	
MARK1	NM_018650	4139	miR-181a
MAT2A	NM_005911	4144	
MBNL2	NM_144778	10150	
MECOM	NM_001105078	2122	
METAP2	NM_006838	10988	miR-181a
MMD	NM_012329	23531	
MOBKL1A	NM_173468	92597	
MRPS6	NM_032476	64968	
MYEF2	NM_016132	50804	
MYPN	NM_032578	84665	
NAT12	NM_001011713	122830	miR-33
NEFL	NM_006158	4747	
ONECUT2	NM_004852	9480	miR-181a
OTUD3	NM_015207	23252	
PAFAH1B1	NM_000430	5048	
PCDHA1	NM_018900	56147	miR-181a
PCDHA2	NM_018905	56146	miR-181a
PCDHA3	NM_018906	56145	miR-181a
PCDHA4	NM_018907	56144	miR-181a
PCDHA5	NM_018908	56143	miR-181a
PCDHA6	NM_031849	56142	miR-181a
PCDHA7	NM_018910	56141	miR-181a
PCDHA8	NM_018911	56140	miR-181a
PCDHA9	NM_031857	9752	miR-181a
PCDHA10	NM_018901	56139	miR-181a
PCDHA11	NM_018902	56138	miR-181a
PCDHA12	NM_018903	56137	miR-181a
PCDHAC1	NM_018898	56135	miR-181a
PCDHAC2	NM_018899	56134	miR-181a
PCK1	NM_002591	5105	

Symbol	GenBank	Entrez Gene ID	miRNAs with Overlapping Targets
PCMT1	NM_005389	5110	
PCTP	NM_001102402	58488	
PHAX	NM_032177	51808	
PLSCR1	NM_021105	5359	
PLSCR4	NM_001128304	57088	
PLXNA2	NM_025179	5362	
PQLC1	NM_025078	80148	
PRKAB2	NM_005399	5565	
PRKCB	NM_002738	5579	
PSD3	NM_015310	23362	
PTBP2	NM_021190	58155	
RAI2	NM_021785	10742	
RAP2A	NM_021033	5911	
RAVER2	NM_018211	55225	
RBM12	NM_152838	10137	
RCAN1	NM_203417	1827	miR-33
RGS10	NM_002925	6001	
RND3	NM_005168	390	
RNF212	NM_194439	285498	
RNF144B	NM_182757	255488	
RUFY2	NM_001042417	55680	
SAMD12	NM_001101676	401474	miR-181a
SCG3	NM_013243	29106	
SCP2	NM_001007100	6342	
SELI	NM_033505	85465	
SEMA4D	NM_006378	10507	
SEP15	NM_203341	9403	
SERINC3	NM_006811	10955	
SFRS1	NM_006924	6426	
SH3TC2	NM_024577	79628	miR-181a
SIN3A	NM_015477	25942	
SLAIN1	NM_001040153	122060	
SLC2A2	NM_000340	6514	
SLC5A3	NM_006933	6526	
SMG7	NM_173156	9887	
SMNDC1	NM_005871	10285	
SNAP23	NM_130798	8773	
SNX2	NM_003100	6643	
SORL1	NM_003105	6653	
SOSTDC1	NM_015464	25928	
STAU1	NM_001037328	6780	
STEAP4	NM_024636	79689	
STK3	NM_006281	6788	
SUDS3	NM_022491	64426	
SUMF1	NM_182760	285362	
TAPT1	NM_153365	202018	
TBL1XR1	NM_024665	79718	miR-181a
TBX5	NM_080717	6910	miR-10b
TEAD1	NM_021961	7003	
TGFBR3	NM_003243	7049	

Symbol	GenBank	Entrez Gene ID	miRNAs with Overlapping Targets
THBS1	NM_003246	7057	
TJP1	NM_003257	7082	
TMEM59	NM_004872	9528	
TNKS	NM_003747	8658	
TNRC6B	NM_001024843	23112	
TNS1	NM_022648	7145	
TOX	NM_014729	9760	miR-181a
TRA2A	NM_013293	29896	
TRIM2	NM_015271	23321	miR-181a
TRIM37	NM_015294	4591	
TRIP12	NM_004238	9320	
TROVE2	NM_004600	6738	
TSFM	NM_005726	10102	
TSHR	NM_000369	7253	
UBE2Q1	NM_017582	55585	
UBTD2	NM_152277	92181	
UBXN4	NM_014607	23190	
USP15	NM_006313	9958	
USP37	NM_020935	57695	
VAPA	NM_003574	9218	
VASH2	NM_024749	79805	
VGLL3	NM_016206	389136	
VPS54	NM_016516	51542	
WDR37	NM_014023	22884	
XK	NM_021083	7504	
YIPF5	NM_001024947	81555	
YTHDC1	NM_133370	91746	
ZBTB34	NM_001099270	403341	miR-181a
ZC3HAV1	NM_020119	56829	
ZCCHC24	NM_153367	219654	
ZFC3H1	NM_144982	196441	
ZFR	NM_016107	51663	
ZNF148	NM_021964	7707	miR-33
ZNF410	NM_021188	57862	
ZNF423	NM_015069	23090	
ZNF490	NM_020714	57474	
ZNF706	NM_016096	51123	
ZNF280D	NM_001002843	54816	
3. miR-181a			
ACSL1	NM_001995	2180	
ACVR2A	NM_001616	92	
ACVR2B	NM_001106	93	
ACYP1	NM_001107	97	
ADAM28	NM_014265	10863	
ADAMTSL1	NM_001040272	92949	
ADRBK1	NM_001619	156	
AFTPH	NM_203437	54812	
AHCTF1	NM_015446	25909	
AKAP5	NM_004857	9495	
AKAP6	NM_004274	9472	

Symbol	GenBank	Entrez Gene ID	miRNAs with Overlapping Targets
AKAP7	NM_004842	9465	
ARHGAP26	NM_015071	23092	
ARHGEF3	NM_001128615	50650	
ARSJ	NM_024590	79642	miR-10b
ATG5	NM_004849	9474	
ATM	NM_138292	472	
ATP11C	NM_173694	286410	
ATP2A2	NM_170665	488	
ATP2B1	NM_001001323	490	miR-330
ATXN3	NM_001127696	4287	
BAG2	NM_004282	9532	
BAG4	NM_004874	9530	
BAI3	NM_001704	577	
BAZ2B	NM_013450	29994	
BBS7	NM_018190	55212	
BEND3	NM_001080450	57673	
BHLHE40	NM_003670	8553	
BIRC6	NM_016252	57448	
BOLL	NM_197970	66037	
BRAP	NM_006768	8315	
BRD1	NM_014577	23774	
BRWD1	NM_033656	54014	
BTBD3	NM_014962	22903	
C10ORF104	NM_173473	119504	
C14ORF129	NM_016472	51527	miR-330
C15ORF29	NM_024713	79768	
C16ORF87	NM_001001436	388272	
C19ORF12	NM_031448	83636	
C20ORF12	NM_001099407	55184	
C21ORF66	NM_016631	94104	
C2ORF69	NM_153689	205327	
C5ORF41	NM_153607	153222	
C5ORF47	XM_376444	133491	
C6ORF89	NM_152734	221477	
CABC1	NM_020247	56997	
CALB1	NM_004929	793	
CALM1	NM_006888	801	
CAMK2D	NM_172128	817	
CAPRIN1	NM_005898	4076	miR-10b
CARD8	NM_014959	22900	
CBX7	NM_175709	23492	
CCAR1	NM_018237	55749	
CCDC14	NM_022757	64770	
CCDC117	NM_173510	150275	
CCNB1	NM_031966	891	
CCNJ	NM_019084	54619	
CCNL2	NM_001039577	81669	
CD302	NM_014880	9936	
CDON	NM_016952	50937	
CHMP2B	NM_014043	25978	

Symbol	GenBank	Entrez Gene ID	miRNAs with Overlapping Targets
CLASP1	NM_015282	23332	
CLDN8	NM_199328	9073	miR-330
CLIP1	NM_002956	6249	
CLVS1	NM_173519	157807	
CNTN4	NM_175612	152330	
CPNE2	NM_152727	221184	
CPOX	NM_000097	1371	
CREB5	NM_004904	9586	
CSF2RB	NM_000395	1439	
CTDSPL	NM_005808	10217	
CTTNBP2NL	NM_018704	55917	
CUL3	NM_003590	8452	
DARS	NM_001349	1615	
DCN	NM_133504	1634	
DDX52	NM_007010	11056	
DDX3X	NM_001356	1654	
DDX3	NM_004660	8653	
DEPDC6	NM_022783	64798	
DIRAS3	NM_004675	9077	
DLGAP2	NM_004745	9228	
DNAJC13	NM_015268	23317	
DNAL1	NM_031427	83544	
DOCK10	NM_014689	55619	
DSC3	NM_001941	1825	miR-33
DYNC1LI2	NM_006141	1783	
E2F5	NM_001951	1875	
E2F7	NM_203394	144455	
EIF4A2	NM_001967	1974	
ENPP1	NM_006208	5167	
EPC2	NM_015630	26122	
ETV6	NM_001987	2120	
EXD1	NM_152596	161829	
FAM13B	NM_001101801	51306	
FAM160A2	NM_032127	84067	
FBXO33	NM_203301	254170	
FBXO34	NM_017943	55030	
FIGN	NM_018086	55137	miR-10b
FKBP1A	NM_000801	2280	
FMNL2	NM_052905	114793	
FNDC3B	NM_022763	64778	
FOXK1	NM_001037165	221937	miR-330
FOXP1	NM_032682	27086	
FUCA1	NM_000147	2517	
FUT9	NM_006581	10690	miR-33
G3BP2	NM_012297	9908	
GABRA1	NM_001127648	2554	
GAPVD1	NM_015635	26130	
GATA6	NM_005257	2627	
GATM	NM_001482	2628	
GCC2	NM_181453	9648	

Symbol	GenBank	Entrez Gene ID	miRNAs with Overlapping Targets
GHITM	NM_014394	27069	
GPD2	NM_000408	2820	
GPD1L	NM_015141	23171	
GPRIN3	NM_198281	285513	
GPX8	NM_001008397	493869	
GRB10	NM_001001555	2887	miR-330
HEY2	NM_012259	23493	
HIC2	NM_015094	23119	
HOOK1	NM_015888	51361	
HOXB4	NM_024015	3214	
HOXC8	NM_022658	3224	
HOXD1	NM_024501	3231	
HRH1	NM_001098212	3269	
HSPC159	NM_014181	29094	
IL2	NM_000586	3558	
INO80D	NM_017759	54891	miR-330
IPO8	NM_006390	10526	
ITGA2	NM_002203	3673	
ITSN1	NM_001001132	6453	
KANK1	NM_015158	23189	
KCNH8	NM_144633	131096	
KCTD3	NM_016121	51133	
KDM5A	NM_005056	5927	
KIAA0528	NM_014802	9847	
KIAA1219	NM_020336	57148	
KIAA1239	XM_940885	57495	
KIAA2022	NM_001008537	340533	
KIF3A	NM_007054	11127	
KLHL2	NM_007246	11275	
KLHL5	NM_199039	51088	
KRAS	NM_033360	3845	
LAMP2	NM_001122606	3920	
LARP4	NM_199190	113251	
LCLAT1	NM_182551	253558	
LHFPL3	NM_199000	375612	
LIFR	NM_002310	3977	
LMO3	NM_001001395	55885	
LOC161527	XM_929030	161527	
LONRF2	NM_198461	164832	
LRRC8D	NM_018103	55144	
MAP1B	NM_005909	4131	
MAPK1	NM_002745	5594	
MARK1	NM_018650	4139	miR-330
MATN3	NM_002381	4148	
MBOAT2	NM_138799	129642	
MED8	NM_201542	112950	
MEGF9	NM_001080497	1955	
METAP1	NM_015143	23173	
METAP2	NM_006838	10988	miR-330
MINA	NM_032778	84864	

Symbol	GenBank	Entrez Gene ID	miRNAs with Overlapping Targets
MKLN1	NM_013255	4289	
MORC3	NM_015358	23515	
MPP5	NM_022474	64398	
MTF2	NM_007358	22823	
MTMR12	NM_001040446	54545	
MTMR15	NM_014967	22909	
MTX3	NM_001010891	345778	
MUC7	NM_152291	4589	
NCOA2	NM_006540	10499	
NFAT5	NM_001113178	10725	
NLN	NM_020726	57486	
NOTCH4	NM_004557	4855	
NOVA1	NM_002515	4857	
NR3C1	NM_001018076	2908	
NR6A1	NM_033334	2649	
NRAS	NM_002524	4893	
NTS	NM_006183	4922	
NUDT12	NM_031438	83594	
ONECUT2	NM_004852	9480	miR-330
OSBPL3	NM_145320	26031	
OSBPL8	NM_001003712	114882	
OTUD4	NM_001102653	54726	
PAM	NM_138822	5066	
PAPD5	NM_001040285	64282	
PAPOLG	NM_022894	64895	
PARK2	NM_013987	5071	
PARP11	NM_020367	57097	
PAWR	NM_002583	5074	
PAX9	NM_006194	5083	
PCDHA1	NM_031411	56147	miR-330
PCDHA2	NM_018905	56146	miR-330
PCDHA3	NM_018906	56145	miR-330
PCDHA4	NM_018907	56144	miR-330
PCDHA5	NM_018908	56143	miR-330
PCDHA6	NM_031849	56142	miR-330
PCDHA7	NM_018910	56141	miR-330
PCDHA8	NM_018911	56140	miR-330
PCDHA9	NM_031857	9752	miR-330
PCDHA10	NM_018901	56139	miR-330
PCDHA11	NM_018902	56138	miR-330
PCDHA12	NM_018903	56137	miR-330
PCDHAC1	NM_018898	56135	miR-330
PCDHAC2	NM_018899	56134	miR-330
PCNP	NM_020357	57092	
PDE5A	NM_001083	8654	
PER3	NM_016831	8863	
PGAP1	NM_024989	80055	
PHC3	NM_024947	80012	
PHLPP2	NM_015020	23035	
PHTF2	NM_001127358	57157	

Symbol	GenBank	Entrez Gene ID	miRNAs with Overlapping Targets
PI4K2B	NM_018323	55300	
PITPNB	NM_012399	23760	
PKNOX2	NM_022062	63876	
PLAC1L	NM_173801	219990	
PLEKHA3	NM_019091	65977	
PNRC2	NM_017761	55629	
POLQ	NM_199420	10721	
POLR3G	NM_006467	10622	
POM121	NM_172020	9883	
POM121C	NM_001099415	100101267	
PPP1R12B	NM_002481	4660	
PPP1R9A	NM_017650	55607	
PPP2R5E	NM_006246	5529	
PRDM4	NM_012406	11108	
PRDX3	NM_006793	10935	
PRH2	NM_001110213	5555	
PRKCD	NM_006254	5580	
PRTG	NM_173814	283659	
PSG5	NM_002781	5673	
PSRC1	NM_001032291	84722	
PTGER3	NM_198715	5733	
RAB3IP	NM_022456	117177	
RAD21	NM_006265	5885	
RAN	NM_006325	5901	
RAP1B	NM_001010942	5908	
RASSF2	NM_014737	9770	
RBM26	NM_022118	64062	
RBM25	NM_021239	58517	
REPS2	NM_001080975	9185	
RFC1	NM_002913	5981	
RIN2	NM_018993	54453	
RLF	NM_012421	6018	
RNF8	NM_183078	9025	
RNF34	NM_025126	80196	
ROD1	NM_005156	9991	
RP5-1022P6.2	NM_019593	56261	
RPAP2	NM_024813	79871	
RPE65	NM_000329	6121	
RPS6KB1	NM_003161	6198	
RRP15	NM_016052	51018	
S1PR1	NM_001400	1901	
SAMD12	NM_207506	401474	miR-330
SCD	NM_005063	6319	
SCOC	NM_032547	60592	
SEC24C	NM_004922	9632	miR-33
SEMA3C	NM_006379	10512	
SEMA4G	NM_017893	57715	
SFRS7	NM_001031684	6432	
SH3TC2	NM_024577	79628	miR-330
SIPA1L2	NM_020808	57568	

