5-29-2018

The Role of ID3 and PCB153 in the Hyperproliferation and Dysregulation of Lung Endothelial Cells

Mayur Arvind Doke
Florida International University, mdoke001@fiu.edu

Recommended Citation

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

THE ROLE OF ID3 AND PCB153 IN THE HYPERPROLIFERATION AND

DYSREGULATION OF LUNG ENDOTHELIAL CELLS

A dissertation submitted in partial fulfillment of the

requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PUBLIC HEALTH

by

Mayur Doke

2018

To: Dean Tomás R. Guilarte
       Robert Stempel College of Public Health & Social Work

This dissertation, written by Mayur Doke, and entitled, The Role of ID3 and PCB153 in the Hyperproliferation and Dysregulation of Lung Endothelial Cells, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

_____

Juan P. Liuzzi

_____

Changwon Yoo

_____

Alok Deoraj

_____

Deodutta Roy, Co-Major Professor

_____

Quentin Felty, Co-Major Professor

Date of Defense: May 29, 2018

The dissertation of Mayur Doke is approved.

_____

Dean Tomás R. Guilarte
Robert Stempel College of Public Health & Social Work

_____

Andres G. Gil
Vice President for Research and Economic Development
and Dean of the University Graduate School

Florida International University, 2018

DEDICATION

I dedicate this dissertation to my parents Mr. Arvind Doke and Mrs. Ujwala Doke.

Both of you will always be my inspiration.

ACKNOWLEDGMENTS

This dissertation would not have been possible without the support of many people. I would like to thank my mentor Dr. Quentin Felty for giving me an opportunity to work under his immense and incredible guidance. It was his support throughout the four years, which helped me in surviving graduate school. Apart from being a good mentor, he always cared and protected me. Thank you Dr. Felty for everything. I would like to specially thank my co-major professor Dr. Deodutta Roy for his incredible guidance and problem solving solutions in my PhD. He treated me as if was part of his family, he gave me positive constructive critical reviews about my work in lab meetings to shape up my presentation and out of the box thinking skills. I would like to also thank all my committee members Dr. Alok Deoraj, Dr. Changwon Yoo, and Dr. Juan P. Liuzzi for helping me in developing a critical thought process. I am thankful to all of you for all the valuable advice you gave me since the qualifying examination and during all the committee meetings. Doing research without money is not possible. I would like to thank the Department of Environmental Health Sciences and Robert Stempel College of Public Health & Social Work for providing me the graduate assistantship. I also want to thank UGS for the DYF scholarship award. My sincere thanks to Dr. Jayanta Das for teaching me everything in the lab and guiding me throughout these four years. More specifically, I want to acknowledge Dr. Charles Packianathan for helping me whenever I needed to use the lab instruments.

Completion of this humongous task would not have been possible without my lovely and supportive parents, Mr. Arvind Doke and Mrs. Ujwala Doke. They always encouraged

ABSTRACT OF THE DISSERTATION

THE ROLE OF ID3 AND PCB153 IN THE HYPERPROLIFERATION AND

DYSREGULATION OF LUNG ENDOTHELIAL CELLS

by

Mayur Doke

Florida International University, 2018

Miami, Florida

Professor Quentin Felty, Co-Major Professor

Professor Deodutta Roy, Co-Major Professor

Uncontrolled growth of vascular stem cells as a result of endothelial-mesenchymal transition is considered to cause hyper-proliferative vascular remodeling in severe pulmonary arterial hypertension (PAH) patients. Hyperplastic intimal growth is one of the causes of closure of the lumen of pulmonary arterioles. This abnormal vessel remodeling leads to the progressive increase in pressure of the pulmonary arterioles causing severe PAH. Environmental factors, including polychlorinated biphenyls (PCBs), are considered to be involved in hyper-proliferative vascular remodeling because genetic makeup can only explain about 10% of severe PAH cases. PCB involvement in lung toxicity has received attention because (i) they have been reported to accumulate in the lung; (ii) PCBs produce pathological vascular remodeling in the experimental model; high levels of PCBs are found in human lung tissue; and (iii) epidemiological studies show the association between lung toxicity and PCBs; and prevalence of hypertension and elevated concentrations of particularly PCB153. ID3 is an important determinant of mitogen and reactive oxygen species-induced G1→S phase cell cycle progression.

Although there is, still there is a critical gap in understanding the molecular mechanism(s) of pulmonary proliferative vascular remodeling associated with PCB exposure in humans and the role of the transcription regulator ID3. Our overall objective was to investigate ID3 mediated transcriptional reprogramming as a driver of PCB153-induced pathological proliferative vascular remodeling. Stable ectopic expression of ID3 in lung endothelial cells contributed to endothelial-mesenchymal transition (EndMT), cell proliferation, and cell migration. To determine the molecular mechanism by which ID3 contributes to hyper-proliferative endothelial cells, we investigated ID3 transcriptional reprogramming using ChIP-Seq and RNA-Seq technology. We show here for the first time that ID3 is part of a more general mechanism of transcriptional regulation. We identified 19 genes whose RNA was differentially expressed in ID3 overexpressing cells. In summary, our data demonstrated that PCB153 and/or ID3 induces proliferation of lung endothelial cells via transcriptional reprogramming. Discoveries from these findings will lay the necessary groundbreaking work for testing the efficacy of ID3 antagonists for the prevention and treatment of pathological vascular remodeling as well as provide a new paradigm by which PCBs may contribute to lung vascular toxicity.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS AND ACRONYMS

AVM	Arteriovenous malformation
AS	Atherosclerosis
BPA	Bisphenol A
CDC	Center for Disease Control & Prevention
CI	Confidence Interval
E2	Estradiol
ECM	Extracellular Matrix
EDC	Endocrine Disrupting Chemical
GEO	Gene Expression Omnibus
GM	Geometric Mean
GSE	Geometric Standard Error
HLH	Helix-loop-helix
ID1	Inhibitor of Differentiation-1
ID2	Inhibitor of Differentiation-2
ID3	Inhibitor of Differentiation-3
IH	Infantile Hemangioma
IL	Interleukin
IPAH	Idiopathic pulmonary arterial hypertension
LDL	Low-density lipoprotein
MetS	Metabolic Syndrome

NHANES	National Health and Nutrition Examination Survey
LAM	Lymphangioleiomyomatosis
OR	Odds Ratio
PAD	Peripheral Arterial Disease
PDR	Proliferative Diabetic Retinopathy
PCB	Polychlorinated Biphenyl
SNP	Single Nucleotide Polymorphism
TNF	Tumor Neurosis Factor
VEGF	Vascular Endothelial Growth Factor
VEGFR	Vascular Endothelial Growth Factor Receptor
VSMC	Vascular Smooth Muscle Cells
EED	Estrogenic Endocrine Disruptors
BSA	Bovine serum albumin
DMSO	Dimethyl sulfoxide
Dpf	Days post fertilization
Dpi	Days post injection
EC	Vascular endothelial cell
Hpf	Hours post fertilization
Hpi	Hours post injection
HBSS	Hanks' balanced salt solution

ChIP	chromatin immunoprecipitation
ENCODE	encyclopedia of DNA elements
ChIP-Seq	ChIP sequencing
FDR	false discovery rate
MEME	multiple Em for motif elicitation
MACS	model-based analysis of ChIP-Seq
RNA-Seq	RNA sequencing

CHAPTER 1 Literature Review

This chapter has been partially published:

Introduction

The overwhelming increase in the global incidence of pulmonary disease and its associated complications such as insulin resistance, atherosclerosis, obesity, and degenerative disorders including dementia constitutes a serious public health problem. Since some of the endocrine disruptors are lipophilic chemicals with long half-lives, they tend to bioaccumulate in the lipid rich tissue of exposed populations. Elevated exposure to these chemicals may predispose susceptible individuals to metabolic perturbative disorders by increasing the proliferation of the cells. The Inhibitor of DNA Binding/Differentiation-3, a member of the ID family of transcriptional regulators, has been shown to play a role in adipogenesis, angiogenesis and vasculogenesis. Therefore, ID3 may influence obesity and vascular dysfunction in response to environmental factors. This literature review will highlight the current understanding of how ID3 may contribute to complex chronic diseases like pulmonary artery hypertension via metabolic perturbations. Based on the increasing number of reports that suggest chronic exposure and accumulation of endocrine disrupting chemicals (EDCs) within the human body are associated with vascular disorders, we will also consider the impact of these chemicals on ID3. Improved understanding of the ID3 pathways by which exposure to EDCs can potentiate complex chronic diseases in populations with vascular disorders and metabolic disorders (obesity, metabolic syndrome, and glucose intolerance) will likely provide useful knowledge in the prevention and control of complex chronic diseases associated with exposure to environmental pollutants.

Toxic EEDs exposure to humans

Estrogenic endocrine disruptors (EEDs) are compounds that alter hormone production or function including phytoestrogens, heavy metals, and anthropogenic chemicals. More specifically, EEDs include compounds such as genistein, arsenic, DES, PCBs, phthalates, and Bisphenol A (BPA). Polychlorinated biphenyls (PCBs) are man-made aromatic chlorinated hydrocarbons. PCBs have been used commercially since 1930 in various industrial applications like transformers, capacitors, plasticizers, paints, lubricants, and surface coatings because of their distinct physicochemical properties like low electrical conductance, fire resistance, and chemical inertness. PCBs consists of 209 congeners (Ballschmiter & Zell, 1980; Rigaudy, Klesney, & International Union of Pure and Applied Chemistry, 1979), which have been manufactured as complex, and multiple mixtures of congeners by the progressive chlorination of batches of biphenyl until a particular PCB congener was achieved. As PCBs are lipophilic, they tend to bioaccumulate in nature. Moreover, the half-life of PCBs is about 2–6 years (Shiraia & Kissel, 1996). PCBs can able to accumulate in lipid-rich tissues of living organisms (ASTDR, 2000a; Matthews & Dedrick, 1984a). In humans, half of dry weight of the brain consist of lipids which make the brain lipid-rich organ after adipose tissue (Betsholtz, 2015). Moreover, lungs also harbor surfactant lipid composition, mainly di-saturated phosphatidylcholine and cholesterol (Veldhuizen, Nag, Orgeig, & Possmayer, 1998). PCBs being lipophilic in nature, conjugate with lipoproteins in plasma and cell membrane lipids and then eventually accumulate in the tissues of liver and lung (ASTDR, 2000a; Elabbas et al., 2014; European Food Safety Authority, 2005). PCBs are lipophilic and chemically stable in nature so they can pass through any cell membranes and enter

into the nuclear compartment to bind estrogen receptors (Kimbrough & Krouskas, 2003; Safe, 1994). Although Stockholm Convention on Persistent Organic Pollutants has banned the production of PCBs in 2001, researchers have shown many evidences of bioaccumulation of PCBs in the ecosystem and food chain. Low chlorinated (1-4Cl) PCBs are volatilized into air, and highly chlorinated (5–9Cl) PCBs accumulate in food, soil, sediments, and humans (Dobson et al., 1993; Selene & Chou, 2003). Aquatic sediments are main reservoirs of high chlorinated PCBs and concentration of PCBs increases in the air due to evaporation from water to air which is always greater in summer than in winter, because these parameters increase with temperature (Per & Anderss, 1987). Lungs are a highly vascularized organ and receives chemicals from blood through food and air; hence, it is vulnerable to chemical toxicants. Bioaccumulation of PCBs can increase the inflammation in vascular injury which has been shown to promote crosstalk between different organ systems, leading to inflammatory disease pathologies (Achour et al., 2017; He et al., 2017; Lasserre et al., 2009; Wahlang et al., 2017). Moreover, several studies have shown higher levels of PCBs in breast, lung, and liver tissues (ASTDR, 2000a; Chao, Wang, Lin, & Chung, 2006; Elabbas et al., 2014; European Food Safety Authority, 2005; Grimm et al., 2015; Polder et al., 2003). In addition to this, our secondary analysis of The National Health and Nutrition Examination Survey (NHANES) data from 1999-2004, we found that exposure to PCBs increases the risk of PAH in the U.S. population ≥ 20 years old. The concern over adverse health effects from human exposure to PCBs has led to investigations of non-dioxin-like polychlorinated biphenyls (NDL-PCBs), dioxin-like PCBs (DL-PCBs), together with polychlorinated dibenzo-p-dioxins (PCDDs) and

4

polychlorinated dibenzofurans (PCDFs) in foods and the effects of the consumption of these contaminated foods. The Joint FAO/WHO Expert Committee on Food Additives (JECFA) committee has identified seven PCBs called as "indicator PCBs" (JECFA, 2007). These PCBs are divided into two groups of NDL-PCBs and DL-PCBs depending on the site of chlorine substitution on the phenyl rings. Six out of these seven are NDL-PCBs (PCB 28, PCB 52, PCB 101, PCB 138, PCB153, and PCB 180), and one is a DL-PCB (PCB 118). DL-PCBs bind strongly to the aryl hydrocarbon receptor (AhR) due to their planar structure (Bjermo et al., 2013). The higher chlorinated NDL-PCBs congeners are more stable and tend to bioaccumulate in greater concentration within the food- chain and particularly foods of animal origin like fish and fish products, milk, eggs, and meat products (Albro & Fishbein, 1972; ASTDR, 2000a; Elabbas et al., 2014; European Food Safety Authority, 2005; JECFA, 2007). Out of six NDL-PCBs, PCB153 showed higher half-life compared to other five PCBs which is about 14.4 years and has a very slow rate of biotransformation (Chen, Luo, Wong, & Chen, 1982; Matthews & Dedrick, 1984b; Ritter et al., 2011a). It has already been shown by epidemiological and experimental studies the association of PCB153 with metabolic disorders (Mullerova et al., 2008; Wahlang et al., 2013). We and others have shown that PCB153 increases oxidative stress or reactive oxygen species (ROS) that mediate hyper-proliferation of endothelial cells (Felty & Porther, 2008a; Mueller, Baudler, Welzel, Böhm, & Nickenig, 2002). However, ubiquitous exposure to weaker endocrine disruptors such as PCBs and BPA are more likely to occur in humans. NHANES data provide information about the U.S population in demographic, dietary, physical examination, laboratory, and questionnaire data (Centers for Disease Control and Prevention, 2017). The analysis of NHANES survey data

5

showed that there is strong association between exposures to EDCs and risk of PAH. Moreover, meta-analysis of Gene Expression Omnibus (GEO) and NHANES data demonstrated PCB geometric mean concentrations found higher levels in people at risk of PAH than people not at risk of PAH (Assaggaf, 2017). Interestingly, the limits of detection (LODs) of PCB153 showed higher risk of PAH than other PCBs based on Registry to Evaluate Early and Long-term PAH Disease Management (REVEAL Registry) scoring methods (Benza et al., 2012) (Assaggaf, 2017).

EEDs signaling mechanism

Epidemiological studies have reported that women who received hormone replacement therapy (HRT) show higher risk to PAH when compared to women with no HRT (Irey, pathology, & 1973, ; Kleiger, Boxer, Ingham, Chest, & 1976, ; Morse, Horn, & Barst, 1999). Researchers have shown the involvement of estrogen in vascular and pulmonary disorders (Khalil, 2013). 17β-estradiol (E2) signaling occurs through both genomic (nuclear) and non-genomic (extra-nuclear) pathways (Heldring et al., 2007; Shanle & Xu, 2010). Nuclear estrogen receptors consist of ERα and ERβ while membrane estrogen receptors (mERs: GPER, GPR30, ER-X, and Gq-mER) are mostly G protein-coupled receptors (Bondar, Kuo, Hamid, & Micevych, 2009; Filardo, Quinn, Bland, & Frackelton, 2000; Gorosito, Lorenzo, & Cambiasso, 2008; Lagrange, Ronnekleiv, & Kelly, 1997; Qiu et al., 2003; Skipper et al., 1993; Toran-Allerand et al., 2002). Relative binding affinities of ortho, para-DDE, hydroxylated PCB, BPA, and DES have been shown to be significantly weaker than E2 in binding to nuclear ERs and mERs like GPR30 (Bonefeld-Jorgensen, Andersen, Rasmussen, & Vinggaard, 2001; Korach, Sarver, Chae, McLachlan, & McKinney, 1988; Nadal et al., 2000; P. Thomas & Dong,

2006; Watson, Bulayeva, Wozniak, & Finnerty, 2005). Moreover, EEDs have been shown to mimic estrogenic activity and interfere with the endocrine system through these receptor signaling pathways (Diamanti-Kandarakis et al., 2009; Rasier, Parent, Gérard, Lebrethon, & Bourguignon, 2007; Zsarnovszky, Le, Wang, & Belcher, 2005). This has been suggested to be a result of the higher exposure to estrogen (Austin et al., 2009). Moreover, many idiopathic pulmonary arterial hypertension (IPAH) registries reported that females are more prone to PAH than men (Badesch et al., 2010b; Humbert et al., 2006). Estrogen causes multiple effects through ERα and/or ERβ in PAH-associated pathways (Massaro, Clerch, & Massaro, 2007; Patrone et al., 2003).

Pulmonary arterial hypertension

Pulmonary artery hypertension is complex and a multifactorial disease characterized by the multifactorial interaction of genetic, behavioral, and environmental factors. PAH is a progressive disease caused by narrowing or tightening of the pulmonary arteries. The right side of the heart becomes enlarged due to the increased strain of pumping blood through the lungs which leads to the common symptoms of PAH (breathlessness, fatigue, weakness, angina) (Genetic and Rare Diseases Information Center (GARD), 2009). In addition to this, hypertension and elevated triglyceride levels, obesity also constitute the metabolic and biochemical risk factors of vascular dysfunction (Kopeć et al., 2017). PAH is characterized by endothelial damage, increased pulmonary vascular resistance, and right heart failure. The estimated prevalence of PAH is 15–50 cases per million (Peacock, Murphy, McMurray, Caballero, & Stewart, 2007) and has an annual incidence of 1–2 cases per million people in the US and Europe (Gaine & Rubin, 1998). Interestingly, it is

2-4 times more prone in women than men (Badesch et al., 2010a). Researchers have shown molecular level interactions associated with PAH pathogenesis include: inflammation, uncontrolled vascular growth, and formation of obliterative vascular lesions (Irey & Norris, 1973). Epidemiological studies have shown that the incidence of PAH is more common in overweight and obese individuals with a BMI greater than 30 kg/m^2 (Badesch et al., 2010a). PAH mainly occurs in pulmonary artery which carries the blood from the right ventricle of the heart to the lungs for oxygenation. The pulmonary artery is divided into three layers- the outermost layer is known as the tunica externa, is mainly made of collagen fibers. Underneath that, the artery is lined by the tunica media, which is made up of smooth muscle cells. The innermost layer is called as the tunica intima. This layer is mainly made up of endothelial cells. These endothelial cells maintain a balance between vasodilation and vasoconstriction. The hollow internal cavity through which the blood flows is called the lumen. Increased proliferation and progressive migration of endothelial cells further results in intimal fibrosis. In the late stage of disease progression, the formation of plexiform lesions as shown in Figure 1. Progressive reduction of the blood flow leads to high pulmonary arterial pressure defined to be at least 25 mm Hg at rest or 30 mm Hg with exertion measured by right heart catheterization. According to the Dana Point classification system, PAH is classified into five groups. Group 1- Pulmonary arterial hypertension, Idiopathic ("primary pulmonary hypertension") and secondary to systemic disorders.

Group 2- Pulmonary hypertension due to left heart disease (pulmonary venous hypertension)

Group 3- Pulmonary hypertension associated with respiratory disease and/or hypoxia

COPD, Interstitial lung disease, OSA

Group 4- Chronic thromboembolic/embolic pulmonary hypertension

Group 5- Pulmonary hypertension from unclear mechanisms.

The Mayo clinic in Florida did retrospective chart review study of patients from January 1990 to October 2013 with PAH characterized by right heart catheterization (N = 640). The study revealed that 39% of PAH patients were associated with some form of metabolic risk factors like diabetes mellitus and obesity (Ussavarungsi, Thomas, & Burger, 2015). Researchers studied the PAH patients lung tissues sample and microarray data of those samples revealed decreased expression or mutation of bone morphogenetic protein receptor type II (BMPR2) (Du et al., 2003; Hamid et al., 2009; Machado et al., 2001). BMPR2 is associated with TGF-beta superfamily of ligands which regulates signaling of downstream genes like the transcription factor peroxisome proliferator-activated receptor γ (PPARγ) and its downstream effectors, apolipoprotein E (ApoE) and adiponectin (Yang et al., 2002). PPARγ is best known for its role in glucose metabolism and adipogenesis (He et al., 2003; Hevener et al., 2003; Lehrke & Lazar, 2005). Earlier studies have also shown that patients with metabolic syndrome are associated with defective BMPR2 signaling. Together these studies help connect PAH with metabolic syndrome through dysregulated BMPR2 pathway. BMPR2 is known to regulate multiple important pathways like endothelin-1, MCP-1, and eNOS (Farber & Loscalzo, 2004; ITOH et al., 2006).
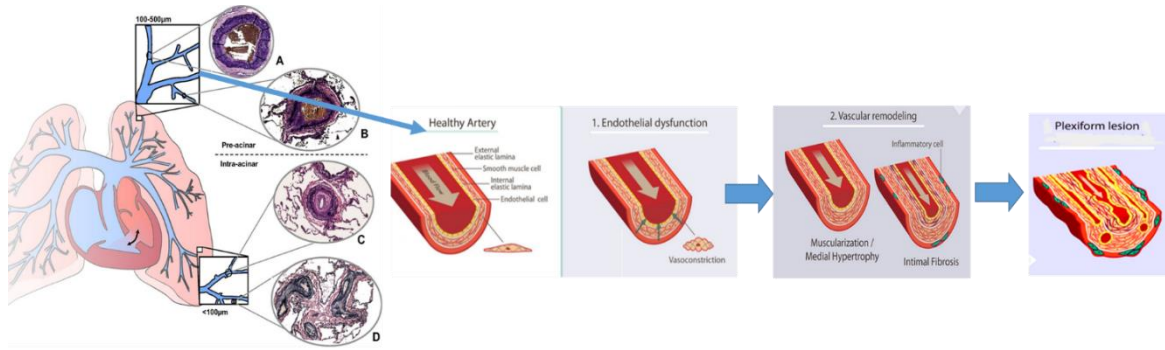
Figure 1. Pathological pulmonary arterial remodeling. (Revised from Gaine S. JAMA 2000;284:3160-3168).

A mechanism of pulmonary arterial hypertension development. Sequence of events that leads from endothelial cell initial proliferation give rise to intimal and medial hypertrophy, neo vascularization of arterioles and complex vascular lesions.

Transcription regulator ID3

Transcriptional regulation by ID3 ultimately functions to increase cell proliferation and preserve multipotency. ID3 mediated gene regulation governing these processes in endothelial stem cells provide a possible explanation for how ID3 promotes vascular perturbations (Kaplan et al., 2015). ID3 protein-protein interactions occur via the helix loop helix (HLH)-motif. ID3 protein interactions block the DNA binding activity of basic HLH (bHLH) transcription factors encoded by the genes TCF3, TCF4, and TCF12. TCF3 gene encodes for E12, E47 proteins. TCF4 gene encodes for E2-2, and TCF12 gene encodes for HEB proteins in humans (Murre, McCaw, & Baltimore, 1989). E12, E47, E2-2 and HEB proteins are class I bHLH proteins that possess the basic DNA binding domain. These E proteins specifically recognize and bind to Ephrussi-box (E-box) sequences (CANNTG) on the DNA (Lazorchak, Jones, & Zhuang, 2005; Sun, 2004). ID3 has been most often reported to interact with proteins encoded by the TCF3

gene (Benezra, Davis, Lockshon, Turner, & Weintraub, 1990). ID3 protein-protein interactions can regulate transcription by E-proteins preventing their binding and subsequent activation of target gene promoters (Benezra et al., 1990). ID3 has frequently been described throughout the literature as an inhibitor of gene expression. For example, ID3 promotes cells to pass through cell cycle checkpoints by inhibiting the expression of cell cycle inhibitor gene p21^{Cip1} as shown in Figure 2 (Taylor et al., 2006a). However, ID3 can also act as a positive transcriptional regulator depending on the cellular context. E-proteins suppress the expression of embryonic genes OCT4, SOX2, and NANOG leading to cell differentiation (Bensellam, Montgomery, Luzuriaga, Chan, & Laybutt, 2015). As shown in Figure 2, ID3 can increase the expression of these embryonic genes by repressing TCF3. We have demonstrated that ectopic overexpression of ID3 in endothelial cells increased OCT4 and SOX2 expression and the acquisition of the vascular stem cell signature: CD133$^+$ VEGFR3$^+$ CD34$^+$ (Das, Voelkel, & Felty, 2015). Based on these findings, ID3 can maintain undifferentiated cells by increasing the expression of embryonic pluripotency factors via repression of TCF3. Since ID3 is a transcription regulator of genes involved in both cell proliferation and stemness, EEDs may facilitate the uncontrolled proliferation of adipocytes through ID3 contributing to pulmonary or metabolic disorders.
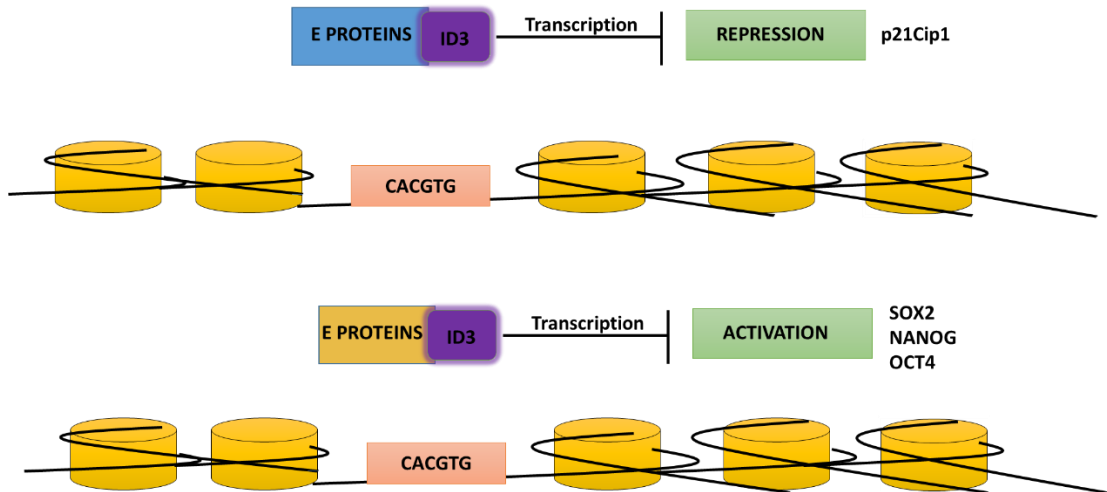
Figure 2 - ID3 regulates a variety of cellular processes

This includes cellular growth, senescence, apoptosis, differentiation, angiogenesis, and neoplastic transformation. This Figure illustrates the ID3 interaction with E proteins. The ID3 protein controls transcription of genes like p21^{Cip1}, OCT4, SOX2, and NANOG by binding to the E-proteins and preventing them from interacting with the E-box sequence on the DNA.

ID3 and EEDs

        The molecular factors that contribute to endocrine disruption have yet to be

fully elucidated. The ID (inhibitor of differentiation) family of small proteins consists of

four proteins (ID1-ID4). ID1 and ID3 have been shown to regulate cell growth, self-

renewal, senescence, angiogenesis, and neurogenesis (Benezra et al., 1990; Lasorella,

Benezra, & Iavarone, 2014; Zebedee & Hara, 2001) (Duncan, DiCicco-Bloom, Xiang,

Benezra, & Chada, 1992; Ellmeier & Weith, 1995; Jen, Manova, & Benezra, 1997).

Depending on the cellular context, ID1 and ID3 have been shown to exhibit overlapping

functions as dual gene knockout combinations have demonstrated redundancy (Lyden et

al., 1999a). ID1 and ID3 have been shown to be co-expressed in early development of

cell cycle progression, angiogenesis, and neurogenesis in the mouse model (Duncan et

al., 1992; Ellmeier & Weith, 1995; Jen et al., 1997). Amongst PCB congeners, PCB153 has been found to be one of the largest contributors to total PCB body burden in humans (Mullerova et al., 2008; Wahlang et al., 2013). We have previously demonstrated that PCB153 modulates ID3 expression and phosphorylation (Das & Felty, 2014c). ID3 is highly expressed in the embryonic tissue and highly proliferating and undifferentiated adult cells (Lyden et al., 1999b). We and others have shown that PCB153 increases oxidative stress or ROS that mediate ID3 expression (Felty & Porther, 2008b; Mueller et al., 2002). Exposure to estrogenic chemicals have been shown to increase ROS in the nucleus in which they modify the surrounding DNA necessary for transcriptional activation of cell growth genes (Okoh, Felty, Parkash, Poppiti, & Roy, 2013; Parkash, Felty, & Roy, 2006c; Perillo et al., 2008). In other words, ROS that we have already shown to be induced by treatment with PCB153 may be involved in ID3 mediated transcription regulation. EEDs have been shown to increase ROS production in adipocytes. Di-(2-ethyl hexyl)phthalate (DEHP) increased ROS in rat adipocytes (Rajesh, Sathish, Srinivasan, Selvaraj, & Balasubramanian, 2013). Mitochondria are a major source of ROS production in mammalian cells (Wang, Zhang, Lu, & Liu, 2015). Although other endogenous ROS sources besides mitochondria such as NADPH oxidase exist, we have shown that estrogenic chemicals increase mitochondrial ROS (Felty et al., 2005). Furthermore, the presence of ERα and ERβ in mitochondria may potentially be targets of EEDs contributing to oxidative stress (Chen, Delannoy, Cooke, & Yager, 2004; Chen, Eshete, Alworth, & Yager, 2004; Monje & Boland, 2001). Although there is evidence linking EEDs exposure to increased mitochondrial ROS, it is unclear whether it is responsible for redox sensitive phosphorylation of ID3 upon exposure to PCB153.

Exposure to EEDs may also exert negative health effects by altering epigenetic marks including DNA methylation and histone acetylation ultimately influencing gene expression in lung tissues. Epigenetic transgenerational effects has been demonstrated in animals exposed to EEDs: BPA and phthalates (Manikkam, Tracey, Guerrero-Bosagna, & Skinner, 2013). Early life exposures to EEDs like DES and PCB153 are known to alter DNA methyltransferase activity (Wu, Zhou, & Ohsako, 2006). Since altered DNA methylation have been found in PCB153 exposed cells (Sales et al., 2013), it is biologically plausible that chromatin modifications including acetylation/deacetylation of histones are another way for ID3 to regulate transcription. ID proteins have been shown to promote acetylation and transcriptional activity by recruiting histone acetyltransferase (HATs): i) CREB-binding protein (CBP), ii) p300 (E1A binding protein p300), and/or P300/CBP-associated factor (PCAF) to the chromatin (Knowell et al., 2013). Moreover, ID proteins interact with another chromatin modifying protein, ZRF1 (Aloia, Gutierrez, Caballero, & Di Croce, 2015). These evidences suggest that ID3 may regulate transcription through interactions with both transcription factors and chromatin modifying proteins.

ID3 and Vascular Diseases

ID3 involvement in vascular disease has been studied together with the lipoxygenase (12/15-LO) which is known to generate pro-inflammatory changes in blood vessels that precede the development of atherosclerosis (Chakrabarti, Cole, Wen, Keller, & Nadler, 2009). 12/15-LO is an important mediator of vascular smooth muscle cells (VSMC) growth and its growth-promoting effects were shown to be mediated by ID3 transcription (Butany, Ahluwalia, Nair, & David, 2004). Increased expression of 12/15-

LO in the vessel wall enhanced ID3-dependent cell proliferation, fibronectin deposition, and neointimal formation. Population-based studies have found SNP (single nucleotide polymorphisms) rs11574 in the coding region of the human ID3 gene associated with subclinical atherosclerosis in the Diabetes Heart Study (Doran et al., 2010). ID3 SNP rs11574 showed a significant association of coronary artery disease for Caucasians and to a lesser extent African Americans and Hispanics (Manichaikul et al., 2014). Ectopic expression of ID3 in VSMC regulates the cell cycle (Yang et al., 2013). ID3 has also been shown to play a complex role with atherosclerosis. ID3 expression is increased by hyperlipidemia and oxidized LDL (Taylor et al., 2006b). ID3 regulates angiotensin II and carotid intima-media thickness. Angiotensin II promotes hyperplasia through upregulating ID3 (Phillip & Iii, 2009). The ID3 SNP could be a potential loss of function mutation if it inhibits the functioning of E proteins, thus being an atheroprotective factor. As shown in Figure 3, ID3 may impact vascular cell dysfunction leading to intimal lesions. ID3 stimulates visceral adipose vascular endothelial growth factor A (VEGFA) expression, depot expansion, and microvascular blood volume (Cutchins et al., 2012). ID3 promotes angiogenesis in high fat diet - (HFD -) induced visceral adiposity (Cutchins et al., 2012). ID3 KO shows a protective effect from HFD induced visceral fat depot expansion. Furthermore, HFD-induced VEGFA expression in visceral adipose tissue was completely abolished by loss of ID3. BMP9 induces both ID1 and ID3 which are necessary for induction of Ephrin B2 (Kim, Peacock, George, & Hughes, 2012). A summary in Figure 4 shows an ID3 signaling pathway involved in vascular malformations. Similar to adipocytes, ID3 induced VEGF overexpression and Ephrin B2 downregulation may be involved in endothelial cells.

15

Epigenetic imprinting of vascular progenitor or endothelial stem cells by ID3 in individuals exposed to EEDs is another plausible mechanism based on the following evidence. Specifically, we postulate that EEDs can induce ROS-mediated ID3 phosphorylation and acetylation, histone acetylation, and DNA base oxidation collectively that control expression of ID3 target genes involved in vascular cells programming. This in turn controls the fate and epigenetic footprints of vascular progenitor cells. Reactive oxygen species like H_2O_2 are highly diffusible molecules. In addition to affecting ID3 signaling pathways, ROS can also facilitate histone acetylation and oxidize nuclear DNA resulting in chromatin modification. These modifications are significant because transcription in eukaryotes occurs in the context of DNA, packaged into chromatin. The basic unit of chromatin is the nucleosome, in which DNA is wrapped around the core histones H2A, H2B, H3, and H4. Acetylation of lysine in the histone tails can facilitate the opening of repressive chromatin structures in promoter regions to provide access for the transcription regulator ID3. In support of our concept that both histone acetylation and ROS-mediated DNA oxidation control the transcription of EED-induced genes, we and others have shown that E2 and PCB153-induced ROS in the nucleus, particularly H_2O_2, modify the surrounding DNA (Okoh, Deoraj, & Roy, 2011; Parkash, Felty, & Roy, 2006a; Perillo et al., 2008). It has been recently shown that DNA oxidation through recruiting 8-oxoguanine–DNA glycosylase triggers chromatin and DNA conformational changes that are essential for estrogen-mediated transcription of genes (Perillo et al., 2008). ROS generating agents and inflammation have been shown to modulate chromatin-bound hSirT1 deacetylase activity on the promoters of several genes (Adcock, Cosio, Tsaprouni, Barnes, & Ito, 2005; Brunet et al., 2004; Moodie et al., 2004;

Pediconi et al., 2009). Taken together, EEDs through induction of ROS may increase histone acetylations by post-translational activation of acetylases and oxidation of DNA bases, which are necessary for ID3-mediated transcription regulation of target genes involved in vascular and metabolic complications.

Given the significant role of ID3 in stemness, it is possible that EEDs exposure may contribute to an increase in vascular progenitor cells. Epigenetic changes mediated by ID3 in the stem cells may ultimately increase the total number of endothelial cells that can be produced by an individual. Subsequent environmental exposures of these susceptible individuals who have a higher number of endothelial stem cells, to begin with, may potentially lead to pathological vascular remodeling including plexiform lesions.
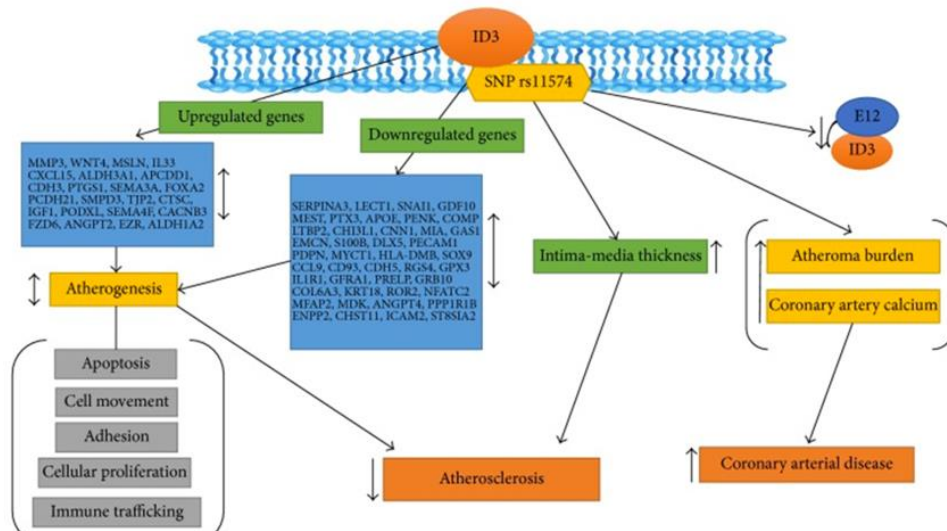
Figure 3- ID3 molecular mechanism and vascular diseases.

Summary illustration of ID3 signaling involved in vascular disease pathogenesis. ID3 signaling can lead to neointimal lesions or vascular remodeling by transcriptional regulation of the target genes described above (up- and down regulated genes by ID3).
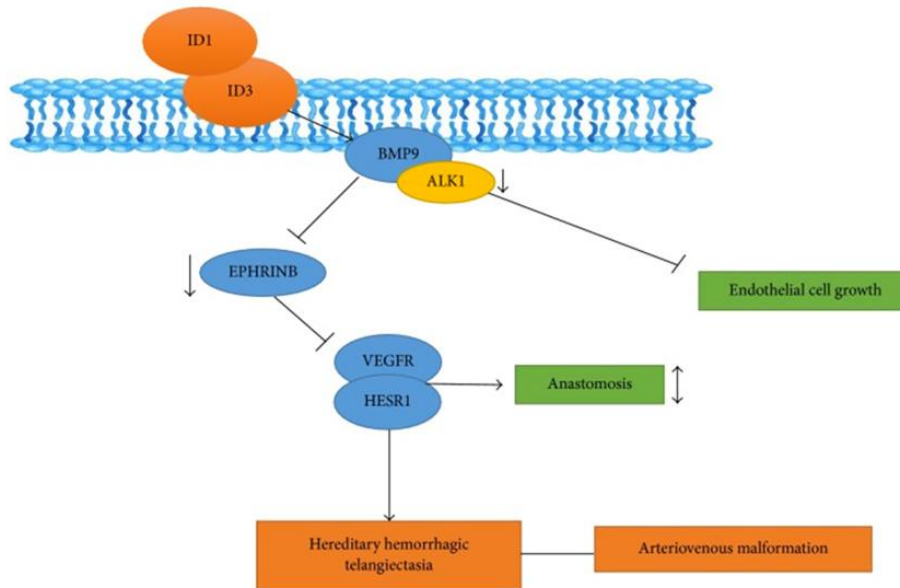
Figure 4 - ID3 involvement in arteriovenous malformation.

Summary of experimental studies showing impact of ID3 in vascular cell dysfunction. ID3 involved in VEGFR higher expression by inducing BMP9 and downregulation of Ephrin B.

Interaction of ID3 and EEDs

In order to investigate how environmental exposures, affect human health at genetic and protein level, we used the Comparative Toxicogenomic Database (CTD). CTD is a public website, and research tool that consists of scientific data illustrating chemical-gene interactions and chemical-disease associations. CTD includes *in vitro* and *in vivo* data studies describing relationships between chemicals, genes, and diseases. The database was used to manually curates information about EEDs–gene/protein interactions, EEDs–disease and gene-disease associations. We used these public databases to investigate the role of ID3, especially in various metabolic pathways. This tool can be used to decipher gene-environment or gene-EEDs interactions involved in the generation of vascular diseases.

We initially selected PCB as a chemical and chose hypertension as the disease. Our search results showed 29 genes associated with both the hypertension and PCBs. Because ID3 is our candidate gene, we created a list of genes related to ID3 and PCBs. We established that 38 genes are associated with this group in Table 1. Furthermore, we show an interaction between these genes, we additionally inputted them into STRING, a database of recognized and predicted protein-protein interactions. The interactions contain direct (physical) and indirect (functional) associations which stem from computational prediction, interactions aggregated from other (primary) databases, and from knowledge transfer between organisms. As seen in Figure 5 STRING provides a network of these proteins and furthermore provides a pathway description for the mutually represented proteins (Szklarczyk et al., 2015a). We created Table 2 with the help of STRING network, and it shows that the involvement of the ID3 protein in the various pathway.

PCBs and Hypertension related genes	ID3 and PCBs related genes
ABCA3	ATF3
BMPR2	CDK2
CAV1	CNOT3
CBLN2	CSK
CCL1	E2F4
CCL2	ELSPBP1
CXCL8	EMX1
EDN1	FHL2
EIF2AK4	GATA4
FBLN5	GNB2
FOXO1	ID3
GUCY1A1	MLX
HDAC1	MYOD1
HDAC4	NPHP4
HDAC5	PUF60
HIF1A	RND1
HTR2B	SMURF2
KCNMA1	SREBF1
LOX	TCF3
NOS3	TCF4
NPPB	UNK
PRKG1	ZNF3
RETNLA	ZNF408
SLC31A1	ZNF626
SLC6A4	FAM74A1
SMAD9	ID4
SOD2	IKBKG
TGFA	NKX2-5
TNF	PAX5
	PRDM14
	SAP30
	TCF12
	VHL
	CBFA2T2
	COPS5
	COPS6
	COPS7A
	ELOC

Table 1- Genes involved ID3 -PCBs and PCBs -hypertension.

CTD network curated information about genes involved in PCBs–ID3 protein interactions, PCBs–hypertension and ID3- hypertension associations.
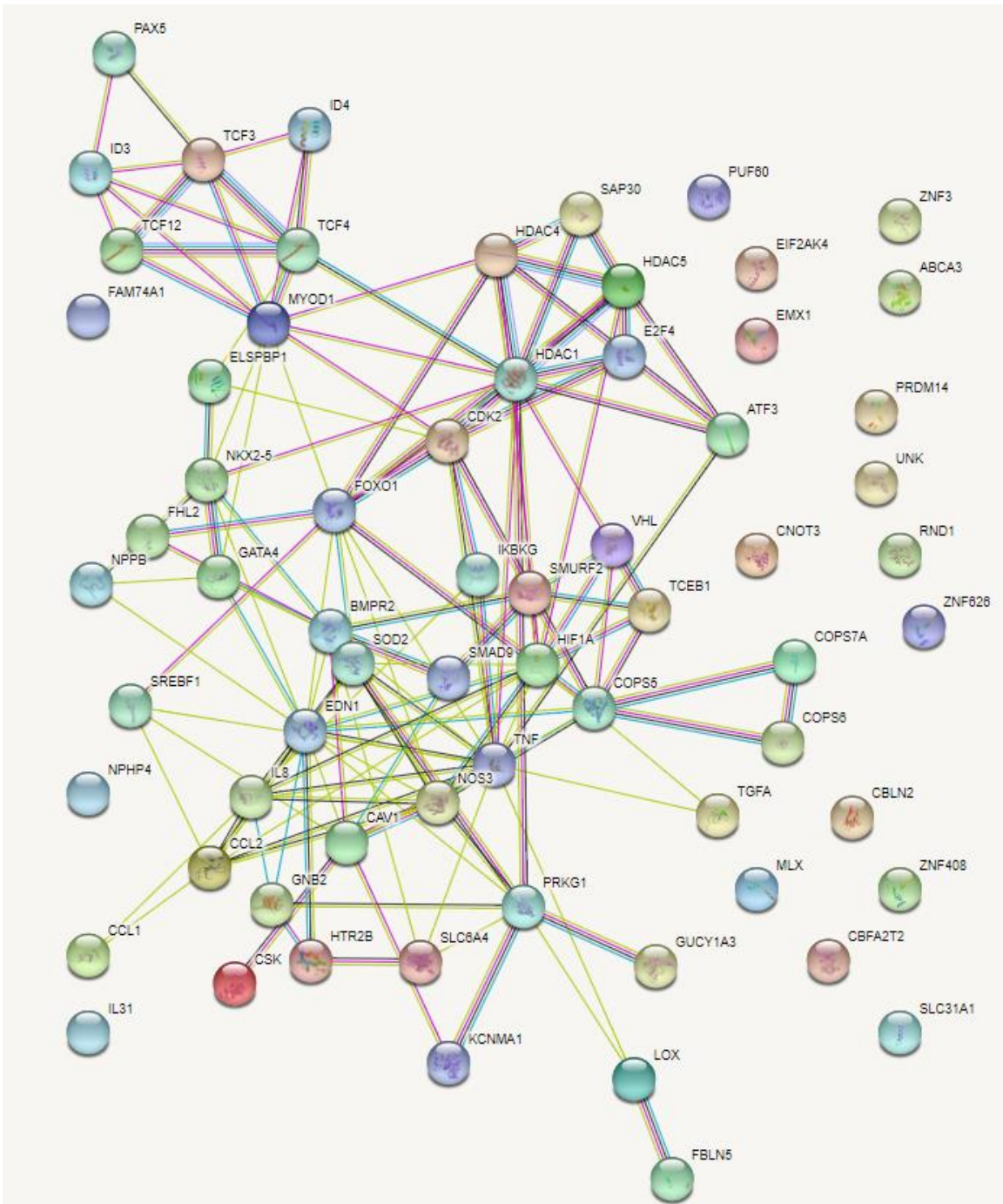
Figure 5 - STRING network illustration of proteins involved in ID3 - PCBs and PCBs -hypertension.

number of nodes: 67	expected number of edges: 61
number of edges: 142	PPI enrichment p-value: < 1.0e-16
average node degree: 4.24	*your network has significantly more interactions*
avg. local clustering coefficient: 0.403	*than expected (what does that mean?)*

Biological Process (GO)

pathway ID	pathway description	count in gene set	false discovery rate
GO:0050880	regulation of blood vessel size	10	1.04e-10
GO:0042493	response to drug	16	6.04e-10
GO:0009891	positive regulation of biosynthetic process	27	7.28e-10
GO:0006357	regulation of transcription from RNA polymerase II promoter	26	1.02e-09
GO:0045935	positive regulation of nucleobase-containing compound metabolic process	26	1.02e-09

Molecular Function (GO)

pathway ID	pathway description	count in gene set	false discovery rate
GO:0000989	transcription factor activity, transcription factor binding	17	3.05e-09
GO:0005515	protein binding	39	3.95e-08
GO:0008134	transcription factor binding	13	3.37e-07
GO:0003712	transcription cofactor activity	14	7.58e-07
GO:0000978	RNA polymerase II core promoter proximal region sequence-specific DNA binding	10	1.32e-05

KEGG Pathways

pathway ID	pathway description	count in gene set	false discovery rate
04350	TGF-beta signaling pathway	7	1.69e-06
04022	cGMP-PKG signaling pathway	6	0.0012
05200	Pathways in cancer	8	0.0012
04066	HIF-1 signaling pathway	5	0.00159
04621	NOD-like receptor signaling pathway	4	0.00195

Table 2- Summary of functional enrichments involved in the genes related to ID3 - PCBs and PCBs - hypertension
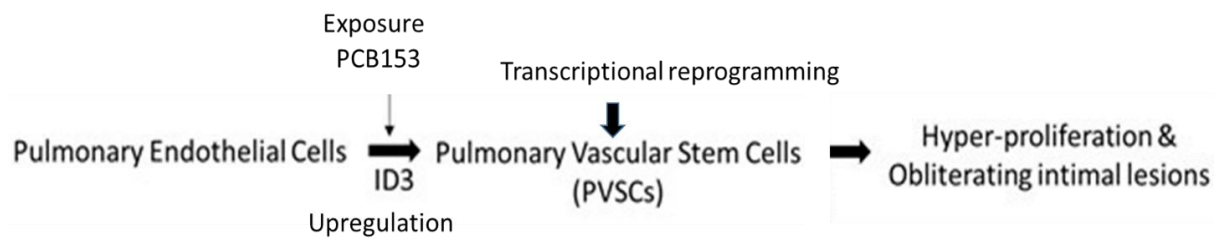
Figure 6 - Central Hypothesis of study

The diagram represents central hypothesis, which is based on background information and earlier studies. We hypothesize that the exposure of PCB153 increases ID3 expression. ID3 may be involved in vascular reprogramming and hyperproliferation of endothelial cells.

ID3 has been shown to promote vascular disorders in experimental models. Studies have reported associations between endothelial or vascular dysfunction in PAH and exposure to EEDs: BPA, DES, and PCBs. Based on the evidence discussed in this review, elevated exposure to EEDs or unopposed increase in the body burden of estrogens may increase the expression of the transcription regulator ID3. We propose that ID3 may be an additional molecular risk factor for vascular perturbations or PAH from environmental exposure to estrogenic chemicals. Emerging evidence demonstrated that ID3 can regulate mitochondrial function and morphology associated with changes in the expression of electron transport chain complex components and transcription factor A, mitochondrial (TFAM) (Bensellam, Montgomery, Luzuriaga, Chan, & Laybutt, 2015b). We have previously shown that ER independent mitochondrial ROS signaling contributes to the growth of cells treated with 17β-estradiol (Felty et al., 2005; Felty & Roy, 2005; Parkash, Felty, & Roy, 2006b). Therefore, ID3 mediated vascular programming of mitochondria may be dysregulated by exposure to EEDs and increase susceptibility to PAH or vascular disorder. In addition, the ID3 dependent production of cytokines and recruitment of macrophages in lung tissue are suggested to play an important role in the inflammatory process to enhance susceptibility to complications.

In closing, we have systematically reviewed the existing evidence to illustrate the association between ID3, EEDs, & PAH. Furthermore, we extended this understanding of how ID3 and vascular perturbations by environmental factors such as EEDs can increase the risk of PAH. Research is warranted to better define the influence of EEDs and ID3 alongside gene-environment interactions on PAH. A better understanding of how ID3 & EEDs affect the risk of PAH may open up new avenues for prevention and treatment of

diseases. We have shown previously that PCB153 and natural estrogen 17β-estradiol increase ID3 expression in endothelial cells. Based on our findings, we postulate that ID3 is a molecular target of EEDs like PCB153 in lung tissue and a better understanding of this relationship may help to explain how EEDs can lead to the transcriptional programming of deviant endothelial cells. Thus, the purpose of this dissertation will be to discuss the current understanding of ID3 in vascular reprogramming and to consider the potential for PCB153 to influence susceptibility to pulmonary or vascular disorders via ID3 signaling as shown in Figure 6.

References-

Achour, A., Derouiche, A., Barhoumi, B., Kort, B., Cherif, D., Bouabdallah, S., Tebourbi, O. (2017). Organochlorine pesticides and polychlorinated biphenyls in human adipose tissue from northern Tunisia: Current extent of contamination and contributions of socio-demographic characteristics and dietary habits. *Environmental Research*, *156*, 635–643. http://doi.org/10.1016/j.envres.2017.04.021

Adcock, I. M., Cosio, B., Tsaprouni, L., Barnes, P. J., & Ito, K. (2005). Redox Regulation of Histone Deacetylases and Glucocorticoid-Mediated Inhibition of the Inflammatory Response. *Antioxidants & Redox Signaling*, *7*(1–2), 144–152. http://doi.org/10.1089/ars.2005.7.144

Albro, P. W., & Fishbein, L. (1972). Intestinal absorption of polychlorinated biphenyls in rats. *Bulletin of Environmental Contamination and Toxicology*, *8*(1), 26–31. http://doi.org/10.1007/BF01684500

Aloia, L., Gutierrez, A., Caballero, J. M., & Di Croce, L. (2015). Direct interaction between Id1 and Zrf1 controls neural differentiation of embryonic stem cells. *EMBO Reports*, *16*(1), 63–70. http://doi.org/10.15252/embr.201439560

Assaggaf, H., & Felty, Q. (2017, April 2). Gender, estrogen, and obliterative lesions in the lung. *International Journal of Endocrinology. Hindawi.* https://doi.org/10.1155/2017/8475701.

ASTDR. (2000). Toxicological Profile for Polychlorinated Biphenyls (PCBs). *Agency for Toxic Substances and Disease Registry*, (November), 1–948. Retrieved from https://www.atsdr.cdc.gov/toxprofiles/tp17.pdf

Austin, E. D., Cogan, J. D., West, J. D., Hedges, L. K., Hamid, R., Dawson, E. P., Phillips, J. A. (2009). Alterations in oestrogen metabolism: implications for higher penetrance of familial pulmonary arterial hypertension in females. *The European Respiratory Journal*, *34*(5), 1093–1099. http://doi.org/10.1183/09031936.00010409

Badesch, D. B., Raskob, G. E., Elliott, C. G., Krichman, A. M., Farber, H. W., Frost, A. E., McGoon, M. D. (2010a). Pulmonary Arterial Hypertension. *Chest*, *137*(2), 376–387. http://doi.org/10.1378/chest.09-1140

Badesch, D. B., Raskob, G. E., Elliott, C. G., Krichman, A. M., Farber, H. W., Frost, A. E., McGoon, M. D. (2010b). Pulmonary arterial hypertension: baseline characteristics from the REVEAL Registry. *Chest*, *137*(2), 376–387. http://doi.org/10.1378/chest.09-1140

Ballschmiter, K., & Zell, M. (1980). Fmenius Zeitschrift fur Analysis of Polychlorinated Biphenyls (PCB) by Glass Capillary Gas Chromatography Composition of Technical Aroclor-and Clophen-PCB Mixtures. *Fresenius Z. Anal. Chem*, *302*, 20–31. Retrieved from https://link.springer.com/content/pdf/10.1007%2FBF00469758.pdf

Bastos Sales, L., Kamstra, J. H., Cenijn, P. H., van Rijt, L. S., Hamers, T., & Legler, J. (2013). Effects of endocrine disrupting chemicals on *in vitro* global DNA methylation and adipocyte differentiation. *Toxicology in vitro*, *27*(6), 1634–1643. http://doi.org/10.1016/j.tiv.2013.04.005

Benezra, R., Davis, R. L., Lockshon, D., Turner, D. L., & Weintraub, H. (1990). The protein Id: A negative regulator of helix-loop-helix DNA binding proteins. *Cell*, *61*(1), 49–59. http://doi.org/10.1016/0092-8674(90)90214-Y

Bensellam, M., Montgomery, M. K., Luzuriaga, J., Chan, J. Y., & Laybutt, D. R. (2015). Inhibitor of differentiation proteins protect against oxidative stress by regulating the antioxidant–mitochondrial response in mouse beta cells. *Diabetologia*, *58*(4), 758–770. http://doi.org/10.1007/s00125-015-3503-1

Benza, R. L., Gomberg-Maitland, M., Miller, D. P., Frost, A., Frantz, R. P., Foreman, A. J., McGoon, M. D. (2012). The REVEAL Registry risk score calculator in patients newly diagnosed with pulmonary arterial hypertension. *Chest*, *141*(2), 354–362. http://doi.org/10.1378/chest.11-0676

Betsholtz, C. (2015). Lipid transport and human brain development. *Nature Genetics*, *47*(7), 699–701. http://doi.org/10.1038/ng.3348

Bjermo, H., Darnerud, O., Lignell, S., Pearson, M., Rantakokko, P., Nälsén, C., … Glynn, A. (2013). Fish intake and breastfeeding time are associated with serum concentrations of organochlorines in a Swedish population. *Environment International*, *51*, 88–96. http://doi.org/10.1016/j.envint.2012.10.010

Bondar, G., Kuo, J., Hamid, N., & Micevych, P. (2009). Estradiol-Induced Estrogen Receptor- Trafficking. *Journal of Neuroscience*, *29*(48), 15323–15330. http://doi.org/10.1523/JNEUROSCI.2107-09.2009

Bonefeld-Jorgensen, E. C., Andersen, H. R., Rasmussen, T. H., & Vinggaard, A. M. (2001). Effect of highly bioaccumulated polychlorinated biphenyl congeners on estrogen and androgen receptor activity. *Toxicology*, *158*(3), 141–153. http://doi.org/10.1016/S0300-483X(00)00368-1

Brunet, A., Sweeney, L. B., Sturgill, J. F., Chua, K. F., Greer, P. L., Lin, Y., Greenberg, M. E. (2004). Stress-Dependent Regulation of FOXO Transcription Factors by the SIRT1 Deacetylase. *Science*, *303*(5666), 2011–2015. http://doi.org/10.1126/science.1094637

Butany, J., Ahluwalia, M. S., Nair, V., & David, T. E. (2004). Cryopreserved pulmonary homograft: Postimplant changes. *Cardiovascular Pathology*, *13*(1), 59–61. http://doi.org/10.1016/S1054-8807(03)00092-9

CDC Centers for Disease Control and Prevention. (2017). Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey Data. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Preventio.

Chakrabarti, S. K., Cole, B. K., Wen, Y., Keller, S. R., & Nadler, J. L. (2009). 12/15-Lipoxygenase products induce inflammation and impair insulin signaling in 3t3-l1 adipocytes. *Obesity*, *17*(9), 1657–1663. http://doi.org/10.1038/oby.2009.192

Chao, H.-R., Wang, S.-L., Lin, T.-C., & Chung, X.-H. (2006). Levels of organochlorine pesticides in human milk from central Taiwan. *Chemosphere*, *62*(11), 1774–1785. http://doi.org/10.1016/j.chemosphere.2005.07.036

Chen, J. Q., Delannoy, M., Cooke, C., & Yager, J. D. (2004). Mitochondrial localization of ER and ER in human MCF7 cells. *AJP: Endocrinology and Metabolism*, *286*(6), E1011–E1022. http://doi.org/10.1152/ajpendo.00508.2003

Chen, J. Q., Eshete, M., Alworth, W. L., & Yager, J. D. (2004). Binding of MCF-7 cell mitochondrial proteins and recombinant human estrogen receptors ? and ? to human mitochondrial dna estrogen response elements. *Journal of Cellular Biochemistry*, *93*(2), 358–373. http://doi.org/10.1002/jcb.20178

Chen, P. H., Luo, M. L., Wong, C. K., & Chen, C. J. (1982). Comparative rates of elimination of some individual polychlorinated biphenyls from the blood of PCB-poisoned patients in Taiwan. *Food and Chemical Toxicology*, *20*(4), 417–425. http://doi.org/10.1016/S0278-6915(82)80107-5

Commission, T. H. E., The, O. F., Communities, E., & Facts, T. H. E. (1999). Official Journal of the European Communities. *Analytical Proceedings*, *21*(6), 196. http://doi.org/10.1039/ap9842100196

Cutchins, A., Harmon, D. B., Kirby, J. L., Doran, A. C., Oldham, S. N., Skaflen, M., … McNamara, C. A. (2012). Inhibitor of Differentiation-3 Mediates High Fat Diet-Induced Visceral Fat Expansion. *Arteriosclerosis, Thrombosis, and Vascular Biology*, *32*(2), 317–324. http://doi.org/10.1161/ATVBAHA.111.234856

Das, J. K., & Felty, Q. (2014). PCB153-induced overexpression of ID3 contributes to the development of microvascular lesions. *PLoS ONE*, *9*(8). http://doi.org/10.1371/journal.pone.0104159

Das, J. K., Voelkel, N. F., & Felty, Q. (2015). ID3 contributes to the acquisition of molecular stem cell-like signature in microvascular endothelial cells: Its implication for understanding microvascular diseases. *Microvascular Research*, *98*, 126–138. http://doi.org/10.1016/j.mvr.2015.01.006

Diamanti-Kandarakis, E., Bourguignon, J.-P., Giudice, L. C., Hauser, R., Prins, G. S., Soto, A. M., … Gore, A. C. (2009). Endocrine-disrupting chemicals: an Endocrine Society scientific statement. *Endocrine Reviews*, *30*(4), 293–342. http://doi.org/10.1210/er.2009-0002

Dobson, S. (Stuart), Esch, G. J. van., International Program on Chemical Safety., United Nations Environment Programme., International Labour Organisation., & World Health Organization. (1993). *Polychlorinated biphenyls and terphenyls*. World Health Organization.

Doran, A. C., Lehtinen, A. B., Meller, N., Lipinski, M. J., Slayton, R. P., Oldham, S. N., McNamara, C. A. (2010). Id3 Is a novel atheroprotective factor containing a functionally significant single-nucleotide polymorphism associated with intima-media thickness in humans. *Circulation Research*, *106*(7), 1303–1311. http://doi.org/10.1161/CIRCRESAHA.109.210294

Du, L., Sullivan, C. C., Chu, D., Cho, A. J., Kido, M., Wolf, P. L., Thistlethwaite, P. A. (2003). Signaling Molecules in Nonfamilial Pulmonary Hypertension. *New England Journal of Medicine*, *348*(6), 500–509. http://doi.org/10.1056/NEJMoa021650

Duncan, M., DiCicco-Bloom, E. M., Xiang, X., Benezra, R., & Chada, K. (1992). The gene for the helix-loop-helix protein, Id, is specifically expressed in neural precursors. *Developmental Biology*, *154*(1), 1–10. http://doi.org/10.1016/0012-1606(92)90042-F

Elabbas, L. E., Esteban, J., Barber, X., Hamscher, G., Nau, H., Bowers, W. J., Håkansson, H. (2014). In Utero and Lactational Exposure to a Mixture of Environmental Contaminants Detected in Canadian Arctic Human Populations Alters Retinoid Levels in Rat Offspring with Low Margins of Exposure. *Journal of Toxicology and Environmental Health, Part A*, *77*(5), 223–245. http://doi.org/10.1080/15287394.2013.861776

Ellmeier, W., & Weith, A. (1995). Expression of the helix loop helix gene Id3 during murine embryonic development. *Developmental Dynamics*, *203*(2), 163–173. http://doi.org/10.1002/aja.1002030205

European Food Safety Authority. (2005). Opinion of the Scientific Panel on Contaminants in the Food Chain on a Request From the Commission Related To the Presence of Non Dioxin-Like Polychlorinated Biphenyls (Pcb) in Feed and Food. *Efsa Journal*, *284*(11), 1–137. http://doi.org/10.2903/j.efsa.2005.284

Farber, H. W., & Loscalzo, J. (2004). Pulmonary Arterial Hypertension. *New England Journal of Medicine*, *351*(16), 1655–1665. http://doi.org/10.1056/NEJMra035488

Felty, Q., & Porther, N. (2008a). Estrogen-induced redox sensitive Id3 signaling controls the growth of vascular cells. *Atherosclerosis*, *198*(1), 12–21. http://doi.org/10.1016/j.atherosclerosis.2007.12.048

Felty, Q., & Porther, N. (2008b). Estrogen-induced redox sensitive Id3 signaling controls the growth of vascular cells. *Atherosclerosis*, *198*(1), 12–21. http://doi.org/10.1016/j.atherosclerosis.2007.12.048

Felty, Q., & Roy, D. (2005). Estrogen, mitochondria, and growth of cancer and non-cancer cells. *Journal of Carcinogenesis*, *4*(1), 1. http://doi.org/10.1186/1477-3163-4-1

Felty, Q., Xiong, W.-C., Sun, D., Sarkar, S., Singh, K. P., Parkash, J., & Roy, D. (2005). Estrogen-induced mitochondrial reactive oxygen species as signal-transducing messengers. *Biochemistry*, *44*(18), 6900–9. http://doi.org/10.1021/bi047629p

Filardo, E. J., Quinn, J. A., Bland, K. I., & Frackelton, A. R. (2000). Estrogen-Induced Activation of Erk-1 and Erk-2 Requires the G Protein-Coupled Receptor Homolog, GPR30, and Occurs via *Trans* -Activation of the Epidermal Growth Factor Receptor through Release of HB-EGF. *Molecular Endocrinology*, *14*(10), 1649–1660. http://doi.org/10.1210/mend.14.10.0532

Gaine, S. P., & Rubin, L. J. (1998). Primary pulmonary hypertension. *The Lancet*, *352*(9129), 719–725. http://doi.org/10.1016/S0140-6736(98)02111-4

Gorosito, S. V., Lorenzo, A. G., & Cambiasso, M. J. (2008). Estrogen receptor α is expressed on the cell-surface of embryonic hypothalamic neurons. *Neuroscience*, *154*(4), 1173–1177. http://doi.org/10.1016/j.neuroscience.2008.05.001

Grimm, F., Hu, D., Kania-Korwel, I., Lehmler, H., Ludewig, G., Hornbuckle, K., … Robertson, L. W. (2015). Metabolism and metabolites of polychlorinated biphenyls (PCBs) HHS Public Access. *Crit Rev Toxicol*, *45*(3), 245–272. http://doi.org/10.3109/10408444.2014.999365

Hamid, R., Cogan, J. D., Hedges, L. K., Austin, E., Phillips, J. A., Newman, J. H., & Loyd, J. E. (2009). Penetrance of pulmonary arterial hypertension is modulated by the expression of normal *BMPR2* allele. *Human Mutation*, *30*(4), 649–654. http://doi.org/10.1002/humu.20922

He, W., Barak, Y., Hevener, A., Olson, P., Liao, D., Le, J., Evans, R. M. (2003). Adipose-specific peroxisome proliferator-activated receptor knockout causes insulin resistance in fat and liver but not in muscle. *Proceedings of the National Academy of Sciences*, *100*(26), 15712–15717. http://doi.org/10.1073/pnas.2536828100

He, Y., Peng, L., Huang, Y., Peng, X., Zheng, S., Liu, C., & Wu, K. (2017). Association of breast adipose tissue levels of polychlorinated biphenyls and breast cancer development in women from Chaoshan, China. *Environmental Science and Pollution Research*, *24*(5), 4778–4790. http://doi.org/10.1007/s11356-016-8208-6

Heldring, N., Pike, A., Andersson, S., Matthews, J., Cheng, G., Hartman, J., Gustafsson, J.-A. (2007). Estrogen Receptors: How Do They Signal and What Are Their Targets. *Physiological Reviews*, *87*(3), 905–931. http://doi.org/10.1152/physrev.00026.2006

Hereditary leiomyomatosis and renal cell cancer | Genetic and Rare Diseases Information Center (GARD) – an NCATS Program. (). Retrieved March 20, 2018, from https://rarediseases.info.nih.gov/diseases/7501/pulmonary-arterial-hypertension

Hevener, A. L., He, W., Barak, Y., Le, J., Bandyopadhyay, G., Olson, P., Olefsky, J. (2003). Muscle-specific Pparg deletion causes insulin resistance. *Nature Medicine*, *9*(12), 1491–1497. http://doi.org/10.1038/nm956

Humbert, M., Sitbon, O., Chaouat, A., Bertocchi, M., Habib, G., Gressin, V., Simonneau, G. (2006). Pulmonary arterial hypertension in France: results from a national registry. *American Journal of Respiratory and Critical Care Medicine*, *173*(9), 1023–1030. http://doi.org/10.1164/rccm.200510-1668OC

International Union of Pure and Applied Chemistry. Commission on the Nomenclature of Organic Chemistry., J., Rigaudy, J., Klesney, S. P., & International Union of Pure and Applied Chemistry. (1979). *Nomenclature of organic chemistry*. Pergamon Press.

Irey, N., pathology, H. M.-A. of, & 1973, undefined. (). Intimal vascular lesions associated with female reproductive steroids. *Popline.org*. Retrieved from https://www.popline.org/node/491301

Irey, N. S., & Norris, H. J. (1973). Intimal vascular lesions associated with female reproductive steroids. *Archives of Pathology*, *96*(4), 227–234.