Symbol	GenBank	Entrez Gene ID	miRNAs with Overlapping Targets
SIRT1	NM_012238	23411	
SLC19A2	NM_006996	10560	
SLC24A1	NM_004727	9187	
SLC25A24	NM_213651	29957	
SLC7A11	NM_014331	23657	
SLITRK1	NM_052910	114798	
SPIN1	NM_006717	10927	
SPOCK1	NM_004598	6695	
SPP1	NM_001040058	6696	
SPRY4	NM_030964	81848	
SRPK2	NM_182692	6733	
ST8SIA4	NM_005668	7903	
STX7	NM_003569	8417	
SUCLG2	NM_003848	8801	
SYNE1	NM_133650	23345	
TADA2B	NM_152293	93624	
TBC1D1	NM_015173	23216	
TBC1D4	NM_014832	9882	
TBL1X	NM_005647	6907	
TBL1XR1	NM_024665	79718	miR-330
TBPL1	NM_004865	9519	
TCERG1	NM_006706	10915	
TET2	NM_017628	54790	
TFEC	NM_012252	22797	
TFRC	NM_003234	7037	
TGFBR1	NM_004612	7046	
TGFBRAP1	NM_004257	9392	
TIFA	NM_052864	92610	
TIMP3	NM_000362	7078	
TLL1	NM_012464	7092	
TMEM26	NM_178505	219623	
TMEM27	NM_020665	57393	
TMEM131	NM_015348	23505	
TMEM165	NM_018475	55858	
TMF1	NM_007114	7110	
TNF	NM_000594	7124	
TNFRSF11B	NM_002546	4982	
TNFSF4	NM_003326	7292	
TNPO1	NM_153188	3842	
TOM1L1	NM_005486	10040	
TOR1AIP2	NM_145034	163590	
TOX	NM_014729	9760	miR-330
TRDMT1	NM_176081	1787	
TRIM2	NM_015271	23321	miR-330
TSPAN8	NM_004616	7103	
TSPYL4	NM_021648	23270	
UBE2A	NM_003336	7319	
UBE2D1	NM_003338	7321	
UBP1	NM_001128160	7342	
USP42	NM_032172	84132	

Symbol	GenBank	Entrez Gene ID	miRNAs with Overlapping Targets
VBP1	NM_003372	7411	
VCAN	NM_001126336	1462	
WHSC2	NM_005663	7469	
WNK1	NM_018979	65125	
XRN1	NM_019001	54464	
ZBTB34	NM_001099270	403341	miR-330
ZBTB44	NM_014155	29068	
ZFAND6	NM_019006	54469	
ZFP36L1	NM_004926	677	
ZFP36L2	NM_006887	678	
ZIC3	NM_003413	7547	
ZNF83	NM_001105549	55769	
ZNF124	NM_003431	7678	
ZNF439	NM_152262	90594	
ZNF440	NM_152357	126070	
ZNF441	NM_152355	126068	
ZNF454	NM_182594	285676	
ZNF468	NM_001008801	90333	
ZNF559	NM_032497	84527	
ZNF594	NM_032530	84622	
ZNF655	NM_138494	79027	
ZNF673	NM_001129898	55634	
ZSCAN23	NM_001012455	222696	
4. miR-10b			
ARG2	NM_001172	384	
ARSJ	NM_024590	79642	miR-181a
BDNF	NM_170731	627	
C1ORF71	XM_001717264	163882	
CAPRIN1	NM_005898	4076	miR-181a
CLCC1	NM_001048210	23155	
EBF2	NM_022659	64641	
FIGN	NM_018086	55137	miR-181a
FNBP1L	NM_001024948	54874	
GALNT1	NM_020474	2589	
HOXA3	NM_153631	3200	
HOXB3	NM_002146	3213	
KLF11	NM_003597	8462	
LCA5	NM_001122769	167691	miR-33
MAP3K7	NM_003188	6885	
NONO	NM_007363	4841	
RB1CC1	NM_014781	9821	
RBM27	NM_018989	54439	
RNF7	NM_183237	9616	
SNX18	NM_001102575	112574	
TBX5	NM_181486	6910	miR-330
TFAP2C	NM_003222	7022	
TRNT1	NM_182916	51095	
USP46	NM_022832	64854	
WDR26	NM_001115113	80232	

Supplementary Table 3: 40 networks associated with the predicted targets of miR-33, miR-330, miR-181a, and miR-10b.

Network No.	p-value	No. of Predicted Transcripts	Molecules in Network
1. miR-33			
1	1E-34	16	ABCA1, ACAA2, Alcohol group acceptor phosphotransferase, CACNA1C, CACNA1D, CACNA1H, CDK6, CROT, CSNK1D, GLO1, HADHB, HMGA2, KCNMA1, LDL, LIMK2, MIR124, MMP16, MSR1, NARG1, NEK2, NFkB (complex), PDGF BB, PGRMC2, PIM1, Pka, PLEKHA1, PPP1R13L, PTPN13, RCAN1, RYR1 (includes EG:114207), SERPINB6, ST18, SUCLG2, TNFRSF9, Trypsin
2	1E-27	14	AHSG, ARD1A, BAT1, CHMP1B, CLSPN, CP110, CPB2, CTNNB1, DSC3, FBXW11, FGB, FH, GLCCI1, GOT1, Groucho, HBS1L, HNF1A, HNF4A, HPX, ING3, LIN7C, MIR9-1 (includes EG:407046), NAT8, NFYB, PCBD1, PLK1, PNMA1, PTPN13, PZP, RIMBP2, SLC26A7, SLITRK3, SPAST, ZNF281, ZNF300
3	1E-26	13	ABCE1, ANP32A, ARID5B, ATXN1, B3GALT2, BACE1, BAT2, BECN1, C11ORF41, CCDC85B, CFL1, CHI3L1, CSF1R, DAZAP2 (includes EG:9802), DZIP3, EEA1, FAM46C(includes EG:54855), FGA, KIAA2018, KIF3C, KRT18, LPA, MIR17 (includes EG:406952), MIR212 (includes EG:406994), MIR29A (includes EG:407021), MIR29B1, MIR29B2, NAT12, NEDD4L, SATB2, SLC39A14, SLU7, TARDBP, TGFB1, TRIP6
4	1E-21	11	ASAP1, CADM1, CCL9, CDC42, CDC42BPA, CDC42BPB, CXCL12, EBF1, EFS, ERK, FCGR2B, FUT9, FZD5, GOPC, IL1B, ITSN2, LIMK2, LPP, MATK, PALLD, POSTN, progesterone, PRRX1, PTK2, RRS1, SEC23A, SEC24C (includes EG:9632), SHPRH, SIX4, SLC25A25, SLC6A4, SSTR3, UBE2V2, VIM, ZNF140
5	1E-13	8	amino acids, APLP2, APP, APPBP2, beta-estradiol, CA2, CXCL5, DYRK3, EN2, EPB41L3, EPO, FLOT1, FSTL1, GRI, GRIA3, GRIA4, GRIP2, HIPK2, HIST4H4 (includes EG:121504), HSP90B1, KISS1, MLLT3, nitric oxide, NSF, PARD6B, PPP3R1, PRDM2, PRKCA, PRKG1, PTPN13, SLC6A4, SYT11, WNT5B, ZNF148, ZNRF1
2. miR-330			
1	1E-40	23	ACVR1, AGTR2, Ap1, APPL1, ATP2B1, BMPR2, C5ORF23, CAPZA1, CCND3, CD247, ERK, Estrogen Receptor, FGFR1, FSH, GNRH, GNRHR, hCG, HSF2, HSPH1, Insulin, KLF10, Lh, MAT2A, MRPS6 (includes EG:64968), PDGF BB, PP2A, RAP2A, RND3, SFRS1, Shc, SMG7, SNAP23, Tgf beta, THBS1, TRA2A
2	1E-37	24	ABCA1, ARFGEF2, BCL11B, BTRC, Calcineurin protein(s), CALCR, CBX5, CHP, cldn, CLDN8, CLDN18, Creb, CYP7A1, GJC1, HELZ, HNRNPU, LDL, N-cor, Nfat (family), NFkB (complex), PCK1, PRKAC, PRKCB, RCAN1, RNA polymerase II, Rxr, SEP15, SLC2A2, STAU1, TBL1XR1, TEAD1, TGFBR3, TJP1, TSHR, Vegf
3	1E-33	20	Akt, BCL9, Caspase, Ck2, EEF1A1, EIF5, EIF4E, ERBB4, ERK1/2, F Actin, FRK, GRB10, Histone h3, Histone h4, ID2, Jnk, KAT2B, KDM4C, MAGED2, Mapk, MECOM, MYEF2, NRG2 (includes EG:381149), P38 MAPK, PAFAH1B1, PCDHA4, PCDHA11, PI3K, Pkc(s), Rb, SIN3A, STK3, SUDS3, TNS1, XK
4	1E-26	17	AFF2, C10ORF10, CBLC, CD19, CD82, CNTFR, CRK, DAB1, EDEM1, EFEMP2, ERBB, FCGR1A/2A/3A, INPPL1, KHDRBS1, KRTAP4-12, LYN, NEDD9, PCDHA1, PCDHA2, PCDHA3, PCDHA5, PCDHA6, PCDHA7, PCDHA8, PCDHA9, PCDHA10, PCDHA12, PLSCR1, PLSCR4, RBM12, SSR1, STX18, XBP1, YIPF5
5	1E-25	16	ADIPOR1, ANGEL2, APOC3, ARL17P1, C9ORF5, C9ORF64, CIAO1, CNBP, EEF2K, FAM107B, FOXA2, glycogen, HNF1B, HNF4A, HNRNPR, KIAA1012, LRP5, MAGOH, METAP2, MLXIPL, NCBP1, ONECUT2, PCK1, PEPCK, PRKAB2, RAI2, RPL31, SLC2A2, SYTL4, TADA3L, TMEM59 (includes EG:9528), TOE1, TROVE2, TTR, USP15
6	1E-25	16	BMP2, C14ORF129, C2CD2, C5ORF15, CRADD, D4S234E, DNM1L, EPM2A (includes EG:7957), FRK, GSK3B, HNRNPA3, HOXA13, ITM2C, KLF10, MIR17 (includes EG:406952), MIR27B (includes EG:407019), MIRN330, MNT, MYPN, NAT12, PAX8, PLXNA2, PPP1CA, PPP1R3C, RBL2, SERPINB2, SFRP1, SFRS5, SMYD2, SOSTDC1, TBX2, TNF, ZBTB34, ZNF410, ZNF423

Network No.	p-value	No. of Predicted Transcripts	Molecules in Network
7	1E-23	15	ATL2, BCL2, BCLAF1, CGN, CMPK1, DCAF7, DERL1, DHX15, DPP10, E2F4, EFHC1, HIST2H2BE, IFNA2, KIF1B, LAPTM4B, MARK1, MIR122, MYCBP2, NDE1, NEFL, OFD1, PTPN3, SCG3, TBC1D4, TP53BP2, TRIM2, TSFM (includes EG:10102), UBXN4, USP18, USP37 (includes EG:57695), VCP, VPS54, YWHAG, YWHAZ, ZC3HAV1
8	1E-23	15	ABL1, AFF4, amino acids, APEX2, ARHGAP12, CSNK1G3, DLGAP5, DNM3, DYRK2, ERLIN2, EWSR1, GNS, hydrogen peroxide, MBNL2, MERTK, MIR133A, MIR133A-1, MIR133A-2, MIRN140, MOBKL1A, NBEA (includes EG:26960), PCDHAC2, PCMT1, PTBP2, RIPK4, SELI, SMNDC1, SNX2, SRF, STK38L (includes EG:23012), TCEB3B, TGFB1, TRAF2, TRIM37, UBTD2
9	1E-22	15	AGT, ANKH, AZIN1, BOK, CLNS1A, DICER1, DOCK5, EIF5, EIF2C2, FMR1, GRIA3, HSF2, LNX2, MIRLET7B (includes EG:406884), MIRN346, NR3C1, OTUD3, PIWIL4, PPP1R3C, PPP2CA, PRKRA, RNF144B, SERINC3, STXBP5, TARBP2, TBX2, TBX5, TNRC6B, UBE2A, UBE2I, UBE2L6, UBE2Q1, UBE2T, UGCG, WDR37
10	1E-21	14	AGR2, ARSA, ARSB, ARSD, ARSE, ARSF, ARSG, ARSI, ARSJ, BFSP1, BMX, CDKN2A, D4S234E, DAG1, DLX1, ERBB2, ERC1, KANK2, MMD, PLXNB2, PQLC1, PTRF, RASA3, RUFY2, SCP2, SLCO1A1, STEAP4, SULF1, SULF2, SUMF1, TLN1, TNF, VIM, ZFR, ZNF148
11	1E-19	13	ASS1, ATP2C1, BTRC, CCDC85B, CDKN1A, CDR2, CLCN5, Cofilin, ERAP1, EXOC7, EXOC8, GCN1L1, HIVEP2, HOXB4, IGF2BP1, IMPACT, KRT20, LRPPRC, MAZ, MBIP, MCM10, MIRLET7A1, MYC, PKN1, PNN, POLR1B, POLR2L (includes EG:5441), PRC1, RFX1, TADA2L, TRIP12, UXT, VASH2, YTHDC1, ZFC3H1
12	1E-19	13	ADAMTS5, AP1S2, API5, DHX15, DRD2, FMO, FMO2, FOXK1, GH1, GPRASP1, HTT, KCNK3, KLF16, MIR103-1 (includes EG:406895), MLH1, PCTP, PSD3, PTPN22, RALB, retinoic acid, RPL6, RPS19, SATB1, SFRS2IP, SLAIN1, SLC5A3, TOX, UBA7, VAMP1, VAPA, VAPB, VGLL3, YWHAB, ZNF133, ZNF706
13	1E-13	10	ACTR5, ACTR8, AP2M1, APEX2, ARHGAP29, ATP6V0E1, ERVK6, FGD1, FIGF, FURIN, GATS, HEBP2, INO80, INO80B, INO80D, INO80E, INS1, ITGB1, JPH1, KATNA1, KLHL24, LDLRAP1, MIR124, MIR124-1, phosphatidylinositol-3, 4, 5-triphosphate, PREX1, Rac, RAVER2, RGS10, SEMA4D, SLC16A1, SORL1, SURF4, TNKS, ZCCHC24
3. miR-181a			
1	1E-43	26	BAG4, BIRC6, CALB1, CAMK2D, DLGAP2, E2F7, ENPP1, FUCA1, HEY2, HRH1, Ifn, MIR1, MTMR12, MUC7, NFkB (complex), NFkB (family), NOTCH4, NTS, PARK2, PAWR, POM121C, PRDX3, PRH2, Proteasome, SPHINGOMYELINASE, TFRC, TIMP3, Tnf receptor, TOM1L1, TRIM2, UBE2, UBE2A, UBE2D1, Ubiquitin, ZFAND6
2	1E-41	25	ACVR2A, ACVR2B, Alcohol group acceptor phosphotransferase, ATM, BHLHE40, DDX3X, EIF4A2, FBXO33, FBXO34, FKBP1A, GABRA1, Histone H1, IFN Beta, Importin beta, IPO8, KLHL2, MAPK1, p70 S6k, PDE5A, POLQ, PP1, PP1-C, PPP1R9A, PPP2R5E, PRKCD, RAN, REPS2, Smad, SPRY4, Tgf beta, TGFBR, TGFBR1, TGFBRAP1, TNFRSF11B, TNPO1
3	1E-34	22	Akt, Ap1, ATXN3, BAG2, BRWD1, CABC1, Cbp/p300, CBX7, CCAR1, E2F5, Estrogen Receptor, ETV6, GRB10, Growth hormone, Histone h4, HOXB4, HOXC8, Hsp90, ITSN1, KDM5A, N-cor, NCOA2, NR3C1, Pias, RAD21, Rxr, SPP1, SRPK2, STAT5a/b, TBC1D4, TBL1X, TBL1XR1, TET2, tyrosine kinase, VitaminD3-VDR-RXR
4	1E-32	21	ACSL1, AKAP, AKAP5, AKAP6, AKAP7, AMPK, ATG5 (includes EG:9474), ATP2A2, CaMKII, Caspase, CCNL2, CDON, Creb, CREB5, Cytochrome c, DNAJC13, DYNC1LI2, FSH, GATA6, hCG, Histone h3, Hsp70, Mmp, P38 MAPK, PCDHA4, Pka, RFC1, RNA polymerase II, SCD, SFRS7, SIRT1, TBPL1, TCERG1, ZBTB44, ZNF83