Itoh, T., Nagaya, N., Ishibashi-Ueda, H., Kyotani, S., Oya, H., Sakamaki, F., Nakanishi, N. (2006). Increased plasma monocyte chemoattractant protein-1 level in idiopathic pulmonary arterial hypertension. *Respirology*, *11*(2), 158–163. http://doi.org/10.1111/j.1440-1843.2006.00821.x

JECFA. (2007). Safety evaluation of certain food additives and contaminants. *WHO Food Additives Series*, (48), 692 p. http://doi.org/10.1021/jf60163a014

Jen, Y., Manova, K., & Benezra, R. (1997). Each member of the Id gene family exhibits a unique expression pattern in mouse gastrulation and neurogenesis. *Developmental Dynamics*, *208*(1), 92–106. http://doi.org/10.1002/(SICI)1097-0177(199701)208:1<92::AID-AJA9>3.0.CO;2-X

Kaplan, J. L., Marshall, M. A., McSkimming, C. C., Harmon, D. B., Garmey, J. C., Oldham, McNamara, C. A. (2015). Adipocyte progenitor cells initiate monocyte chemoattractant protein-1-mediated macrophage accumulation in visceral adipose tissue. *Molecular Metabolism*, *4*(11), 779–794. http://doi.org/10.1016/j.molmet.2015.07.010

Khalil, R. A. (2013). Estrogen, vascular estrogen receptor and hormone therapy in postmenopausal vascular disease. *Biochemical Pharmacology*, *86*(12), 1627–1642. http://doi.org/10.1016/j.bcp.2013.09.024

Kim, J. H., Peacock, M. R., George, S. C., & Hughes, C. C. W. (2012). BMP9 induces EphrinB2 expression in endothelial cells through an Alk1-BMPRII/ActRII-ID1/ID3-dependent pathway: Implications for hereditary hemorrhagic telangiectasia type II. *Angiogenesis*, *15*(3), 497–509. http://doi.org/10.1007/s10456-012-9277-x

Kimbrough, R. D., & Krouskas, C. A. (2003). Human exposure to polychlorinated biphenyls and health effects: a critical synopsis. *Toxicological Reviews*, *22*(4), 217–33. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15189045

Kleiger, R., Boxer, M., Ingham, R., Chest, D. H.-, (1976). Pulmonary hypertension in patients using oral contraceptives: a report of six cases. *Journal.chestnet.org*. Retrieved from http://journal.chestnet.org/article/S0012-3692(16)56125-X/abstract

Knowell, A. E., Patel, D., Morton, D. J., Sharma, P., Glymph, S., & Chaudhary, J. (2013). Id4 dependent acetylation restores mutant-p53 transcriptional activity. *Molecular Cancer*, *12*(1), 161. http://doi.org/10.1186/1476-4598-12-161

Kopeć, G., Waligóra, M., Tyrka, A., Jonas, K., Pencina, M. J., Zdrojewski, T., … Podolec, P. (2017). Low-density lipoprotein cholesterol and survival in pulmonary arterial hypertension. *Scientific Reports*, *7*, 41650. http://doi.org/10.1038/srep41650

Korach, K. S., Sarver, P., Chae, K., McLachlan, J. A., & McKinney, J. D. (1988). Estrogen receptor-binding activity of polychlorinated hydroxybiphenyls: conformationally restricted structural probes. *Molecular Pharmacology*, *33*(1), 120–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3122017

Lagrange, A. H., Ronnekleiv, O. K., & Kelly, M. J. (1997). Modulation of G protein-coupled receptors by an estrogen receptor that activates protein kinase A. *Molecular Pharmacology*, *51*(4), 605–12. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9106625

Lasorella, A., Benezra, R., & Iavarone, A. (2014). The ID proteins: master regulators of cancer stem cells and tumour aggressiveness. *Nature Reviews Cancer*, *14*(2), 77–91. http://doi.org/10.1038/nrc3638

Lasserre, J. P., Fack, F., Revets, D., Planchon, S., Renaut, J., Hoffmann, L., Bohn, T. (2009). Effects of the endocrine disruptors atrazine and PCB153 on the protein expression of MCF-7 human cells. *Journal of Proteome Research*, *8*(12), 5485–5496. http://doi.org/10.1021/pr900480f

Lazorchak, A., Jones, M. E., & Zhuang, Y. (2005, June 1). New insights into E-protein function in lymphocyte development. *Trends in Immunology*. Elsevier Current Trends. http://doi.org/10.1016/j.it.2005.03.011

Lehrke, M., & Lazar, M. A. (2005). The Many Faces of PPARγ. *Cell*, *123*(6), 993–999. http://doi.org/10.1016/j.cell.2005.11.026

Lyden, D., Young, A. Z., Zagzag, D., Yan, W., Gerald, W., O'Reilly, R., Benezra, R. (1999a). Id1 and Id3 are required for neurogenesis, angiogenesis and vascularization of tumour xenografts. *Nature*, *401*(6754), 670–677. http://doi.org/10.1038/44334

Lyden, D., Young, A. Z., Zagzag, D., Yan, W., Gerald, W., O'Reilly, R., … Benezra, R. (1999b). Id1 and Id3 are required for neurogenesis, angiogenesis and vascularization of tumour xenografts. *Nature*, *401*(6754), 670–677. http://doi.org/10.1038/44334

Machado, R. D., Pauciulo, M. W., Thomson, J. R., Lane, K. B., Morgan, N. V, Wheeler, L., Nichols, W. C. (2001). BMPR2 haploinsufficiency as the inherited molecular mechanism for primary pulmonary hypertension. *American Journal of Human Genetics*, *68*(1), 92–102. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11115378

Manichaikul, A., Rich, S. S., Perry, H., Yeboah, J., Law, M., Davis, M., Taylor, A. M. (2014). A Functionally Significant Polymorphism in ID3 Is Associated with Human Coronary Pathology. *PLoS ONE*, *9*(3), e90222. http://doi.org/10.1371/journal.pone.0090222

Manikkam, M., Tracey, R., Guerrero-Bosagna, C., & Skinner, M. K. (2013). Plastics Derived Endocrine Disruptors (BPA, DEHP and DBP) Induce Epigenetic Transgenerational Inheritance of Obesity, Reproductive Disease and Sperm Epimutations. *PLoS ONE*, *8*(1), e55387. http://doi.org/10.1371/journal.pone.0055387

Massaro, D., Clerch, L. B., & Massaro, G. D. (2007). Estrogen receptor-alpha regulates pulmonary alveolar loss and regeneration in female mice: morphometric and gene expression studies. *American Journal of Physiology. Lung Cellular and Molecular Physiology*, *293*(1), L222-228. http://doi.org/10.1152/ajplung.00384.2006

Matthews, H. B., & Dedrick, R. L. (1984a). Pharmacokinetics of PCBs. *Annual Review of Pharmacology and Toxicology*, *24*(1), 85–103. http://doi.org/10.1146/annurev.pa.24.040184.000505

Matthews, H. B., & Dedrick, R. L. (1984b). Pharmacokinetics of PCBs. *Annual Review of Pharmacology and Toxicology*, *24*(2), 85–103. http://doi.org/10.1146/annurev.pa.24.040184.000505

Monje, P., & Boland, R. (2001). Subcellular distribution of native estrogen receptor alpha and beta isoforms in rabbit uterus and ovary. *Journal of Cellular Biochemistry*, *82*(3), 467–79. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11500923

Moodie, F. M., Marwick, J. A., Anderson, C. S., Szulakowski, P., Biswas, S. K., Bauter, M. R., … Rahman, I. (2004). Oxidative stress and cigarette smoke alter chromatin remodeling but differentially regulate NF- B activation and proinflammatory cytokine release in alveolar epithelial cells. *The FASEB Journal*, *18*(15), 1897–9. http://doi.org/10.1096/fj.04-1506fje

Morse, J. H., Horn, E. M., & Barst, R. J. (1999). Hormone Replacement Therapy: A Possible Risk Factor in Carriers of Familial Primary Pulmonary Hypertension. *Chest*, *116*(3), 847. http://doi.org/10.1378/CHEST.116.3.847

Mueller, C., Baudler, S., Welzel, H., Böhm, M., & Nickenig, G. (2002). Identification of a novel redox-sensitive gene, Id3, which mediates angiotensin II-induced cell growth. *Circulation*, *105*(20), 2423–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12021231

Mullerova, D., Kopecky, J., Matejkova, D., Muller, L., Rosmus, J., Racek, J., Matejovic, M. (2008). Negative association between plasma levels of adiponectin and polychlorinated biphenyl 153 in obese women under non-energy-restrictive regime. *International Journal of Obesity*, *32*(12), 1875–8. http://doi.org/10.1038/ijo.2008.169

Murre, C., McCaw, P. S., & Baltimore, D. (1989). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell*, *56*(5), 777–783. http://doi.org/10.1016/0092-8674(89)90682-X

Nadal, A., Ropero, A. B., Laribi, O., Maillet, M., Fuentes, E., & Soria, B. (2000). Nongenomic actions of estrogens and xenoestrogens by binding at a plasma membrane receptor unrelated to estrogen receptor alpha and estrogen receptor beta. *Proceedings of the National Academy of Sciences*, *97*(21), 11603–11608. http://doi.org/10.1073/pnas.97.21.11603

Okoh, V., Deoraj, A., & Roy, D. (2011). Estrogen-induced reactive oxygen species-mediated signalings contribute to breast cancer. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, *1815*(1), 115–133. http://doi.org/10.1016/j.bbcan.2010.10.005

Okoh, V. O., Felty, Q., Parkash, J., Poppiti, R., & Roy, D. (2013). Reactive Oxygen Species via Redox Signaling to PI3K/AKT Pathway Contribute to the Malignant Growth of 4-Hydroxy Estradiol-Transformed Mammary Epithelial Cells. *PLoS ONE*, *8*(2), e54206. http://doi.org/10.1371/journal.pone.0054206

Parkash, J., Felty, Q., & Roy, D. (2006). Estrogen Exerts a Spatial and Temporal Influence on Reactive Oxygen Species Generation that Precedes Calcium Uptake in High-Capacity Mitochondria: Implications for Rapid Nongenomic Signaling of Cell Growth . *Biochemistry*, *45*(9), 2872–2881. http://doi.org/10.1021/bi051855x

Parkash, J., Felty, Q., & Roy, D. (2006). Estrogen Exerts a Spatial and Temporal Influence on Reactive Oxygen Species Generation that Precedes Calcium Uptake in High-Capacity Mitochondria: Implications for Rapid Nongenomic Signaling of Cell Growth . *Biochemistry*, *45*(9), 2872–2881. http://doi.org/10.1021/bi051855x

Parkash, J., Felty, Q., & Roy, D. (2006). Estrogen exerts a spatial and temporal influence on reactive oxygen species generation that precedes calcium uptake in high-capacity mitochondria: Implications for rapid nongenomic signaling of cell growth. *Biochemistry*, *45*(9), 2872–2881. http://doi.org/10.1021/bi051855x

Patrone, C., Cassel, T. N., Pettersson, K., Piao, Y.-S., Cheng, G., Ciana, P., Nord, M. (2003). Regulation of postnatal lung development and homeostasis by estrogen receptor beta. *Molecular and Cellular Biology*, *23*(23), 8542–8552.

Peacock, A. J., Murphy, N. F., McMurray, J. J. V., Caballero, L., & Stewart, S. (2007). An epidemiological study of pulmonary arterial hypertension. *European Respiratory Journal*, *30*(1), 104–109. http://doi.org/10.1183/09031936.00092306

Pediconi, N., Guerrieri, F., Vossio, S., Bruno, T., Belloni, L., Schinzari, V., Levrero, M. (2009). hSirT1-dependent regulation of the PCAF-E2F1-p73 apoptotic pathway in response to DNA damage. *Molecular and Cellular Biology*, *29*(8), 1989–98. http://doi.org/10.1128/MCB.00552-08

Per, L., & Anderss, S. (1987). Transport of Polychlorinated Biphenyls (PCBs) In Freshwater From Sediment To Water, *36*, 33–46.

Perillo, B., Ombra, M. N., Bertoni, A., Cuozzo, C., Sacchetti, S., Sasso, A., Avvedimento, E. V. (2008). DNA Oxidation as Triggered by H3K9me2 Demethylation Drives Estrogen-Induced Gene Expression. *Science*, *319*(5860), 202–206. http://doi.org/10.1126/science.1147674

Perillo, B., Ombra, M. N., Bertoni, A., Cuozzo, C., Sacchetti, S., Sasso, A., Avvedimento, E. V. (2008). DNA Oxidation as Triggered by H3K9me2 Demethylation Drives Estrogen-Induced Gene Expression. *Science*, *319*(5860). Retrieved from http://science.sciencemag.org/content/319/5860/202

Phillip, A., & Iii, O. (2009). Angiotensin Ii Induction Of Regional Effects In Murine Vasculature. *University of Kentucky Doctoral Dissertations*. Retrieved from https://uknowledge.uky.edu/gradschool_diss/683

Polder, A., Odland, J. O., Tkachev, A., Føreid, S., Savinova, T. N., & Skaare, J. U. (2003). Geographic variation of chlorinated pesticides, toxaphenes and PCBs in human milk from sub-arctic and arctic locations in Russia. *Science of The Total Environment*, *306*(1–3), 179–195. http://doi.org/10.1016/S0048-9697(02)00492-8

Qiu, J., Bosch, M. A., Tobias, S. C., Grandy, D. K., Scanlan, T. S., Ronnekleiv, O. K., & Kelly, M. J. (2003). Rapid signaling of estrogen in hypothalamic neurons involves a novel G-protein-coupled estrogen receptor that activates protein kinase C. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *23*(29), 9529–40. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14573532

Rajesh, P., Sathish, S., Srinivasan, C., Selvaraj, J., & Balasubramanian, K. (2013). Diethyl Hexyl Phthalate (DEHP) is associated with insulin resistance in adipose tissue of male rat: Protective role of antioxidant vitamins (C & E). *Journal of Cellular Biochemistry*, *114*(3), 558–569. http://doi.org/10.1002/jcb.24399

Rasier, G., Parent, A.-S., Gérard, A., Lebrethon, M. C., & Bourguignon, J.-P. (2007). Early Maturation of Gonadotropin-Releasing Hormone Secretion and Sexual Precocity after Exposure of Infant Female Rats to Estradiol or Dichlorodiphenyltrichloroethane1. *Biology of Reproduction*, *77*(4), 734–742. http://doi.org/10.1095/biolreprod.106.059303

Ritter, R., Scheringer, M., MacLeod, M., Moeckel, C., Jones, K. C., & Hungerbühler, K. (2011). Intrinsic human elimination half-lives of polychlorinated biphenyls derived from the temporal evolution of cross-sectional biomonitoring data from the United Kingdom. *Environmental Health Perspectives*, *119*(2), 225–31. http://doi.org/10.1289/ehp.1002211

Safe, S. H. (1994). Polychlorinated Biphenyls (PCBs): Environmental Impact, Biochemical and Toxic Responses, and Implications for Risk Assessment. *Critical Reviews in Toxicology*, *24*(2), 87–149. http://doi.org/10.3109/10408449409049308

Selene, C. H., & Chou, J. (2003). Concise International Chemical Assessment Document 53: Hydrogen sulfide: Human health aspects. *IPCS Concise International Chemical Assessment Documents*. Retrieved from http://www.who.int/ipcs/publications/cicad/en/cicad55.pdf

Shanle, E. K., & Xu, W. (2010). Selectively targeting estrogen receptors for cancer treatment. *Advanced Drug Delivery Reviews*, *62*(13), 1265–1276. http://doi.org/10.1016/j.addr.2010.08.001

Shiraia, J. H., & Kissel, J. C. (1996). Uncertainty in estimated half-lives of PCBS in humans: impact on exposure assessment. *The Science of the Total Environment*, *187*, 199–210. Retrieved from https://ac.els-cdn.com/004896979605142X/1-s2.0-004896979605142X-main.pdf?_tid=c0c9d064-c7f8-11e7-80cb-00000aab0f26&acdnat=1510525865_b7eeeb174a31e95e6ec9166133d25ff4

Skipper, J. K., Young, L. J., Bergeron, J. M., Tetzlaff, M. T., Osborn, C. T., & Crews, D. (1993). Identification of an isoform of the estrogen receptor messenger RNA lacking exon four and present in the brain. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(15), 7172–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8346231

Sun, X. H. (2004, January 1). Multitasking of helix-loop-helix proteins in lymphopoiesis. *Advances in Immunology*. Academic Press. http://doi.org/10.1016/S0065-2776(04)84002-1

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., … Von Mering, C. (2015). STRING v10: Protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Research*, *43*(D1), D447–D452. http://doi.org/10.1093/nar/gku1003

Taylor, A. M., Li, F., Thimmalapura, P., Gerrity, R. G., Sarembock, I. J., Forrest, S., McNamara, C. A. (2006a). Hyperlipemia and Oxidation of LDL Induce Vascular Smooth Muscle Cell Growth: An Effect Mediated by the HLH Factor Id3. *Journal of Vascular Research*, *43*(2), 123–130. http://doi.org/10.1159/000090131

Taylor, A. M., Li, F., Thimmalapura, P., Gerrity, R. G., Sarembock, I. J., Forrest, S., McNamara, C. A. (2006b). Hyperlipemia and oxidation of LDL induce vascular smooth muscle cell growth: an effect mediated by the HLH factor Id3. *Journal of Vascular Research*, *43*(2), 123–30. http://doi.org/10.1159/000090131

Thomas, P., & Dong, J. (2006). Binding and activation of the seven-transmembrane estrogen receptor GPR30 by environmental estrogens: A potential novel mechanism of endocrine disruption. *Journal of Steroid Biochemistry and Molecular Biology*, *102*(1–5 SPEC. ISS.), 175–179. http://doi.org/10.1016/j.jsbmb.2006.09.017

Toran-Allerand, C. D., Guan, X., MacLusky, N. J., Horvath, T. L., Diano, S., Singh, M., Tinnikov, A. A. (2002). ER-X: a novel, plasma membrane-associated, putative estrogen receptor that is regulated during development and after ischemic brain injury. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *22*(19), 8391–401. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12351713

Ussavarungsi, K., Thomas, C. S., & Burger, C. D. (2015). Prevalence of metabolic syndrome in patients with pulmonary hypertension. *The Clinical Respiratory Journal*, n/a-n/a. http://doi.org/10.1111/crj.12406

Veldhuizen, R., Nag, K., Orgeig, S., & Possmayer, F. (1998). The role of lipids in pulmonary surfactant. *Biochimica et Biophysica Acta*, *1408*(2–3), 90–108. http://doi.org/10.1016/S0925-4439(98)00061-1

Wahlang, B., Barney, J., Thompson, B., Wang, C., Hamad, O. M., Hoffman, J. B., Hennig, B. (2017). PCB126 Exposure Increases Risk for Peripheral Vascular Diseases in a Liver Injury Mouse Model. *Toxicological Sciences*, 1–12. http://doi.org/10.1093/toxsci/kfx180

Wahlang, B., Falkner, K. C., Gregory, B., Ansert, D., Young, D., Conklin, D. J., Cave, M. (2013). Polychlorinated biphenyl 153 is a diet-dependent obesogen that worsens nonalcoholic fatty liver disease in male C57BL6/J mice. *Journal of Nutritional Biochemistry*, *24*(9), 1587–1595. http://doi.org/10.1016/j.jnutbio.2013.01.009

Wang, W., Zhang, Y., Lu, W., & Liu, K. (2015). Mitochondrial Reactive Oxygen Species Regulate Adipocyte Differentiation of Mesenchymal Stem Cells in Hematopoietic Stress Induced by Arabinosylcytosine. *PLOS ONE*, *10*(3), e0120629. http://doi.org/10.1371/journal.pone.0120629

Watson, C., Bulayeva, N., Wozniak, A., & Finnerty, C. (2005). Signaling from the membrane via membrane estrogen receptor-α: Estrogens, xenoestrogens, and phytoestrogens. *Steroids*, *70*(5–7), 364–371. http://doi.org/10.1016/j.steroids.2005.03.002

Wu, Q., Zhou, Z.-J., & Ohsako, S. (2006). Effect of environmental contaminants on DNA methyltransferase activity of mouse preimplantation embryos. *Wei Sheng Yan Jiu = Journal of Hygiene Research*, *35*(1), 30–2. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16598927

Yang, J., Li, X., Li, Y., Southwood, M., Ye, L., Long, L., Morrell, N. W. (2013). Id proteins are critical downstream effectors of BMP signaling in human pulmonary arterial smooth muscle cells. *American Journal of Physiology-Lung Cellular and Molecular Physiology*, *305*(4), L312–L321. http://doi.org/10.1152/ajplung.00054.2013

Yang, W.S., Jeng, C.Y., Wu, T. J., Tanaka, S., Funahashi, T., Matsuzawa, Y., Chuang, L.M. (2002). Synthetic peroxisome proliferator-activated receptor-gamma agonist, rosiglitazone, increases plasma levels of adiponectin in type 2 diabetic patients. *Diabetes Care*, *25*(2), 376–80. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11815513

Zebedee, Z., & Hara, E. (2001). Id proteins in cell cycle control and cellular senescence. *Oncogene*, *20*(58), 8317–8325. http://doi.org/10.1038/sj.onc.1205092

Zsarnovszky, A., Le, H. H., Wang, H. S., & Belcher, S. M. (2005). Ontogeny of Rapid Estrogen-Mediated Extracellular Signal-Regulated Kinase Signaling in the Rat Cerebellar Cortex: Potent Nongenomic Agonist and Endocrine Disrupting Activity of the Xenoestrogen Bisphenol A. *Endocrinology*, *146*(12), 5388–5396. http://doi.org/10.1210/en.2005-0565

CHAPTER 2

Lung endothelial cell dysfunction from exposure to polychlorinated biphenyl 153 (PCB153) contributes to lung toxicity.

Abstract

ID3, a transcription regulator, has been shown to contribute to the aggressive spread of endothelial cells by facilitating the generation of tumor blood vessels. Tumor vascular cells characterized as CD133$^+$ endothelial stem/progenitor cells are associated with poor prognosis in lung cancer and pulmonary artery hypertension. Previously, we have shown that ID3 overexpression reprograms adult endothelial cells to acquire vascular stem markers including CD133. However, little is known regarding the molecular contribution of ID3 to tumor vascular reprogramming and how this may impact the aggressive growth of lung cancer. In the current project, we evaluated ID3 mediated vascular reprogramming upon PCB exposure. This idea is consistent with the literature, which shows that PCBs accumulate in the human lung; PCBs produce pathological vascular remodeling; high levels of PCBs are found in human lung tissue; and epidemiological studies associating lung toxicity with PCBs. Recent studies identify PCB153 as one of the largest contributors for total PCB body burden in humans. Using human lung endothelial and smooth muscle cells, we exposed ID3 overexpressing and vector control cells to PCB153. We observed a significant increase in cell proliferation as determined by the BrdU incorporation assay and Fluorescence activated cell sorting (FACS) analysis. Similarly, a 3D HuBiogel model, which mimics *in vivo* conditions, showed a significant increase in size and number of vascular spheres upon PCB153 treatment. Pluripotent vascular stem cells showed the loss of VE-cadherin and gain of MMP9, N-cadherin, and vimentin, which are markers of endothelial-mesenchymal transition. In summary, our findings show that persistent environmental pollutant PCB153 induces ID3 expression that plays a critical role in regulating survival and hyper-proliferation of lung endothelial

cells. Implications of our research include a new paradigm by which PCBs may contribute to lung vascular toxicity through pathological vascular remodeling and/or angiogenesis.

Key words- ID3, endocrine disruptors, Lung endothelial cells, plexiform lesions.

Introduction-

PCBs are ubiquitous man-made aromatic chlorinated hydrocarbons considered as an organic persistent pollutants. Being lipophilic in nature, PCBs have been shown to bio-accumulate in the food chain and tissues of living organisms. High levels of PCBs are found in human lung tissue and epidemiological studies have shown the association of lung toxicity by chronic exposure to PCBs. PCB153 is one of the 209 congeners of the PCBs most often found in human foods and lipid-rich human tissues. PCB153 has been shown to be involved in pathological vascular remodeling and selective bioaccumulation in lung in the experimental models. Our earlier study on investigation of the effects of exposure of PCB153 to primary cells, human umbilical vein endothelial cells (HUVECs) and normal human dermal fibroblasts (NHDF) demonstrated PCB-induced ROS contributed to neovascularization in endothelial cells dysfunction (Das & Felty, 2015; Doke, Felty, & Das, 2017). Epidemiological studies have reported that women who received hormone replacement therapy (HRT) show higher risk to PAH when compared to women with no HRT (Irey et al., 1973 ; Kleiger et al., 1976; Morse et al., 1999). In addition to this, women are 2-4 times more prone to pulmonary disorder or PAH than men (Badesch et al., 2010a). So, the extra estrogen or synthetic estrogen present in the form environmental pollutants maybe responsible for pulmonary vascular disease in

humans. Since, the aerosols or vapors containing highly lipophilic PCBs may enter through inhalation and tend to accumulate in lipid rich tissues like lungs, brain and breast (ASTDR, 2000b). Epidemiological studies have associated PCBs with cardiovascular disease (Perkins, Petriello, Newsome, & Hennig, 2016). Therefore, the present study aims to investigate the effect of environmental pollutants on neo-vascularization in human pulmonary lung endothelial cells, which are exposed to environmentally relevant concentrations (ppb) of PCB153, and physiological concentrations of 17β-estradiol.

Materials and Methods
Cell culture and treatment conditions

Primary cells, Human pulmonary microvascular endothelial cells or human lung endothelial cells (HPMEC-ST1.6R). HPMEC cells were maintained in DMEM-F12 medium supplemented with 5% FBS. Cells were cultured at 37 ℃ in a humidified atmosphere with 5% CO_2. We also used pulmonary smooth muscle cell line (SMC) purchased from Lonza Inc.

Estrogenic chemical treatments

Stock solutions of PCB153 and 17β-estradiol were prepared in dimethyl sulfoxide (DMSO). The same amount of DMSO as in PCB and estradiol-treated cells were added to control cultures. The level of DMSO in experimental media was less than 0.1%. PCB blood levels have been reported to reach approximately 1000 ng/mL (~3 μM) in occupationally exposed individuals (Wassermann, Cucos, & Miller, 1979). Our unpublished data showed a significant increase in PCB- induced vascularization with

PCB153 concentrations of 10–100 ng/mL. Based on known PCB blood levels from occupational exposure and our preliminary results; we chose a PCB dose of 100 ng/mL (∼0.3 μM) to expose endothelial cells. Our previous studies also showed endothelial cell proliferation and vascular tube formation at physiological doses of 17β-estradiol (Das & Felty, 2014a; Felty, 2006; Felty & Porther, 2008b) therefore, HPMEC and HPMEC ID3[+] were exposed to estradiol at 100 ng/mL (∼3.6 nM). PCB congeners 2,2′,4,4′,5,5′-hexachlorobiphenyl (PCB153) were purchased from AccuStandard (NewHaven, CT) and dissolved in dimethyl sulfoxide (DMSO). All other chemicals and reagents were purchased from Sigma (St. Louis, MO).

ID3 overexpression

The HPMEC cells were stably transfected with either Precision LentiORF for ID3 (Thermo Scientific Open Biosystems) or empty vector lentiviral pLEX-JRED/TurboGFP by the trans-lentiviral packaging kit with Express-in transfection reagent according to the manufacturer's instructions. We used the MOI (multiplicity of infection) of 25 and selected cells that overexpressed ID3 with blasticidin S (5 mg/mL) as per manufacturer's instructions. Cells expressing TurboGFP were identified by fluorescence microscopy.

Cell migration assay

The cell migration assay was performed to quantify and compare the rate of migration of cells between wild-type lung endothelial cells and ID3 overexpressing lung endothelial cells. Lung EC and lung EC ID3[+] cells were cultured in DMEM-F12 media with 5% fetal bovine serum in six-well plates to create a confluent monolayer, and an equal size scratch was created by manually scraping the cell monolayer with a sterile plastic p200-pipette

tip in the middle area of the well as a line. Then lung EC and lung EC ID3[+] cells were treated with 0.1% DMSO as a control, E2 (100 ng/mL), PCB153 (100 ng/mL) and combination treatment of E2 + PCB153 in separate wells and incubated in growth medium for 12h. The initial cell migration quantification was performed on images collected 0h after scratching when the scratch size had stabilized. Additional images were collected randomly from the scratched areas at 6 and 12h after scratching. Pictures were taken under a microscope at 40X magnification and the results were analyzed using ImageJ Launcher. This experiment was carried out in triplicate.

Endothelial spheroid assay

HPMEC and HPMEC ID3[+] cells were suspended in serum-free DMEM/F12 (1:1) culture medium supplemented with B27. For EC spheroid formation, approximately 100–150 HPMEC and HPMEC ID3[+] cells were seeded separate well in an ultra-low attachment 96-well plate (Corning Inc. Lowell, MA) followed the published protocol of our laboratory (Das & Felty, 2014a). The effect of estrogenic chemical PCB153 was determined by pre-treating the cells with 100 ng/mL E2 2h prior to exposure with 100 ng/mL (~0.3 µM) PCB153 on the day of seeding cells. Endothelial spheroids were grown for 10 days in liquid culture in the absence or presence of PCB153. A total of 15 endothelial spheroids with a minimum diameter of 50 µm were counted in each experimental group. Pictures were taken on 5d and 10d after seeding. We counted number and size of spheroids on 5d and 10d. Data were analyzed by ANOVA; Turkey HSD test for multiple comparisons.

MTT assay

To determine the effect of E2 and extra estrogen environmental pollutant PCB153 on the metabolic activity of lung endothelial cells, we performed cell viability and cell proliferation assays. For cell viability measurement, we performed MTT assay which is colorimetric assay and measures the metabolic activity of cells. Actually, MTT is a tetrazolium dye known as 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide. NAD(P)H-dependent oxidoreductase enzymes present in mitochondria of metabolically active and dividing cells. This enzyme catalyses tetrazolium dye and converts into insoluble formazan, which gives a purple colour (Berridge & Tan, 1993; Berridge, Herst, & Tan, 2005). The intensity of the purple colour that can be measured at an optical density (OD) of 590 nm and its proportional to a number of metabolic active cells. We seeded HPMEC and HPMEC ID3$^+$ cells with 50% confluence in 96 well clear walled plate. We allowed the cells to grow for 24h in 5% FBS culture medium. After 24 hours, we gave treatments, 0.1% DMSO (control), E2 (100 ng/mL), PCB153 (100 ng/mL) and combination treatment of E2 + PCB153 to make total volume 100 µL in each well. We measured the metabolic activity of cells after 48h with the addition of MTT dye and evaluating the purple colour intensity at 590 nm.

SRB assay

The sulforhodamine B (SRB) assay is used for measurement of the total cellular protein content of adherent cells in a 96-well format. This assay is useful in the determination of cellular proliferation and growth. We seeded HPMEC and HPMEC ID3$^+$ cells with 50% confluence in 96 well clear-walled plate. We allowed the cells to grow for 24h in 5%

FBS culture medium. After 24 hours, we gave treatments, 0.1% DMSO (control), E2 (100 ng/mL), PCB153 (100 ng/mL) and combination treatment of E2 + PCB153 to make total volume 100 µL in each well. After an incubation of 48h, cell monolayers were fixed with 10% (wt/vol) trichloroacetic acid and stained for 30 min, after which the excess dye is removed by washing repeatedly with 1% (vol/vol) acetic acid. The protein-bound dye is dissolved in 10 mM Tris base solution for OD determination at 590 nm using a microplate reader (Vichai & Kirtikara, 2006).

BrdU assay- BrdU assay is a direct method for evaluation of cell proliferation. In this rapid method, Bromodeoxyuridine (5-bromo-2'-deoxyuridine, BrdU) a synthetic nucleoside that is an analogue of thymidine gets incorporated into DNA synthesis of proliferating cells. It is a suitable method for detection and quantification of the stimulatory effect of environmental chemicals (Russo et al., 1984). For BrdU assay, we used Roche's kit and followed the manual instructions.

We seeded HPMEC and HPMEC ID3[+] cells in 5% FBS DMEM F-12 media with 50% confluence and cultured with addition of BrdU in 96 well clear-walled plate. After 24 hours, we gave treatments, 0.1% DMSO (control), E2 (100 ng/mL), PCB153 (100 ng/mL) and combination treatment of E2 + PCB153 to make total volume 100 µL in each well. After an incubation of 48h, cell monolayers fixed with FixDenat. Followed by fixation we added anti- BrdU-POD, which binds to BrdU in newly synthesized DNA. This immune complex detected by a plate reader at 450 nm.

Western blot

Whole cell lysates were prepared with lysis buffer containing 25 mM Tris-HCl buffer (pH 8.0), 150 mM NaCl, 0.2% NP-40, 10% glycerol, 8 mM b-glycerophosphate, 2.5 mM sodium pyro- phosphate, 10 mM NaF, 0.2 mM Na_3VO_4, 1 mM DTT and 10 ml/ml protease inhibitor cocktail (Sigma-Aldrich) followed the published protocol of our laboratory (Das & Felty, 2014a). Proteins were quantified using the Bradford assay reagent (Bio-Rad) according to the manufacturer's instructions. Proteins (35–75 mg) were separated by 15%, 7.5% and 10% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to polyvinylidene fluoride (PVDF) membranes (Millipore). Membranes were blocked with 5% nonfat milk and incubated with the following antibodies: ID3 (Cal BioReagent), VE-cadherin-2 antibody (F-4): sc-515467 mouse monoclonal Ab, vimentin antibody (V9): sc-6260 mouse monoclonal Ab, N-cadherin antibody (H-63): sc-7939 and B-actin (Cell Signaling). Antibody dilutions used were according to manufacturer's recommendations for detection by immunoblot. Membranes were then incubated with horseradish peroxidase-conjugated secondary IgG antibodies and visualized with ECL plus Western blot reagents (GE Healthcare, Amersham). The membranes were re-probed for b-actin as a loading control. Electrochemiluminescence (ECL) intensity of detected target proteins was imaged and quantified with a Bio-Rad Versa Doc instrument. All immunoblots have completed a minimum of three times for each experiment.

Immunofluorescence

We seeded approximately 200–250 HPMEC and HPMEC ID3$^+$ cells in 8 chambered slide. We allowed the cells to grow for 24h. Then we gave treatment 0.1% DMSO (control), E2 (100 ng/mL), PCB153 (100 ng/mL) in both the cell lines. Then we fixed samples in 2% paraformaldehyde made in PBS for 15 minutes at room temperature. Washed monolayer 2 times in PBS to remove residual paraformaldehyde, store samples submerged in PBS at 4°C until use. Permeabilized cells with 0.1% Triton X-100 made in PBS solution for 15 min. Washed monolayer 5 times with a PBS + 0.5% BSA (PBB). Then blocked with 2% BSA for 45 minutes. Cells were then incubated with primary antibodies and alexa fluor-conjugated secondary antibodies. The confocal fluorescence images were scanned on a Nikon TE2000U inverted microscope. The DRAQ-5 was used to label the nucleus. The confocal images were acquired on a Nikon C1 laser scanning confocal microscope (Melville, NY, USA). MMP-9 was determined by immunofluorescent labelling with rabbit anti-MMP9 polyclonal antibody bs-4593R (Alexa Fluor 566-conjugated secondary antibody). VE-Cadherin was determined by immunofluorescent labelling with mouse monoclonal IgG1 sc-9989 (Alexa Fluor 633-conjugated secondary antibody). Vimentin was determined by immunofluorescent labelling with polyclonal goat IgG AF2105.

Zebrafish Xenograft

We purchased xenotransplantation Zebrafish embryos from Zebrafish International Resource Center (ZIRC), University of Oregon, Eugene. Albino embryos are completely transparent. Prior to injecting HPMEC and HPMEC ID3$^+$ cells on 2 days post fertilization (dpf), albino zebrafish were maintained at 28 °C. During injection, zebrafish were

anaesthetized with MESAB (0.5 mM 3-aminobenzoic acid ethyl ester, 2 mM Na_2HPO) and then placed in a 15 mm Petri dish on their sides on a ramp comprised of 1% agarose. Approximately 50 HPMEC and HPMEC ID3[+] cells were injected into the yolk sac and zebrafish were then transferred in fresh fish water to an incubator and maintained at 37 ºC for up to 7 days post injection (dpi) when the experiments were terminated and animals were euthanized by overexposure to MESAB.

Assessment of temperature effect on zebrafish development

Normally, zebrafish are incubated at 28 ºC. To accommodate human EC proliferation, an incubation temperature close to 37 ºC was desirable. We purchased xenotransplantation Zebrafish embryos from Zebrafish International Resource Center (ZIRC), University of Oregon, Eugene. Albino embryos are completely transparent. Prior to injecting HPMEC and HPMEC ID3[+] cells on 2 dpf, albino zebrafish were maintained at 28 ºC. During injection, zebrafish were anaesthetized with MESAB (0.5 mM 3-aminobenzoic acid ethyl ester, 2 mM Na_2HPO) and then placed in a 15 mm Petri dish on their sides on a ramp comprised of 1% agarose. Approximately 50 HPMEC and HPMEC ID3[+] cells were injected into the yolk sac and zebrafish were then transferred in fresh fish water to an incubator and maintained at 37 ºC for up to 7 dpi when the experiments were terminated and animals were euthanized by overexposure to MESAB.

Cell preparation for transplant

Cells were detached from the culture flask by incubating with 2 mL dispase (Invitrogen) at 37 ºC for 45 min, pelleted by centrifugation at 2500 rpm for 5 min, and resuspended in 2.5 mL HBSS. Cells were labeled with CM-DiI, a lipophilic fluorescent tracking dye

(Invitrogen). According to manufacturer's instructions, CM-DiI is transferred from mother to daughter cells but not between other cells. The cells were incubated with CM-DiI dissolved in DMSO (final concentration: CM-DiI: 4.8 lg/mL, DMSO: 0.4%) for 4 min at 37 ºC, followed by 15 min at 4 ºC. CM-DiI labeled cells were loaded into a pulled glass micropipette that was drawn on an electrode puller and then trimmed to form a needle with a 15 μm internal and a ~18 μm external diameter. The microneedle was attached to an air driven Cell Tram micro injector. The tip of the needle was inserted into the animals and cells were delivered in a single injection. A sharp needle cut in the yolk will reclose with little to no loss of contents. Any injection site in the yolk is suitable. Using 30–100 milliseconds pulse time and 8 pounds per square inch (psi) positive pressure, we optimized the number of cells injected at ~50, which we confirmed by dispensing cells onto a microscope slide and visually counting. The volume of material injected is approximately 20–50 nL. After injection, zebrafish were maintained for 1 h at 28 ºC, examined for the presence of a fluorescent cell mass of ~50 cells localized at the injection site in the yolk sac, and then transferred to an incubator and maintained at 37 ºC. On 1 dpi, zebrafish were examined by fluorescence microscopy (Haldi, Ton, Seng, & McGrath, 2006).

Flow Cytometry

For cell cycle analysis, cells were harvested, rinsed with PBS, and fixed in suspension in 73% ethanol for 20h at −20 °C. After incubation at −20 °C, the cells were washed with PBS containing 1% BSA, stained with PI (10 μg/mL) in PBS containing RNase A (250 μg/mL), and incubated at 37 °C for 30 min in the dark before FACS analysis. Cell cycle

distribution was analyzed with the Guava easyCyte™ using the CytoSoft software program according to the manufacturer's instructions. The percent DNA content was determined by FACS analysis using PI staining.

Statistical analysis

All experiments were performed at least three times and data are expressed as mean ± standard deviation Graphpad prism version 5.01 was used to perform graphics and the two-tailed student's t-test was used to compare statistical significance. $P < 0.05$ was set as a statistical significance. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.

Results:

ID3 overexpression in wild type lung endothelial
We previously demonstrated that E2 treatment increased ID3 phosphorylation and ID3 protein stability; and that E2-induced cell proliferation was dependent on oxidant sensitive ID3 signalling (Felty & Porther, 2008b). We have also demonstrated that exogenous estrogen or estrogenic PCBs like PCB153 increases ID3 protein expression which eventually gives stemness markers to the human cerebral microvascular endothelial cell line HCMEC/D3 (Das & Felty, 2014c). Given the biological significance of ID3 proteins for the maintenance of stemness, we asked whether overexpression of ID3 could induce a molecular stem-cell signature in adult lung ECs HPMEC-ST1.6R cell line. To direct this question, human HPMECs were transduced using an MOI of 25 with either the Precision LentiORF for ID3 overexpression or control empty lentiviral vector pLEX-JRED/TurboGFP according to the manufacturer's instructions. Cells were then

seeded at clonal density and selected with blasticidin S (5 μg/mL). Cells that overexpress

ID3 will be selected for culture and referred to as HPMEC ID3[+] as shown in Figure 7.

Confocal microscope image
of HPMEC ID3[+]

Phase contrast image of
HPMEC ID3[+]
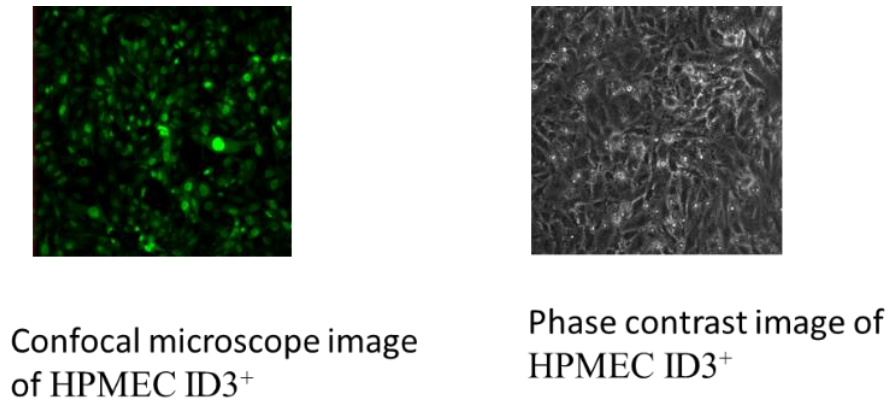
Figure 7. ID3 overexpression in Lung EC

HPMEC cells were stably transfected with Precision LentiORF for ID3 (Thermo Scientific Open Biosystems). Selected cells that overexpressed ID3 with blasticidin S (5 mg/ml) as per manufacturer's instructions. Cells expressing TurboGFP in HPMEC ID3+ cells were identified by fluorescence microscopy. We also shown phase contrast image of HPMEC ID3+ cells along with it.

Cell migration (Wound healing) assay
In order to study and distinguish between the wild-type lung endothelial cell line HPMEC

and newly transfected stable ID3 overexpressed HPMEC ID3[+] cell line, we did some

experiments to detect the phenotypic behaviour in these two cell lines. We evaluated the

effect of ID3 overexpression had on cell migration using a wound healing or cell

migration assay. A cell monolayer consisting of HPMEC wild type or HPMEC ID3[+]

overexpressing cell was mechanically scratched to create a wound. Then HPMEC and

HPMEC ID3$^+$ cells with wound were treated with 0.1% DMSO as a control, E2 (100 ng/mL), PCB153 (100 ng/mL) in separate wells and incubated in growth medium for 12 h, we examined the extent of wound closure up to 12h. We took images of wound healing at 6h and 12h intervals. We analyzed the cell migration rate by using ImageJ Launcher software. By measuring the wound area compared with the initial wound area, we measured the number of cells migrated into scratched area. By measuring the migration rate of both the cell lines, we found that the ability of HPMEC ID3$^+$ cells to migrate to the wound area was significantly greater than HPMEC cells. Images were captured after the wound at the 6h interval. The number of HPMEC ID3$^+$ cells migrated in the scratched area were significantly ($p < 0.0001$) higher than HPMEC cells in 0.1% DMSO (control) itself (Figure 8). Migration of HPMEC ID3$^+$ cells was higher when compared to HPMEC cells. Likewise, we saw that treatments of E2 (100 ng/mL) and PCB153 (100 ng/mL) significantly increase the number of HPMEC ID3$^+$ cells migrated in the scratched area ($p < 0.0001$) compared to HPMEC cells. At the 12h interval, HPMEC ID3$^+$ cells completely migrated to the scratched area and filled out the wound in all treatments except 0.1%DMSO (control). In case of HPMEC cells, all treatments E2 (100 ng/mL), PCB153 (100 ng/mL) showed significant ($p < 0.0001$) results compared to 0.1% DMSO (control) in terms of cell migration into scratched area (Fig. 8 and 9).
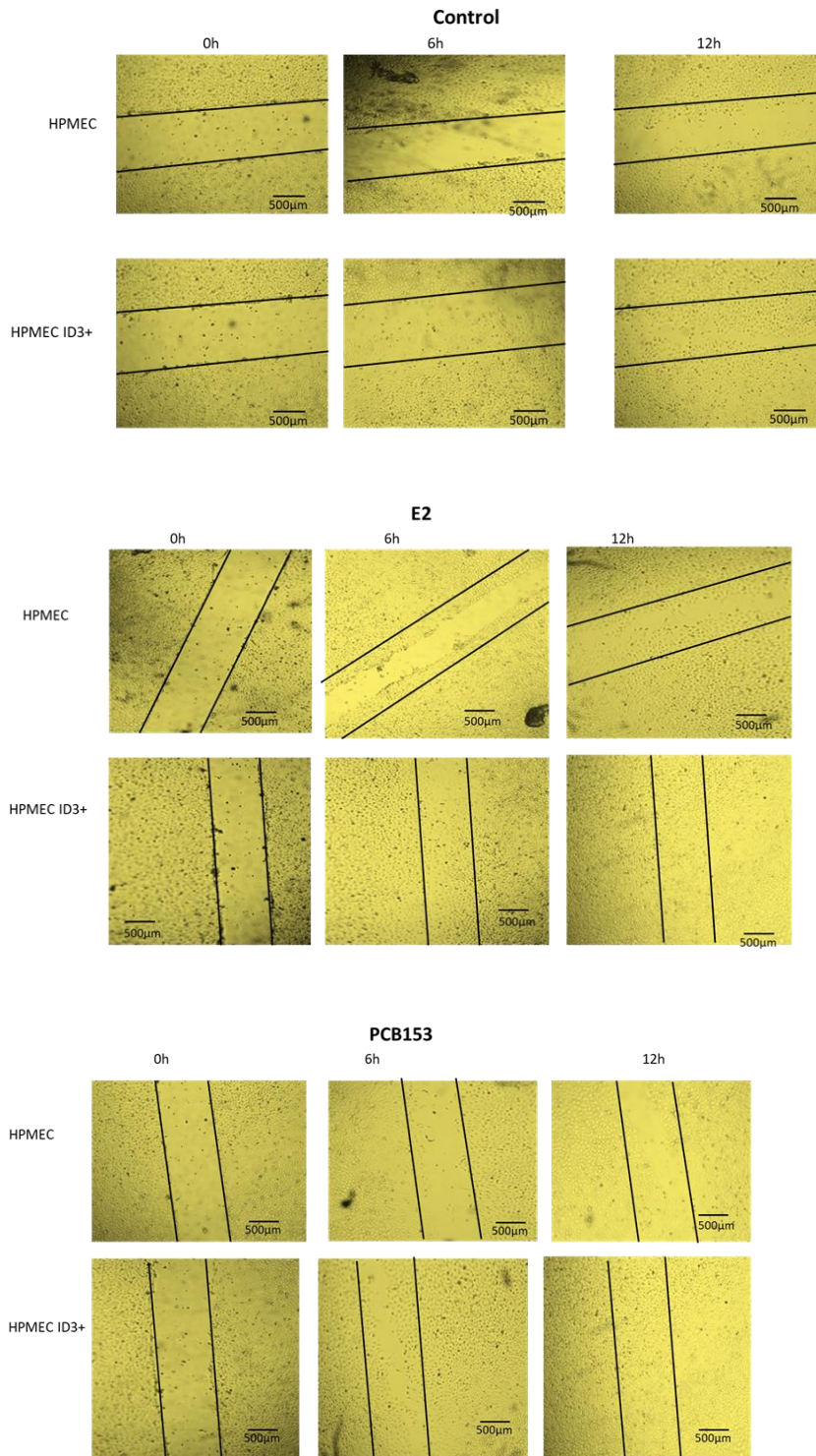
Figure 8- Wound healing assay

Representative images showing migration of both lung EC & EC ID3$^+$ cells. When both the cells lines treated with control, E2 and PCB153. E2 and PCB153 increased the migration rate of cells compared to control in 6h and 12h intervals. Magnification X200
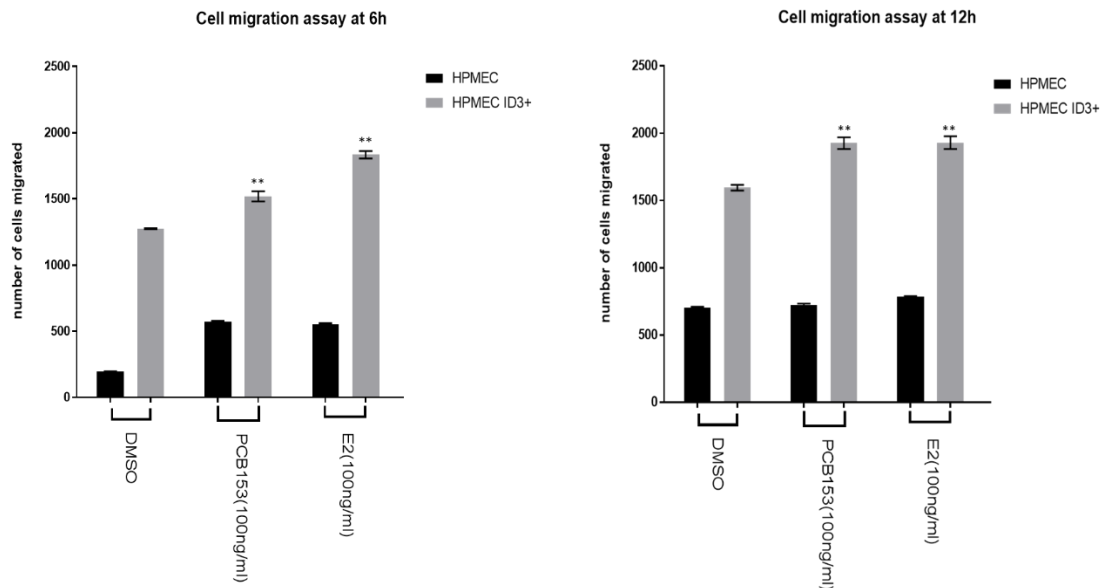
Figure 9- Cell migration

Higher migration of lung ECs overexpressing ID3. Experimental treatments: vehicle, E2, PCB153. Treatments with E2 and PCB153 increased the migration of ID3 overexpressing lung ECs compared to control at 6h and 12h intervals. Graph shows number of cells migrated in 6h and 12h intervals. Error bars represent the mean number of cells migrated ± SD in 5 wells. E2 & PCB153= 100 ng/mL. **p<0.01 PCB153 & E2 vs DMSO for lung EC ID3+. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.

MTT, SRB and BrdU assay of HPMEC and HPMEC ID3+ cells

We tested the lung endothelial cell proliferation in presence of estrogenic chemicals by MTT, SRB and BrdU assay. Since researchers have shown that lung endothelial cells possess estrogen receptors, the HPMEC and HPMEC ID3+ cells were showing a stimulatory effect in terms of cell proliferation and growth when they treated with E2 (100 ng/mL) and PCB153 (100 ng/mL). Our data revealed that treatments with E2 (100 ng/mL) and PCB153 (100 ng/mL) significantly increased cell growth when compared to

vehicle control. Cell growth was determined by the MTT and SRB assays and corroborated by BrdU assay. The data from SRB and BrdU assays reported that treatments E2 (100 ng/mL), PCB153 (100 ng/mL) showed a significant increase in numbers of cells compared to control. In terms of SRB assay, the total protein content of treated cells was significantly higher than control. The BrdU assay data results showed the synthesis of new DNA in proliferating cells was significantly higher in treatments than control. These results show that PCB153 (100 ng/mL) drives endothelial cell proliferation. In addition to this, we determined the contribution of ID3 to endothelial proliferation upon treatment with E2 (100 ng/mL), PCB153 (100 ng/mL) than HPMEC cells in the Figures 10, 11, and 12.

Since estrogenic chemicals like PCB153 shows their growth stimulating endothelial cell survival effects on HPMEC and HPMEC ID3$^+$ cells, we investigated signaling mechanism action of PCB153. Felty et al. showed that PCB153 increases microvascular lesions in HUVEC cells through ROS pathway. To address this question we assessed the effects of ER blocker and ROS scavengers on HPMEC and HPMEC ID3$^+$ cells in presence of E2 and PCB153 treatment.

To investigate whether ER was involved in the E2 and PCB153-mediated hyperproliferation of HPMEC ID3$^+$ cell, cells were treated with the known ER antagonist, tamoxifen. Tamoxifen is a nonsteroidal triphenylethylene derivative that binds to the estrogen receptor (Jordan, et al., 1980.). It acts as a selective estrogen receptor modulator. Although tamoxifen prevents the growth of estrogen-sensitive tissue, the researchers have shown that tamoxifen treatment has been clinically effective in triple

negative breast cancer patients (Adjuvant Trial Organisation, 1988). We treated HPMEC and HPMEC ID3+ cells with single treatment of tamoxifen (5 µM) dose as wells as co-treatment with E2 + tamoxifen, PCB153 + tamoxifen. We carried out MTT, SRB and BrdU assays and analyzed the data. We found that tamoxifen (5 µM) significantly (P<0.01) inhibit the growth of cells in both the HPMEC and HPMEC ID3+ cell lines compare to control. In addition to this, our data showed that co-treatment of E2 +tamoxifen and PCB153 + tamoxifen showed significant (P<0.01) reduction in growth of cells compared to E2 (100 ng/mL) and PCB153 (100 ng/mL) respectively. These data suggests that tamoxifen may be blocking metabolism and redox cycling of estrogen, PCB153 and acts as free radical scavengers as previously reported by researchers in MDA cell line which is triple negative breast cancer cell line (Arteaga, Villaseca, Bianchi, Rojas, & Marshall, 2003) as shown in Figure 13, 14 and 15.

To determine whether E2 and PCB153 increases the number of HPMEC and HPMEC ID3+ cells by either estrogen receptor or ROS signaling pathway. Fulvestrant is drug treatment for hormone receptor-positive metastatic breast cancer in postmenopausal women with disease progression following anti-estrogen therapy. It is a complete estrogen receptor antagonist with no agonist effects, which in addition, accelerates the proteasomal degradation of the estrogen receptor. We found the similar results like Fulvestrant drug treatment. We carried out MTT, SRB and BrdU assays and analyzed the data. We found that Fulvestrant (5 µM) significantly (P<0.01) inhibit the growth of cells in both the HPMEC and HPMEC ID3+ cell lines compare to control. In addition to this, our data showed that co-treatment of E2 + Fulvestrant, PCB153 + Fulvestrant showed significant (P<0.01) reduction in growth of cells compared to E2 (100 ng/mL) and

PCB153 (100 ng/mL) respectively (as shown in Figure 10, 11 and 12).

To investigate the E2 and PCB153 increases number and new DNA synthesis of HPMEC and HPMEC ID3[+] cells through ROS signaling, we used antioxidant like ebselen. It is an organoselenium compound and acts as a mimic of glutathione peroxidase (Kumar, Bullet, & Tomar, 2014). Interestingly, we found the similar results like tamoxifen and Fulvestrant drug treatment. We carried out MTT, SRB and BrdU assays and analyzed the data. We found that Ebselen (10 µM) significantly (P<0.01) inhibit the growth of cells in both the HPMEC and HPMEC ID3[+] cell lines compare to control. In addition to this, our data showed that co-treatment of E2 + Ebselen, PCB153+Ebselen showed significant (P<0.01) reduction in growth of cells compared to E2 (100 ng/mL) and PCB153 (100 ng/mL) (as shown in Figure 16, 17 and 18). ID3 overexpression increased the growth and DNA synthesis of lung ECs upon exposure to mitogens PCB153 and E2. ID3 overexpression make these cells resistant to Tamoxifen, Ebselen and fulvestrant when compared to vector alone. ID3 give these cells advantage so they grow or proliferate more than vector.

Fulvestrant

Figure 10- Measuring the effect of Fulvestrant on HPMEC and HPMEC ID3$^+$ cells using MTT assay

In both HPMEC and HPMEC ID3$^+$ cells, E2 and PCB153 treatment increased cell growth. HPMEC and HPMEC ID3$^+$ cells exposed to 17β-estradiol and/or PCB153. E2 & PCB153= 100 ng/mL. Graph of cell survival determined by MTT assay 24 h after exposure. All treatment groups showed significant increase in growth b**p<0.01 for HPMEC vs control and a**p<0.01 for HPMEC ID3$^+$ vs control. Fulvestrant (1 µM) is estrogen receptor agonist significantly (P<0.01) inhibit the growth of both HPMEC and HPMEC ID3$^+$ cells compared to control. Co-treatment of E2+Fulvestrant (c**p<0.01), PCB153+Fulvestrant (d**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in HPMEC cells. Co-treatment of E2+ Fulvestrant (e**p<0.01), PCB153+Fulvestrant (f**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in HPMEC ID3$^+$ cells. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.

Figure 11- Measuring the effect of Fulvestrant on HPMEC and HPMEC ID3+ cells using SRB assay

In both HPMEC and HPMEC ID3+ cells, E2 and PCB153 treatment increased the survival and proliferation of cells. HPMEC and HPMEC ID3+ cells exposed to17β-estradiol and/or PCB153. E2 & PCB153= 100 ng/mL. Graph of cell survival determined by SRB assay 24 h after exposure. All the treatment showed significant increase b**p<0.01 for HPMEC vs control and a**p<0.01 for HPMEC ID3+ vs control. Fulvestrant (1 µM) is estrogen receptor agonist significantly (P<0.01) inhibit the growth of both HPMEC and HPMEC ID3+ cells compared to control. Co-treatment of E2+ Fulvestrant (c**p<0.01), PCB153 + Fulvestrant (d**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in HPMEC cells. Co-treatment of E2+ Fulvestrant (e**p<0.01), PCB153 + Fulvestrant (f**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in HPMEC ID3+ cells. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.

Figure 12- Measuring the effect of Fulvestrant on HPMEC and HPMEC ID3+ cells using BrdU assay

In both HPMEC and HPMEC ID3+ cells, E2 and PCB153 treatment increased the survival and proliferation of cells. HPMEC and HPMEC ID3+ cells exposed to17β-estradiol and/or PCB153. E2 & PCB153= 100 ng/mL. Graph of cell survival determined by BrdU assay 24 h after exposure. All the treatment showed significant increase b**p<0.01 for HPMEC vs control and a**p<0.01 for HPMEC ID3+ vs control. Fulvestrant (1 µM) is estrogen receptor agonist significantly (P<0.01) inhibit the growth of both HPMEC and HPMEC ID3+ cells compared to control. Co-treatment of E2 + Fulvestrant (c**p<0.01), PCB153 + Fulvestrant (d**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in HPMEC cells. Co-treatment of E2 + Fulvestrant (e**p<0.01), PCB153 + Fulvestrant (f**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in HPMEC ID3+ cells. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.

Tamoxifen

MTT assay at 48h

Figure 13- Measuring the effect of Tamoxifen on HPMEC and HPMEC ID3[+] cells using MTT assay

In both HPMEC and HPMEC ID3[+] cells, E2 and PCB153 treatment increased the survival and proliferation of cells. HPMEC and HPMEC ID3[+] cells exposed to17β-estradiol and/or PCB153. E2 & PCB153= 100 ng/mL. Graph of cell survival determined by MTT assay 24 h after exposure. All the treatment showed significant increase b**p<0.01 for HPMEC vs control and a**p<0.01 for HPMEC ID3[+] vs control. Tamoxifen (1 µM) is estrogen receptor agonist significantly (P<0.01) inhibit the growth of both HPMEC and HPMEC ID3[+] cells compared to control. Co-treatment of E2+ Tamoxifen (c**p<0.01), PCB153 + Tamoxifen (d**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in HPMEC cells. Co-treatment of E2+ Tamoxifen (e**p<0.01), PCB153 + Tamoxifen (f**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in HPMEC ID3+ cells. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.
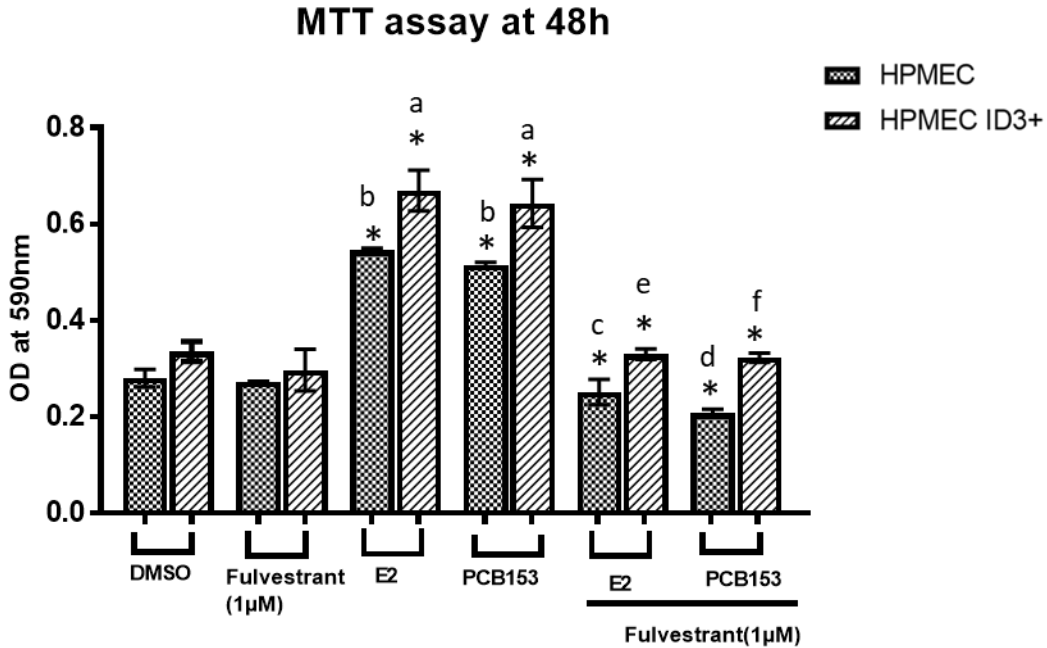
Figure 14- Measuring the effect of Fulvestrant on HPMEC and HPMEC ID3+ cells using SRB assay

In both HPMEC and HPMEC ID3+ cells, E2 and PCB153 treatment increased the survival and proliferation of cells. HPMEC and HPMEC ID3+ cells exposed to17β-estradiol and/or PCB153. E2 & PCB153= 100 ng/mL. Graph of cell survival determined by SRB assay 24 h after exposure. All the treatment showed significant increase b**p<0.01 for HPMEC vs control and a**p<0.01 for HPMEC ID3+ vs control. Tamoxifen (1 µM) is estrogen receptor agonist significantly (P<0.01) inhibit the growth of both HPMEC and HPMEC ID3+ cells compared to control. Co-treatment of E2+ Tamoxifen (c**p<0.01), PCB153 + Tamoxifen (d**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in HPMEC cells. Co-treatment of E2+ Tamoxifen (e**p<0.01), PCB153 + Tamoxifen (f**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in HPMEC ID3+ cells. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.
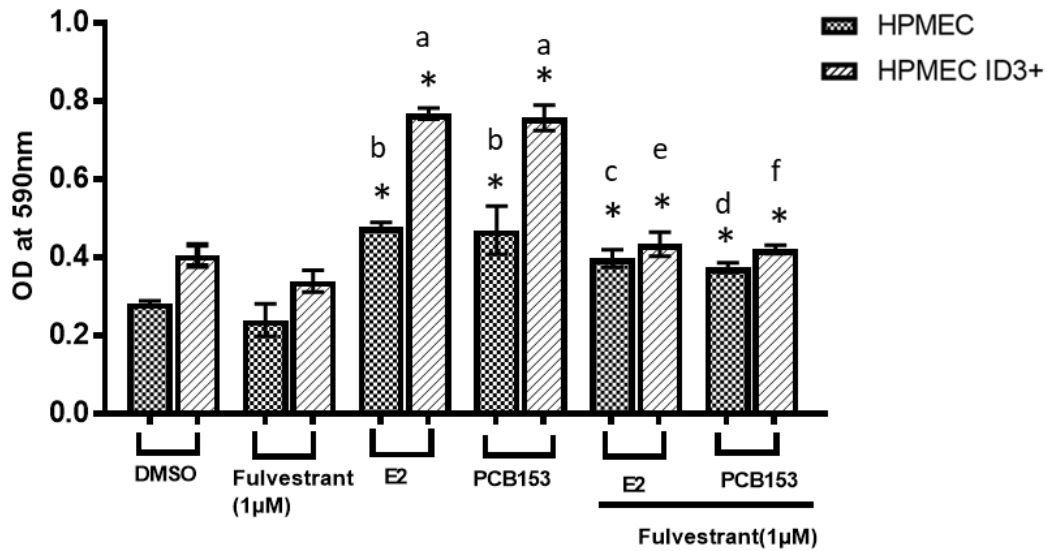
Figure 15- Measuring the effect of Fulvestrant on HPMEC and HPMEC ID3+ cells using BrdU assay

In both HPMEC and HPMEC ID3+ cells, E2 and PCB153 treatment increased the survival and proliferation of cells. HPMEC and HPMEC ID3+ cells exposed to17β-estradiol and/or PCB153. E2 & PCB153= 100 ng/mL. Graph of cell survival determined by BrdU assay 24 h after exposure. All the treatment showed significant increase b**p<0.01 for HPMEC vs control and a**p<0.01 for HPMEC ID3+ vs control. Tamoxifen (1 µM) is estrogen receptor agonist significantly (P<0.01) inhibit the growth of both HPMEC and HPMEC ID3+ cells compared to control. Co-treatment of E2+ Tamoxifen (c**p<0.01), PCB153 + Tamoxifen (d**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in HPMEC cells. Co-treatment of E2+ Tamoxifen (e**p<0.01), PCB153 + Tamoxifen (f**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in HPMEC ID3+ cells. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.