Network No.	p-value	No. of Predicted Transcripts	Molecules in Network
5	1E-21	16	ADAM11, ADAM28, C2ORF69, C5ORF41, CCNJ, CD40LG, CTDSPL, FBXO33, FRK, ILK, KCNH8, MEGF9, MIR140 (includes EG:406932), MIR153-1, MIR153-2, MIR181B1, MIR181B2, MIR21 (includes EG:406991), MIR217 (includes EG:406999), MT1G, NEFH, NFIB, OTUD4, PAPOLA, PAPOLG, PAX8, PCDHAC2, PHLPP2, RAB9A, RB1, SEMA4G, SLITRK1, SPIN1, TERT, ZIC3
6	1E-20	17	14-3-3, Angiotensin II receptor type 1, Calcineurin A, Calcineurin protein(s), CARD8, CD3, CLASP1, CLIP1, DOCK10, DSC3, Eotaxin, Fcer1, GPD2, Ifn gamma, IKK (complex), Importin alpha, LRRC8D, MAP2K1/2, MEF2, MHC Class II, MKLN1, NFAT5, NFAT (complex), Nfat (family), NR6A1, PLEKHA3, Ptger, PTGER3, SEMA3C, TCR, TIFA, TNF, TNFSF4, VAV, ZFP36L1
7	1E-20	15	ADRBK1, CRK, EFNA1, ENG, ERBB4, FLNA, GNB2L1, INHBA, KAT2B, MIR34A (includes EG:407040), MMP13, MTF2, NEDD9, NFYC, NKX2-1, PCDHA1, PCDHA2, PCDHA3, PCDHA5, PCDHA6, PCDHA7, PCDHA8, PCDHA9, PCDHA10, PCDHA11, PCDHA12, PDPK1, PRDM4, SMAD3, sphingosine-1-phosphate, TMEM165, TP53BP2, WEE1, ZEB2
8	1E-20	15	ACTR5, ACTR8, ALDOB, BBS7, BEND3, CD302, CDCA7L, CHMP2B, ERVK6, FOXA2, GATS, HNF1A, HOXD1, INO80B, INO80D, INO80E, KANK1, METAP1, MIR124, MIR124-1, ONECUT2, OSBPL8, PCNP, POLR3G, POU2F1, retinoic acid, RPE65, RUVBL1, SUCLG1, SUCLG2, SUMO1, TMEM109, TNFRSF21, TWIST2, ZNF673
9	1E-19	15	ACTB, AFTPH, ANKH, APBB1, beta-estradiol, C19ORF12, CPNE2, EIF3M, ENC1, FUT9, GH1, Glutathione peroxidase, GPX8, GSTM3 (includes EG:2947), GSTT1, hydrogen peroxide, MGST2, PARP, PARP4, PARP11, PER3, PNRC2 (includes EG:55629), progesterone, PSG5, PTGER3, PTP4A2, ROD1, SLC7A11, SMARCA4, SMPDL3A, SPOCK1, SRD5A2, TOR1AIP2, TOX, TSC22D1
10	1E-18	14	ATYPICAL PROTEIN KINASE C, BCR, C1q, Collagen type I, Collagen(s), DCN, ERK, Focal adhesion kinase, Ige, IgG, Igm, Integrin, ITGA2, KIF3A, Laminin, LAMP2, LIFR, Mek, NRAS, Pdgf, PDGF BB, PLA2, Pld, PP2A, RAB3IP, Rac, Rap1, Ras, RASSF2, RIN2, RPS6KB1, S1PR1, TLL1, TSPAN8, VCAN
11	1E-18	14	AP3B2, ARRB1, ATP11C, BRD1, BRF1, cldn, CLDN4, CLDN6, CLDN8, CNTN4, CSDA, DNAJB6, FIGN, FOXK1, GAPVD1, HNRNPAB, KCTD3, LOC161527, MIR31 (includes EG:407035), MIRN341, MTX3, NFYB, OGG1, OSBPL3, PLEC1, RAB8B, RBM26, SON, SSTR3, ST8SIA4, TJP1, VIM, VPS39, YWHAB, ZNF468
12	1E-18	14	AHCTF1, ARHGEF3, BTBD3, CDC2L1 (includes EG:984), COBRA1, DDX52, FAM13B, HEATR1, IPO7, KIAA1239, KIAA2022, KPNB1, MED8, MED25, MED26, MED28, MED29, MIR195, MIR373, MIR181C (includes EG:406957), MIR199A1, MLH1, MTMR15, MYC, NUP133, PAPD5, PHTF2, POLR2C, POLR2D, RDBP, Rfc, RPAP2, SR140, WHSC2, XRN1
13	1E-16	13	BRAP, C15ORF29, C20ORF12, CDC14A, CDK5, CTSB, DNAL1, FAM160A2, GFI1B, GHITM, GOLGA2, GPS2, HIC2, HLA-B, HNF4A, LARP4, MINA, MIRN336, MYST2, NDRG1, ONECUT1, PDK2, PPARGC1A, PPP1CA, PPP1R12B, PRDM5, PRKCE, RAB33B, RABAC1, ROCK1, SEC23A, SEC24C (includes EG:9632), SLC6A4, TMF1, ZNF594
14	1E-16	13	ADAMTS4, ADAMTS5, ARF4, C21ORF33, CAPRIN1, CASP14, CCDC14, CCDC85B, COMP, DNAJB4, FNDC3B, GCC2, MATN3, MATN4, NNMT, PARVA, PAX3, PHGDH, PI4K2B, RAB18, RBM25 (includes EG:58517), RNF8, RNH1, SDPR, SF3A3, SLC2A4, STOML2, STX7, TBC1D1, TGFB1, TSPYL4, VPS11, WBP4, WNK1, ZFP36L2
15	1E-16	13	ADAMTSL1, AHSG, ALP, ALPI, ATP2A1, B4GALNT1, BMYO, C16ORF87, CD274, CD1A, CSF1, DBT, DDX3Y (includes EG:8653), DGKA, EPC2, GATM, GRB2, HMGN3, HSPC159, IL4, L-triiodothyronine, LIF, MIRN328, PAX9, RP5-1022P6.2, SHH, SLC24A1, SLC29A1, ST5, TFEC, TMEM131, VEGFA, ZBTB34, ZFP36, ZNF124
16	1E-15	13	ACADVL, AIMP2, ALB, BAI3, BAZ2B, BBC3, C10ORF104, CBX4, CPOX, CTBP2, DENR (includes EG:8562), EEF2K, EIF4A3 (includes EG:9775), GPRIN3, HIPK2, HP, ITGA2, LMO3, METAP2, MORC3, MTA1, PAX5, PCGF2, PDE4B, PEG3, PHC2, PHC3, PSRC1, RBBP6 (includes EG:5930), S100A4, SERPING1, SLC19A2, SUPT3H (includes EG:8464), TADA2B, TP53

Network No.	p-value	No. of Predicted Transcripts	Molecules in Network
17	1E-15	12	ACYP1, APOBEC3B, C14ORF129, CDC5L, CHIC2, CMBL, CRB3, CYP2D9, DISC1, ETNK2, EXOSC4, G3BP2, GSK3B, HNF4A, HOOK1, Hydrolase, KIF3C, LTA4H, MDFI (includes EG:4188), MPP1, MPP5, NLN, NPDC1, PARP4, PRCC, PTPN7, REXO2, SCOC, SFI1, SYNE1, UBP1, VPS29, ZNF439, ZNF440, ZNF559
18	1E-14	12	ADCY, ADRBK1, ARHGAP26, ATP2B1, CCNB1, Cdc2, CSF2RB, CUL3, Cyclin A, Cyclin B, Cyclin E, DARS, DIRAS3, ERK1/2, G protein beta gamma, Gpcr, IL1, IL2, Insulin, Interferon alpha, Jnk, KRAS, LDL, Lh, Mapk, PI3K, PITPNB, Pkc(s), PLC, RAP1B, Ras homolog, Sapk, Shc, STAT, Vegf
19	1E-11	10	amino acids, ASPH, C7ORF16, CDKL1, CTTNBP2NL, DAPK2 (includes EG:23604), DEPDC6 (includes EG:64798), EEF1A1, FMNL2 (includes EG:114793), FNBP4, FOXP1, HAT1, HTT, IL12 (complex), MARK1, MBOAT2, MLST8, MTOR, NXN, PDK1, PGGT1B, PPP2CA, PQBP1, PRPF40A, PTPN7, RIPK4, RLF, RNF34, RP6-213H19.1, STAT4, STRADA, TPP2, TRDMT1, TRIM30, ZNF655
20	1E-11	10	Actin, AFAP1, AKAP5, Alpha tubulin, APLNR, ARPP-21, C21ORF66, CALM1, Calmodulin, CAMK2N2, Ck2, CYB5R3, F Actin, FGD4, GABRR1, GRIN1, KCNQ2, KCNQ3, KIAA1219, MAP1B, NOVA1, PAM, PCP4, PFDN4, PFDN6, POM121, PPEF2, sodium chloride, SRC, TMEM27, TRIM13, Tubulin, UNC13A, USP6, VBP1
4. miR-10b			
1	1E-23	10	ACT1, ANKRD1, ARPP-21, ASB2, BDNF, CALB2, CNGA2, EBF2, EDARADD, FASLG, FIGN, FNBP1L, GRIN2B, GRIN2C, HTT, IP6K2, KLF11, LCA5, MAP3K7, MBTPS1, MIR124-1, NAPB, NEUN, NFkB (complex), PELI1, PELI3, RBM27, RNF7, SGSM2, SNX18, TPH1, UBC, UBE2V1, WDR34, YY1
2	1E-22	10	AGT, ANKRD1, ARG2, Arginase, beta-estradiol, CAPRIN1, CREB1, G3BP1, GALNT1, GRIN2C, GRM3, HOXA2, HOXA3 (includes EG:3200), HOXB3, HOXD3, LOR, MAFB, MYLPF, NR3C1, PBX1, PPP3R1, RB1CC1, RPS6KA5, SLC19A1, STAT3, TBX2, TBX3, TBX5, TFAP2C, TGFB1, TP53, TRNT1, WDR26, WWOX, WWTR1

Supplementary Table 4: Canonical pathways involving at least three molecules present in top networks significantly associated with predicted targets of miR-33, miR-330, miR-181a, and miR-10b.

Canonical Pathways	Present in miR-33 Network?	Present in miR-330 Network?	Present in miR-181a Network?	Present in miR-10b Network?	Total miRNA Networks with Canonical Pathway
Molecular Mechanisms of Cancer	Yes	Yes	Yes	Yes	4
Acute Phase Response Signaling	Yes	Yes	Yes	No	3
Colorectal Cancer Metastasis Signaling	Yes	Yes	Yes	No	3
Glucocorticoid Receptor Signaling	Yes	Yes	Yes	No	3
GNRH Signaling	Yes	Yes	Yes	No	3
G-Protein Coupled Receptor Signaling	Yes	Yes	Yes	No	3
Hepatic Cholestasis	Yes	Yes	Yes	No	3
NF-kB Signaling	Yes	No	Yes	Yes	3
PPAR Signaling	Yes	Yes	Yes	No	3
Production of Nitric Oxide and Reactive Oxygen Species in Macrophages	Yes	Yes	Yes	No	3
Role of Macrophages, Fibroblasts and Endothelial Cells in Rheumatoid Arthritis	Yes	Yes	Yes	No	3
Tight Junction Signaling	Yes	Yes	Yes	No	3
Type I Diabetes Mellitus Signaling	No	Yes	Yes	Yes	3
Actin Cytoskeleton Signaling	Yes	Yes	No	No	2
Androgen Signaling	Yes	Yes	No	No	2
Aryl Hydrocarbon Receptor Signaling	Yes	Yes	No	No	2
Axonal Guidance Signaling	Yes	Yes	No	No	2
B Cell Receptor Signaling	No	Yes	Yes	No	2
BMP signaling pathway	Yes	Yes	No	No	2
Breast Cancer Regulation by Stathmin1	Yes	Yes	No	No	2
Calcium Signaling	Yes	Yes	No	No	2
Cardiac Hypertrophy Signaling	Yes	Yes	No	No	2
Ceramide Signaling	No	Yes	Yes	No	2
Ephrin Receptor Signaling	Yes	Yes	No	No	2
Glioblastoma Multiforme Signaling	Yes	Yes	No	No	2
Glioma Signaling	Yes	Yes	No	No	2
Hepatic Fibrosis / Hepatic Stellate Cell Activation	No	Yes	Yes	No	2
HMGB1 Signaling	No	Yes	Yes	No	2
Huntington's Disease Signaling	No	Yes	No	Yes	2
iCOS-iCOSL Signaling in T Helper Cells	Yes	No	Yes	No	2
IL-12 Signaling and Production in Macrophages	Yes	Yes	No	No	2
IL-6 Signaling	No	Yes	Yes	No	2
IL-8 Signaling	Yes	Yes	No	No	2
ILK Signaling	No	Yes	Yes	No	2
LXR/RXR Activation	Yes	No	Yes	No	2
PAK Signaling	Yes	Yes	No	No	2
PPARalpha/RXRalpha Activation	Yes	Yes	No	No	2
Protein Kinase A Signaling	Yes	No	Yes	No	2
PTEN Signaling	No	Yes	Yes	No	2

Canonical Pathways	Present in miR-33 Network?	Present in miR-330 Network?	Present in miR-181a Network?	Present in miR-10b Network?	Total miRNA Networks with Canonical Pathway
RAR Activation	Yes	Yes	No	No	2
Renin-Angiotensin Signaling	Yes	Yes	No	No	2
Role of NFAT in Cardiac Hypertrophy	Yes	Yes	No	No	2
Role of NFAT in Regulation of the Immune Response	Yes	Yes	No	No	2
Role of Osteoblasts, Osteoclasts and Chondrocytes in Rheumatoid Arthritis	No	Yes	Yes	No	2
Thrombin Signaling	No	Yes	Yes	No	2
Type II Diabetes Mellitus Signaling	Yes	No	Yes	No	2
4-1BB Signaling in T Lymphocytes	Yes	No	No	No	1
Acute Myeloid Leukemia Signaling	Yes	No	No	No	1
AMPK Signaling	No	Yes	No	No	1
Amyloid Processing	Yes	No	No	No	1
Amyotrophic Lateral Sclerosis Signaling	Yes	No	No	No	1
Apoptosis Signaling	No	No	Yes	No	1
Atherosclerosis Signaling	Yes	No	No	No	1
Autoimmune Thyroid Disease Signaling	No	Yes	No	No	1
Cardiac B-adrenergic Signaling	Yes	No	No	No	1
CCR5 Signaling in Macrophages	No	Yes	No	No	1
CD28 Signaling in T Helper Cells	No	Yes	No	No	1
CD40 Signaling	No	No	No	Yes	1
Cdc42 Signaling	No	Yes	No	No	1
Cellular Effects of Sildenafil (Viagra)	Yes	No	No	No	1
Cholecystokinin/Gastrin-mediated Signaling	No	Yes	No	No	1
Chronic Myeloid Leukemia Signaling	Yes	No	No	No	1
Crosstalk between Dendritic Cells and Natural Killer Cells	No	No	Yes	No	1
CXCR4 Signaling	No	Yes	No	No	1
Death Receptor Signaling	No	No	Yes	No	1
Dendritic Cell Maturation	No	No	Yes	No	1
EGF Signaling	No	Yes	No	No	1
EIF2 Signaling	No	Yes	No	No	1
Endothelin-1 Signaling	No	Yes	No	No	1
ERK/MAPK Signaling	No	Yes	No	No	1
Erythropoietin Signaling	No	Yes	No	No	1
Estrogen Receptor Signaling	No	Yes	No	No	1
Estrogen-Dependent Breast Cancer Signaling	No	Yes	No	No	1
Factors Promoting Cardiogenesis in Vertebrates	No	Yes	No	No	1
Germ Cell-Sertoli Cell Junction Signaling	No	Yes	No	No	1
Human Embryonic Stem Cell Pluripotency	No	Yes	No	No	1
Hypoxia Signaling in the Cardiovascular System	No	No	Yes	No	1
IGF-1 Signaling	No	Yes	No	No	1
IL-1 Signaling	Yes	No	No	No	1
IL-2 Signaling	No	Yes	No	No	1
IL-3 Signaling	No	Yes	No	No	1

Canonical Pathways	Present in miR-33 Network?	Present in miR-330 Network?	Present in miR-181a Network?	Present in miR-10b Network?	Total miRNA Networks with Canonical Pathway
Induction of Apoptosis by HIV1	No	No	Yes	No	1
Inositol Phosphate Metabolism	Yes	No	No	No	1
Insulin Receptor Signaling	No	Yes	No	No	1
Integrin Signaling	No	Yes	No	No	1
mTOR Signaling	No	Yes	No	No	1
Natural Killer Cell Signaling	No	Yes	No	No	1
Neuropathic Pain Signaling in Dorsal Horn Neurons	No	No	No	Yes	1
Neurotrophin/TRK Signaling	No	Yes	No	No	1
Nicotinate and Nicotinamide Metabolism	Yes	No	No	No	1
Nitric Oxide Signaling in the Cardiovascular System	Yes	No	No	No	1
Ovarian Cancer Signaling	No	Yes	No	No	1
P13K/AKT Signaling	No	Yes	No	No	1
p70S6K Signaling	No	Yes	No	No	1
PDGF Signaling	No	Yes	No	No	1
Phospholipase C Signaling	No	Yes	No	No	1
Prolactin Signaling	No	Yes	No	No	1
Protein Ubiquitination Pathway	No	No	Yes	No	1
PXR/RXR Activation	Yes	No	No	No	1
Rac Signaling	Yes	No	No	No	1
Regulation of eIF4 and p70S6K Signaling	No	Yes	No	No	1
Regulation of IL-2 Expression in Activated and Anergic T Lymphocytes	No	Yes	No	No	1
Relaxin Signaling	Yes	No	No	No	1
Renal Cell Carcinoma	No	Yes	No	No	1
Role of NANOG in Mammalian Embryonic Stem Cell Pluripotency	No	Yes	No	No	1
Role of PKR in Interferon Induction and Antiviral Response	No	No	Yes	No	1
SAPK/JNK Signaling	No	Yes	No	No	1
Small Cell Lung Cancer Signaling	Yes	No	No	No	1
Sphingosine-1-phosphate Signaling	No	Yes	No	No	1
Synaptic Long Term Potentiation	Yes	No	No	No	1
Systemic Lupus Erythematosus Signaling	No	Yes	No	No	1
T Cell Receptor Signaling	No	Yes	No	No	1
TGF-B Signaling	No	Yes	No	No	1
Thrombopoietin Signaling	No	Yes	No	No	1
VDR/RXR Activation	Yes	No	No	No	1
Wnt/B-catenin Signaling	No	Yes	No	No	1
Xenobiotic Metabolism Signaling	No	No	Yes	No	1

Supplementary Table 5: Biological functions significantly (p-value < 0.005) associated with predicted transcriptional targets of miR-33, miR-330, miR-181a, and miR-10b.