Ebselen

Figure 16- Measuring the effect of Ebselen on HPMEC and HPMEC ID3[+] cells using MTT assay

In both HPMEC and HPMEC ID3[+] cells, E2 and PCB153 treatment increased the survival and proliferation of cells. HPMEC and HPMEC ID3[+] cells exposed to 17β-estradiol and/or PCB153. E2 & PCB153= 100 ng/mL. Graph of cell survival determined by MTT assay 24 h after exposure. All the treatment showed significant increase b**$p<0.01$ for HPMEC vs control and a**$p<0.01$ for HPMEC ID3[+] vs control. Ebselen (1 μM) is estrogen receptor agonist significantly ($P<0.01$) inhibit the growth of both HPMEC and HPMEC ID3[+] cells compared to control. Co-treatment of E2+ Ebselen (c**$p<0.01$), PCB153 + Ebselen (d**$p<0.01$) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in HPMEC cells. Co-treatment of E2 + Ebselen (e**$p<0.01$), PCB153 + Ebselen (f**$p<0.01$) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in HPMEC ID3[+] cells. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.
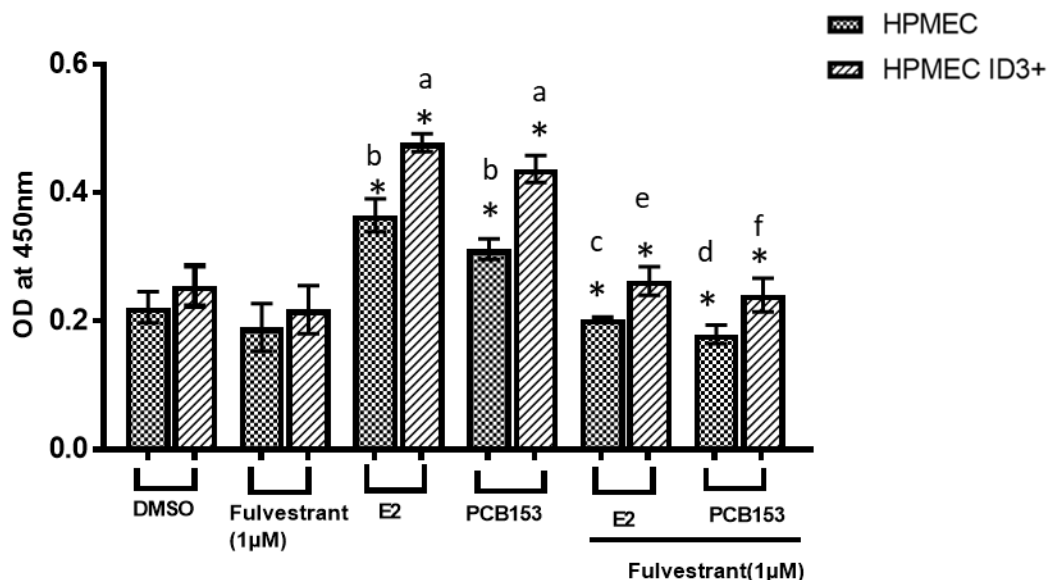
Figure 17- Measuring the effect of Ebselen on HPMEC and HPMEC ID3$^+$ cells using SRB assay

In both HPMEC and HPMEC ID3$^+$ cells, E2 and PCB153 treatment increased the survival and proliferation of cells. HPMEC and HPMEC ID3$^+$ cells exposed to17β-estradiol and/or PCB153. E2 & PCB153= 100 ng/mL. Graph of cell survival determined by SRB assay 24 h after exposure. All the treatment showed significant increase b**p<0.01 for HPMEC vs control and a**p<0.01 for HPMEC ID3$^+$ vs control. Ebselen (1 μM) is estrogen receptor agonist significantly (P<0.01) inhibit the growth of both HPMEC and HPMEC ID3$^+$ cells compared to control. Co-treatment of E2+ Ebselen (c**p<0.01), PCB153 + Ebselen (d**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100ng/mL) respectively in HPMEC cells. Co-treatment of E2 + Ebselen (e**p<0.01), PCB153 + Ebselen (f**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in HPMEC ID3$^+$ cells. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.
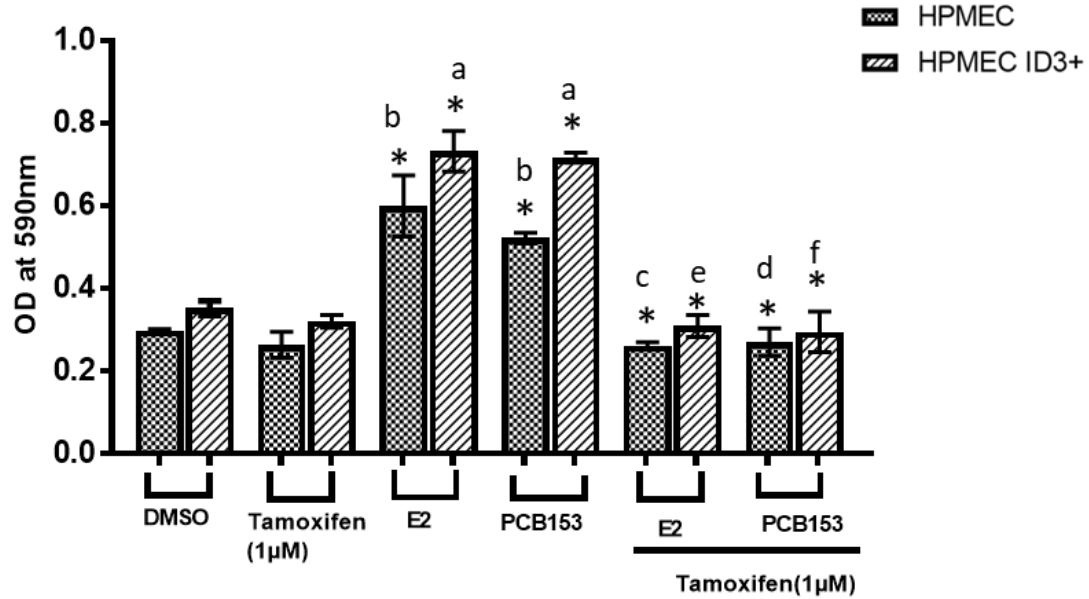
64

Figure 18- Measuring the effect of Ebselen on HPMEC and HPMEC ID3+ cells using BrdU assay

In both HPMEC and HPMEC ID3+ cells, E2 and PCB153 treatment increased the survival and proliferation of cells. HPMEC and HPMEC ID3+ cells exposed to17β-estradiol and/or PCB153. E2 & PCB153= 100 ng/mL. Graph of cell survival determined by BrdU assay 24 h after exposure. All the treatment showed significant increase b**p<0.01 for HPMEC vs control and a**p<0.01 for HPMEC ID3+ vs control. Ebselen (1 µM) is estrogen receptor agonist significantly (P<0.01) inhibit the growth of both HPMEC and HPMEC ID3+ cells compared to control. Co-treatment of E2+ Ebselen (c**p<0.01), PCB153 + Ebselen (d**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in HPMEC cells. Co-treatment of E2 + Ebselen (e**p<0.01), PCB153 + Ebselen (f**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in HPMEC ID3+ cells. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.
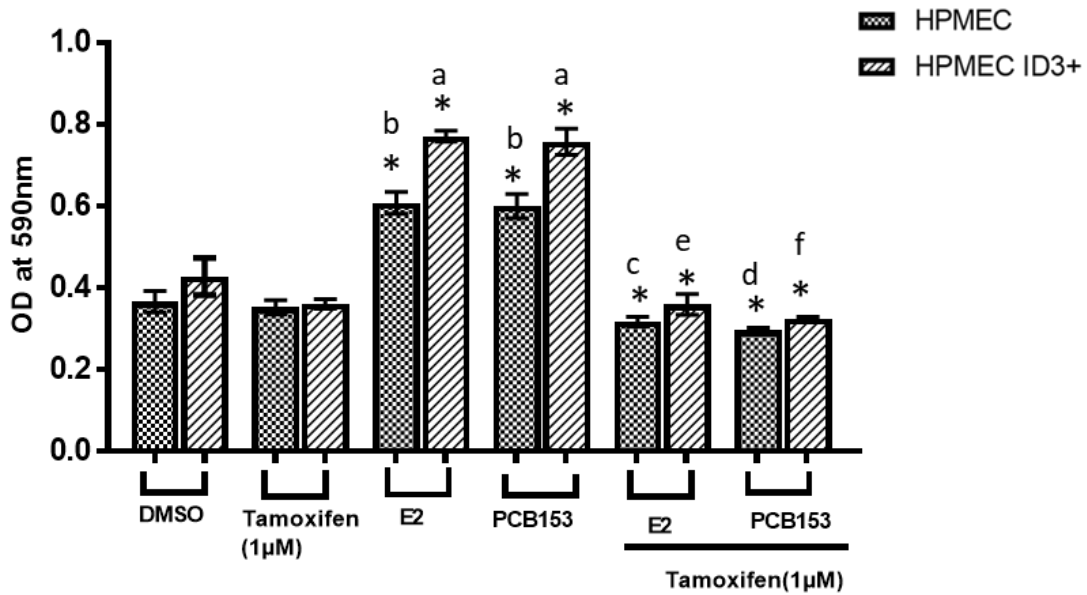
Smooth muscle cells (SMC)
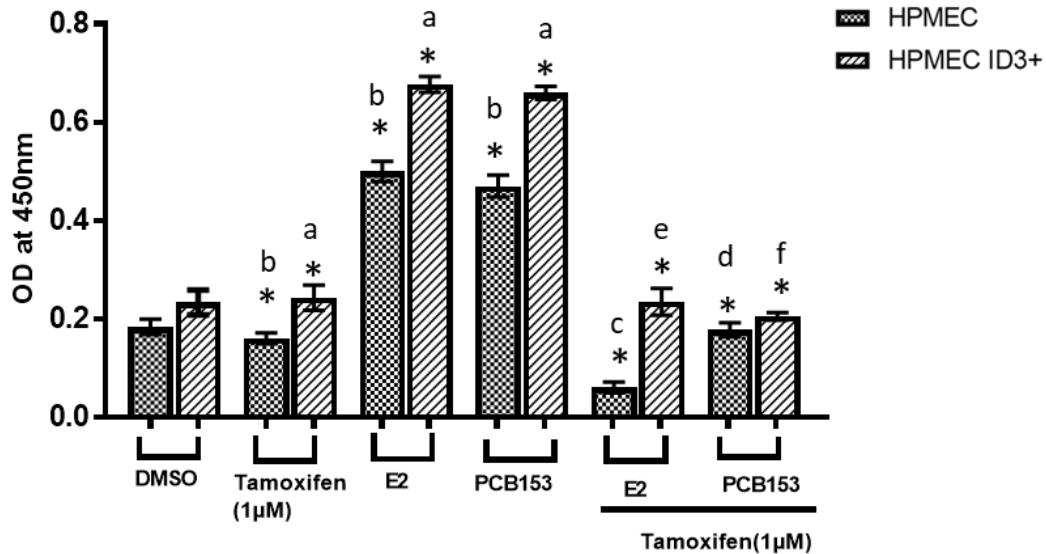
Researchers have shown that presence of ERs in SMC and involvement of SMCs in plexiform lesions and atherosclerotic plaque along with lung endothelial cells (Bennett, 1999; Doran, Meller, & McNamara, 2008; Rosenthal, 2001). We evaluated the combination effect of E2 and PCB153 on SMC using MTT, SRB and BrdU assays. Similar to lung endothelial cells, we seeded SMCs in a 96 well plate and gave treatments E2 (100 ng/mL), PCB153 (100 ng/mL). The results showed that all treatments induced significant (P<0.01) increase in numbers of SMCs compared to control. We also checked the effects of tamoxifen, Ebselen and Fulvestrant treatment SMCs. We found out that tamoxifen, Ebselen and Fulvestrant treatments significantly (P<0.01) inhibit the growth of SMCs compare to control as wells as in treatments E2 (100 ng/mL) and PCB153 (100 ng/mL). Given the biological significance of ID3 proteins for the maintenance of stemness, we questioned whether overexpression of ID3 could induce a molecular stem-cell like a signature in adult lung SMCs cell line. To direct this question, human SMCs were transduced using an MOI of 25 with either the Precision LentiORF for ID3 overexpression or control empty lentiviral vector pLEX-JRED/TurboGFP according to the manufacturer's instructions. Cells were then seeded at clonal density and selected with blasticidin S (5 μg/mL). Cells that overexpress ID3 will be selected for culture and referred to as SMC ID3[+] for Ebselen as shown in Figure 20, 21 and 22. Similarly, for tamoxifen (as shown in Figure 23, 24 and 25) and Fulvestrant (as shown in Figure 26, 27 and 28).

Phase Contrast
Image of SMC ID3+

Confocal microscope
Image of SMC ID3+

Figure 19- ID3 overexpression in Lung SMC

SMC cells were stably transfected with Precision LentiORF for ID3 (Thermo Scientific Open Biosystems). Selected cells that overexpressed ID3 with blasticidin S (5 mg/ml) as per manufacturer's instructions. Cells expressing TurboGFP in SMC ID3$^+$ cells were identified by fluorescence microscopy. We also shown phase contrast image of SMC ID3+ cells along with it.
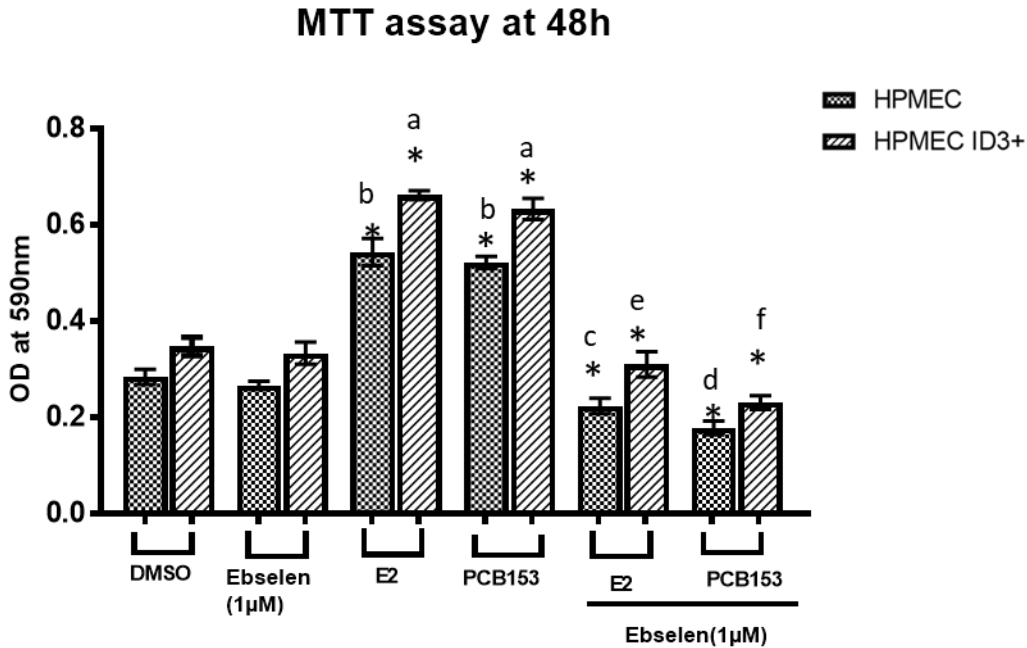
Ebselen

MTT assay at 48h

Figure 20- Measuring the effect of Ebselen on SMC and SMC ID3[+] cells using MTT assay

In both SMC and SMC ID3[+] cells, E2 and PCB153 treatment increased the survival and proliferation of cells. SMC and SMC ID3[+] cells exposed to17β-estradiol and/or PCB153. E2 & PCB153= 100 ng/mL. Graph of cell survival determined by MTT assay 24 h after exposure. All the treatment showed significant increase b**p<0.01 for SMC vs control and a**p<0.01 for SMC ID3[+] vs control. Ebselen (1 μM) is estrogen receptor agonist significantly (P<0.01) inhibit the growth of both SMC and SMC ID3[+] cells compared to control. Co-treatment of E2 + Ebselen (c**p<0.01), PCB153 + Ebselen (d**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in SMC cells. Co-treatment of E2 + Ebselen (e**p<0.01), PCB153 + Ebselen (f**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in SMC ID3[+] cells. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.
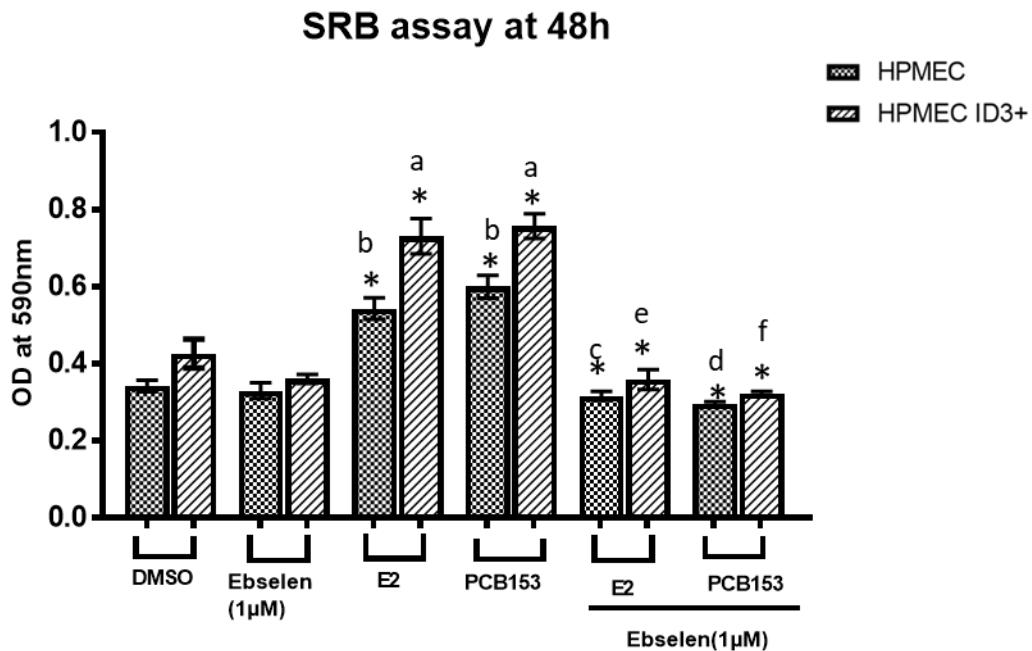
Figure 21- Measuring the effect of Ebselen on SMC and SMC ID3$^+$ cells using SRB assay

In both SMC and SMC ID3$^+$ cells, E2 and PCB153 treatment increased the survival and proliferation of cells. SMC and SMC ID3$^+$ cells exposed to17β-estradiol and/or PCB153. E2 & PCB153= 100 ng/mL. Graph of cell survival determined by SRB assay 24 h after exposure. All the treatment showed significant increase b**p<0.01 for SMC vs control and a**p<0.01 for SMC ID3$^+$ vs control. Ebselen (1 µM) is estrogen receptor agonist significantly (P<0.01) inhibit the growth of both SMC and SMC ID3$^+$ cells compared to control. Co-treatment of E2 + Ebselen (c**p<0.01), PCB153 + Ebselen (d**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in SMC cells. Co-treatment of E2+ Ebselen (e**p<0.01), PCB153 + Ebselen (f**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in SMC ID3$^+$ cells. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.
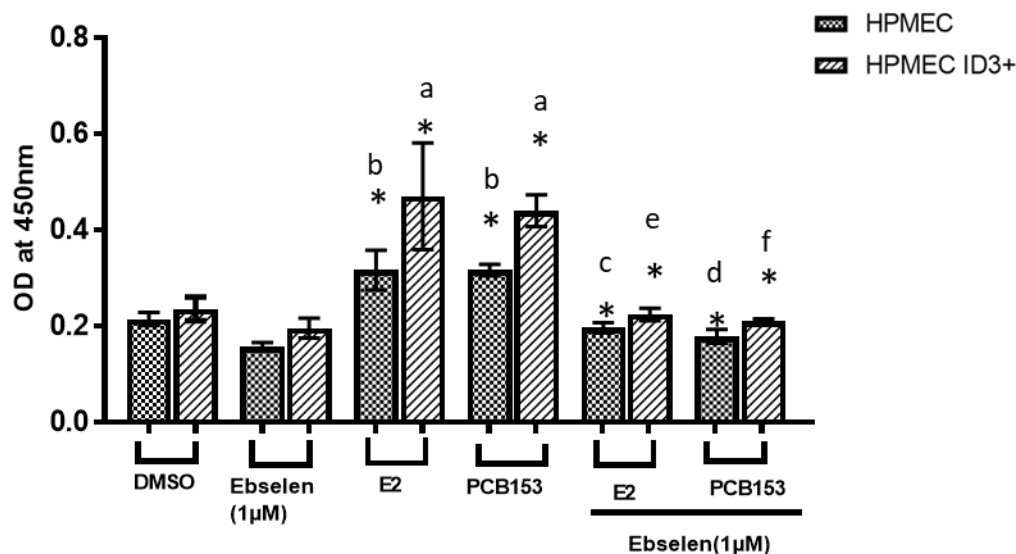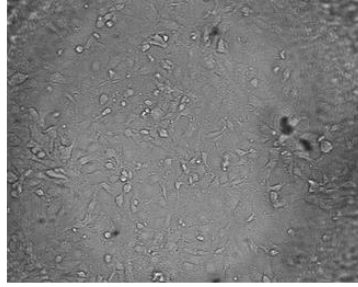
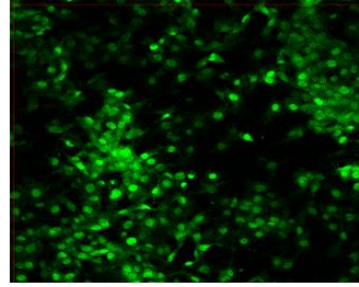**Figure 22- Measuring the effect of Ebselen on SMC and SMC ID3+ cells using BrdU assay**

In both SMC and SMC ID3+ cells, E2 and PCB153 treatment increased the survival and proliferation of cells. SMC and SMC ID3+ cells exposed to17β-estradiol and/or PCB153. E2 & PCB153= 100 ng/mL. Graph of cell survival determined by BrdU assay 24 h after exposure. All the treatment showed significant increase b**p<0.01 for SMC vs control and a**p<0.01 for SMC ID3+ vs control. Ebselen (1 µM) is estrogen receptor agonist significantly (P<0.01) inhibit the growth of both SMC and SMC ID3+ cells compared to control. Co-treatment of E2+ Ebselen (c**p<0.01), PCB153 + Ebselen (d**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in SMC cells. Co-treatment of E2+ Ebselen (e**p<0.01), PCB153 + Ebselen (f**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in SMC ID3+ cells. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.

Tamoxifen

MTT assay at 48h

Figure 23- Measuring the effect of Tamoxifen on SMC and SMC ID3[+] cells using MTT assay

In both SMC and SMC ID3[+] cells, E2 and PCB153 treatment increased the survival and proliferation of cells. SMC and SMC ID3[+] cells exposed to17β-estradiol and/or PCB153. E2 & PCB153= 100 ng/mL. Graph of cell survival determined by MTT assay 24 h after exposure. All the treatment showed significant increase b**p<0.01 for SMC vs control and a**p<0.01 for SMC ID3[+] vs control. Tamoxifen (1 μM) is estrogen receptor agonist significantly (P<0.01) inhibit the growth of both SMC and SMC ID3[+] cells compared to control. Co-treatment of E2+ Tamoxifen (c**p<0.01), PCB153 + Tamoxifen (d**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in SMC cells. Co-treatment of E2+ Tamoxifen (e**p<0.01), PCB153 + Tamoxifen (f**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in SMC ID3[+] cells. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.

Figure 24- Measuring the effect of Tamoxifen on SMC and SMC ID3$^+$ cells using SRB assay

In both SMC and SMC ID3$^+$ cells, E2 and PCB153 treatment increased the survival and proliferation of cells. SMC and SMC ID3$^+$ cells exposed to17β-estradiol and/or PCB153. E2 & PCB153= 100 ng/mL. Graph of cell survival determined by SRB assay 24 h after exposure. All the treatment showed significant increase b**p<0.01 for SMC vs control and a**p<0.01 for SMC ID3$^+$ vs control. Tamoxifen (1 µM) is estrogen receptor agonist significantly (P<0.01) inhibit the growth of both SMC and SMC ID3$^+$ cells compared to control. Co-treatment of E2+ Tamoxifen (c**p<0.01), PCB153 + Tamoxifen (d**p<0.01) showed significant reduction in growth of cells compared to E2 (100ng/mL), PCB153 (100 ng/mL) respectively in SMC cells. Co-treatment of E2+ Tamoxifen (e**p<0.01), PCB153 + Tamoxifen (f**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in SMC ID3$^+$ cells. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.
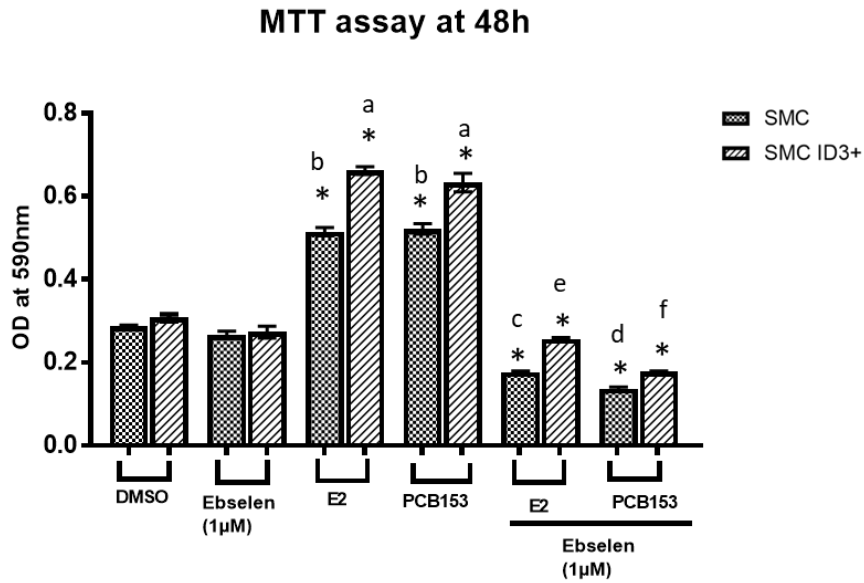
Figure 25- Measuring the effect of Tamoxifen on SMC and SMC ID3$^+$ cells using BrdU assay

In both SMC and SMC ID3$^+$ cells, E2 and PCB153 treatment increased the survival and proliferation of cells. SMC and SMC ID3$^+$ cells exposed to 17β-estradiol and/or PCB153. E2 & PCB153= 100 ng/mL. Graph of cell survival determined by BrdU assay 24 h after exposure. All the treatment showed significant increase b**p<0.01 for SMC vs control and a**p<0.01 for SMC ID3$^+$ vs control. Tamoxifen (1 μM) is estrogen receptor agonist significantly (P<0.01) inhibit the growth of both SMC and SMC ID3$^+$ cells compared to control. Co-treatment of E2+ Tamoxifen (c**p<0.01), PCB153 + Tamoxifen (d**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in SMC cells. Co-treatment of E2+ Tamoxifen (e**p<0.01), PCB153 + Tamoxifen (f**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in SMC ID3$^+$ cells. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.
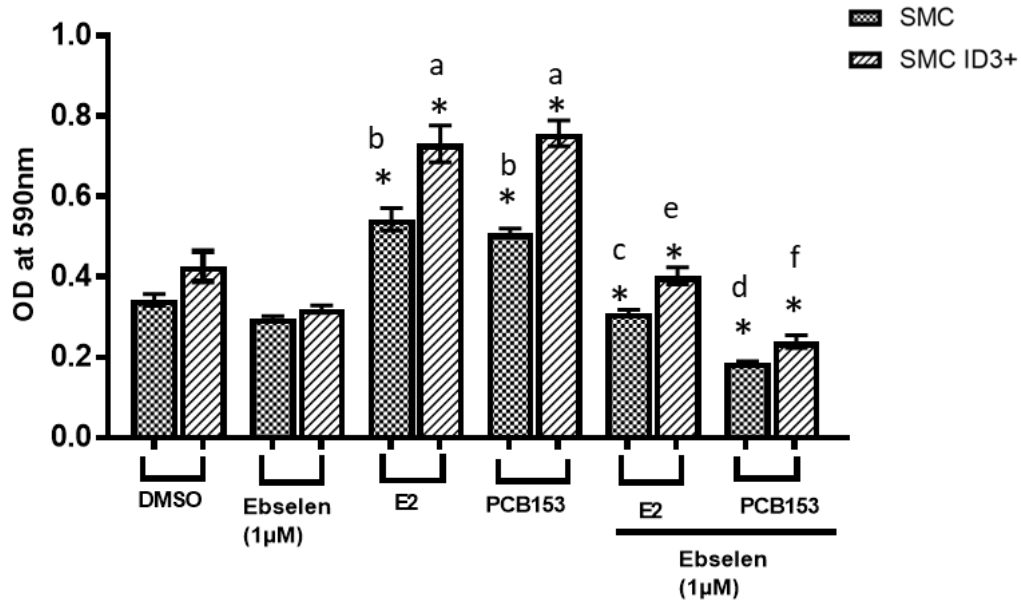
73

Fulvestrant

MTT assay at 48h

Figure 26- Measuring the effect of Fulvestrant on SMC and SMC ID3$^+$ cells using MTT assay

In both SMC and SMC ID3$^+$ cells, E2 and PCB153 treatment increased the survival and proliferation of cells. SMC and SMC ID3$^+$ cells exposed to17β-estradiol and/or PCB153. E2 & PCB153= 100 ng/mL. Graph of cell survival determined by MTT assay 24 h after exposure. All the treatment showed significant increase b**p<0.01 for SMC vs control and a**p<0.01 for SMC ID3$^+$ vs control. Fulvestrant (1 μM) is estrogen receptor agonist significantly (P<0.01) inhibit the growth of both SMC and SMC ID3$^+$ cells compared to control. Co-treatment of E2+ Fulvestrant (c**p<0.01), PCB153 + Fulvestrant (d**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in SMC cells. Co-treatment of E2+ Fulvestrant (e**p<0.01), PCB153 + Fulvestrant (f**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in SMC ID3$^+$ cells. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.

Figure 27- Measuring the effect of Fulvestrant on SMC and SMC ID3$^+$ cells using SRB assay

In both SMC and SMC ID3$^+$ cells, E2 and PCB153 treatment increased the survival and proliferation of cells. SMC and SMC ID3$^+$ cells exposed to17β-estradiol and/or PCB153. E2 & PCB153= 100 ng/mL. Graph of cell survival determined by SRB assay 24 h after exposure. All the treatment showed significant increase b**p<0.01 for SMC vs control and a**p<0.01 for SMC ID3$^+$ vs control. Fulvestrant (1 μM) is estrogen receptor agonist significantly (P<0.01) inhibit the growth of both SMC and SMC ID3$^+$ cells compared to control. Co-treatment of E2+ Fulvestrant (c**p<0.01), PCB153 + Fulvestrant (d**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in SMC cells. Co-treatment of E2+ Fulvestrant (e**p<0.01), PCB153 + Fulvestrant (f**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in SMC ID3$^+$ cells. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.
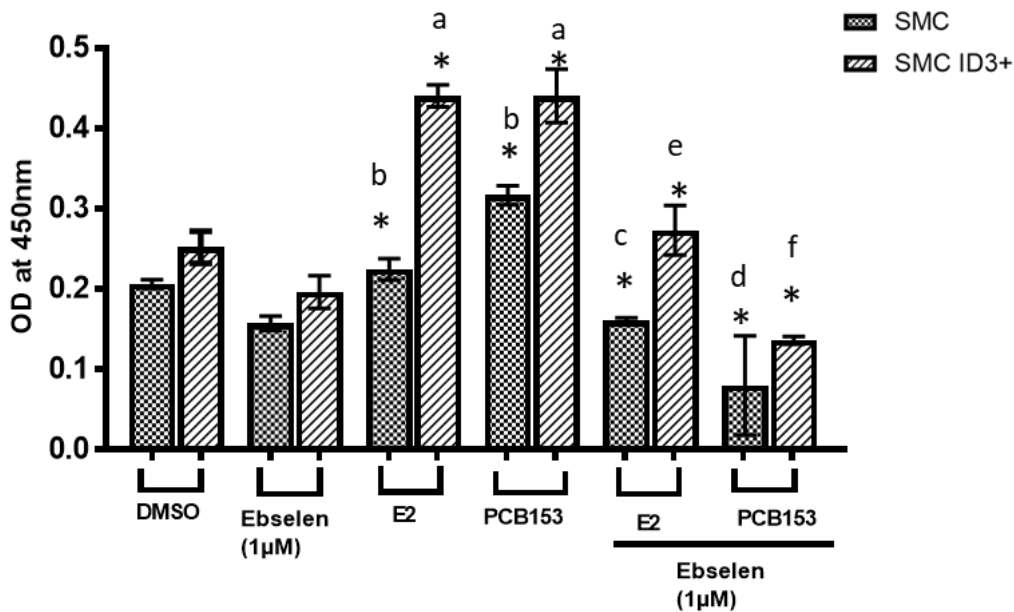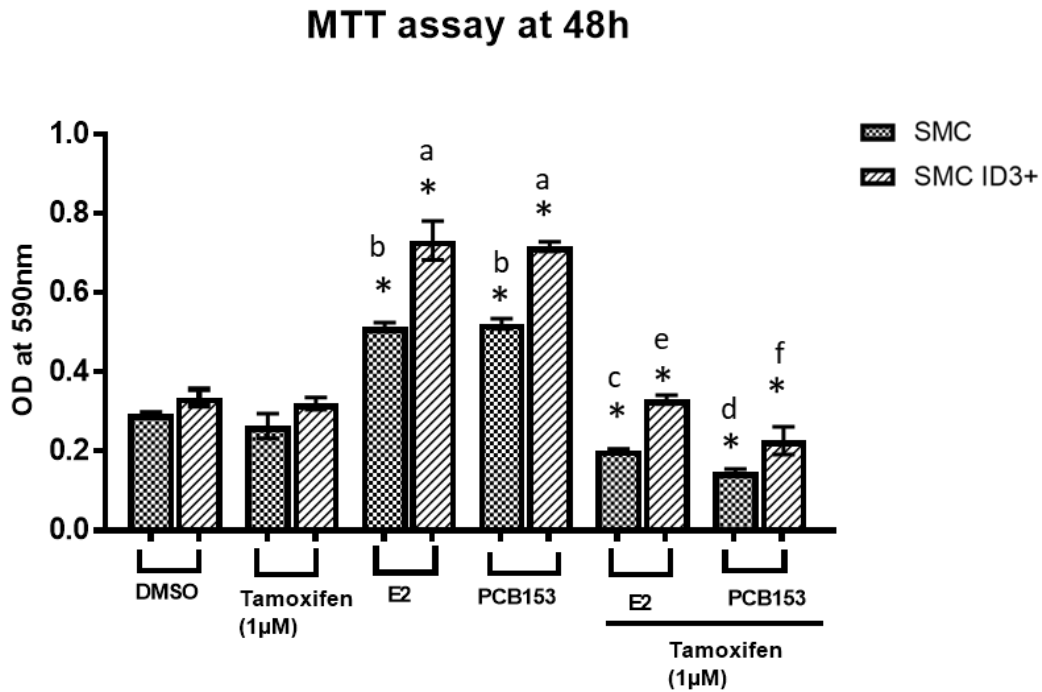
BrdU assay at 48h

Figure 28- Measuring the effect of Fulvestrant on SMC and SMC ID3⁺ cells using BrdU assay

In both SMC and SMC ID3$^+$ cells, E2 and PCB153 treatment increased the survival and proliferation of cells. SMC and SMC ID3$^+$ cells exposed to17β-estradiol and/or PCB153. E2 & PCB153= 100 ng/mL. Graph of cell survival determined by BrdU assay 24 h after exposure. All the treatment showed significant increase b**p<0.01 for SMC vs control and a**p<0.01 for SMC ID3$^+$ vs control. Fulvestrant (1 μM) is estrogen receptor agonist significantly (P<0.01) inhibit the growth of both SMC and SMC ID3$^+$ cells compared to control. Co-treatment of E2+ Fulvestrant (c**p<0.01), PCB153 + Fulvestrant (d**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in SMC cells. Co-treatment of E2+ Fulvestrant (e**p<0.01), PCB153 + Fulvestrant (f**p<0.01) showed significant reduction in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in SMC ID3$^+$ cells. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.
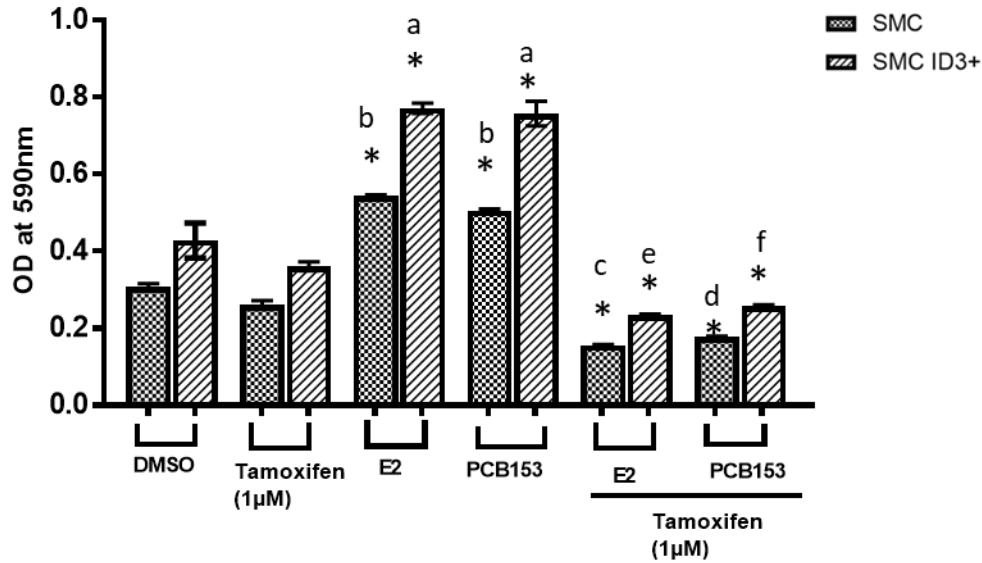
SU5416 treatment to HPMEC ID3$^+$ and SMCs

Researchers have shown *in vivo* and *in vitro* studies of exposure to SU5416 reported to produce endothelial stem cell markers in lung endothelial cells which give rise to lung vascular lesions (Das et al., 2015). In addition to this, many PAH studies showed the upregulated levels of ID3 protein in plexiform lesions (Yu, Wilson, Taylor, & Polgar, 2015). Therefore, we compared the endothelial types HPMEC and HPMEC ID3$^+$ to determine the effect of SU5416 exposure which is a chemical inducer of molecular stem cell-like signature. We tested the lung endothelial cell proliferation in presence of estrogenic chemicals and SU5416 by MTT, SRB and BrdU assay. We treated HPMEC and HPMEC ID3$^+$ cells with single treatment of SU5416 (1 µM) dose as wells as co-treatment with E2 + SU5416, PCB153 + SU5416. We carried out MTT, SRB and BrdU assays and analyzed the data. We found that SU5416 (5µM) significantly (P<0.01) induce the growth of cells in both the HPMEC and HPMEC ID3+ cell lines compare to control. In addition to this, our data showed that co-treatment of E2 + SU5416, PCB153 + SU5416, showed significant (P<0.01) increase in growth of cells compared to E2 (100 ng/mL) and PCB153 (100 ng/mL) respectively as shown in Figure 29, 30 and 31.

Figure 29- Measuring the effect of SU5416 on HPMEC and HPMEC ID3[+] cells using MTT assay

In both HPMEC and HPMEC ID3[+] cells, E2 and PCB153 treatment increased the survival and proliferation of cells. HPMEC and HPMEC ID3[+] cells exposed to17β-estradiol and/or PCB153. E2 & PCB153= 100 ng/mL. Graph of cell survival determined by MTT assay 24 h after exposure. All the treatment showed significant increase b**p<0.01 for HPMEC vs control and a**p<0.01 for HPMEC ID3[+] vs control. SU5416 (1 μM) is VEGF receptor antagonist. Co-treatment of E2+ SU5416 (c**p<0.01), PCB153 + SU5416 (d**p<0.01) showed significant increase in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in HPMEC cells. Co-treatment of E2+ SU5416 (f**p<0.01), PCB153 + SU5416 (g**p<0.01), showed significant increase in growth of cells compared to E2 (100 ng/mL) and PCB153 (100 ng/mL) respectively in HPMEC ID3[+] cells. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.
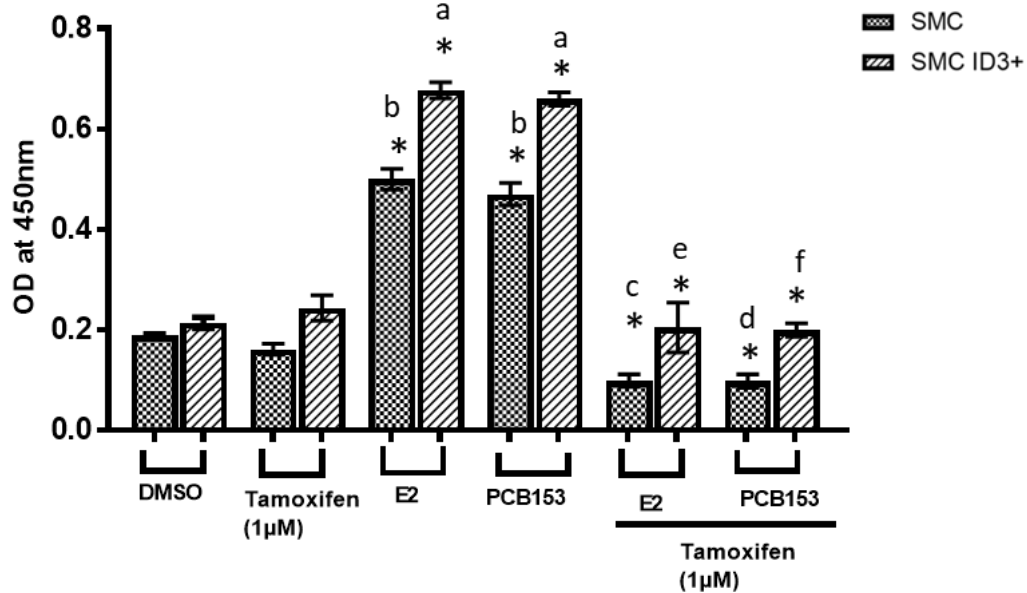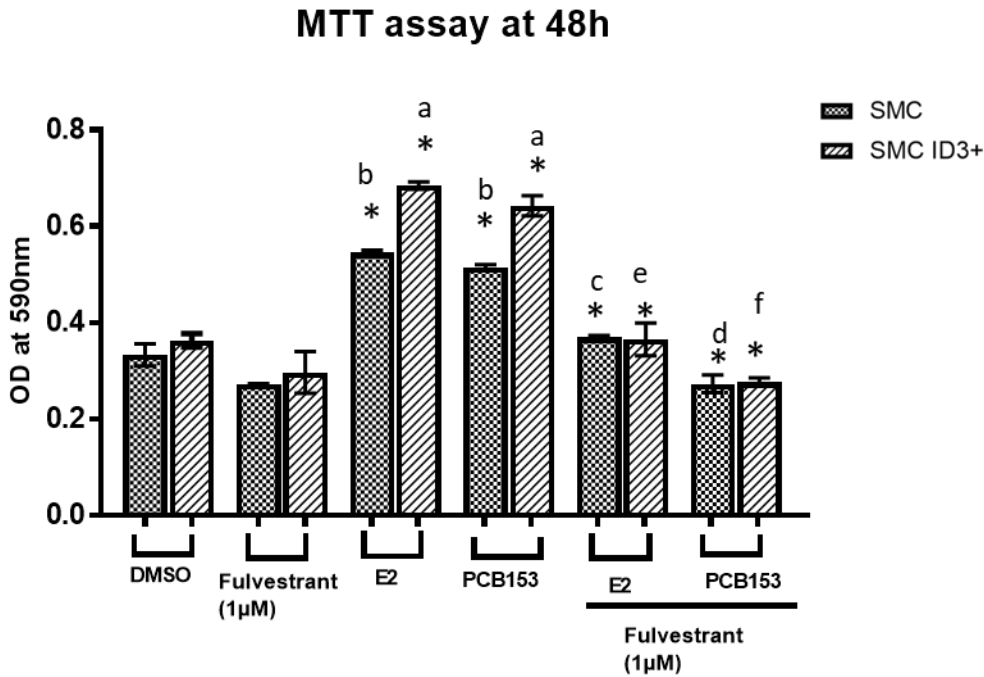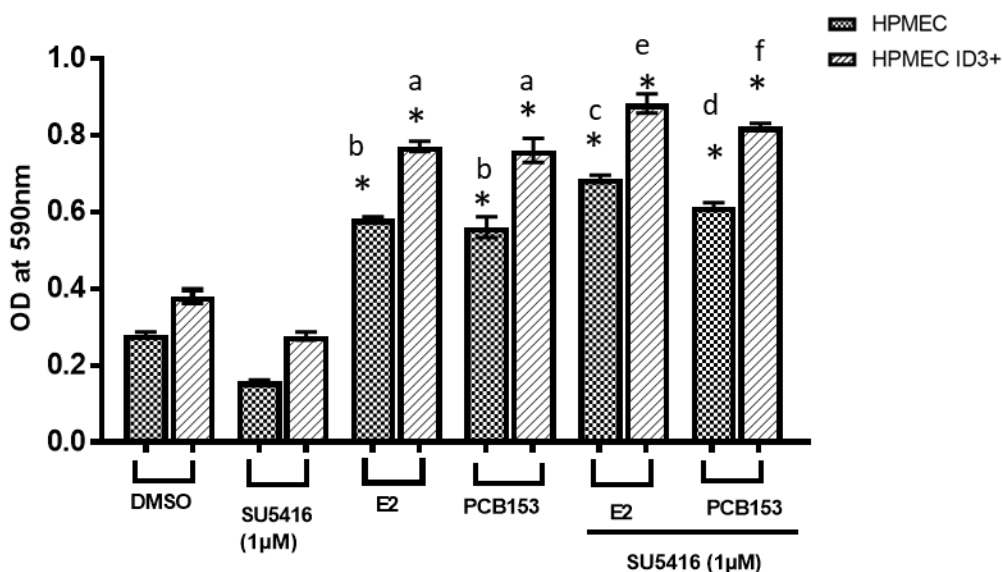
Figure 30- Measuring the effect of SU5416 on HPMEC and HPMEC ID3[+] cells using SRB assay

In both HPMEC and HPMEC ID3[+] cells, E2 and PCB153 treatment increased the survival and proliferation of cells. HPMEC and HPMEC ID3[+] cells exposed to17β-estradiol and/or PCB153. E2 & PCB153= 100ng/mL. Graph of cell survival determined by SRB assay 24 h after exposure. All the treatment showed significant increase b**p<0.01 for HPMEC vs control and a**p<0.01 for HPMEC ID3+ vs control. SU5416 (1 μM) is VEGF receptor antagonist. Co-treatment of E2+ SU5416 (c**p<0.01), PCB153 + SU5416 (d**p<0.01) showed significant increase in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in HPMEC cells. Co-treatment of E2+ SU5416 (f**p<0.01), PCB153 + SU5416 (g**p<0.01), E2+PCB153 + SU5416 (h**p<0.01) showed significant increase in growth of cells compared to E2 (100 ng/mL) and PCB153 (100 ng/mL) respectively in HPMEC ID3+ cells. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.

BrdU assay at 48h

Figure 31- Measuring the effect of SU5416 on HPMEC and HPMEC ID3+ cells using BrdU assay

In both HPMEC and HPMEC ID3+ cells, E2 and PCB153 treatment increased the survival and proliferation of cells. HPMEC and HPMEC ID3+ cells exposed to17β-estradiol and/or PCB153. E2 & PCB153= 100 ng/mL. Graph of cell survival determined by BrdU assay 24 h after exposure. All the treatment showed significant increase b**p<0.01 for HPMEC vs control and a**p<0.01 for HPMEC ID3+ vs control. SU5416 (1 μM) is VEGF receptor antagonist. Co-treatment of E2+ SU5416 (c**p<0.01), PCB153 + SU5416 (d**p<0.01) showed significant increase in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in HPMEC cells. Co-treatment of E2+ SU5416 (f**p<0.01), PCB153 + SU5416 (g**p<0.01), E2+PCB153 + SU5416 (h**p<0.01) showed significant increase in growth of cells compared to E2 (100 ng/mL) and PCB153 (100 ng/mL) respectively in HPMEC ID3+ cells. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.
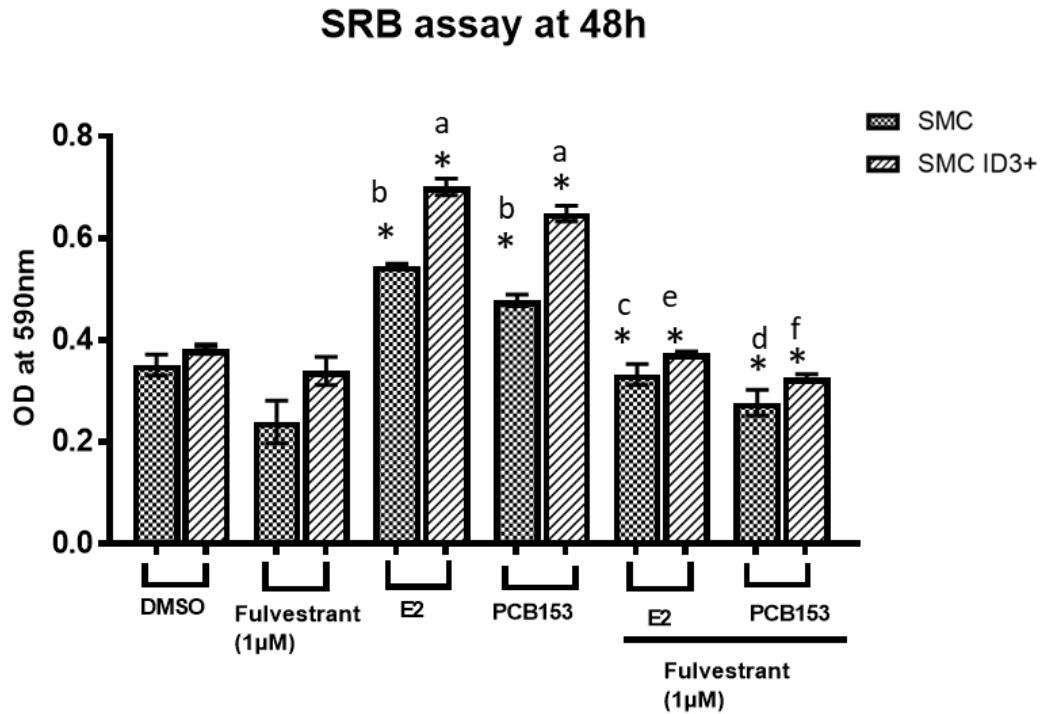
Similarly, we tested the effect of SU5416 drug effect on SMCs. We found that SU5416 (5 μM) significantly (P<0.01) induce the growth of cells in the SMCs compare to control. In addition to this, our data showed that co-treatment of E2 + SU5416, PCB153 + SU5416, showed significant (P<0.01) increase in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively as shown in Figure 32, 33 and 34.

SU5416

MTT assay at 48h
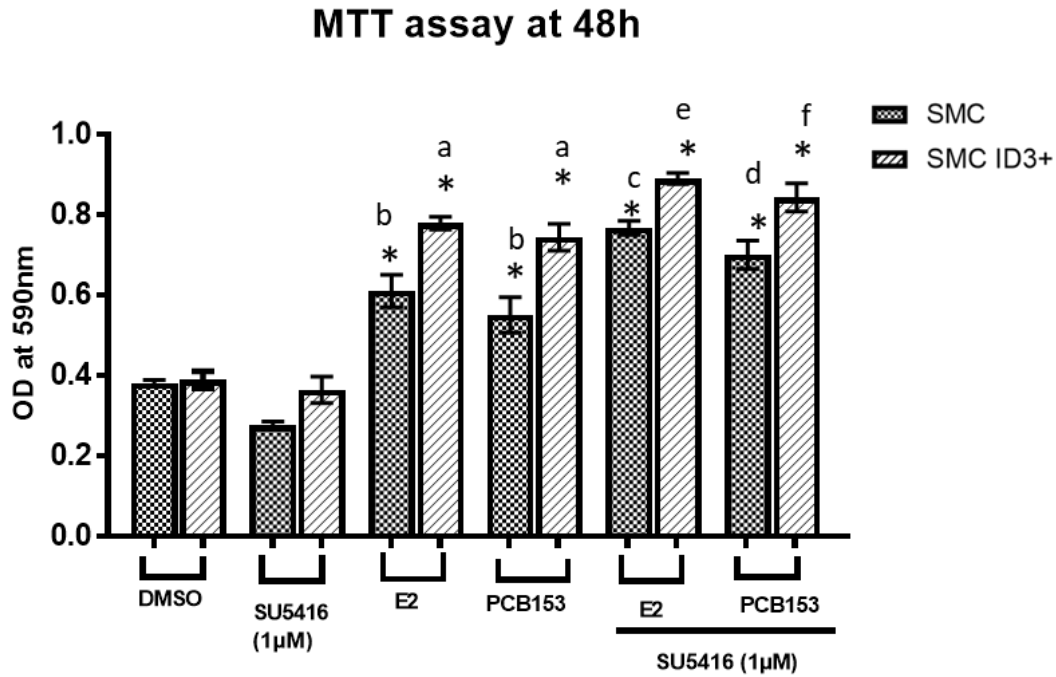
Figure 32- Measuring the effect of SU5416 on SMC and SMC ID3⁺ cells using MTT assay

In both SMC and SMC ID3⁺ cells, E2 and PCB153 treatment increased the survival and proliferation of cells. SMC and SMC ID3⁺ cells exposed to17β-estradiol and/or PCB153E2 & PCB153= 100 ng/mL. Graph of cell survival determined by MTT assay 24 h after exposure. All the treatment showed significant increase b**p<0.01 for SMC vs control and a**p<0.01 for SMC ID3⁺ vs control. SU5416 (1 µM) is VEGF receptor antagonist. Co-treatment of E2+ SU5416 (c**p<0.01), PCB153 + SU5416 (d**p<0.01) showed significant increase in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in SMC cells. Co-treatment of E2+ SU5416 (f**p<0.01), PCB153 + SU5416 (g**p<0.01), E2+PCB153 + SU5416 (h**p<0.01) showed significant increase in growth of cells compared to E2 (100 ng/mL) and PCB153 (100 ng/mL) respectively in SMC ID3⁺ cells. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.

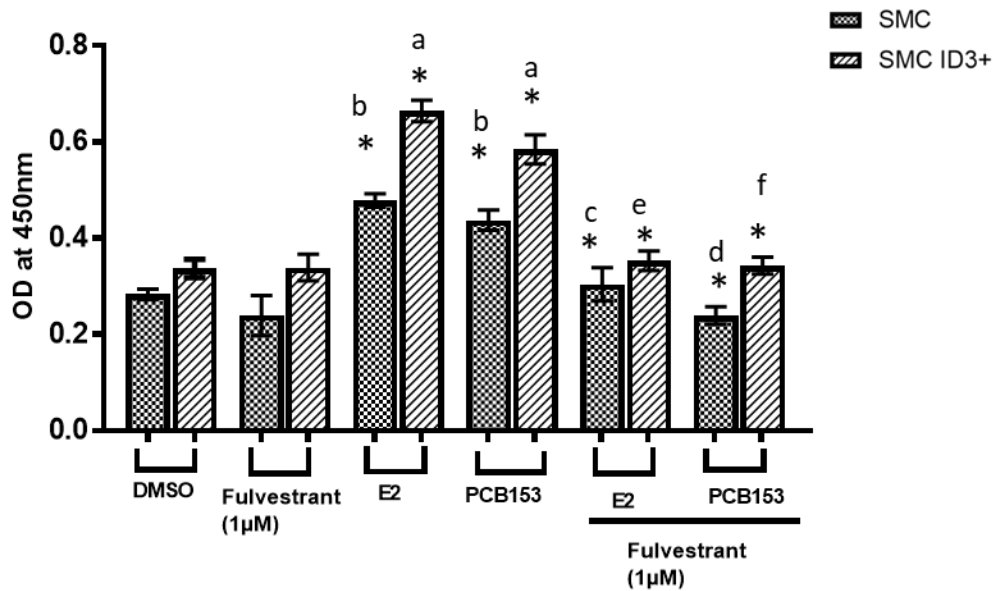**Figure 33- Measuring the effect of SU5416 on SMC and SMC ID3⁺ cells using SRB assay**

In both SMC and SMC ID3⁺ cells, E2 and PCB153 treatment increased the survival and proliferation of cells. SMC and SMC ID3⁺ cells exposed to17β-estradiol and/or PCB153. E2 & PCB153= 100 ng/mL. Graph of cell survival determined by SRB assay 24 h after exposure. All the treatment showed significant increase b**p<0.01 for SMC vs control and a**p<0.01 for SMC ID3⁺ vs control. SU5416 (1μM) is VEGF receptor antagonist. Co-treatment of E2+ SU5416 (c**p<0.01), PCB153 + SU5416 (d**p<0.01) showed significant increase in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in SMC cells. Co-treatment of E2+ SU5416 (f**p<0.01), PCB153 + SU5416 (g**p<0.01), E2+PCB153 + SU5416 (h**p<0.01) showed significant increase in growth of cells compared to E2 (100 ng/mL) and PCB153 (100 ng/mL) respectively in SMC ID3⁺ cells. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.
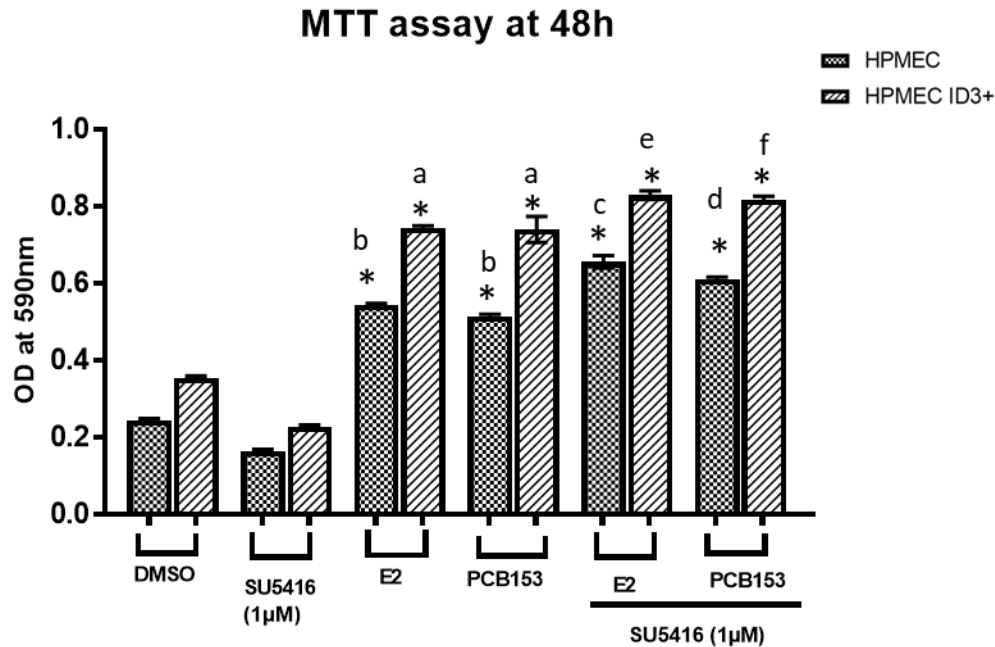
Figure 34- Measuring the effect of SU5416 on SMC and SMC ID3$^+$ cells using BrdU assay

In both SMC and SMC ID3$^+$ cells, E2 and PCB153 treatment increased the survival and proliferation of cells. SMC and SMC ID3$^+$ cells exposed to 17β-estradiol and/or PCB153. E2 & PCB153= 100 ng/mL. Graph of cell survival determined by BrdU assay 24 h after exposure. All the treatment showed significant increase b**p<0.01 for SMC vs control and a**p<0.01 for SMC ID3$^+$ vs control. SU5416 (1μM) is VEGF receptor antagonist. Co-treatment of E2+ SU5416 (c**p<0.01), PCB153 + SU5416 (d**p<0.01) showed significant increase in growth of cells compared to E2 (100 ng/mL), PCB153 (100 ng/mL) respectively in SMC cells. Co-treatment of E2+ SU5416 (f**p<0.01), PCB153 + SU5416 (g**p<0.01), E2+PCB153 + SU5416 (h**p<0.01) showed significant increase in growth of cells compared to E2 (100 ng/mL) and PCB153 (100 ng/mL) respectively in SMC ID3$^+$ cells. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.

**Endothelial spheroid assay**

In spite of endothelial cells lined the innermost layer of artery grows in monolayer, cells in plexiform lesions grow in the three-dimension structure resembles spheroid. These cells do not follow characteristics of the cells, which grow in a monolayer. The cells derived from tumour tissue alter their both genotype and phenotype characteristics once they cultured in monolayer (Lee et al., 2006). Therefore, we tried to imitate the conditions of tumour *in vitro.* In addition to this, recent data from our laboratory showed that ID3 induced a molecular stem cell-like signature—CD133$^+$ VEGFR3$^+$ CD34$^+$ in microvascular ECs. We grew the cells in no FBS DMEM F-12 medium supplemented with B-27 reduced growth supplement. We used a stem cell sphere-forming assay to determine the effect of E2 and PCB153 on HPMEC ID3$^+$ cells. HPMEC ID3$^+$ spheroids were bigger in diameter than HPMEC spheroids. We also found the treatment with PCB153 (100 ng/mL) was statistically significant (P<0.01) than control HPMEC ID3$^+$ spheroids in terms spheroid size and spheroid number. We found the similar results on the 10th day, all the treatments the treatment with PCB153 (100 ng/mL) was statistically significant (P<0.01) than control HPMEC ID3$^+$ spheroids in terms spheroid size and spheroid number as shown in Figures 35, 36 and 37.

cultured in non-adherent round bottom plates

DMEM-F12 media with B 27 supplement

Endothelial cells

Spheroid at 5d & 10d
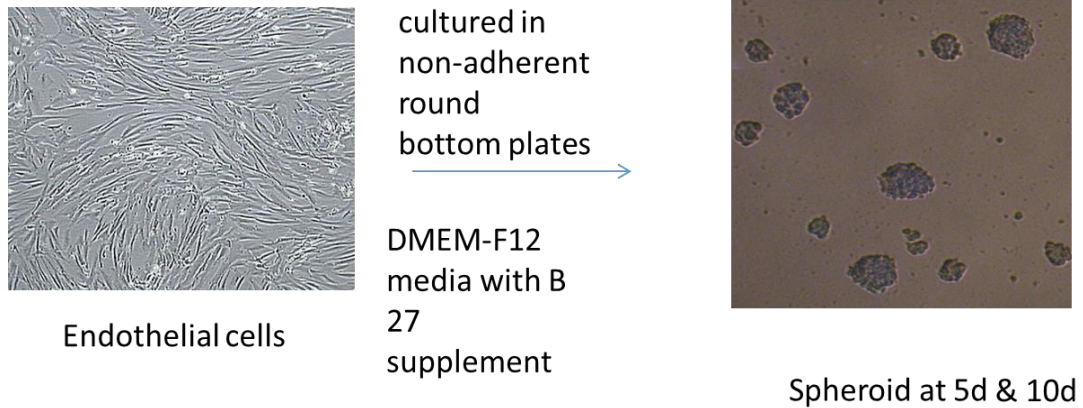
Fig 28. Schematic diagram of spheroid formation.
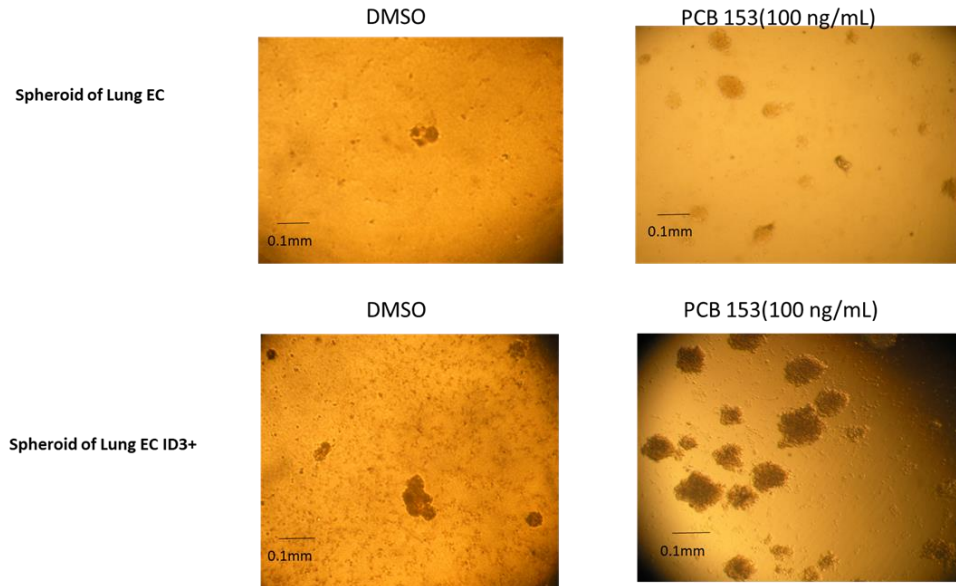
Figure 35- PCB153 increases spheroid formation

It shows pictures of spheroid. PCB153 increased the size and number of spheroid compared to control in lung EC & lung EC ID3[+]. Scale bar= 10 µm.
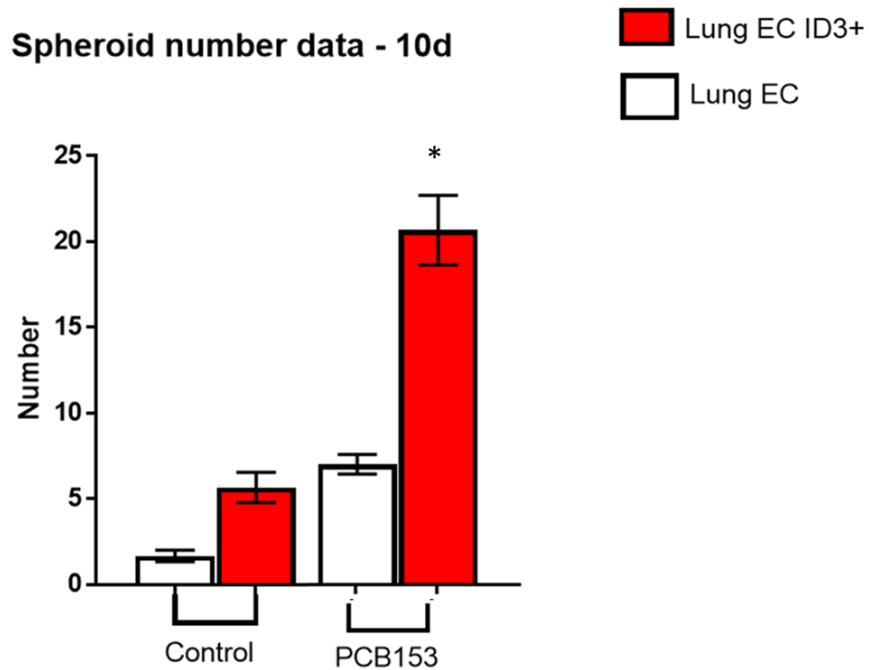
Figure 36- PCB153 increases spheroid number in HPMEC ID3$^+$

In both lung EC & EC ID3$^+$, PCB153 increased the number of spheroids compared to control. Magnification X200. Graph shows quantification of spheroid number. Error bars represent the mean sphere number ± SD in 5 wells. PCB153= 100 ng/mL. a **p<0.01 for lung EC; b **p<0.01 for lung EC ID3$^+$ vs control at day 10 as well as *p<0.05 for lung EC vs control at day 10. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.
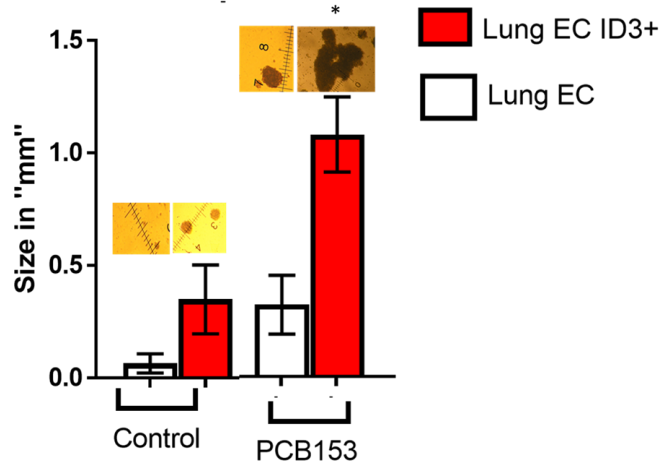
Figure 37- PCB153 increases spheroid size in HPMEC ID3+

In both lung EC & EC ID3[+], PCB153 increased the diameter of spheres compared to control. Magnification X200. Graph shows quantification of spheroid diameter. Error bars represent the mean sphere size ± SD in 5 spheres. PCB153= 100 ng/mL. Representative microphotographs of treatment groups are shown inset. Error bars represent the mean diameter of 5 spheroids ± SD. a **$p<0.01$ for lung EC; b **$p<0.01$ for lung EC ID3[+] vs control at day 10 as well as *$p<0.05$ for Lung EC ID3[+] vs control at day 10. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.