Functions	p-value
1. miR-33	
Amino Acid Metabolism	0.0034
Behavior	0.0001
Cancer	0.0041
Carbohydrate Metabolism	0.0041
Cardiovascular Disease	0.0046
Cell Cycle	0.0009
Cell Death	0.0041
Cell Morphology	0.0003
Cell-mediated Immune Response	0.0041
Cell-To-Cell Signaling and Interaction	0.0036
Cellular Assembly and Organization	0.0036
Cellular Compromise	0.0041
Cellular Development	0.0003
Cellular Function and Maintenance	0.0041
Cellular Growth and Proliferation	0.0041
Cellular Movement	0.0041
Connective Tissue Development and Function	0.0041
Drug Metabolism	0.0041
Embryonic Development	0.0003
Endocrine System Development and Function	0.0041
Gene Expression	0.0041
Genetic Disorder	0.0002
Hair and Skin Development and Function	0.0017
Hematological Disease	0.0041
Hematological System Development and Function	0.0041
Hematopoiesis	0.0041
Immunological Disease	0.0041
Inflammatory Response	0.0041
Lipid Metabolism	0.0041
Metabolic Disease	0.0041
Molecular Transport	0.0041
Nervous System Development and Function	0.0041
Neurological Disease	0.0002
Organ Development	0.0041
Organ Morphology	0.0041
Organismal Injury and Abnormalities	0.0041
Post-Translational Modification	0.0034
Protein Synthesis	0.0025
Psychological Disorders	0.0002
Renal and Urological System Development and Function	0.0009
Reproductive System Development and Function	0.0041
Skeletal and Muscular Disorders	0.0014
Skeletal and Muscular System Development and Function	0.0040
Small Molecule Biochemistry	0.0034

Functions	p-value
Tissue Development	0.0040
Vitamin and Mineral Metabolism	0.0041
2. miR-330	
Carbohydrate Metabolism	0.0006
Cell Cycle	0.0043
Cell Morphology	0.0038
Cell-mediated Immune Response	0.0038
Cell-To-Cell Signaling and Interaction	0.0011
Cellular Development	0.0028
Cellular Function and Maintenance	0.0002
Cellular Growth and Proliferation	0.0028
Cellular Movement	0.0002
Embryonic Development	0.0019
Endocrine System Development and Function	0.0028
Gene Expression	0.0028
Hematological System Development and Function	0.0038
Immune Cell Trafficking	0.0038
Inflammatory Response	0.0019
Lipid Metabolism	0.0006
Molecular Transport	0.0018
Nervous System Development and Function	0.0004
Organismal Development	0.0013
RNA Trafficking	0.0028
Small Molecule Biochemistry	0.0006
Tissue Development	0.0019
3. miR-181a	
Cancer	0.0012
Carbohydrate Metabolism	0.0032
Cardiovascular Disease	0.0000
Cell Cycle	0.0001
Cell Death	0.0001
Cell Morphology	0.0012
Cell-mediated Immune Response	0.0038
Cell-To-Cell Signaling and Interaction	0.0004
Cellular Assembly and Organization	0.0015
Cellular Development	0.0012
Cellular Function and Maintenance	0.0012
Cellular Growth and Proliferation	0.0004
Cellular Movement	0.0012
Connective Tissue Development and Function	0.0012
Connective Tissue Disorders	0.0012
Dermatological Diseases and Conditions	0.0012
DNA Replication, Recombination, and Repair	0.0038
Drug Metabolism	0.0036
Endocrine System Development and Function	0.0002
Endocrine System Disorders	0.0020
Gastrointestinal Disease	0.0005
Gene Expression	0.0001

Functions	p-value
Genetic Disorder	0.0003
Hair and Skin Development and Function	0.0012
Hematological Disease	0.0012
Hematological System Development and Function	0.0038
Hematopoiesis	0.0038
Immune Cell Trafficking	0.0038
Immunological Disease	0.0012
Infection Mechanism	0.0010
Inflammatory Disease	0.0005
Inflammatory Response	0.0012
Lipid Metabolism	0.0023
Metabolic Disease	0.0020
Molecular Transport	0.0012
Nervous System Development and Function	0.0003
Neurological Disease	0.0005
Nucleic Acid Metabolism	0.0012
Organ Development	0.0017
Organismal Development	0.0030
Organismal Functions	0.0023
Organismal Injury and Abnormalities	0.0019
Organismal Survival	0.0020
Respiratory Disease	0.0023
Respiratory System Development and Function	0.0004
Skeletal and Muscular Disorders	0.0012
Skeletal and Muscular System Development and Function	0.0002
Small Molecule Biochemistry	0.0004
Tissue Development	0.0002
Tissue Morphology	0.0004
Tumor Morphology	0.0012
4. miR-10b	
Amino Acid Metabolism	0.0015
Auditory and Vestibular System Development and Function	0.0046
Behavior	0.0030
Cardiovascular Disease	0.0015
Cardiovascular System Development and Function	0.0001
Cell Death	0.0015
Cell Morphology	0.0015
Cell-To-Cell Signaling and Interaction	0.0015
Cellular Assembly and Organization	0.0015
Cellular Compromise	0.0017
Cellular Development	0.0000
Cellular Function and Maintenance	0.0015
Cellular Growth and Proliferation	0.0013
Cellular Movement	0.0015
Connective Tissue Development and Function	0.0016
Connective Tissue Disorders	0.0030
Developmental Disorder	0.0046
Digestive System Development and Function	0.0015

Functions	p-value
Embryonic Development	0.0010
Endocrine System Development and Function	0.0002
Endocrine System Disorders	0.0046
Gastrointestinal Disease	0.0046
Gene Expression	0.0003
Genetic Disorder	0.0046
Hair and Skin Development and Function	0.0030
Hematological Disease	0.0015
Hematological System Development and Function	0.0015
Hematopoiesis	0.0015
Humoral Immune Response	0.0015
Immune Cell Trafficking	0.0046
Immunological Disease	0.0046
Inflammatory Disease	0.0030
Inflammatory Response	0.0046
Lymphoid Tissue Structure and Development	0.0030
Nervous System Development and Function	0.0015
Neurological Disease	0.0002
Organ Development	0.0002
Organ Morphology	0.0015
Organismal Development	0.0002
Organismal Functions	0.0015
Organismal Injury and Abnormalities	0.0015
Reproductive System Development and Function	0.0046
RNA Post-Transcriptional Modification	0.0030
Skeletal and Muscular Disorders	0.0030
Skeletal and Muscular System Development and Function	0.0016
Small Molecule Biochemistry	0.0015
Tissue Development	0.0010
Tissue Morphology	0.0001
Visual System Development and Function	0.0015

Supplementary Table 6: Biological functions significantly (p-value < 0.005) associated with formaldehyde-responsive genes, as identified through pathway analysis of the Li et. al. 2007 genomic database.

Functions	p-value
Cell Death	1.65E-22
Cellular Growth and Proliferation	1.17E-19
Cancer	1.87E-19
Gene Expression	4.40E-16
Cell Cycle	1.51E-15
Cellular Development	6.49E-13
Developmental Disorder	7.45E-12
Reproductive System Disease	2.93E-11
Cardiovascular System Development and Function	2.38E-10
Organismal Development	2.38E-10
Organismal Survival	3.94E-10
Hematological System Development and Function	1.93E-09
Hematopoiesis	1.93E-09
Tissue Development	5.73E-09
Cell-mediated Immune Response	1.09E-08
Cellular Function and Maintenance	1.09E-08
Connective Tissue Disorders	1.10E-08
Immunological Disease	1.10E-08
Inflammatory Disease	1.10E-08
Skeletal and Muscular Disorders	1.10E-08
Tissue Morphology	5.86E-08
Gastrointestinal Disease	5.86E-08
Skeletal and Muscular System Development and Function	7.68E-08
DNA Replication, Recombination, and Repair	1.47E-07
Cardiovascular Disease	3.01E-07
Organ Development	4.23E-07
Genetic Disorder	4.39E-07
Neurological Disease	4.39E-07
Embryonic Development	4.92E-07
Reproductive System Development and Function	4.92E-07
Connective Tissue Development and Function	7.51E-07
Cellular Compromise	9.51E-07
Cellular Movement	9.84E-07
Cell Morphology	1.74E-06
Hematological Disease	4.32E-06
Infection Mechanism	5.28E-06
Nervous System Development and Function	5.28E-06
Hair and Skin Development and Function	1.06E-05
Molecular Transport	1.53E-05

Functions	p-value
Protein Synthesis	1.53E-05
Post-Translational Modification	2.79E-05
Protein Folding	2.79E-05
Cell Signaling	5.05E-05
Small Molecule Biochemistry	5.05E-05
Digestive System Development and Function	8.26E-05
Hepatic System Development and Function	8.26E-05
Cell-To-Cell Signaling and Interaction	1.31E-04
Behavior	1.40E-04
Organ Morphology	1.40E-04
Organismal Injury and Abnormalities	1.47E-04
Cellular Assembly and Organization	1.59E-04
Lymphoid Tissue Structure and Development	2.00E-04
Respiratory Disease	2.03E-04
Metabolic Disease	2.40E-04
Tumor Morphology	2.84E-04
Endocrine System Disorders	3.15E-04
Dermatological Diseases and Conditions	3.39E-04
Lipid Metabolism	6.94E-04
RNA Trafficking	6.94E-04
Renal and Urological Disease	6.94E-04

Supplementary Table 7: Transcriptional targets predicted to be regulated by miR-125b. P_{CT} refers to the probability of preferentially conserved targeting.

Target	Full Name	Context Score	P_{CT}
GCNT1	glucosaminyl (N-acetyl) transferase 1, core 2 (beta-1,6-N-acetylglucosaminyltransferase)	-0.67	> 0.99
TMEM86A	transmembrane protein 86A	-0.26	> 0.99
PODXL	podocalyxin-like	-0.47	> 0.99
ARID3B	AT rich interactive domain 3B (BRIGHT-like)	-0.51	> 0.99
FLJ20309	hypothetical protein FLJ20309	-0.37	0.99
SH3TC2	SH3 domain and tetratricopeptide repeats 2	-0.73	0.99
PHF15	PHD finger protein 15	-0.31	0.98
GJC1	gap junction protein, gamma 1, 45kDa	-0.37	0.98
KLF13	Kruppel-like factor 13	-0.51	0.97
OLFML2A	olfactomedin-like 2A	-0.49	0.97
MFHAS1	malignant fibrous histiocytoma amplified sequence 1	-0.52	0.97
IRF4	interferon regulatory factor 4	-0.58	0.97
RAPGEF5	Rap guanine nucleotide exchange factor (GEF) 5	-0.39	0.97
LBH	limb bud and heart development homolog (mouse)	-0.37	0.97
ENPP1	ectonucleotide pyrophosphatase/phosphodiesterase 1	-0.4	0.97
UBN1	ubinuclein 1	-0.33	0.97
FAM176A	family with sequence similarity 176, member A	-0.33	0.97
MXD4	MAX dimerization protein 4	-0.07	0.97
SMURF1	SMAD specific E3 ubiquitin protein ligase 1	-0.24	0.97
TLE3	transducin-like enhancer of split 3 (E(sp1) homolog, Drosophila)	-0.2	0.97
TRPS1	trichorhinophalangeal syndrome I	-0.19	0.97
KIAA1522	KIAA1522	-0.37	0.97
ASAH3L	N-acylsphingosine amidohydrolase 3-like	-0.34	0.97
UBE2R2	ubiquitin-conjugating enzyme E2R 2	-0.36	0.97
SEMA4D	sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4D	-0.42	0.97
LIN28	lin-28 homolog (C. elegans)	-0.27	0.97
CPSF6	cleavage and polyadenylation specific factor 6, 68kDa	-0.34	0.97
TGOLN2	trans-golgi network protein 2	-0.39	0.97
OSBPL9	oxysterol binding protein-like 9	-0.45	0.97
TBC1D1	TBC1 (tre-2/USP6, BUB2, cdc16) domain family, member 1	-0.36	0.97
SLC39A9	solute carrier family 39 (zinc transporter), member 9	-0.47	0.96
ENPEP	glutamyl aminopeptidase (aminopeptidase A)	-0.52	0.96
ST8SIA4	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 4	-0.26	0.96
GRB10	growth factor receptor-bound protein 10	-0.43	0.96
MYT1	myelin transcription factor 1	-0.39	0.96
STARD13	StAR-related lipid transfer (START) domain containing 13	-0.99	0.96
PTAR1	protein prenyltransferase alpha subunit repeat containing 1	-0.26	0.96
BMF	Bcl2 modifying factor	-0.38	0.96
ZNRF3	zinc and ring finger 3	-0.24	0.96

Target	Full Name	Context Score	P_{CT}
ZSCAN29	zinc finger and SCAN domain containing 29	-0.58	0.96
SPTB	spectrin, beta, erythrocytic (includes spherocytosis, clinical type I)	-0.29	0.96
NIN	ninein (GSK3B interacting protein)	-0.38	0.96
SMG1	PI-3-kinase-related kinase SMG-1	-0.31	0.96
NUP210	nucleoporin 210kDa	-0.55	0.96
GGA2	golgi associated, gamma adaptin ear containing, ARF binding protein 2	-0.31	0.96
DUS1L	dihydrouridine synthase 1-like (S. cerevisiae)	-0.48	0.96
MTF1	metal-regulatory transcription factor 1	-0.33	0.96
C14orf43	chromosome 14 open reading frame 43	-0.13	0.96
PPAT	phosphoribosyl pyrophosphate amidotransferase	-0.43	0.96
SMEK1	SMEK homolog 1, suppressor of mek1 (Dictyostelium)	-0.61	0.96
ORC2L	origin recognition complex, subunit 2-like (yeast)	-0.33	0.96
ACHE	acetylcholinesterase (Yt blood group)	-0.57	0.96
CGN	cingulin	-0.43	0.95
LRP4	low density lipoprotein receptor-related protein 4	-0.28	0.95
C6orf47	chromosome 6 open reading frame 47	-0.3	0.95
TMEM77	transmembrane protein 77	-0.5	0.95
KCNA1	potassium voltage-gated channel, shaker-related subfamily, member 1 (episodic ataxia with myokymia)	-0.24	0.95
KIAA0174	KIAA0174	-0.29	0.95
LFNG	LFNG O-fucosylpeptide 3-beta-N-acetylglucosaminyltransferase	-0.48	0.95
HIC2	hypermethylated in cancer 2	-0.02	0.95
TNFAIP3	tumor necrosis factor, alpha-induced protein 3	-0.29	0.95
CASP2	caspase 2, apoptosis-related cysteine peptidase (neural precursor cell expressed, developmentally down-regulated 2)	-0.25	0.95
CRB2	crumbs homolog 2 (Drosophila)	-0.44	0.95
KIAA0317	KIAA0317	-0.13	0.95
SCARB1	scavenger receptor class B, member 1	-0.34	0.95
ANPEP	alanyl (membrane) aminopeptidase (aminopeptidase N, aminopeptidase M, microsomal aminopeptidase, CD13, p150)	-0.3	0.95
PSTPIP2	proline-serine-threonine phosphatase interacting protein 2	-0.39	0.95
CYP24A1	cytochrome P450, family 24, subfamily A, polypeptide 1	-0.41	0.95
ZSWIM5	zinc finger, SWIM-type containing 5	-0.4	0.95
NCAN	neurocan	-0.23	0.95
TNFSF4	tumor necrosis factor (ligand) superfamily, member 4 (tax-transcriptionally activated glycoprotein 1, 34kDa)	-0.49	0.95
NECAB3	N-terminal EF-hand calcium binding protein 3	-0.39	0.95
SLC6A17	solute carrier family 6, member 17	-0.22	0.95
LRFN2	leucine rich repeat and fibronectin type III domain containing 2	-0.25	0.95
FAM134A	family with sequence similarity 134, member A	-0.25	0.95
C11orf57	chromosome 11 open reading frame 57	-0.28	0.94
CACNB1	calcium channel, voltage-dependent, beta 1 subunit	-0.08	0.94
NIPA1	non imprinted in Prader-Willi/Angelman syndrome 1	-0.18	0.94
GTPBP2	GTP binding protein 2	-0.39	0.94

Target	Full Name	Context Score	P_{CT}
EBF4	early B-cell factor 4	-0.24	0.94
CDR2L	cerebellar degeneration-related protein 2-like	-0.2	0.94
ETV6	ets variant gene 6 (TEL oncogene)	-0.31	0.94
RAPGEFL1	Rap guanine nucleotide exchange factor (GEF)-like 1	-0.22	0.94
CGREF1	cell growth regulator with EF-hand domain 1	-0.35	0.94
CCNJ	cyclin J	-0.46	0.94
KCNIP3	Kv channel interacting protein 3, calsenilin	-0.25	0.94
PSCD1	pleckstrin homology, Sec7 and coiled-coil domains 1(cytohesin 1)	-0.3	0.93
CTF8	chromosome transmission fidelity factor 8 homolog (S. cerevisiae)	-0.38	0.93
MSI1	musashi homolog 1 (Drosophila)	-0.21	0.93
TAF9B	TAF9B RNA polymerase II, TATA box binding protein (TBP)-associated factor, 31kDa	-0.39	0.93
SGPL1	sphingosine-1-phosphate lyase 1	-0.25	0.93
TMPRSS13	transmembrane protease, serine 13	-0.31	0.93
SEMA4C	sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4C	-0.28	0.93
CDH5	cadherin 5, type 2, VE-cadherin (vascular epithelium)	-0.43	0.93
ABHD6	abhydrolase domain containing 6	-0.31	0.93
ENTPD4	ectonucleoside triphosphate diphosphohydrolase 4	-0.22	0.92
ZFYVE1	zinc finger, FYVE domain containing 1	-0.37	0.92
MTUS1	mitochondrial tumor suppressor 1	-0.23	0.92
ATXN1	ataxin 1	-0.19	0.92
LNPEP	leucyl/cystinyl aminopeptidase	-0.25	0.92
DIRAS1	DIRAS family, GTP-binding RAS-like 1	-0.3	0.92
ESRRA	estrogen-related receptor alpha	-0.37	0.92
SH3BP4	SH3-domain binding protein 4	-0.28	0.92
ICHTHYIN	ichthyin protein	-0.51	0.92
SUV39H1	suppressor of variegation 3-9 homolog 1 (Drosophila)	-0.3	0.92
MTMR3	myotubularin related protein 3	-0.2	0.92
FAM118A	family with sequence similarity 118, member A	-0.27	0.92
RBM38	RNA binding motif protein 38	-0.2	0.92
EIF5A2	eukaryotic translation initiation factor 5A2	-0.15	0.92
FAM116A	family with sequence similarity 116, member A	-0.33	0.92
CDC42SE1	CDC42 small effector 1	-0.29	0.92
HCN3	hyperpolarization activated cyclic nucleotide-gated potassium channel 3	-0.34	0.92
USP37	ubiquitin specific peptidase 37	-0.19	0.91
KIAA0644	KIAA0644 gene product	-0.33	0.91
FUT4	fucosyltransferase 4 (alpha (1,3) fucosyltransferase, myeloid-specific)	-0.65	0.91
RND2	Rho family GTPase 2	-0.31	0.91
SOX11	SRY (sex determining region Y)-box 11	-0.41	0.91
FBXO45	F-box protein 45	-0.29	0.91
PCSK7	proprotein convertase subtilisin/kexin type 7	-0.28	0.91
TP53INP1	tumor protein p53 inducible nuclear protein 1	-0.29	0.91

Target	Full Name	Context Score	P_{CT}
MCL1	myeloid cell leukemia sequence 1 (BCL2-related)	-0.42	0.91
EAF1	ELL associated factor 1	-0.12	0.91
NFIB	nuclear factor I/B	-0.09	0.91
LIFR	leukemia inhibitory factor receptor alpha	-0.28	0.91
MAN1B1	mannosidase, alpha, class 1B, member 1	-0.3	0.91
NEU1	sialidase 1 (lysosomal sialidase)	-0.4	0.91
BAK1	BCL2-antagonist/killer 1	-0.56	0.9
SLC35A4	solute carrier family 35, member A4	-0.34	0.9
KIAA1244	KIAA1244	-0.18	0.9
MAP2K7	mitogen-activated protein kinase kinase 7	-0.26	0.9
VDR	vitamin D (1,25- dihydroxyvitamin D3) receptor	-0.26	0.9
SLC39A13	solute carrier family 39 (zinc transporter), member 13	-0.24	0.9

Supplementary Table 8: Transcriptional targets predicted to be regulated by miR-142-3p. P_{CT} refers to the probability of preferentially conserved targeting.

Target	Full Name	Context Score	P_{CT}
ASH1L	ash1 (absent, small, or homeotic)-like (Drosophila)	-0.59	> 0.99
RICTOR	rapamycin-insensitive companion of mTOR	-0.57	> 0.99
ITGB8	integrin, beta 8	-0.62	0.98
C20orf194	chromosome 20 open reading frame 194	-0.72	0.97
SNF1LK	SNF1-like kinase	-0.39	0.97
FAM44B	family with sequence similarity 44, member B	-0.65	0.95
C10orf18	chromosome 10 open reading frame 18	-0.87	0.95
ZCCHC14	zinc finger, CCHC domain containing 14	-0.5	0.94
AFF1	AF4/FMR2 family, member 1	-0.25	0.94
EML4	echinoderm microtubule associated protein like 4	-0.42	0.93
LCOR	ligand dependent nuclear receptor corepressor	-0.33	0.93
CCDC6	coiled-coil domain containing 6	-0.33	0.92
BNC2	basonuclin 2	-0.32	0.9

Supplementary Table 9: Pathways significantly associated with the predicted targets of miR-125b.

Canonical Pathways	p-value	miR-125b Predicted Targets
Sphingolipid Metabolism	0.003	NEU1, SGPL1, ACER2, FUT4
Apoptosis Signaling	0.003	MAP2K7, CASP2, BAK1, MCL1
Glycosphingolipid Biosynthesis - Globoseries	0.012	ST8SIA4, FUT4
Glycosphingolipid Biosynthesis - Neolactoseries	0.012	ST8SIA4, FUT4
Glycosphingolipid Biosynthesis - Ganglioseries	0.014	ST8SIA4, FUT4
N-Glycan Degradation	0.014	NEU1, MAN1B1
O-Glycan Biosynthesis	0.017	GCNT1, FUT4
N-Glycan Biosynthesis	0.037	MAN1B1, FUT4
Sphingosine-1-phosphate Signaling	0.039	RND2, ACER2, CASP2
TNFR1 Signaling	0.042	CASP2, TNFAIP3
Semaphorin Signaling in Neurons	0.048	RND2, SEMA4D

Supplementary Table 10: Pathways significantly associated with the predicted targets of miR-142-3p.

Canonical Pathways	p-value	miR-142-3p Predicted Targets
ILK Signaling	0.008	ITGB8, RICTOR
Role of IL-17F in Allergic Inflammatory Airway Diseases	0.031	SIK1
Macropinocytosis Signaling	0.048	ITGB8

Supplementary Table 11: All formaldehyde-responsive miRNAs within the rat. FC (fold change = exposed / unexposed) values are listed by miRNAs identified as signficantly altered at the expression level by formaldehyde exposure. If a miRNA was not significantly altered at the expression level in a particular exposure condition or tissue, this is indicated with ns (not significant).

miRNAs	Nose 7-Day FC	Nose 28-Day FC	Nose 28-Day + Recovery FC	WBC 7-Day FC	WBC 28-Day FC	WBC 28-Day + Recovery FC	BM 7-Day FC	BM 28-Day FC	BM 28-Day + Recovery FC
miR-34b*	-3.07	ns	ns	ns	ns	ns	ns	ns	ns
miR-144	-2.69	ns	ns	ns	ns	ns	ns	ns	ns
miR-142-3p	-2.17	ns	ns	ns	ns	ns	ns	ns	ns
miR-126*	-1.85	ns	ns	ns	ns	ns	ns	ns	ns
miR-34c*	-1.85	ns	ns	ns	ns	ns	ns	ns	ns
let-7b	-1.82	ns	ns	ns	ns	ns	ns	ns	ns
miR-672	-1.80	ns	ns	ns	ns	ns	ns	ns	ns
miR-214	-1.79	ns	ns	ns	ns	ns	ns	ns	ns
miR-375	-1.78	ns	ns	ns	ns	ns	ns	ns	ns
miR-145	-1.72	ns	ns	ns	ns	ns	ns	ns	ns
miR-135b	-1.69	ns	ns	ns	ns	ns	ns	ns	ns
miR-202*	-1.55	ns	ns	ns	ns	ns	ns	ns	ns
miR-27a	-1.55	ns	ns	ns	ns	ns	ns	ns	ns
miR-29a	-1.54	ns	ns	ns	ns	ns	ns	ns	ns
miR-129-1*	1.52	ns	ns	ns	ns	ns	ns	ns	ns
miR-211	1.52	ns	ns	ns	ns	ns	ns	ns	ns
miR-218	1.54	ns	ns	ns	ns	ns	ns	ns	ns
miR-128	1.57	ns	ns	ns	ns	ns	ns	ns	ns
miR-20a	1.57	ns	ns	ns	ns	ns	ns	ns	ns
miR-20b-5p	1.57	ns	ns	ns	ns	ns	ns	ns	ns
miR-17-5p	1.58	ns	ns	ns	ns	ns	ns	ns	ns
miR-129	1.59	ns	ns	ns	ns	ns	ns	ns	ns
miR-378	1.61	ns	ns	ns	ns	ns	ns	ns	ns
miR-7a-1*	1.64	ns	ns	ns	ns	ns	ns	ns	ns
miR-335	1.65	ns	ns	ns	ns	ns	ns	ns	ns
miR-204	1.67	ns	ns	ns	ns	ns	ns	ns	ns
miR-344a-3p	1.67	ns	ns	ns	ns	ns	ns	ns	ns
miR-425	1.67	ns	ns	ns	ns	ns	ns	ns	ns
miR-872	1.68	ns	ns	ns	ns	ns	ns	ns	ns
miR-191*	1.69	ns	ns	ns	ns	ns	ns	ns	ns
miR-19a	1.72	ns	ns	ns	ns	ns	ns	ns	ns
miR-340-5p	1.73	ns	ns	ns	ns	ns	ns	ns	ns
miR-350	1.77	ns	ns	ns	ns	ns	ns	ns	ns

miRNAs	Nose 7-Day FC	Nose 28-Day FC	Nose 28-Day + Recovery FC	WBC 7-Day FC	WBC 28-Day FC	WBC 28-Day + Recovery FC	BM 7-Day FC	BM 28-Day FC	BM 28-Day + Recovery FC
miR-18a	1.78	ns	ns	ns	ns	ns	ns	ns	ns
miR-31*	1.82	ns	ns	ns	ns	ns	ns	ns	ns
miR-137	1.84	ns	ns	ns	ns	ns	ns	ns	ns
miR-592	2.00	ns	ns	ns	ns	ns	ns	ns	ns
miR-130b	2.04	ns	ns	ns	ns	ns	ns	ns	ns
miR-488	2.27	ns	ns	ns	ns	ns	ns	ns	ns
miR-9*	2.41	ns	ns	ns	ns	ns	ns	ns	ns
miR-9	2.45	ns	ns	ns	ns	ns	ns	ns	ns
miR-96*	3.43	ns	ns	ns	ns	ns	ns	ns	ns
miR-182	5.10	ns	ns	ns	ns	ns	ns	ns	ns
miR-183	6.32	ns	ns	ns	ns	ns	ns	ns	ns
miR-124	ns	-19.20	ns	ns	ns	ns	ns	ns	ns
miR-140*	ns	-3.64	ns	ns	ns	ns	ns	ns	ns
miR-1	ns	-2.93	ns	ns	ns	ns	ns	ns	ns
miR-127	ns	-2.79	ns	ns	ns	ns	ns	ns	ns
miR-434	ns	-2.29	ns	ns	ns	ns	ns	ns	ns
miR-221	ns	-2.27	ns	ns	ns	ns	ns	ns	ns
miR-551b	ns	-2.26	ns	ns	ns	ns	ns	ns	ns
miR-878	ns	-2.23	ns	ns	ns	ns	ns	ns	ns
miR-455*	ns	-2.01	ns	ns	ns	ns	ns	ns	ns
miR-26a	ns	-1.97	ns	ns	ns	ns	ns	ns	ns
miR-129-2*	ns	-1.89	ns	ns	ns	ns	ns	ns	ns
miR-511*	ns	-1.73	ns	ns	ns	ns	ns	ns	ns
miR-547	ns	-1.69	ns	ns	ns	ns	ns	ns	ns
miR-455	ns	-1.60	ns	ns	ns	ns	ns	ns	ns
miR-146b	ns	-1.52	ns	ns	ns	ns	ns	ns	ns
miR-664	ns	-1.52	ns	ns	ns	ns	ns	ns	ns
miR-324-3p	ns	-1.52	ns	ns	ns	ns	ns	ns	ns
miR-141*	ns	1.52	ns	ns	ns	ns	ns	ns	ns
miR-99b	ns	1.81	ns	ns	ns	ns	ns	ns	ns
miR-181c	ns	2.13	ns	ns	ns	ns	ns	ns	ns
miR-34a	ns	2.47	ns	ns	ns	ns	ns	ns	ns
miR-30a	ns	2.51	ns	ns	ns	ns	ns	ns	ns
miR-29c*	ns	2.64	ns	ns	ns	ns	ns	ns	ns
miR-106b	ns	3.01	ns	ns	ns	ns	ns	ns	ns
miR-365	-2.79	-3.26	ns	ns	ns	ns	ns	ns	ns
let-7c	-2.64	-3.75	ns	ns	ns	ns	ns	ns	ns
miR-10b	-2.45	-6.89	ns	ns	ns	ns	ns	ns	ns

miRNAs	Nose 7-Day FC	Nose 28-Day FC	Nose 28-Day + Recovery FC	WBC 7-Day FC	WBC 28-Day FC	WBC 28-Day + Recovery FC	BM 7-Day FC	BM 28-Day FC	BM 28-Day + Recovery FC
miR-21	-2.45	-3.51	ns	ns	ns	ns	ns	ns	ns
miR-146a	-2.25	-2.04	ns	ns	ns	ns	ns	ns	ns
miR-199a-5p	-2.22	-3.27	ns	ns	ns	ns	ns	ns	ns
miR-322	-2.17	-2.23	ns	ns	ns	ns	ns	ns	ns
miR-133b	-2.10	-3.04	ns	ns	ns	ns	ns	ns	ns
let-7a	-2.04	-5.98	ns	ns	ns	ns	ns	ns	ns
miR-125b-5p	-1.91	-4.93	ns	ns	ns	ns	ns	ns	ns
miR-223	-1.88	-1.98	ns	ns	ns	ns	ns	ns	ns
miR-195	-1.85	-2.33	ns	ns	ns	ns	ns	ns	ns
miR-450a	-1.84	-2.34	ns	ns	ns	ns	ns	ns	ns
miR-23b	-1.82	3.12	ns	ns	ns	ns	ns	ns	ns
miR-23a	-1.80	-4.71	ns	ns	ns	ns	ns	ns	ns
miR-322*	-1.78	-2.77	ns	ns	ns	ns	ns	ns	ns
miR-1949	-1.72	-2.89	ns	ns	ns	ns	ns	ns	ns
miR-199a-3p	-1.72	-2.17	ns	ns	ns	ns	ns	ns	ns
miR-10a-5p	-1.65	-2.54	ns	ns	ns	ns	ns	ns	ns
miR-203	-1.64	-17.11	ns	ns	ns	ns	ns	ns	ns
let-7f	-1.55	-3.46	ns	ns	ns	ns	ns	ns	ns
miR-200c	1.92	1.52	ns	ns	ns	ns	ns	ns	ns
miR-200b	2.41	1.66	ns	ns	ns	ns	ns	ns	ns
miR-200a	2.55	1.68	ns	ns	ns	ns	ns	ns	ns
miR-200a*	2.60	1.66	ns	ns	ns	ns	ns	ns	ns
miR-429	2.68	2.70	ns	ns	ns	ns	ns	ns	ns
miR-598-3p	3.03	2.03	ns	ns	ns	ns	ns	ns	ns
miR-200b*	4.03	2.09	ns	ns	ns	ns	ns	ns	ns
miR-96	4.26	2.44	ns	ns	ns	ns	ns	ns	ns
miR-183*	5.24	2.62	ns	ns	ns	ns	ns	ns	ns
miR-503	-1.61	-1.61	ns	1.57	ns	ns	ns	ns	ns
miR-143	-2.16	-3.00	ns	ns	1.62	ns	ns	ns	ns
miR-126	-1.77	-3.02	ns	ns	1.53	ns	ns	ns	ns
miR-497	-1.73	-1.95	ns	ns	1.60	ns	ns	ns	ns
miR-31	1.99	3.73	ns	ns	1.51	ns	ns	ns	ns
miR-451	-2.71	ns	ns	-1.51	ns	ns	ns	ns	ns
miR-142-5p	-2.19	ns	ns	1.69	ns	ns	ns	ns	ns
miR-193	-2.01	ns	ns	1.55	ns	ns	ns	ns	ns
miR-542-3p	-1.61	ns	ns	1.68	ns	ns	ns	ns	ns
miR-150	-1.52	ns	ns	ns	-2.83	ns	ns	ns	ns