Co-culture of HPMEC ID3+ and SMC ID3+ spheroid assay-

Earlier, researchers have shown the involvement of SMCs along with lung endothelial cells in formation of plexiform lesions and tumors (Doran et al., 2008;. Lee et al., 1998). We decided to culture HPMEC ID3[+] cells with SMCs in 1:8 ratio to get the best results (in press). We seeded and co-culture HPMEC ID3[+] cells with SMC ID3[+] in 1:8 ratio in an ultra-low attachment 96-well plate with no FBS DMEM F-12 medium supplemented with B-27 reduced growth supplement. Then we treated cells with E2 (100 ng/mL), PCB153 (100 ng/mL) in separate wells. We analyzed the data on 5th and 10th day. We found that all the treatments E2 (100 ng/mL), PCB153 (100 ng/mL) were statistically significant

(P<0.01) than control spheroids in terms spheroid size and spheroid number as shown in Figures 38-43.
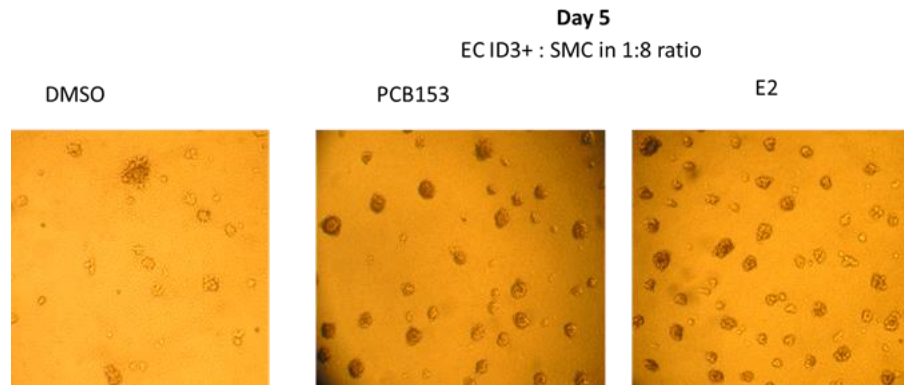
Figure 38- Images of co-culture of HPMEC ID3$^+$ and SMC ID3$^+$ spheroid at day5

PCB153 and E2 increased the size and number of spheroid compared to control. Scale bar= 10µm.

Spheroid size data-5d
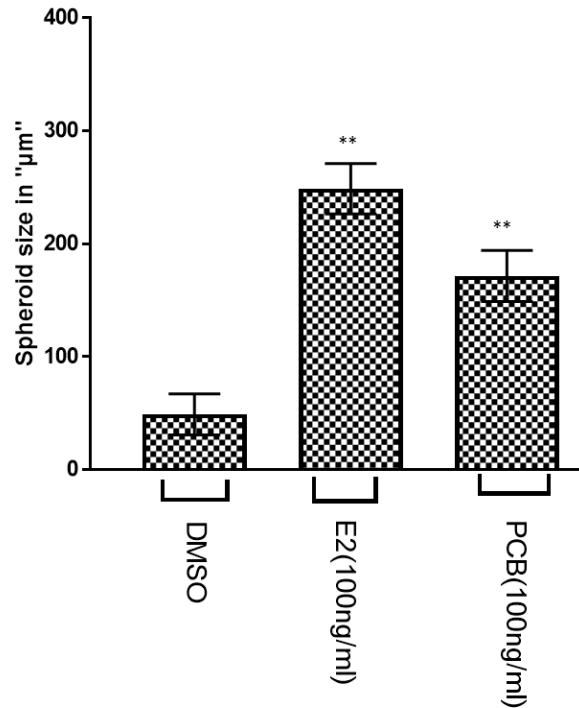
Figure 39- PCB153 and E2 increased the size of spheroid co-culture of HPMEC ID3$^+$ cells with SMC ID3$^+$ at Day5

Co-culture of HPMEC ID3$^+$ cells with SMC ID3$^+$ cells in 1:8 ratio in an ultra-low attachment 96-well plate with no FBS DMEM F-12 medium supplemented with B-27 reduced growth supplement. Then we treated cells with E2 (100 ng/mL), PCB153 (100 ng/mL) in separate wells. E2 and PCB153 increased the diameter of spheres compared to control. Magnification X200. Graph shows quantification of spheroid diameter. Error bars represent the mean sphere size ± SD in 15 spheres. E2 & PCB153= 100 ng/mL. Representative microphotographs of treatment groups are shown inset. E2 and PCB153= 100 ng/mL. **p<0.01 for Lung EC ID3$^+$: SMC ID3$^+$vs control at day 5 Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.

Spheroid number data-5d
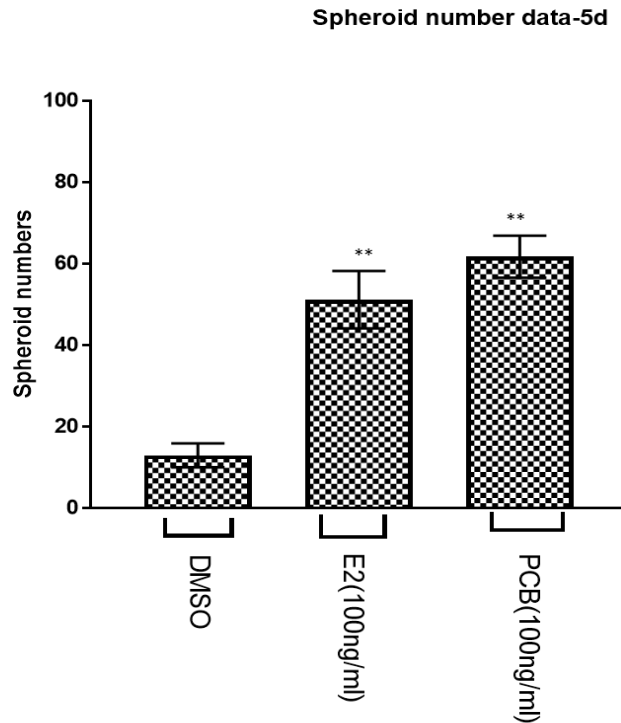
Figure 40- PCB153 and E2 increased the number of spheroid co-culture of HPMEC ID3+ cells with SMC ID3+ at Day5

Co-culture of HPMEC ID3+ cells with SMC ID3+ cells in 1:8 ratio in an ultra-low attachment 96-well plate with no FBS DMEM F-12 medium supplemented with B-27 reduced growth supplement. Then we treated cells with E2 (100 ng/mL), PCB153 (100 ng/mL) in separate wells. E2, PCB153 treatments increased the number of spheroids compared to control. Magnification X200. Graph shows quantification of spheroid number. Error bars represent the mean sphere number ± SD in 5 wells. E2 & PCB153= 100 ng/mL. **p<0.01 for lung EC ID3+: SMC ID3+ vs control at day 5 Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.

Figure 41- It shows pictures of co-culture of HPMEC ID3+ and SMC spheroid at Day10

PCB153 and E2 increased the size and number of spheroid compared to control. Scale bar= 10μm.

90

Spheroid size data-10d
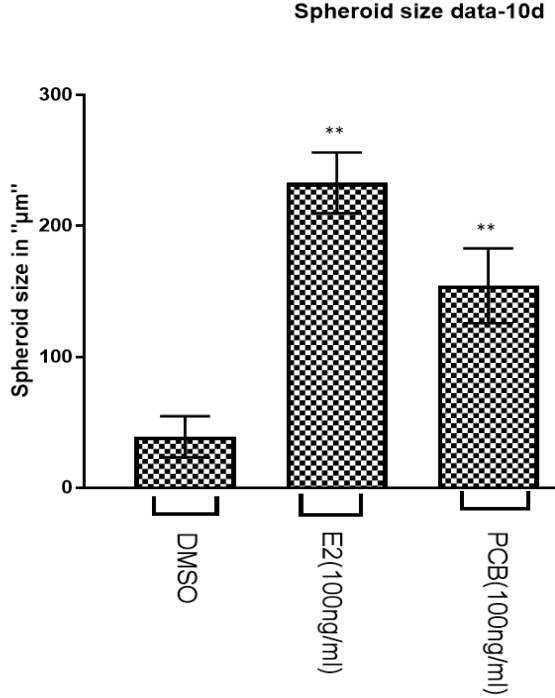
Figure 42- PCB153 and E2 increased the size of spheroid co-culture of HPMEC ID3+ cells with SMC ID3$^+$ at day10

Co-culture of HPMEC ID3$^+$ cells with SMC ID3$^+$ cells in 1:8 ratio in an ultra-low attachment 96-well plate with no FBS DMEM F-12 medium supplemented with B-27 reduced growth supplement. Then we treated cells with E2 (100 ng/mL), PCB153 (100 ng/mL) in separate wells. E2 and PCB153 increased the diameter of spheres compared to control. Magnification X200. Graph shows quantification of spheroid diameter. Error bars represent the mean sphere size ± SD in 15 spheres. E2 & PCB153= 100 ng/mL. Representative microphotographs of treatment groups are shown inset. E2 & PCB153= 100 ng/mL. **p<0.01 for lung EC ID3$^+$: SMC ID3$^+$vs control at day 10 Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.

Figure 43- PCB153 and E2 increased the number of spheroid co-culture of HPMEC ID3+ cells with SMC ID3$^+$ at day10

Co-culture of HPMEC ID3$^+$ cells with SMC ID3$^+$ cells in 1:8 ratio in an ultra-low attachment 96-well plate with no FBS DMEM F-12 medium supplemented with B-27 reduced growth supplement. Then we treated cells with E2 (100 ng/mL), PCB153 (100 ng/mL) in separate wells. E2, PCB153 treatments increased the number of spheroids compared to control. Magnification X200. Graph shows quantification of spheroid number. Error bars represent the mean sphere number ± SD in 5 wells. E2 & PCB153= 100 ng/mL. **p<0.01 for lung EC ID3$^+$: SMC ID3$^+$ vs control at day 10 Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.

Spheroid assay- MTT, SRB, BrdU assay

Proliferative vascular lesions have been studied extensively using the VEGF─receptor antagonist, Sugen 5416 (SU5416), plus chronic hypoxia (SuHx) rodent model as shown in Figure 44. Therefore, the aim of this study was to determine whether SU5416 exposure of ECs become more sensitive to estrogen-induced cell growth. We exposed human pulmonary endothelial cell line HPMEC ID3$^+$ to SU5416 to select for a sub-population of cells, which were then treated with 17β-estradiol and/or PCB153. We observed a significant increase of 2-3 fold change in cell growth and proliferation as determined by MTT, SRB, BrdU, analysis when SU5416 treated cells were exposed to estrogenic chemicals as shown in Figure 45, 46 and 47.
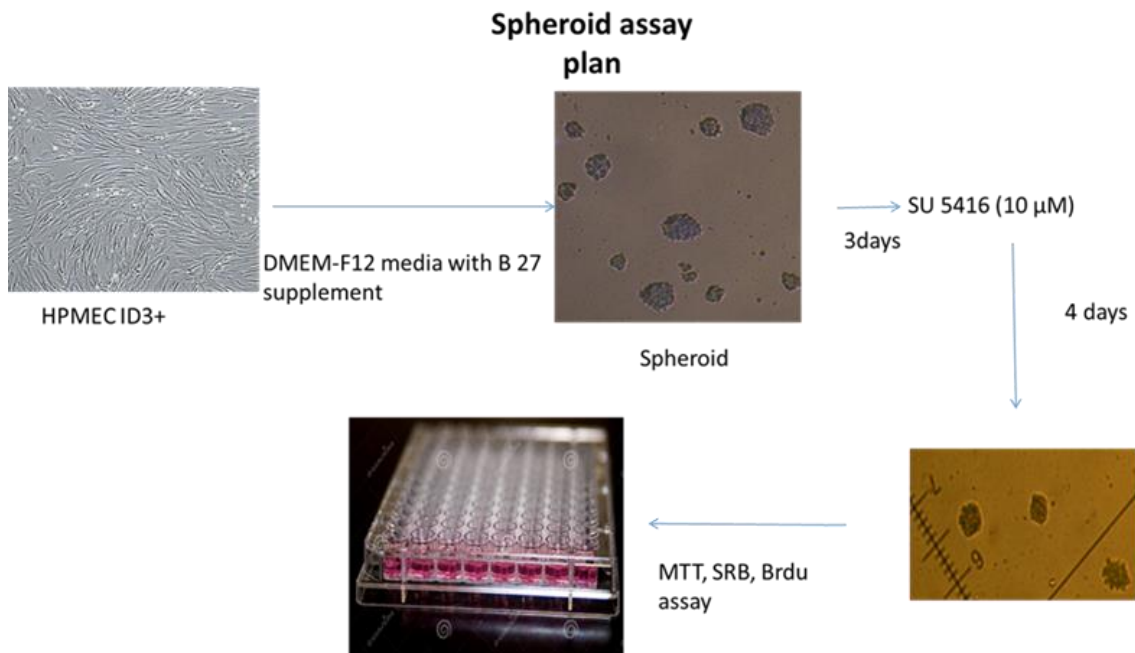
Figure 44- Procedure of formation of spheroid by treatment of ID3$^+$ overexpressing lung endothelial cells with SU5416 and then cultured these cells to measure survival and growth.

MTT assay at 48h
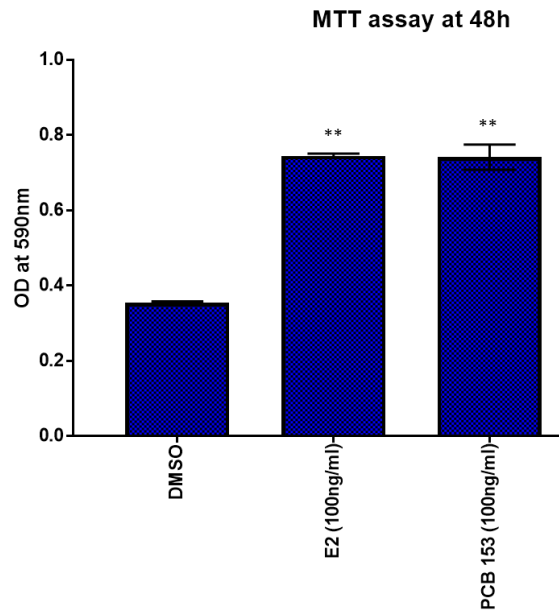
Figure 45- SU5416 treatment increased the survival and proliferation of HPMEC ID3⁺ cells.

HPMEC ID3⁺ cells exposed to SU5416 for 72 h to select for a sub-population of cells, which were then grown with no FBS DMEM F-12 medium supplemented with B-27 reduced growth supplement for 4 days. Then these cells treated with 17β-estradiol and/or PCB153. Graph of cell survival determined by MTT assay of endothelial spheroids 24 h after exposure. All the treatment showed significant increase**p<0.01 for HPMEC ID3⁺ vs control

SRB assay at 48h
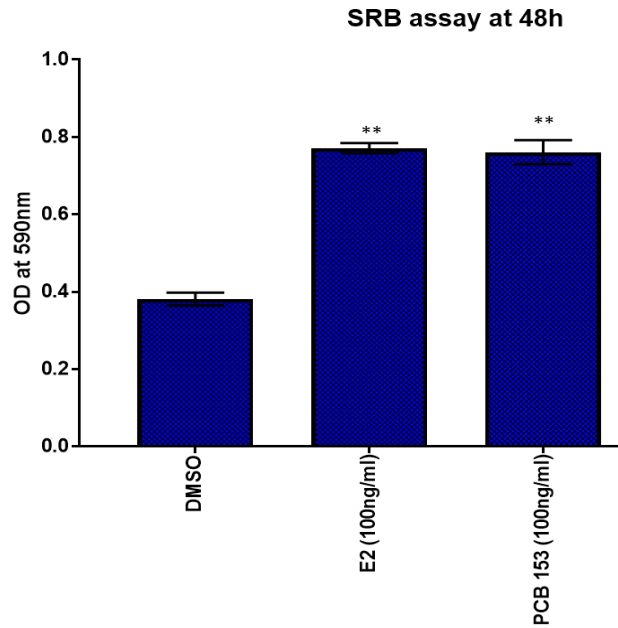
Figure 46- SU5416 treatment increased the survival and proliferation of HPMEC ID3+ cells.

HPMEC ID3+ cells exposed to SU5416 for 72 h to select for a sub-population of cells which were then grown with no FBS DMEM F-12 medium supplemented with B-27 reduced growth supplement for 4 days. Then these cells treated with 17β-estradiol and/or PCB153. Graph of cell survival determined by SRB assay of endothelial spheroids 24 h after exposure. All the treatment showed significant increase**p<0.01 for HPMEC ID3+ vs control
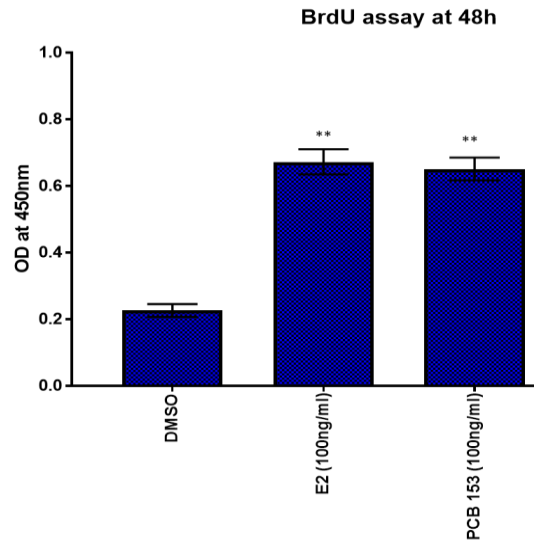
Figure 47- SU5416 treatment increased the survival and proliferation of HPMEC ID3⁺ cells.

HPMEC ID3⁺ cells exposed to SU5416 for 72 h to select for a sub-population of cells which were then grown with no FBS DMEM F-12 medium supplemented with B-27 reduced growth supplement for 4 days. Then these cells treated with 17β-estradiol and/or PCB153. Graph of cell survival determined by BrdU assay of endothelial spheroids 24 h after exposure. All the treatment showed significant increase**p<0.01 for HPMEC ID3⁺ vs control

Flow Cytometry
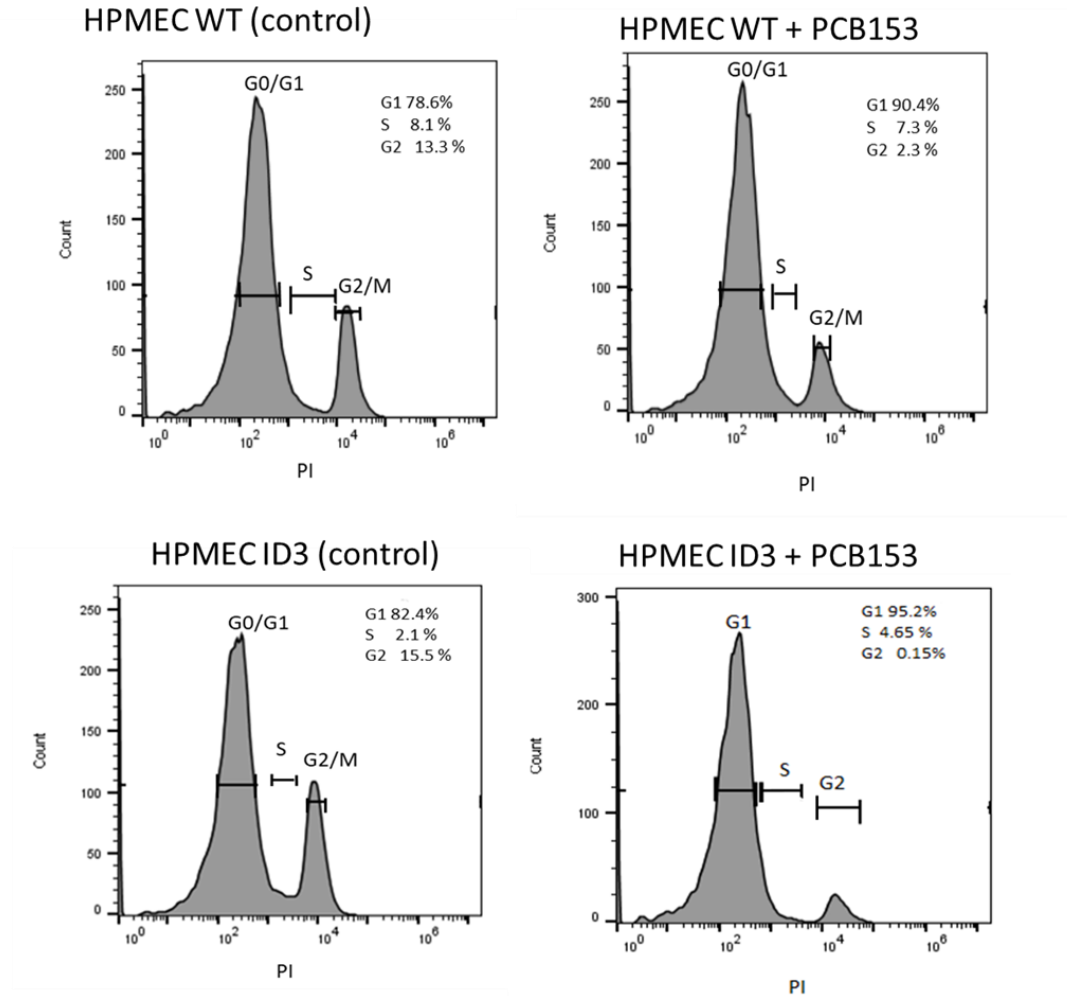
Figure 48- Analysis by flow cytometry of cell cycle profiles of HPMEC and HPMEC ID3+

The distribution of cells in the different phases of the cell cycle was analyzed in absence (control) or presence of PCB153 in both HPMEC and HPMEC ID3+ cell types.

The cell cycle is process in which cell undergo its division and duplication of its DNA (DNA replication) to produce two daughter cells. The cell cycle consists of four distinct phases: G1 phase, S phase (synthesis), G2 phase (collectively known as interphase) and M phase (mitosis). G0 is a resting phase where the cell has left the cycle and has stopped dividing. The cell cycle starts with this phase. G1 is first phase of interphase also called as growth phase where cell prepares itself to go under cell division. The cell increases its supply of proteins, increases the number of organelles (such as mitochondria, ribosomes), and grows in size. The S phase starts when DNA synthesis commences; when it is complete, all of the chromosomes have been replicated, i.e., each chromosome has two (sister) chromatids. G2 phase occurs after DNA replication and is a period of protein synthesis and rapid cell growth to prepare the cell for mitosis.

Our flow cytometry experiments revealed that approximately 82.4% of the DNA content in HPMEC ID3$^+$ resided in the G0/G1 phase compared to only 78.6% of the wild-type cells. Thus, HPMEC ID3$^+$ showed a statistically significant 4.7205% of difference and 1.04 -fold increase in G0/G1 phase cells (Fig. 48). The percentage of DNA content in S phase decreased from approximately 8.1% in HPMEC to 2.1% in HPMEC ID3$^+$. Thus, ID3 significantly reduced the number of cells in S phase by more than 2-fold (Fig. 48). The HPMEC ID3$^+$ showed increase in percentage of DNA content in G2/M phase from approximately 13.3% in wild-type cells to 15.5% in the HPMEC ID3$^+$ population (Fig. 48).Our data suggest that in the HPMEC ID3$^+$ population there is a significant increase in the number of cells that reside in the G0/G1 phase which is indicative of proliferative quiescence. Since a higher proliferative capacity may exhaust a stem cell population, this observed increase in G0/G1 phase by HPMEC ID3$^+$ may help to maintain a stem cell-like

state by slowing cell cycle progression. Similarly, we found same trend in case of treatment with PCB153 to HPMEC and HPMEC ID3[+.] When both the cells treated with PCB153, we found that cells showed more DNA percentage reside in G0/G1 phase compared to the control. In case of HPMEC cells, PCB153 exposure increased DNA percentage from 78.6% to 90.4%. In case of HPMEC ID3[+] cells, PCB153 exposure increased DNA percentage from 82.4% to 95.2%. Since Felty et al. already shown, ID3 overexpression gives stemness character to the cells and stem cells reside longer in the G0/G1 phase of the cell cycle. Proliferative quiescence is a state observed in stem cells which reside longer in the G0/G1 phase of the cell cycle; and it has been hypothesized that quiescence may prevent against the depletion of the stem cell population. Similarly, we found the same effect in SMC and SMC ID3[+] cells. SMC ID3[+] cells showed that their percentage of DNA reside more in G0/G1 phase and cells prepare for synthesis in S phase. In addition, we found that the exposure to PCB153 increases percentage of DNA in G0/G1 phase in SMC ID3[+] compared to control (Figure 49).
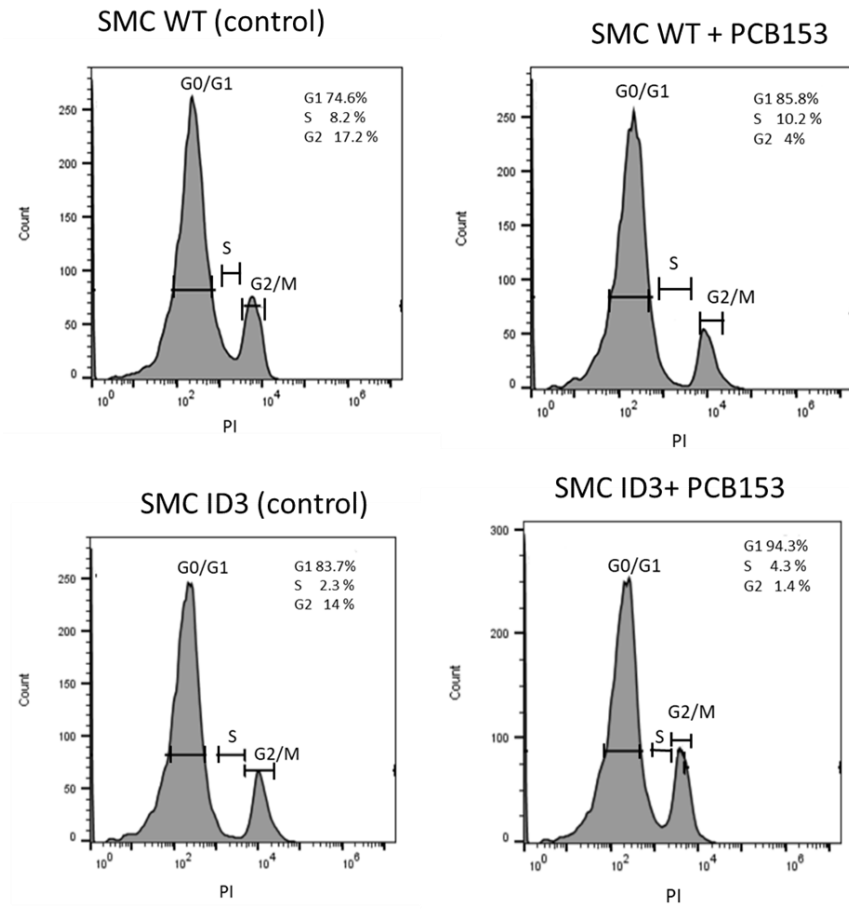
Figure 49- Analysis by flow cytometry of cell cycle profiles of SMC and SMC ID3⁺

The distribution of cells in the different phases of the cell cycle was analyzed in absence (control) or presence of PCB153 in both SMC and SMC ID3⁺ cell types.

Western blot

Aim- To determine the effect of PCB153 on ID3 expression.

We seeded the HPMEC and HPMEC ID3$^+$ cells in four 100 mm petri dishes each. We added 0.1% DMSO (control), and PCB153 (100 ng/mL) to separate petri dishes. We analyzed that combination treatment of PCB153 increase 2-3-fold change in ID3 expression in HPMEC ID3$^+$ cells compared to control. PCB153 (100 ng/mL) treatment showed significant (P<0.01) increased than control as shown in Figure 50. We have also checked the expression of ID3 protein in SMC and SMC ID3$^+$ cell line in Figure 51.
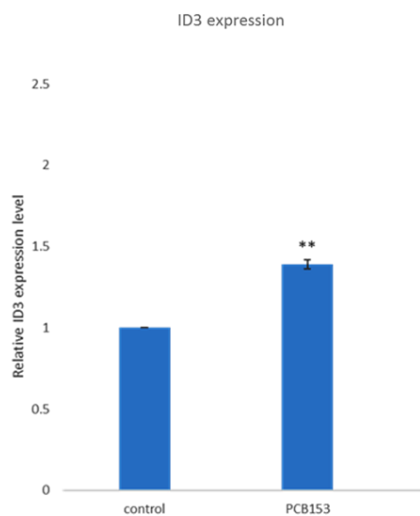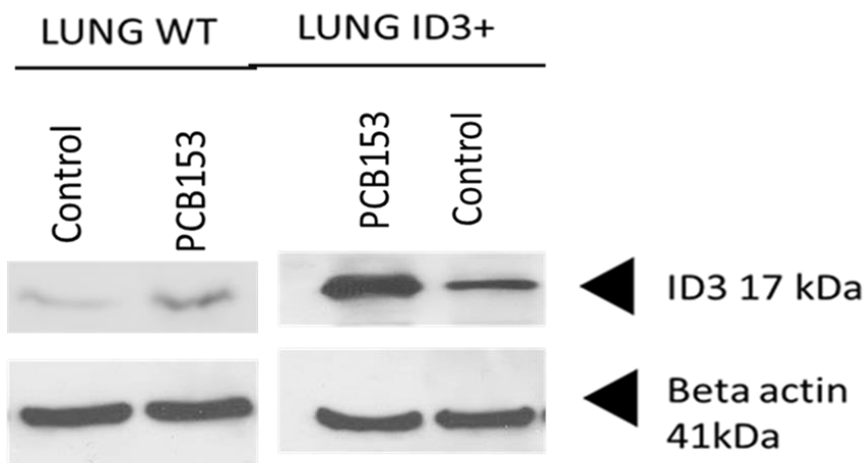
Figure 50- PCB153 increases ID3 expression in HPMEC cells.

Detection of ID3 protein 17 KDa and Beta actin as a control 41 KDa in lung EC ID3[+] cell line. Gel percentage 15%. Primary antibody ID3 monoclonal Ab- 1:2500 and Secondary Ab- 1:10000. PCB153 = 100 ng/mL. HPMEC and HPMEC ID3[+] cells treated with 0.1% DMSO (control), PCB153 (100 ng/mL). All the treatment showed significant increase**p<0.01 for HPMEC ID3[+] vs control. Antibody concentration = 0.1 mg
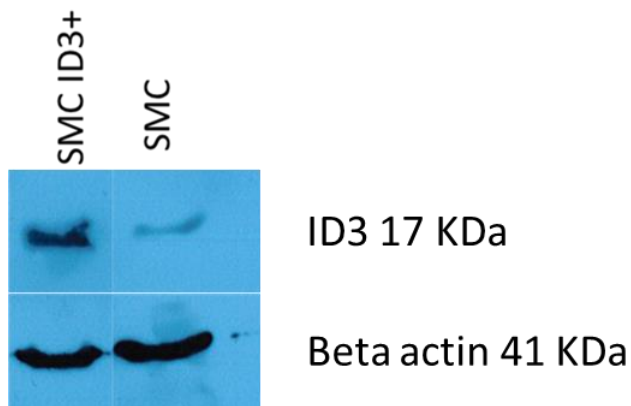
Figure 51- ID3 expression in SMC and SMC ID3⁺ cells

Detection of ID3 protein 17 KDa and Beta actin as a control 41 KDa in lung SMC ID3⁺ cell line. Gel percentage 15%. Primary antibody ID3 monoclonal Ab- 1:2500 and Secondary Ab- 1:10000. SMC and SMC ID3⁺ cells treated with 0.1% DMSO (control). Antibody concentration = 0.1 mg.

EndMT markers expression in lung EC ID3⁺

Endothelial to mesenchymal transition (EndMT) is major event arrive when endothelial cells dramatically alter their shape and motile behavior as they differentiate into mesenchymal cells. In this process, cells detached from cell-cell junctions of neighboring cells and basement membrane (Larue & Bellacosa, 2005). During this transition, cells express and loose some protein markers. Interestingly, EndMT process is an integral part of tumor cells metastasis in which cells show alteration in some protein markers. Expression of VE cadherin, Vimentin, and Matrix metalloproteinase 9 proteins are the hallmark of EndMT process. VE-cadherin (CD144) is a cell-adhesion protein that maintains the cell-cell contacts (Kowalski, 2003). Matrix metalloproteinase 9 (MMP9) is a zinc finger dependent enzyme cuts cell-cell and basement membrane contact and promotes invasion, metastasis, and angiogenesis. Vimentin is cytoskeletal intermediate filaments (IFs) protein and is expressed more in mesenchymal transition. It has been

shown that Vimentin expression is also upregulated in wound healing *in vitro* culture system (Hendrix, Seftor, Seftor, & Trevort, 1997; Liu, Lin, Tang, & Wang, 2015). Moreover, Felty et al. 2014 showed that PCB153 along with ID3 over expression induced cell migration and endothelial spheroid growth in HUVEC cells. They also showed that ID3 overexpression regulates the upregulation of VEGFR3, which is marker of angiogenesis and neovascularization. Hence, we investigated the role of ID3 in case endothelial to mesenchymal transition of cells in presence of estrogenic chemical.