miRNAs	Nose 7-Day FC	Nose 28-Day FC	Nose 28-Day + Recovery FC	WBC 7-Day FC	WBC 28-Day FC	WBC 28-Day + Recovery FC	BM 7-Day FC	BM 28-Day FC	BM 28-Day + Recovery FC
miR-326	ns	ns	ns	1.54	1.74	ns	ns	ns	ns
miR-212	ns	ns	ns	-2.04	ns	-1.91	ns	ns	ns
miR-1224	ns	ns	ns	-3.39	ns	ns	ns	ns	ns
miR-494	ns	ns	ns	-2.75	ns	ns	ns	ns	ns
miR-196c	ns	ns	ns	-2.64	ns	ns	ns	ns	ns
miR-327	ns	ns	ns	-2.32	ns	ns	ns	ns	ns
miR-331*	ns	ns	ns	-2.27	ns	ns	ns	ns	ns
miR-188*	ns	ns	ns	-1.97	ns	ns	ns	ns	ns
miR-3593-5p	ns	ns	ns	-1.97	ns	ns	ns	ns	ns
miR-290	ns	ns	ns	-1.90	ns	ns	ns	ns	ns
miR-30c-1*	ns	ns	ns	-1.89	ns	ns	ns	ns	ns
miR-3582	ns	ns	ns	-1.74	ns	ns	ns	ns	ns
miR-10a-3p	ns	ns	ns	-1.66	ns	ns	ns	ns	ns
miR-188	ns	ns	ns	-1.65	ns	ns	ns	ns	ns
miR-3580-3p	ns	ns	ns	-1.61	ns	ns	ns	ns	ns
miR-500	ns	ns	ns	-1.59	ns	ns	ns	ns	ns
miR-92a-2*	ns	ns	ns	-1.59	ns	ns	ns	ns	ns
miR-760-3p	ns	ns	ns	-1.53	ns	ns	ns	ns	ns
miR-331	ns	ns	ns	1.50	ns	ns	ns	ns	ns
miR-151*	ns	ns	ns	1.57	ns	ns	ns	ns	ns
miR-29b	ns	ns	ns	1.65	ns	ns	ns	ns	ns
miR-34b	ns	ns	ns	1.67	ns	ns	ns	ns	ns
miR-141	ns	ns	ns	1.67	ns	ns	ns	ns	ns
miR-383	ns	ns	ns	1.68	ns	ns	ns	ns	ns
miR-32	ns	ns	ns	2.06	ns	ns	ns	ns	ns
miR-33	ns	ns	ns	2.62	ns	ns	ns	ns	ns
miR-342-3p	ns	ns	ns	ns	-1.88	ns	ns	ns	ns
miR-196b	ns	ns	ns	ns	1.60	ns	ns	ns	ns
miR-652*	ns	ns	ns	ns	ns	-1.89	ns	ns	ns
miR-877	ns	ns	ns	ns	ns	-1.59	ns	ns	ns

Supplementary Table 12: 42 genes differentially expressed by formaldehyde within the nose of rats in the 28-day group.

Affymetrix Transcript ID	Gene Symbol	Fold Change (Exposed / Unexposed)	p-value
10827201	*Clca3*	-18.82	0.00893
10851581	*Slpi*	-2.56	0.00920
10749757	*Cd7*	-1.91	0.00039
10859090	*LOC689800*	-1.90	0.00236
10905453	*Kdelr3*	-1.89	0.00333
10714103	*Mpeg1*	-1.79	0.00283
10876208	*Ccl21b*	-1.79	0.00514
10798610	*Sfrp4*	-1.78	0.00542
10733056	*Ifi47*	-1.75	0.00441
10791250	*Lpl*	-1.74	0.00726
10799552	*Ucma*	-1.72	0.00619
10868627	*Glipr2*	-1.71	0.00361
10896541	*Nov*	-1.70	0.00162
10766869	*Cd34*	-1.70	0.00096
10798135	*Serpinb9*	-1.66	0.00432
10769771	*Fcgr2b*	-1.65	0.00883
10917770	*Mpi*	-1.64	0.00918
10721339	*Clec11a*	-1.62	0.00003
10823548	*Lxn*	-1.61	0.00543
10773180	*Hs3st1*	-1.60	0.00181
10866019	*Clec7a*	-1.59	0.00366
10705874	*Tyrobp*	-1.59	0.00331
10936072	*Lage3*	-1.59	0.00334
10937624	*Mid1*	-1.58	0.00134
10786532	*Mustn1*	-1.58	0.00880
10768376	*Pla2g4a*	-1.57	0.00248
10718954	*Lilrb4*	-1.56	0.00095
10939129	*Pof1b*	-1.55	0.00984
10907869	*Mmp12*	-1.54	0.00612
10736795	*Slfn2*	-1.54	0.00481
10712477	*Lsp1*	-1.54	0.00523
10870837	*Txndc12*	-1.53	0.00119
10819523	*Gbp2*	-1.53	0.00169
10710051	*Far1*	-1.53	0.00442
10771655	*Cxcl10*	-1.53	0.00749
10810503	*Slc10a7*	-1.52	0.00105
10848733	*Ehd4*	-1.51	0.00145
10927780	*Slc40a1*	-1.51	0.00133
10821698	*Osmr*	-1.51	0.00800
10931159	*Tmem37*	1.53	0.00696
10937867	*LOC100362769*	1.67	0.00523
10897666	*Dmc1*	1.68	0.00251

Supplementary Table 13: 130 genes differentially expressed by formaldehyde within the WBC of rats in the 28-day group.

Affymetrix Transcript ID	Gene Symbol	Fold Change (Exposed / Unexposed)	p-value
10930624	*LOC310926*	-2.24	0.00320
10849700	*Mal*	-2.01	0.00703
10889177	*Rhob*	-1.90	0.00352
10897852	*Tnrc6b*	-1.76	0.00673
10784117	*Gjb2*	-1.65	0.00185
10873419	*Ubr4*	-1.56	0.00779
10834719	*Rxra*	-1.53	0.00625
10862527	*Skap2*	1.50	0.00395
10855925	*Mmrn1*	1.51	0.00417
10737359	*Cuedc1*	1.51	0.00497
10812021	*Itgb1*	1.51	0.00168
10829816	*Reep3*	1.52	0.00906
10917969	*Arih1*	1.52	0.00369
10831236	*Ly6g6c*	1.52	0.00588
10814484	*Tbl1xr1*	1.53	0.00666
10911250	*Rora*	1.54	0.00170
10905307	*C1qtnf6*	1.54	0.00945
10855163	*Cul1*	1.54	0.00821
10921141	*Fyco1*	1.54	0.00768
10918718	*Tmod3*	1.54	0.00761
10862327	*Tpk1*	1.54	0.00462
10887831	*Fez2*	1.54	0.00782
10861033	*Ndufa4*	1.55	0.00246
10761375	*Vkorc1l1*	1.55	0.00915
10803520	*Slc39a6*	1.55	0.00330
10765102	*Pigc*	1.55	0.00168
10786163	*Ppif*	1.55	0.00161
10795689	*Edaradd*	1.56	0.00832
10773098	*Tapt1*	1.56	0.00256
10825209	*Cd160*	1.56	0.00264
10708672	*Prcp*	1.56	0.00386
10853171	*Ptpn12*	1.57	0.00762
10896353	*Oxr1*	1.57	0.00242
10935064	*Plp1*	1.58	0.00866
10846762	*Calcrl*	1.58	0.00695
10818823	*Pde5a*	1.58	0.00605
10927809	*Hibch*	1.58	0.00403
10921086	*RGD1311745*	1.59	0.00203
10835817	*Ptgs1*	1.59	0.00963
10782028	*Dnajc3*	1.59	0.00649
10860858	*Bet1*	1.59	0.00999

Affymetrix Transcript ID	Gene Symbol	Fold Change (Exposed / Unexposed)	p-value
10754735	*Muc4*	1.59	0.00465
10937391	*Nxt2*	1.60	0.00907
10798459	*Hist1h2bb*	1.60	0.00553
10853797	*Tes*	1.61	0.00701
10923338	*Coq10b*	1.61	0.00432
10914481	*Kif15*	1.62	0.00546
10770197	*Akt3*	1.62	0.00363
10880583	*Clic4*	1.62	0.00431
10909091	*Siae*	1.63	0.00427
10885006	*Tmx1*	1.64	0.00236
10916920	*Atp5l*	1.65	0.00430
10864590	*Srgap3*	1.65	0.00969
10765937	*Kmo*	1.68	0.00977
10826985	*Dapp1*	1.68	0.00718
10798135	*Serpinb9*	1.68	0.00098
10751636	*Muc20*	1.68	0.00924
10932646	*Ap1s1*	1.69	0.00337
10877943	*Ptplad2*	1.69	0.00831
10812779	*LOC100361629*	1.69	0.00198
10800497	*Galnt1*	1.69	0.00570
10853995	*Asb15*	1.69	0.00061
10851670	*Pltp*	1.70	0.00788
10778558	*Plek*	1.70	0.00184
10825736	*Adora3*	1.72	0.00380
10792592	*Agpat5*	1.73	0.00615
10790581	*LOC290595*	1.75	0.00950
10845298	*Arl5a*	1.75	0.00392
10791545	*Hprt1*	1.76	0.00578
10803359	*Dsc2*	1.76	0.00694
10923782	*Abi2*	1.78	0.00858
10938874	*Zdhhc15*	1.78	0.00543
10716634	*Pcmt1*	1.79	0.00612
10912058	*Cyb5r4*	1.79	0.00957
10923198	*Nab1*	1.80	0.00655
10782234	*Adk*	1.81	0.00971
10821207	*Kif2a*	1.81	0.00231
10937641	*LOC100363276*	1.81	0.00532
10912718	*Cpne4*	1.85	0.00838
10935555	*Fhl1*	1.87	0.00464
10776034	*Mobkl1a*	1.87	0.00579
10870146	*Alg6*	1.87	0.00781
10821377	*Gzmk*	1.89	0.00544
10847728	*RGD1309730*	1.89	0.00882

Affymetrix Transcript ID	Gene Symbol	Fold Change (Exposed / Unexposed)	p-value
10936917	*Xk*	1.90	0.00945
10861213	*Tspan12*	1.92	0.00798
10905356	*Lgals2*	1.92	0.00312
10939480	*Morc4*	1.93	0.00329
10814415	*Trim55*	1.94	0.00695
10854446	*Cald1*	1.96	0.00892
10906533	*Twf1*	1.97	0.00796
10836638	*Klhl23*	1.97	0.00601
10846259	*Lnp*	1.98	0.00934
10937624	*Mid1*	1.99	0.00957
10770159	*Chml*	1.99	0.00821
10865300	*Klrg1*	2.00	0.00416
10710135	*Nucb2*	2.03	0.00831
10926769	*Rhag*	2.05	0.00379
10794836	*Serpinb6a*	2.05	0.00926
10854108	*Calu*	2.05	0.00110
10939310	*Gla*	2.06	0.00645
10784378	*RGD1306437*	2.06	0.00694
10787765	*Psd3*	2.08	0.00782
10791474	*Mfap3l*	2.09	0.00814
10856092	*LOC100362003*	2.10	0.00369
10795611	*RGD1564129*	2.11	0.00659
10903529	*Angpt1*	2.12	0.00234
10932773	*Chrdl1*	2.13	0.00286
10898879	*Tmem117*	2.13	0.00593
10904587	*LOC300024*	2.14	0.00489
10779225	*Spetex-2D*	2.15	0.00653
10833659	*Amd1*	2.16	0.00585
10821415	*Itga2*	2.16	0.00551
10838312	*Mpped2*	2.18	0.00585
10763367	*Serpinb2*	2.19	0.00316
10925695	*Pam*	2.26	0.00523
10862876	*Il12rb2*	2.26	0.00615
10907962	*Trpc6*	2.38	0.00649
10779243	*Spetex-2H*	2.40	0.00544
10779265	*Spetex-2B*	2.45	0.00517
10823819	*Rxfp1*	2.45	0.00759
10779253	*Spetex-2F*	2.47	0.00412
10939505	*Nup62cl*	2.48	0.00345
10779260	*Spetex-2C*	2.48	0.00692
10779233	*Spetex-2G*	2.50	0.00397
10779293	*Spetex-2A*	2.50	0.00447
10787757	*Csgalnact1*	2.61	0.00926

Affymetrix Transcript ID	Gene Symbol	Fold Change (Exposed / Unexposed)	p-value
10812734	*Serf1*	2.67	0.00429
10801683	*Prr16*	2.78	0.00811
10763351	*Serpinb11*	2.85	0.00303

Supplementary Table 14: Individual networks constructed using formaldehyde-associated mRNAs predicted to be regulated by formaldehyde-responsive miRNAs.

Network Constructed	Molecules in Network	p-value
Nose Network 1	5(S)-HETE, 15(S)-HETE, CAP1, CARD9, CLEC11A, CLEC7A, DMC1, EHD4, ERK, FAR1, GLIPR2, HNF4A, Igfbp, IL5, IL31, Jnk, LBP, LTB4R2, LTB4R, Mac1, MID1, MID1IP1, miR-155-5p, NOV, OSMR, P38 MAPK, PDGFC, PMP22, PTH, SEMA7A, SFRP4, STYXL1, TACR1, urea, Vegf	1.0E-27
Nose Network 2	beta-estradiol, BTBD2, CLEC7A, CREB3L4, EHD1, FGB, FXR1, GBP2, HBP1, HS3ST1, HSPE1, IFNGR2, KDELR3, KLHL20, KRT17, LDHA, LSM2, ORM1, OSMR, PAFAH1B3, PPP1R8, S100A6, SEPHS1, SERPINB9, SERPINC1, SMYD3, STK3, sulfotransferase, TACR1, TFG, TMEM37, TNF, UBC, UXT, ZYX	1.0E-15
WBC Network 1	ARAP1, ARL5A, BET1, C18orf8, CALU, CCZ1/CCZ1B, CEP192, CHML, CHMP1A, CTNNAL1, FGFR1OP, FYCO1, GGCX, GOSR2, IFNA2, KIAA1715, MUC20, NME3, PMPCA, PPP2R3C, PRKRIR, PSD3, RAB27A, RABGGTA, SEC22B, SEC23B, SEC24B, SKAP2, TES, TMX1, TNF, TWF1, UBC, UBR4, YKT6	1.0E-36
WBC Network 2	AKT3, Akt, AKT1S1, APPL1, ATF6, beta-estradiol, BMF, CD151, CDH13, CIB1, Collagen Type VI, COX1, COX2, DNAJC3, EDN3, FYB, Integrin alpha 4 beta 1, ITGA2, LYL1, MPPED2, MTORC2, NEDD9, NME3, NOV, PHLPP1, PHLPP2, Psg16, Ptk, SEC24B, SEMA6D, TCL1A, TRPC1, TTC3, Vla-4, ZYX	1.0E-09
WBC Network 3	CLDN7, thiamine diphosphokinase, TPK1	1.0E-03

REFERENCES

Abba MC, Sun H, Hawkins KA, Drake JA, Hu Y, Nunez MI, et al. 2007. Breast Cancer Molecular Signatures as Determined by SAGE: Correlation with Lymph Node Status. Mol Cancer Res 5:881-890.

Ahn KH, Kim SK, Lee JM, Jeon HJ, Lee DH, Kim DK. 2010. Proteomic Analysis of Bronchoalveolar Lavage Fluid Obtained from Rats Exposed to Formaldehyde. Journal of Health Science 56:287-295.

Aliotta JM, Pereira M, Johnson KW, de Paz N, Dooner MS, Puente N, et al. 2010. Microvesicle entry into marrow cells mediates tissue-specific changes in mRNA by direct delivery of mRNA and induction of transcription. Exp Hematol 38:233-245.

Andersen ME, Clewell HJr, Bermudez E, Willson GA, Thomas RS. 2008. Genomic signatures and dose-dependent transitions in nasal epithelial responses to inhaled formaldehyde in the rat. Toxicol Sci 105:368-383.

Andersen ME, Clewell HJ, 3rd, Bermudez E, Dodd DE, Willson GA, Campbell JL, et al. 2010. Formaldehyde: integrating dosimetry, cytotoxicity, and genomics to understand dose-dependent transitions for an endogenous compound. Toxicol Sci 118:716-731.

Bachand AM, Mundt KA, Mundt DJ, Montgomery RR. 2010. Epidemiological studies of formaldehyde exposure and risk of leukemia and nasopharyngeal cancer: a meta-analysis. Crit Rev Toxicol 40:85-100.

Bakand S, Winder C, Khalil C, Hayes A. 2005. Toxicity Assessment of Industrial Chemicals and Airborne Contaminants: Transition from In Vivo to In Vitro Test Methods: A Review. Inhalation Toxicol 17:775-787.

Bartel DP. 2004. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. Cell 116:281-297.

Bartke N, Hannun YA. 2009. Bioactive sphingolipids: metabolism and function. J Lipid Res 50 Suppl:S91-S96.

Beane Freeman LE, Blair A, Lubin JH, Stewart PA, Hayes RB, Hoover RN, et al. 2009. Mortality from lymphohematopoietic malignancies among workers in formaldehyde industries: the National Cancer Institute Cohort. J Natl Cancer Inst 101:751-761.

Benjamini Y, Hochberg Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc B 57:289-300.

Bianchi M, Martucci C, Biella G, Ferrario P, Sacerdote P. 2004. Increased substance P and tumor necrosis factor-α level in the paws following formalin injection in rat tail. Brain Res 1019:255-258.

Blank F, Rothen-Rutishauser BM, Schurch S, Gehr P. 2006. An Optimized In Vitro Model of the Respiratory Tract Wall to Study Particle Cell Interactions. J Aerosol Med 19:392-405.

Bloemen K, Verstraelen S, Van Den Heuvel R, Witters H, Nelissen I, Schoeters G. 2007. The allergic cascade: Review of the most important molecules in the asthmatic lung. Immunology Letters 113:6-18.

Bollati V, Baccarelli A. 2010. Environmental epigenetics. Heredity 105:105-112.

Boobis AR. 2010. Mode of action considerations in the quantitative assessment of tumour responses in the liver. Basic Clin Pharmacol Toxicol 106:173-179.

Bousquet M, Quelen C, Rosati R, Mansat-De Mas V, La Starza R, Bastard C, et al. 2008. Myeloid cell differentiation arrest by miR-125b-1 in myelodysplastic syndrome and acute myeloid leukemia with the t(2;11)(p21;q23) translocation. J Exp Med 205:2499-2506.

Calin GA, Croce CM. 2006. MicroRNA signatures in human cancers. Nat Rev Cancer 6:857-866.

Calvano SE, Xiao W, Richards DR, Felciano RM, Baker HV, Cho RJ, et al. 2005. A network-based analysis of systemic inflammation in humans. Nature 437:1032-1037.

Cammarata G, Augugliaro L, Salemi D, Agueli C, La Rosa M, Dagnino L, et al. 2010. Differential expression of specific microRNA and their targets in acute myeloid leukemia. Am J Hematol 85:331-339.

Camussi G, Deregibus MC, Bruno S, Cantaluppi V, Biancone L. 2010. Exosomes/microvesicles as a mechanism of cell-to-cell communication. Kidney Int 78:838-848.

Casanova M, Heck HD, Everitt JI, Harrington WWJr, Popp JA. 1988. Formaldehyde concentrations in the blood of rhesus monkeys after inhalation exposure. Food Chem Toxicol 26:715-716.

Casanova M, Deyo DF, Heck HD. 1989. Covalent binding of inhaled formaldehyde to DNA in the nasal mucosa of Fischer 344 rats: analysis of formaldehyde and DNA by high-performance liquid chromatography and provisional pharmacokinetic interpretation. Fundam Appl Toxicol 12:397-417.

Casanova M, Morgan KT, Steinhagen WH, Everitt JI, Popp JA, Heck HD. 1991. Covalent binding of inhaled formaldehyde to DNA in the respiratory tract of rhesus monkeys: pharmacokinetics, rat-to-monkey interspecies scaling, and extrapolation to man. Fundam Appl Toxicol 17:409-428.

Casanova M, Morgan KT, Gross EA, Moss OR, Heck HA. 1994. DNA-protein cross-links and cell replication at specific sites in the nose of F344 rats exposed subchronically to formaldehyde. Fundam Appl Toxicol 23:525-536.

Cedar H, Bergman Y. 2009. Linking DNA methylation and histone modification: patterns and paradigms. Nat Rev Genet 10:295-304.

Chang JC, Gross EA, Swenberg JA, Barrow CS. 1983. Nasal cavity deposition, histopathology, and cell proliferation after single or repeated formaldehyde exposures in B6C3F1 mice and F-344 rats. Toxicol Appl Pharmacol 68:161-176.

Chen C-Z, Lodish HF. 2005. MicoRNAs as regulators of mammalian hematopoiesis. Seminars in Immunology 17:155-165.

Chen HC, Chen GH, Chen YH, Liao WL, Liu CY, Chang KP, et al. 2009. MicroRNA deregulation and pathway alterations in nasopharyngeal carcinoma. Br J Cancer 100:1002-1011.

Cho S, Ryu J. 2002. Classifying gene expression data of cancer using classifier ensemble with mutually exclusive features. Proceedings of the IEEE 90:1744-1753.

Cohen SM, Arnold LL. 2008. Cell Proliferation and Carcinogenesis. J Toxicol Pathol 21:1-7.

Cole P, Axten C. 2004. Formaldehyde and leukemia: an improbable causal relationship. Regul Toxicol Pharmacol 40:107-112.

Costinean S, Zanesi N, Pekarsky Y, Tili E, Volinia S, Heerema N, et al. 2006. Pre-B cell proliferation and lymphoblastic leukemia/high-grade lymphoma in E(mu)-miR155 transgenic mice. Proc Natl Acad Sci USA 103:7024-7029.

Dhanasekaran SM, Barrette TR, Ghosh D, Shah R, Varambally S, Kurachi K, et al. 2001. Delineation of prognostic biomarkers in prostate cancer. Nature 412:822-826.

Doyle M, Sexton KG, Jeffries H, Bridge K, Jaspers I. 2004. Effects of 1,3-butadiene, isoprene, and their photochemical degradation products on human lung cells. Environ Health Perspect 112:1488-1495.

Doyle M, Sexton KG, Jeffries H, Jaspers I. 2007. Atmospheric photochemical transformations enhance 1,3-butadiene-induced inflammatory responses in human epithelial cells: The role of ozone and other photochemical degradation products. Chem Biol Interact 166:163-169.

Duhayon S, Hoet P, Van Maele-Fabry G, Lison D. 2008. Carcinogenic potential of formaldehyde in occupational settings: a critical assessment and possible impact on occupational exposure levels. Int Arch Occup Environ Health 81:695-710.

Edgar R, Domrachev M, Lash AE. 2002. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 30:207-210.

Filipowicz W, Bhattacharyya SN, Sonenberg N. 2008. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nat Rev Genet 9:102-114.

Fjellbirkeland L, Cambier S, Broaddus VC, Hill A, Brunetta P, Dolganov G, et al. 2003. Integrin alphavbeta8-mediated activation of transforming growth factor-beta inhibits human airway epithelial proliferation in intact bronchial tissue. Am J Pathol 163:533-542.

Friedman RC, Farh KK, Burge CB, Bartel DP. 2009. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res 19:92-105.

Garcia DM, Baek D, Shin C, Bell GW, Grimson A, Bartel DP. 2011. Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs. Nat Struct Mol Biol 18:1139-1146.

Garzon R, Volinia S, Liu CG, Fernandez-Cymering C, Palumbo T, Pichiorri F, et al. 2008. MicroRNA signatures associated with cytogenetics and prognosis in acute myeloid leukemia. Blood 111:3183-3189.

Garzon R, Heaphy CE, Havelange V, Fabbri M, Volinia S, Tsao T, et al. 2009. MicroRNA 29b functions in acute myeloid leukemia. Blood 114:5331-5341.

Ghosh AK, Secreto CR, Knox TR, Ding W, Mukhopadhyay D, Kay NE. 2009. Circulating microvesicles in B-cell chronic lymphocytic leukemia can stimulate marrow stromal cells: implications for disease progression. Blood 115:1755-1764.

Gilcrease MZ. 2007. Integrin signaling in epithelial cells. Cancer Lett 247:1-25.

Golden R, Pyatt D, Shields PG. 2006. Formaldehyde as a potential human leukemogen: an assessment of biological plausibility. Crit Rev Toxicol 36:135-153.

Goldstein BD. 2011. Hematological and toxicological evaluation of formaldehyde as a potential cause of human leukemia. Hum Exp Toxicol 30:725-735.

Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, et al. 1999. Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring. Science 286:531-537.

Goncharova EA, Lim PN, Chisolm A, Fogle HW, Taylor JH, Goncharov DA, et al. 2010. Interferons modulate mitogen-induced protein synthesis in airway smooth muscle. Am J Physiol Lung Cell Mol Physiol 299:L25-35.

Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP. 2007. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol Cell 27:91-105.

Grubbs FE. 1969. Procedures for Detecting Outlying Observations in Samples. Technometrics 11:1-21.

Han YC, Park CY, Bhagat G, Zhang J, Wang Y, Fan JB, et al. 2010. microRNA-29a induces aberrant self-renewal capacity in hematopoietic progenitors, biased myeloid development, and acute myeloid leukemia. J Exp Med 207:475-489.

Hanahan D, Weinberg RA. 2011. Hallmarks of cancer: the next generation. Cell 144:646-674.

Harkema JR, Carey SA, Wagner JG. 2006. The nose revisited: a brief review of the comparative structure, function, and toxicologic pathology of the nasal epithelium. Toxicol Pathol 34:252-269.

Hauptmann M, Stewart PA, Lubin JH, Beane Freeman LE, Hornung RW, Herrick RF, et al. 2009. Mortality from lymphohematopoietic malignancies and brain cancer among embalmers exposed to formaldehyde. J Natl Cancer Inst 101:1696-1708.

Heck H, Casanova M. 1999. Pharmacodynamics of formaldehyde: applications of a model for the arrest of DNA replication by DNA-protein cross-links. Toxicol Appl Pharmacol 160:86-100.

Heck H, Casanova M. 2004. The implausibility of leukemia induction by formaldehyde: a critical review of the biological evidence on distant-site toxicity. Regul Toxicol Pharmacol 40:92-106.

Heck HD, Casanova-Schmitz M, Dodd PB, Schachter EN, Witek TJ, Tosun T. 1985. Formaldehyde (CH2O) concentrations in the blood of humans and Fischer-344 rats exposed to CH2O under controlled conditions. Am Ind Hyg Assoc J 46:1-3.

Hester SD, Benavides GB, Yoon L, Morgan KT, Zou F, Barry W, et al. 2003. Formaldehyde-induced gene expression in F344 rat nasal respiratory epithelium. Toxicology 187:13-24.

IARC. 2006. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Formaldehyde, 2-butoxyethanol and 1-tert-butoxypropan-2-ol, Volume 88.: International Agency for Research on Cancer. Available: http://monographs.iarc.fr/ENG/Monographs/vol88/index.php.

Im H, Oh E, Mun J, Khim JY, Lee E, Kang HS, et al. 2006. Evaluation of toxicological monitoring markers using proteomic analysis in rats exposed to formaldehyde. J Proteome Res 5:1354-1366.

Incoronato M, Garofalo M, Urso L, Romano G, Quintavalle C, Zanca C, et al. 2010. miR-212 increases tumor necrosis factor-related apoptosis-inducing ligand sensitivity in non-small cell lung cancer by targeting the antiapoptotic protein PED. Cancer Res 70:3638-3646.

Iorio MV, Ferracin M, Liu CG, Veronese A, Spizzo R, Sabbioni S, et al. 2005. MicroRNA Gene Expression Deregulation in Human Breast Cancer. Cancer Res 65:7065-7070.

Iorio MV, Piovan C, Croce CM. 2010. Interplay between microRNAs and the epigenetic machinery: an intricate network. Biochim Biophys Acta 1799:694-701.

Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. 2003. Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res 31:e15.

Ivanovska I, Ball AS, Diaz RL, Magnus JF, Kibukawa M, Schelter JM, et al. 2008. MicroRNAs in the miR-106b family regulate p21/CDKN1A and promote cell cycle progression. Mol Cell Biol 28:2167-2174.

Izzotti A, Calin GA, Arrigo P, Steele VE, Croce CM, De Flora S. 2009. Downregulation of microRNA expression in the lungs of rats exposed to cigarette smoke. The FASEB Journal 23:806-812.

Jackson SP, Bartek J. 2009. The DNA-damage response in human biology and disease. Nature 461:1071-1078.

Jaspers I, Flescher E, Chen LC. 1997. Ozone-induced IL-8 expression and transcription factor binding in respiratory epithelial cells. Am J Physiol Lung Cell Mol Physiol 272:L504-511.

Jirouskova M, Shet AS, Johnson GJ. 2007. A guide to murine platelet structure, function, assays, and genetic alterations. J Thromb Haemost 5:661-669.

Jones MR, Quinton LJ, Blahna MT, Neilson JR, Fu S, Ivanov AR, et al. 2009. Zcchc11-dependent uridylation of microRNA directs cytokine expression. Nat Cell Biol 11:1157-1163.

Jorens P, Damme JV, Backer WD, Bossaert L, DeJongh RF, Herman A, et al. 1992. Interleukin 8 (IL-8) in the bronchoalveolar lavage fluid from patients with the adult respiratory distress syndrome (ARDS) and patients at risk for ARDS. Cytokine 4:592-597.

Karin M, Greten FR. 2005. NF-κB: linking inflammation and immunity to cancer development and progression. Nat Rev Immunol 5:749-759.

Katoh M, Katoh M. 2007. WNT signaling pathway and stem cell signaling network. Clin Cancer Res 13:4042-4045.

Kepler GM, Richardson RB, Morgan KT, Kimbell JS. 1998. Computer simulation of inspiratory nasal airflow and inhaled gas uptake in a rhesus monkey. Toxicol Appl Pharmacol 150:1-11.

Kerns WD, Pavkov KL, Donofrio DJ, Gralla EJ, Swenberg JA. 1983. Carcinogenicity of formaldehyde in rats and mice after long-term inhalation exposure. Cancer Res 43:4382-4392.

Kimbell JS, Subramaniam RP, Gross EA, Schlosser PM, Morgan KT. 2001. Dosimetry modeling of inhaled formaldehyde: comparisons of local flux predictions in the rat, monkey, and human nasal passages. Toxicol Sci 64:100-110.

Kitamura H, Cambier S, Somanath S, Barker T, Minagawa S, Markovics J, et al. 2011. Mouse and human lung fibroblasts regulate dendritic cell trafficking, airway inflammation, and fibrosis through integrin αvβ8-mediated activation of TGF-β. J Clin Invest 121:2863-2875.

Kumar S. 2009. Caspase 2 in apoptosis, the DNA damage response and tumour suppression: enigma no more? Nat Rev Cancer 9:897-903.

Kweon MH, Adhami VM, Lee JS, Mukhtar H. 2006. Constitutive overexpression of Nrf2-dependent heme oxygenase-1 in A549 cells contributes to resistance to apoptosis induced by epigallocatechin 3-gallate. J Biol Chem 281:33761-33772.

Lee KH, Chen YL, Yeh SD, Hsiao M, Lin JT, Goan YG, et al. 2009. MicroRNA-330 acts as tumor suppressor and induces apoptosis of prostate cancer cells through E2F1-mediated suppression of Akt phosphorylation. Oncogene 28:3360-3370.

Lee YS, Dutta A. 2007. The tumor suppressor microRNA let-7 represses the HMGA2 oncogene. Genes Dev 21:1025-1030.

Lewis BP, Burge CB, Bartel DP. 2005. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120:15-20.

Li GY, Lee HY, Shin HS, Kim HY, Lim CH, Lee BH. 2007. Identification of gene markers for formaldehyde exposure in humans. Environ Health Perspect 115:1460-1466.

Li T, Chen JX, Fu XP, Yang S, Zhang Z, Chen KhH, et al. 2011. microRNA expression profiling of nasopharyngeal carcinoma. Oncol Rep 25:1353-1363.

Lino-dos-Santos-Franco A, Domingos HV, de Oliveira AP, Breithaupt-Faloppa AC, Peron JP, Bolonheis S, et al. 2010. Differential effects of formaldehyde exposure on the cell influx and vascular permeability in a rat model of allergic lung inflammation. Toxicol Lett 197:211-218.

Liteplo RG, Meek ME. 2003. Inhaled formaldehyde: exposure estimation, hazard characterization, and exposure-response analysis. J Toxicol Environ Health B Crit Rev 6:85-114.

Liu Y, Yuan Y, Li Y, Zhang J, Xiao G, Vodovotz Y, et al. 2009. Interacting neuroendocrine and innate and acquired immune pathways regulate neutrophil mobilization from bone marrow following hemorrhagic shock. J Immunol 182:572-580.

Lovschall H, Eiskjaer M, Arenholt-Bindslev D. 2002. Formaldehyde cytotoxicity in three human cell types assessed in three different assays. Toxicol In Vitro 16:63-69.

Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, et al. 2005. MicroRNA expression profiles classify human cancers. Nature 435:834-838.

Lu K, Boysen G, Gao L, Collins LB, Swenberg JA. 2008. Formaldehyde-induced histone modifications in vitro. Chem Res Toxicol 21:1586-1593.

Lu K, Ye W, Gold A, Ball LM, Swenberg JA. 2009. Formation of S-[1-(N2-deoxyguanosinyl)methyl]glutathione between glutathione and DNA induced by formaldehyde. J Am Chem Soc 131:3414-3415.

Lu K, Moeller B, Doyle-Eisele M, McDonald J, Swenberg JA. 2011. Molecular Dosimetry of N2-Hydroxymethyl-dG DNA Adducts in Rats Exposed to Formaldehyde. Chem Res Toxicol 24:159-161.

Ma L, Teruya-Feldstein J, Weinberg RA. 2007. Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. Nature 449:682-688.

Marcucci G, Radmacher MD, Mrózek K, Bloomfield CD. 2009. MicroRNA expression in acute myeloid leukemia. Curr Hematol Malig Rep 4:83-88.

Marcucci G, Mrózek K, Radmacher MD, Garzon R, Bloomfield CD. 2011. The prognostic and functional role of microRNAs in acute myeloid leukemia. Blood 117:1121-1129.

Marsh GM, Youk AO. 2004. Reevaluation of mortality risks from leukemia in the formaldehyde cohort study of the National Cancer Institute. Regul Toxicol Pharmacol 40:113-124.

Mavrakis KJ, Van Der Meulen J, Wolfe AL, Liu X, Mets E, Taghon T, et al. 2011. A cooperative microRNA-tumor suppressor gene network in acute T-cell lymphoblastic leukemia (T-ALL). Nat Genet 43:673-678.

McClintock D, Zhuo H, Wickersham N, Matthay M, Ware L. 2008. Biomarkers of inflammation, coagulation and fibrinolysis predict mortality in acute lung injury. Critical Care 12:R41.

McGregor D, Bolt H, Cogliano V, Richter-Reichhelm HB. 2006. Formaldehyde and glutaraldehyde and nasal cytotoxicity: case study within the context of the 2006 IPCS Human Framework for the Analysis of a cancer mode of action for humans. Crit Rev Toxicol 36:821-835.

McHale CM, Zhang L, Smith MT. 2012. Current understanding of the mechanism of benzene-induced leukemia in humans: implications for risk assessment. Carcinogenesis 33:240-252.

Merk O, Speit G. 1998. Significance of formaldehyde-induced DNA-protein crosslinks for mutagenesis. Environ Mol Mutagen 32:260-268.

Miko E, Margitai Z, Czimmerer Z, Várkonyi I, Dezso B, Lányi A, et al. 2011. miR-126 inhibits proliferation of small cell lung cancer cells by targeting SLC7A5. FEBS Lett 585:1191-1196.

Ming G, Song H. 2005. Adult Neurogenesis in the Mammalian Central Nervous System. Annu Rev Neurosci 28:223-250.

Moeller B, Lu K, Doyle-Eisele M, McDonald J, Gigliotti A, Swenberg JA. 2011. Determination of N2-Hydroxymethyl-dG Adducts in the Nasal Epithelium and Bone Marrow of Nonhuman Primates Following 13CD2-Formaldehyde Inhalation Exposure. Chem Res Toxicol 24:162-164.

Montag J, Hitt R, Opitz L, Schulz-Schaeffer WJ, Hunsmann G, Motzkus D. 2009. Upregulation of miRNA hsa-miR-342-3p in experimental and idiopathic prion disease. Mol Neurodegener 27:36.

Monticello TM, Morgan KT, Everitt JI, Popp JA. 1989. Effects of formaldehyde gas on the respiratory tract of rhesus monkeys. Am J Pathol 134:515-527.

Monticello TM, Miller FJ, Morgan KT. 1991. Regional increases in rat nasal epithelial cell proliferation following acute and subchronic inhalation of formaldehyde. Toxicol Appl Pharmacol 111:409-421.

Monticello TM, Swenberg JA, Gross EA, Leininger JR, Kimbell JS, Seilkop S, et al. 1996. Correlation of regional and nonlinear formaldehyde-induced nasal cancer with proliferating populations of cells. Cancer Res 56:1012-1022.

Mukherjee D, Coates PJ, Lorimore SA, Wright EG. 2012. The in vivo expression of radiation-induced chromosomal instability has an inflammatory mechanism. Radiat Res 177:18-24.

Murrell W, Féron F, Wetzig A, Cameron N, Splatt K, Bellette B, et al. 2005. Multipotent stem cells from adult olfactory mucosa. Dev Dyn 233:496-515.

NCBI. 2010. GEO: Gene Expression Omnibus. Available: http://www.ncbi.nlm.nih.gov/geo/ [accessed 4 January 2012].

Noda T, Takahashi A, Kondo N, Mori E, Okamoto N, Nakagawa Y, et al. 2011. Repair pathways independent of the Fanconi anemia nuclear core complex play a predominant role in mitigating formaldehyde-induced DNA damage. Biochem Biophys Res Commun 404:206-210.

NTP. 2011. Report on Carcinogens, Twelfth Edition. Research Triangle Park, NC: U.S. Department of Health and Human Services, Public Health Service, National Toxicology Program. Available: http://ntp.niehs.nih.gov/?objectid=03C9AF75-E1BF-FF40-DBA9EC0928DF8B15.

O'Connell RM, Rao DS, Chaudhuri AA, Boldin MP, Taganov KD, Nicoll J, et al. 2008. Sustained expression of microRNA-155 in hematopoietic stem cells causes a myeloproliferative disorder. J Exp Med 205:585-594.

Overton JH, Kimbell JS, Miller FJ. 2001. Dosimetry modeling of inhaled formaldehyde: the human respiratory tract. Toxicol Sci 64:122-134.

Pataer A, Fang B, Yu R, Kagawa S, Hunt KK, McDonnell TJ, et al. 2000. Adenoviral Bak overexpression mediates caspase-dependent tumor killing. Cancer Res 60:788-792.

Persoz C, Achard S, Leleu C, Momas I, Seta N. 2010. An in vitro model to evaluate the inflammatory response after gaseous formaldehyde exposure of lung epithelial cells. Toxicol Lett 195:99-105.

Pinkerton LE, Hein MJ, Stayner LT. 2004. Mortality among a cohort of garment workers exposed to formaldehyde: an update. Occup Environ Med 61:193-200.

Pothof J, Verkaik NS, van IJcken W, Wiemer EA, Ta VT, van der Horst GT, et al. 2009. MicroRNA-mediated gene silencing modulates the UV-induced DNA-damage response. EMBO 28:2090-2099.

Pyatt D, Natelson E, Golden R. 2008. Is inhalation exposure to formaldehyde a biologically plausible cause of lymphohematopoietic malignancies? Regul Toxicol Pharmacol 51:119-133.

Quesenberry PJ, Aliotta JM. 2008. The paradoxical dynamism of marrow stem cells: considerations of stem cells, niches, and microvesicles. Stem Cell Rev 4:137-147.

Quievryn G, Zhitkovich A. 2000. Loss of DNA-protein crosslinks from formaldehyde-exposed cells occurs through spontaneous hydrolysis and an active repair process linked to proteosome function. Carcinogenesis 21:1573-1580.

Rabinowits G, Gerçel-Taylor C, Day JM, Taylor DD, Kloecker GH. 2009. Exosomal microRNA: a diagnostic marker for lung cancer. Clin Lung Cancer 10:42-46.

Rager JE, Lichtveld K, Ebersviller S, Smeester L, Jaspers I, Sexton KG, et al. 2011a. A Toxicogenomic Comparison of Primary and Photochemically Altered Air Pollutant Mixtures. Environ Health Perspect 119:1583-1589.

Rager JE, Smeester L, Jaspers I, Sexton KG, Fry RC. 2011b. Epigenetic changes induced by air toxics: formaldehyde exposure alters miRNA expression profiles in human lung cells. Environ Health Perspect 119:494-500.

Ramamoorthy A, Li L, Gaedigk A, Bradford LD, Benson EA, Flockhart DA, et al. 2012. In silico and in vitro identification of microRNAs that regulate hepatic nuclear factor 4α expression. Drug Metab Dispos 40:726-733.

Ratajczak J, Miekus K, Kucia M, Zhang J, Reca R, Dvorak P, et al. 2006. Embryonic stem cell-derived microvesicles reprogram hematopoietic progenitors: evidence for horizontal transfer of mRNA and protein delivery. Leukemia 20:847-856.

Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP. 2006. GenePattern 2.0. Nat Genet 38:500-501.

Roemer E, Anton HJ, Kindt R. 1993. Cell proliferation in the respiratory tract of the rat after acute inhalation of formaldehyde or acrolein. J Appl Toxicol 13:103-107.

Roisen FJ, Klueber KM, Lu CL, Hatcher LM, Dozier A, Shields CB, et al. 2001. Adult human olfactory stem cells. Brain Res 890:11-22.

Ross WE, Shipley N. 1980. Relationship between DNA damage and survival in formaldehyde-treated mouse cells. Mutat Res 79:277-283.

Roush S, Slack FJ. 2008. The let-7 family of microRNAs. Trends Cell Biol 18:505-516.

Rumchev KB, Spickett JT, Bulsara MK, Phillips MR, Stick SM. 2002. Domestic exposure to formaldehyde significantly increases the risk of asthma in young children. Eur Respir J 20:403-408.

Sawyers CL, Denny CT, Witte ON. 1991. Leukemia and the disruption of normal hematopoiesis. Cell 64:337-350.

Schetter AJ, Heegaard NH, Harris CC. 2010. Inflammation and cancer: interweaving microRNA, free radical, cytokine and p53 pathways. Carcinogenesis 31:37-49.

Schmidt WM, Spiel AO, Jilma B, Wolzt M, Müller M. 2009. In vivo profile of the human leukocyte microRNA response to endotoxemia. Biochem Biophys Res Commun 380:437-441.

Sengupta S, den Boon JA, Chen IH, Newton MA, Stanhope SA, Cheng YJ, et al. 2008. MicroRNA 29c is down-regulated in nasopharyngeal carcinomas, up-regulating mRNAs encoding extracellular matrix proteins. Proc Natl Acad Sci USA 105:5874-5878.

Sexton KG, Jeffries HE, Jang M, Kamens RM, Doyle M, Voicu I, et al. 2004. Photochemical Products in Urban Mixtures Enhance Inflammatory Responses in Lung Cells. Inhalation Toxicol 16:107-114.

Shames DS, Minna JD, Gazdar AF. 2007. DNA methylation in health, disease, and cancer. Curr Mol Med 7:85-102.

Shi XB, Xue L, Yang J, Ma AH, Zhao J, Xu M, et al. 2007. An androgen-regulated miRNA suppresses Bak1 expression and induces androgen-independent growth of prostate cancer cells. Proc Natl Acad Sci USA 104:19983-19988.

Silva J, García V, Zaballos Á, Provencio M, Lombardía L, Almonacid L, et al. 2011. Vesicle-related microRNAs in plasma of nonsmall cell lung cancer patients and correlation with survival. Eur Respir J 37:617-623.

Speit G, Schmid O, Neuss S, Schütz P. 2008. Genotoxic effects of formaldehyde in the human lung cell line A549 and in primary human nasal epithelial cells. Environ Mol Mutagen 49:300-307.

Speit G, Neuss S, Schmid O. 2010. The human lung cell line A549 does not develop adaptive protection against the DNA-damaging action of formaldehyde. Environ Mol Mutagen 51:130-137.

Storey JD. 2003. The positive false discovery rate: a Bayesian interpretation and the q-value. The Annals of Statistics 31:2013-2035.

Swenberg JA, Lu K, Moeller BC, Gao L, Upton PB, Nakamura J, et al. 2011. Endogenous versus exogenous DNA adducts: their role in carcinogenesis, epidemiology, and risk assessment. Toxicol Sci 120:s130-145.

Tanaka M, Oikawa K, Takanashi M, Kudo M, Ohyashiki J, Ohyashiki K, et al. 2009. Down-regulation of miR-92 in human plasma is a novel marker for acute leukemia patients. PLoS One 4:e5532.

Théry C. 2011. Exosomes: secreted vesicles and intercellular communications. F1000 Biol Rep 3:15.

Thomas RS, Allen BC, Nong A, Yang L, Bermudez E, Clewell HJ, et al. 2007. A method to integrate benchmark dose estimates with genomic data to assess the functional effects of chemical exposure. Toxicol Sci 98:240-248.

Thompson CM, Grafström RC. 2011. Considerations for the implausibility of leukemia induction by formaldehyde. Toxicol Sci 120:230-232.

Tuthill RW. 1984. Woodstoves, formaldehyde, and respiratory disease. Am J Epidemiol 120:952-955.

USDL. 2011. Occupational Exposure to Formaldehyde. OSHA Fact Sheet no. 92-27. Available: http://ehs.okstate.edu/training/OSHAFHYD.HTM [accessed Aug 8 2011].

Valadi H, Ekström K, Bossios A, Sjöstrand M, Lee JJ, Lötvall JO. 2007. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. Nat Cell Biol 9:654-659.

Walker SJ. 2008. Human and Macaque Transcriptomes: A Comparison. In: Encyclopedia of Life Sciences (eLS):John Wiley & Sons, DOI: 10.1002/9780470015902.a0020771.

Wang F, Wang XS, Yang GH, Zhai PF, Xiao Z, Xia LY, et al. 2011. miR-29a and miR-142-3p downregulation and diagnostic implication in human acute myeloid leukemia. Mol Biol Rep [Epub ahead of print] doi:10.1007/s11033-011-1026-5.

Wang J-W, Li K, Hellermann G, Lockey RF, Mohapatra S, Mohapatra S. 2012. MIR-150 Suppresses Lung Inflammation in a Mouse Model of Experimental Asthma. World Allergy Organization Journal 5:doi: 10.1097/1001.WOX.0000411771.0000444473.bd.

Wang X. 2008. miRDB: A microRNA target prediction and functional annotation database with a wiki interface. RNA 14:1012-1017.

Whitehead. 2011. TargetScanHuman: Prediction of microRNA targets, release 5.2. Available: http://www.targetscan.org/ [accessed 26 October 2011].

Whitehead. 2012. TargetScanHuman: Prediction of microRNA targets, release 6.2. Available: http://www.targetscan.org/ [accessed 5 July 2012].

WHO. 2001. Chapter 5.8 Formaldehyde. Regional Office for Europe. Copenhagen, Denmark.

Wieslander G, Norbäck D, Björnsson E, Janson C, Boman G. 1997. Asthma and the indoor environment: the significance of emission of formaldehyde and volatile organic compounds from newly painted indoor surfaces. Int Arch Occup Environ Health 69:115-124.

Wong AM, Kong KL, Tsang JW, Kwong DL, Guan XY. 2012. Profiling of Epstein-Barr virus-encoded microRNAs in nasopharyngeal carcinoma reveals potential biomarkers and oncomirs. Cancer 118:698-710.

Yanaihara N, Caplen N, Bowman E, Seike M, Kumamoto K, Yi M, et al. 2006. Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. Cancer Cell 9:189-198.

Yendamuri S, Calin G. 2009. The role of microRNA in human leukemia: a review. Leukemia 23:1257-1263.

Yoshida K, Kondoh G, Matsuda Y, Habu T, Nishimune Y, Morita T. 1998. The mouse RecA-like gene Dmc1 is required for homologous chromosome synapsis during meiosis. Mol Cell 1:707-718.

Yue J, Sheng Y, Orwig KE. 2008. Identification of novel homologous microRNA genes in the rhesus macaque genome. BMC Genomics 9.

Zeller J, Neuss S, Mueller J, Kühner S, Holzmann K, Högel J, et al. 2011. Assessment of genotoxic effects and changes in gene expression in humans exposed to formaldehyde by inhalation under controlled conditions. Mutagenesis 26:555-561.

Zhang L, Steinmaus C, Eastmond DA, Xin XK, Smith MT. 2009. Formaldehyde exposure and leukemia: a new meta-analysis and potential mechanisms. Mutat Res 681:150-168.

Zhang L, Tang X, Rothman N, Vermeulen R, Ji Z, Shen M, et al. 2010. Occupational Exposure to Formaldehyde, Hematotoxicity, and Leukemia-Specific Chromosome Changes in Cultured Myeloid Progenitor Cells. Cancer Epidemiol Biomarkers Prev 19:80-88.

Zhang Y, Liu D, Chen X, Li J, Li L, Bian Z, et al. 2010. Secreted monocytic miR-150 enhances targeted endothelial cell migration. Mol Cell 39:133-144.

Zhou M, Liu Z, Zhao Y, Ding Y, Liu H, Xi Y, et al. 2010. MicroRNA-125b confers the resistance of breast cancer cells to paclitaxel through suppression of pro-apoptotic Bcl-2 antagonist killer 1 (Bak1) expression. J Biol Chem 285:21496-21507.