We seeded the HPMEC and HPMEC ID3$^+$ cells in eight chambered slide and added the treatments 0.1% DMSO (control) and PCB153 (100 ng/mL). We found that VE cadherin expressed more in HPMEC cells compared to HPMEC ID3$^+$ in control. Vimentin and MMP-9 expressed more in HPMEC ID3$^+$ cells compared to HPMEC cells. We also found that vimentin and MMP-9 expression of HPMEC ID3$^+$ cells was significantly higher in PCB153 (100 ng/mL) than control as shown in Figure 52. We have also showed the changes in expression of EndMT markers in ID3 overexpressing HPMEC cell line. We observed gain of N-Cadherin, Vimentin and loss of VE-Cadherin in Figures 53, 54 and 55.
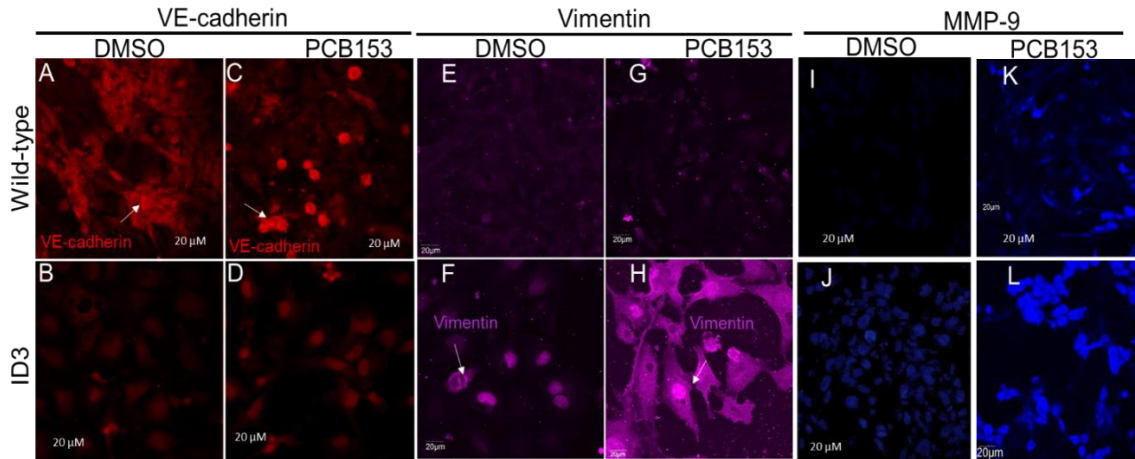
Figure 52- Immunofluorescence. ID3 altered expression of VE-cadherin, vimentin, and MMP-9 in lung endothelial cell line HPMEC ID3+.

Cells were exposed to vehicle or PCB153 100 ng/mL for 24h. (**A-B**) Stable expression of ID3 showed a loss of endothelial marker VE-cadherin a characteristic of endothelial-to-mesenchymal transition (EndMT). (**E-F**) ID3 increased expression of mesenchymal marker vimentin. (**I-J**) ID3 increased expression of mesenchymal marker matrix metalloproteinases 9 (MMP9).
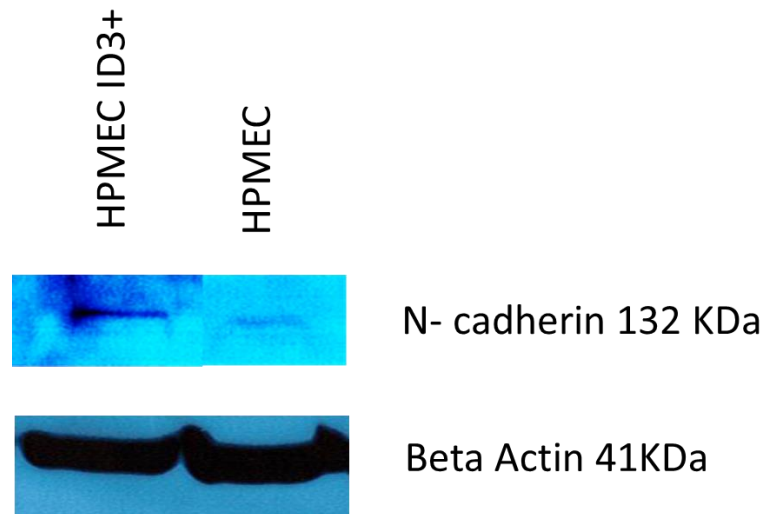
Figure 53- Detection of N-Cadherin protein

Detection of N-Cadherin protein 132 KDa in HPMEC and HPMEC ID3+ cell line. Gel percentage 7.5%. Sample loaded per well = 50 microgram. N-cadherin- (H-63): sc-7939 mouse monoclonal Ab- 1:2500 and Secondary Ab- 1:10000. Antibody concentration = 10 µg/50 µl
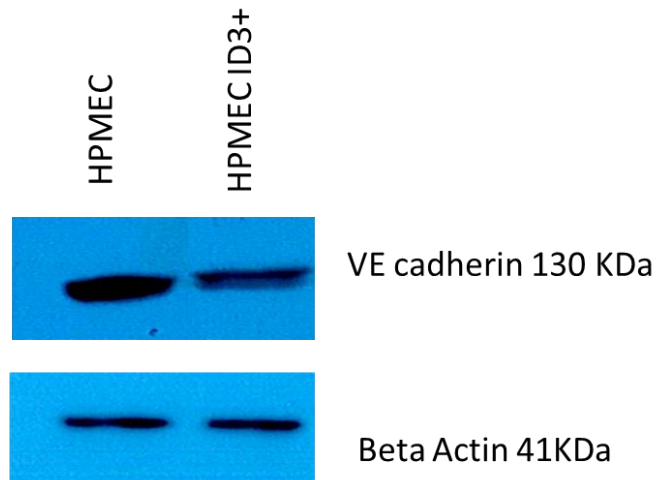
Figure 54- Detection of VE-Cadherin protein

Detection of VE-Cadherin protein 130 KDa in HPMEC and HPMEC ID3[+] cell line. Gel percentage 7.5%. Sample loaded per well = 50 microgram. VE-cadherin-2 Antibody (F-4): sc-515467 mouse monoclonal Ab- 1:2500 and Secondary Ab- 1:10000. Antibody concentration = 10 µg/50 µl
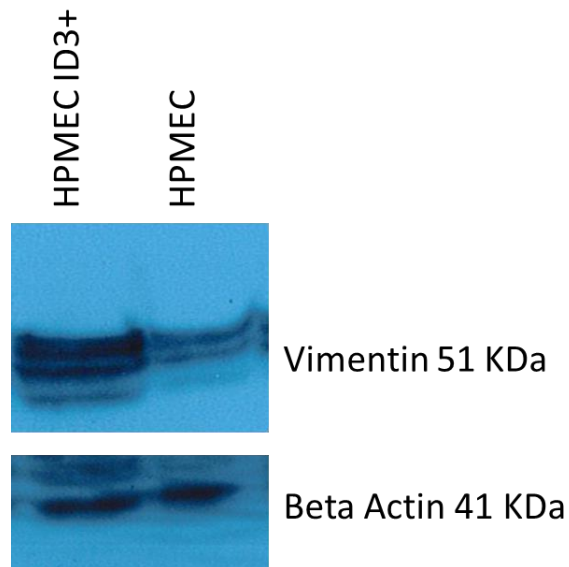
Figure 55- Detection of Vimentin protein

Detection of Vimentin protein 51 KDa in HPMEC and HPMEC ID3[+] cell line. Gel percentage 10%. Sample loaded per well = 50 microgram. Vimentin Antibody (V9): sc-6260 mouse monoclonal Ab- 1:2500 and Secondary Ab- 1:10000. Antibody concentration = 10 µg/50 µl

Zebrafish model to study role of ID3 in EC proliferation

Our main objective of this research is to investigate the *in vivo* role of ID3 in the pathogenesis of proliferative vascular lesions in early stages of zebrafish embryos and in xenograft tumors. The purpose of this study is to develop a zebrafish xenograft model of metastasis by injecting Cancer Stem Cells (CSCs), or Tumor Initiating Cells (TICs) isolated from ID3 positive or negative or ID3 target gene(s) expressing *in vitro* cultured lung endothelial cells in early stages of zebrafish embryos. Endothelial cells in plexiform vascular lesions (a pathological hallmark of PAH) have been characterized as hyper-proliferative and apoptosis resistant, which is consistent with a neoplastic hypothesis. This lesion is a major cause of high morbidity and mortality in severe end-stage IPAH patients because the pulmonary vascular lesion (PVL) progressively obliterate pulmonary arterioles and impinge upon the unobstructed vessels resulting in heart failure. Our research over the past 10 years has uncovered that the disruption of mitochondrial ROS production and redox signaling to nuclear regulatory pathways by estrogen or its metabolite, 4- hydroxy-estradiol contribute to the development of vascular lesions. In addition, our preliminary results in 2D and 3D model showed that ID3 seems to be essential for generation of stem-like cells and EndMT and cell migration, but the targets of ID3 gene are unknown. The present research is centered on the concept that activation of dynamic "rheostat" molecules in CSCs, or TICs regulate gene networks of chromosomal/genomic instability and they drive vascular lesion in pulmonary arterioles involving oxidative stress, apoptosis, and stemness. In line with this overall paradigm, we propose mechanistic studies targeting the redox response, involving ID3 gene shown to induce oxidative stress, anti-apoptosis, stemness and genomic instability. The

significance of the use of zebrafish model in this study is to identify the ID3 targets using both genetic and biochemical analyses for their ability to promote or inhibit non-malignant vascular lesions *in vivo*. Zebrafish embryos will be used to develop xenograft vascular lesions to determine if essential effectors of ID3 and their potential targets are relevant drug targets that may epigenetically drives the formation of non-malignant vascular lesions from exposure to PCBs *in vivo*. Our data showed that HPMEC ID3$^+$ cells after injection show migration through blood vessels to thoracic and head part of the zebrafish as shown in Figure 56.
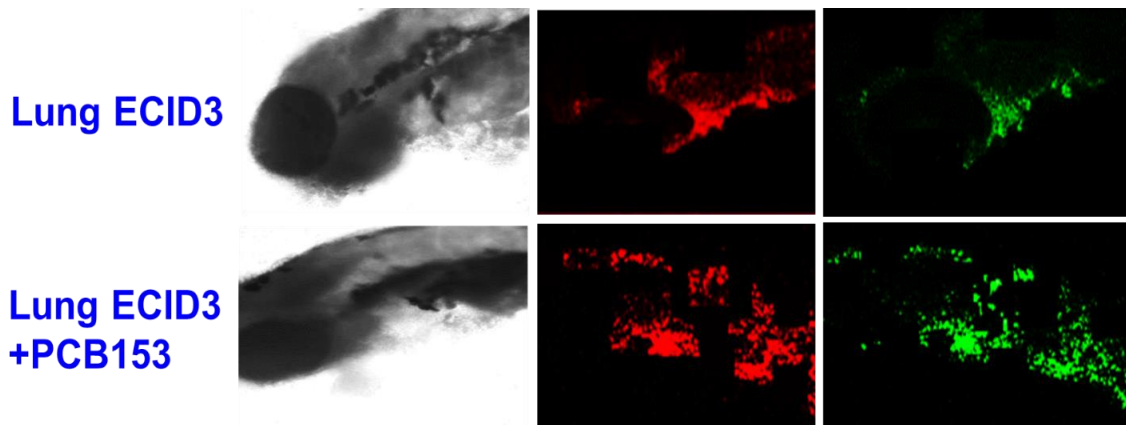
Figure 56- Proliferation and migration of a human lung EC ID3$^+$ cell line in zebrafish

The lung ECs ID3$^+$ of xenotransplantation in zebra fish embryo (2 dpf) showed that spread of the GFP fluorescent dye labeled lung ECs moved through blood vessels from the site of injection to the region of secondary target organ for metastasis lesions. The photomicrographs were captured by Nikon confocal microscopy. The treatment of PCB153 on lung ECs ID3$^+$ group showed the more migratory cells for adhesion, micro invasion and metastasis. The red color showed injected all cells stained with the fluorescent dye, CellTracker™ CM-DiI (stained both GFP and non GFP cells) which is well suited for monitoring all live cell movement in the blood vessel of the zebra fish embryo. 400X).

Discussion

In the present study, we determined PCB-induced neovascularization with the aim of elucidating the role of environmental PCBs in lung endothelial dysfunction while focused on ID3. Enhanced neovascularization has been implicated in the development of plexiform lesions in idiopathic pulmonary arterial hypertension, and atherosclerotic lesions in cardiovascular disease. It is evident that humans have been and will continue to be exposed to PCBs due to deposition of harmful chemical contaminants in the environment through chemical industries waste, landfills and incinerators (Tanabe, 1988). PCB153 is included as a persistent non-dioxin-like polychlorinated biphenyl in PCB congeners, which are abundantly present in the food-chain and the environment. PCB153 may cause deleterious public health effects because of their tendency to bioaccumulate to high concentrations due to their slow elimination from the human body. In adults, the estimated elimination half-life for PCB153 is 14.4 years (Ritter et al., 2011b).

Our lab earlier showed that the PCB153 induced ROS formation increases ID3 protein expression in human microvascular endothelial cell line hCMEC/D3. We also showed in our earlier studies that the higher expression of ID3 protein increases neovascularization in human microvascular endothelial cells. The lung endothelial cells come into direct contact with blood in lung microenvironment. The human blood carries many dissolved gases, chemicals from environment, and these all directly come in contact with the innermost layer of lung endothelial cells, which lead to form lesions. Therefore, in presence of chemical inducers or inhibitors lung endothelial cells may undergo an alteration in the phenotype. In the study, we investigated the effect of PCB153 on lung

endothelial cells while focused on ID3. We exposed lung endothelial cells with PCB153 and studied the behavior of cells. In wound healing assay, our results showed that in presence of PCB153 the rate of cells migration to the wound area was significantly higher than control. In addition to this, in case of overexpressed ID3 lung endothelial cells showed higher rate of cell migration when treated with PCB153 compared to wild type cells. These results showed that overexpressed ID3 lung endothelial cells in presence of environmental contaminant of PCB153 showed higher cell proliferation behavior than wild type. To strengthen our claim regarding change in behavior of overexpressed ID3 lung endothelial cells, we did endothelial spheroid assay, which mimics plexiform lesion formation *in vivo* model. We found that PCB153 increased the size and number of spheroid significantly in lung ECs ID3$^+$ compared to wild type lung EC. We also cultured the spheroid from lung EC ID3$^+$ cells and measured ell growth and proliferation determined by MTT, SRB, BrdU, analysis. We found that PCB153 increased growth and proliferation of lung EC ID3$^+$ cells. Our western blot results also demonstrated the growth and proliferation effect at molecular level. ID3 protein expressed significantly higher when lung ECs ID3$^+$ cells treated with PCB153. We also observed cell migration behavior at molecule level with immunofluorescence assay. In this process, cells detached from cell-cell junctions of neighboring cells and basement membrane (Larue & Bellacosa, 2005). During this transition, cells express and loose some protein markers. Interestingly, EndMT process is integral part of tumor cells metastasis in which cells shows alteration in some protein markers. Expression of VE cadherin, vimentin and Matrix metalloproteinase 9 (MMP9) proteins are hallmark of EndMT process. When we exposed lung ECs ID3$^+$ cells with PCB153, we found that higher expression of vimentin

and MMP9 proteins. These results corroborate our hypothesis that environmental chemical PCB153 increases proliferation of cells by increasing ID3 expression as well as instigate stem cells like properties to lung EC cells.

References:

Arteaga, E., Villaseca, P., Bianchi, M., Rojas, A., & Marshall, G. (2004). Raloxifene is a better antioxidant of low-density lipoprotein than estradiol or tamoxifen in postmenopausal women *in vitro*. *Menopause (New York, N.Y.)*, *10*(2), 142–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12627039

ASTDR. (2000). Toxicological Profile for Polychlorinated Biphenyls (PCBs). *Agency for Toxic Substances and Disease Registry*, (November), 1–948. Retrieved from https://www.atsdr.cdc.gov/toxprofiles/tp17.pdf

Badesch, D. B., Raskob, G. E., Elliott, C. G., Krichman, A. M., Farber, H. W., Frost, A. E., McGoon, M. D. (2010). Pulmonary Arterial Hypertension. *Chest*, *137*(2), 376–387. http://doi.org/10.1378/chest.09-1140

Bennett, M. R. (1999). Apoptosis of vascular smooth muscle cells in vascular remodelling and atherosclerotic plaque rupture. *Cardiovascular Research*, *41*, 361–368. Retrieved from https://oup.silverchair-cdn.com/oup/backfile/Content_public/Journal/cardiovascres/41/2/10.1016/S0008-6363(98)00212-0/2/41-2-361.

Berridge, M. V., & Tan, A. S. (1993). Characterization of the Cellular Reduction of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT): Subcellular Localization, Substrate Dependence, and Involvement of Mitochondrial Electron Transport in MTT Reduction. *Archives of Biochemistry and Biophysics*, *303*(2), 474–482. http://doi.org/10.1006/abbi.1993.1311

Berridge, M. V., Herst, P. M., & Tan, A. S. (2005). Tetrazolium dyes as tools in cell biology: New insights into their cellular reduction (pp. 127–152). http://doi.org/10.1016/S1387-2656(05)11004-7

Controlled trial of tamoxifen as a single adjuvant agent in the management of early breast cancer Analysis at Eight Years by "Nolvadex"* Adjuvant Trial Organisation.t. (1988). *Br. J. Cancer*, *57*, 608–611. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2246455/pdf/brjcancer00140-0084.pdf

Das, J. K., & Felty, Q. (2014a). Microvascular Lesions by Estrogen-Induced ID3: Its Implications in Cerebral and Cardiorenal Vascular Disease. *Journal of Molecular Neuroscience*, 618–631. http://doi.org/10.1007/s12031-014-0401-9

Das, J. K., & Felty, Q. (2014b). PCB153-induced overexpression of ID3 contributes to the development of microvascular lesions. *PLoS ONE*, *9*(8). http://doi.org/10.1371/journal.pone.0104159

Das, J. K., & Felty, Q. (2015). Microvascular Lesions by Estrogen-Induced ID3: Its Implications in Cerebral and Cardiorenal Vascular Disease. *Journal of Molecular Neuroscience*, *55*(3), 618–631. http://doi.org/10.1007/s12031-014-0401-9

Das, J. K., Voelkel, N. F., & Felty, Q. (2015). ID3 contributes to the acquisition of molecular stem cell-like signature in microvascular endothelial cells: Its implication for understanding microvascular diseases. *Microvascular Research*, *98*, 126–138. http://doi.org/10.1016/j.mvr.2015.01.006

Doke, M., Felty, Q., & Das, J. (2017). The role of environmental chemical PCB153 in aggressive growth of plexiform lesions associated with lung cancer. *European Journal of Surgical Oncology*, *43*(11), 2218. http://doi.org/10.1016/j.ejso.2017.10.119

Doran, A. C., Meller, N., & McNamara, C. A. (2008). Role of smooth muscle cells in the initiation and early progression of atherosclerosis. *Arteriosclerosis, Thrombosis, and Vascular Biology*, *28*(5), 812–819. http://doi.org/10.1161/ATVBAHA.107.159327

Felty, Q. (2006). Estrogen-induced DNA synthesis in vascular endothelial cells is mediated by ROS signaling. *BMC Cardiovascular Disorders*, *6*, 16. http://doi.org/10.1186/1471-2261-6-16

Felty, Q., & Porther, N. (2008). Estrogen-induced redox sensitive Id3 signaling controls the growth of vascular cells. *Atherosclerosis*, *198*(1), 12–21. http://doi.org/10.1016/j.atherosclerosis.2007.12.048

Haldi, M., Ton, C., Seng, W. L., & McGrath, P. (2006). Human melanoma cells transplanted into zebrafish proliferate, migrate, produce melanin, form masses and stimulate angiogenesis in zebrafish. *Angiogenesis*. http://doi.org/10.1007/s10456-006-9040-2

Hendrix, M. J. C., Seftor, E. A., Seftor, R. E. B., & Trevort, K. T. (1997). Experimental Co-Expression of Vimentin and Keratin Intermediate Filaments in Human Breast Cancer Cells Results in Phenotypic Interconversion and Increased Invasive Behavior. *American Journal ofPathology*, *150*(2). Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1858294/pdf/amjpathol00026-0099.pdf

Irey (1973). Intimal vascular lesions associated with female reproductive steroids. *Popline.org*. Retrieved from https://www.popline.org/node/491301

Jordan, V. C., Allen, K. E., & Dix, C. J (2002) Pharmacology of tamoxifen in laboratory animals. *Cancer Treatment Reports*, *64*(6–7), 745–59. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6775807

Kleiger, R., Boxer, M., Ingham, R., Chest, D. H.-, & 1976, undefined. (). Pulmonary hypertension in patients using oral contraceptives: a report of six cases. *Journal.chestnet.org*. Retrieved from http://journal.chestnet.org/article/S0012-3692(16)56125-X/abstract

Kumar, G., Bullet, A., & Tomar, R. S.. Ebselen, a promising antioxidant drug: mechanisms of action and targets of biological pathways. http://doi.org/10.1007/s11033-014-3417-x

Larue, L., & Bellacosa, A.. Epithelial–mesenchymal transition in development and cancer: role of phosphatidylinositol 3 0 kinase/AKT pathways. http://doi.org/10.1038/sj.onc.1209091

Lee, J., Kotliarova, S., Kotliarov, Y., Li, A., Su, Q., Donin, N. M., Fine, H. A. (2006). Tumor stem cells derived from glioblastomas cultured in bFGF and EGF more closely mirror the phenotype and genotype of primary tumors than do serum-cultured cell lines. *Cancer Cell*, *9*(5), 391–403. http://doi.org/10.1016/j.ccr.2006.03.030

Lee, S.-D., Shroyer, K. R., Markham, N. E., Cool, C. D., Voelkel, N. F., & Tuder, R. M. (1998). Monoclonality of Endothelial Cells Rapid Publication Monoclonal Endothelial Cell Proliferation Is Present in Primary but not Secondary Pulmonary Hypertension. *J. Clin. Invest*, *101*(5), 927–934. Retrieved from http://www.jci.org

Liu, Lin , Tang, & Wang, Y.-K. (2015). Vimentin contributes to epithelial-mesenchymal transition cancer cell mechanics by mediating cytoskeletal organization and focal adhesion maturation. *Oncotarget*, *6*(18). Retrieved from www.impactjournals.com/oncotarget

Morse, J. H., Horn, E. M., & Barst, R. J. (1999). Hormone Replacement Therapy: A Possible Risk Factor in Carriers of Familial Primary Pulmonary Hypertension. *Chest*, *116*(3), 847. http://doi.org/10.1378/CHEST.116.3.847

Paul J Kowalski1, M. A. R. and C. G. K. (2003). E-cadherin expression in primary carcinomas of the breast and its distant metastases. http://doi.org/10.1186/bcr651

Perkins, J. T., Petriello, M. C., Newsome, B. J., & Hennig, B. (2016). Polychlorinated biphenyls and links to cardiovascular disease. *Environmental Science and Pollution Research*, *23*(3), 2160–2172. http://doi.org/10.1007/s11356-015-4479-6

Ritter, R., Scheringer, M., MacLeod, M., Moeckel, C., Jones, K. C., & Hungerbühler, K. (2011). Intrinsic human elimination half-lives of polychlorinated biphenyls derived from the temporal evolution of cross-sectional biomonitoring data from the United Kingdom. *Environmental Health Perspectives*, *119*(2), 225–31. http://doi.org/10.1289/ehp.1002211

Rosenthal, N. (2001). Muscle cell differentiation. *Current Opinion in Cell Biology*, *89*, 39–46. http://doi.org/10.1016/S0955-0674(89)80056-0

Russo, A., Gianni, L., Kinsella, T. J., Klecker, R. W., Jenkins, J., Rowland, J., … Myers, C. (1984). Pharmacological evaluation of intravenous delivery of 5-bromodeoxyuridine to patients with brain tumors. *Cancer Research*, *44*(4), 1702–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6704976

Tanabe, S. (1988). PCB problems in the future: foresight from current knowledge. *Environmental Pollution (Barking, Essex : 1987)*, *50*(1–2), 5–28. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15092651

Vichai, V., & Kirtikara, K. (2006). Sulforhodamine B colorimetric assay for cytotoxicity screening. *Nature Protocols*, *1*(3), 1112–1116. http://doi.org/10.1038/nprot.2006.179

Wassermann, M., Wassermann, D., Cucos, S., & Miller, H. J. (1979). World Pcbs Map: Storage And Effects In Man And His Biologic Environment In The 1970s. *Annals of the New York Academy of Sciences*, *320*(1 Health Effect), 69–124. http://doi.org/10.1111/j.1749-6632.1979.tb56594.x

Zhang, H., Lawson, W. E., Polosukhin, V. V, Pozzi, A., Blackwell, T. S., Litingtung, Y., … Trembath, R. (2007). Inhibitor of differentiation 1 promotes endothelial survival in a bleomycin model of lung injury in mice. *The American Journal of Pathology*, *171*(4), 1113–26. http://doi.org/10.2353/ajpath.2007.070226

CHAPTER 3

The genome-wide mapping of ID3 target genes reveals the role of ID3 in the proliferation and dysregulation of lung endothelial cells.

Abstract

Background: ID3 is a key gene implicated in maintaining the stemness of embryonic and adult stem cells, which lacks the basic DNA-binding domain. ID3 appears to re-activate in several human cancers like NSCLC and in diseases such as PAH. However, the detailed response programming of ID3 in HPMEC has not yet been defined.

Results: We show that overexpression of the ID3 gene in HPMEC cells increases cell proliferation. We then comprehensively characterize the ID3 response program by an integrated analysis using advanced genomic technologies including Chromatin Immunoprecipitation Sequencing (ChIP-Seq) and RNA sequencing. Using ChIP-Seq technology, we identified 2336 ID3 binding regions (ChIP peaks) in the HPMEC ID3+ (ID3 overexpressing HPMEC cell line) cells genome. ID3 binding regions with the highest peak score demonstrated the sequence GAGAGAGAGAGAGAG that occurred a total of 900 instances in 2493 ID3 binding regions. We then further with the help of Genomic Regions Enrichment of Annotations Tool (GREAT) assigned biological meaning to 2493 regions. GREAT associated 2336 regions with 2834 genes. RNA-Seq analysis identified total of 1061 genes whose expression altered in response to ID3 overexpression in HPMEC cells. Out of 1061 genes, 63 genes showed significant (>1 log2 (FC)) fold change. RNA Seq data demonstrated HSPA8, SH3BP5L, EEF2K proteins were down-regulated and proteins like EGR1, TNFAIP3, and KLF10 were up-regulated. Furthermore, the combination of ChIP-Seq and transcriptome data analysis used as a compelling approach to decipher the regulation of gene expression with the help of Binding and expression target analysis (BETA) software. Integration of ChIP Seq with

differential gene expression data of RNA Seq to infer direct target genes of ID3 revealed interesting findings, which includes total of 26 genes significantly regulated by ID3.

Conclusions: We present an integrated analysis of ChIP-Seq and RNA sequencing representing the ID3 response program responsible for HPMEC cells proliferation. The insights gained from our integrated analysis further our understanding of the potential actions of ID3 in carcinogenesis and serve as a useful resource to further develop pharmaceutical drugs targeting the ID3 protein.

Key words: endocrine disruptors, ID3 Lung endothelial cells, PAH, plexiform lesions.

Introduction-

ID3 is one of the members of the ID protein family, which consists of four genes (ID1–ID4). ID3 does not have DNA binding domain and interacts with other proteins via helix-loop-helix (HLH) motif. It has been reported that ID3 acts as a transcription regulator which regulate transcription in a dominant-negative manner by dimerizing with basic HLH transcription factors like E12, E47, E2-2, and HEB (Benezra, Davis, Lockshon, Turner, & Weintraub, 1990; Doke, Avecilla, & Felty, 2018; Norton, 2000). ID3 is a pleiotropic protein involved in the modulation of a variety of biological processes such as cell differentiation, cell proliferation, cell cycle control, senescence, apoptosis or angiogenesis, and metastasis (Lyden et al., 1999; Lyden et al., 1999b; Ruzinova & Benezra, 2003). Previously, we have shown that ID3 overexpression reprograms adult endothelial cells to acquire vascular stem markers including CD133. ID3 is mostly

implicated in maintaining the stemness of embryonic and adult stem cells while its expression is down-regulated as tissues mature (Lyden et al., 1999b). However, research studies showed reactivation of ID3 expression in mature prostate and breast tissues involved in aggressive growth of cancer (Gupta et al., 2007; Sharma, Patel, & Chaudhary, 2012). The proliferative ECs have been reported to be an aggressive endothelial phenotype associated with a poor prognosis in non-small cell lung cancer and severe idiopathic pulmonary arterial hypertension (Tanaka et al., 2003; Tuder, Groves, Badesch, & Voelkel, 1994). We have also shown the exposure of E2 and EED like PCB153 to ECs increases ROS. Since ID3 is redox sensitive gene, it acts an important determinant of ROS-induced proliferation of E2 and PCB153 exposed ECs (Das & Felty, 2014a, 2014d; Felty & Porther, 2008b; Mueller et al., 2002). However, little is known regarding the molecular contribution of ID3 to vascular reprogramming and how it interacts with other transcription factors (TFs), which bind to DNA and regulate the transcriptional machinery to activate or repress the expression of target genes. In the current study, we used ChIP-Seq technology and conducted a genome-wide target identification for ID3 binding in lung EC cells. In addition to this, we used RNA sequencing for accurate detection and measure of gene expression of differential expressed (DE) genes between HPMEC and ID3 overexpressed HPMEC. Further, we performed an integrated analysis of ChIP and RNA Seq data, which revealed key response programs that potentially may play important role in lung ECs proliferation.

Materials and Methods
Cell culture and treatment conditions

Primary cells, Human pulmonary microvascular endothelial cells or Human lung endothelial cells (HPMEC-ST1.6R). HPMEC cells were maintained in Dulbecco's modified Eagle's medium (DMEM-F12) supplemented with 5% (v/v) fetal bovine serum (FBS). Cells were cultured at 37 ºC in a humidified atmosphere with 5% CO_2.

Estrogenic chemical treatments

Stock solutions of PCB153 and 17β-estradiol were prepared in dimethyl sulfoxide (DMSO). The same amount of DMSO as in PCB and estradiol-treated cells were added to control cultures. The level of DMSO in experimental media was less than 0.1%. PCB blood levels have been reported to reach approximately 1000 ng/mL ($\sim$3 μM) in occupationally exposed individuals (Wassermann et al., 1979). Our unpublished data showed a significant increase in PCB-induced vascularization with PCB153 concentrations of 10–100 ng/mL. Based on known PCB blood levels from occupational exposure and our preliminary results; we chose a PCB dose of 100 ng/mL ($\sim$0.3 μM) to expose endothelial cells. Our previous studies also showed endothelial cell proliferation and vascular tube formation at physiological doses of 17β-estradiol (Das & Felty, 2014a; Felty, 2006; Felty & Porther, 2008b) therefore, HPMEC and HPMEC ID3[+] were exposed to estradiol at 100 ng/mL ($\sim$3.6 nM). PCB congeners 2,2′,4,4′,5,5′-hexachlorobiphenyl (PCB153) were purchased from AccuStandard (Newhaven, CT) and dissolved in dimethyl sulfoxide (DMSO). All other chemicals and reagents were purchased from Sigma (St. Louis, MO).

ID3 overexpression

The HPMEC cells were stably transfected with either Precision LentiORF for ID3 (Thermo Scientific Open Biosystems) or empty vector lentiviral pLEX-JRED/TurboGFP by the trans-lentiviral packaging kit with Express-in transfection reagent according to the manufacturer's instructions. We used the MOI (multiplicity of infection) of 25 and selected cells that overexpressed ID3 with blasticidin S (5 mg/ml) as per manufacturer's instructions. Cells expressing TurboGFP were identified by fluorescence microscopy.

Experiment procedure for Chromatin Immunoprecipitation (ChIP) Sequencing (ChIP-Seq)

Chromatin-immunoprecipitation followed by sequencing of the immuno-precipitated DNA is an expeditiously becoming a very important method for understanding the mechanisms of protein: DNA interactions. It was first introduced and described in 2007 (Robertson et al., 2007). It is a powerful tool for understanding the mechanisms of gene regulation by transcription factors and modified histones. ChIP-Seq process is very tedious and time-consuming to perform, involving many steps and variables that must be optimized by each investigator in their model system. To perform ChIP-Seq, chromatin is isolated from cells or tissues and fragmented. Specific antibodies against chromatin-associated proteins are used to enrich for specific chromatin fragments. The DNA is recovered, sequenced and aligned to a reference genome to determine specific protein binding loci.

ChIP sequencing makes use of the power of massively parallel or next-generation sequencing (NGS) which have increased our knowledge of transcription factor biology, DNA methylation and histone modifications. ChIP- Seq is a counting assay that uses only short reads to align to the genome, but requires millions of them to provide meaningful data.

In order to perform ChIP-Seq analysis, we used QIAGEN Company made EpiTect ChIP OneDay Kit, which is designed to isolate and purify high-quality DNA bound to our protein of interest. The EpiTect ChIP OneDay Kit contains all buffers and components needed for immunoprecipitation, crosslink reversal, and genomic DNA purification for detection. To begin with chromatin immunoprecipitation, we seeded the cells and treated with treatment for 48 hours. Then, we cross-linked cells with formaldehyde. The chromatin containing covalent complexes between genomic DNA and all nuclear factors is isolated and sheared by sonication into 100-300 base pair (bp) sizes. Immunoprecipitation pulls down not only the target nuclear factor of interest but also any specifically bound genomic DNA sequences. Reversal of the chemical cross-linking and nucleic acid purification prepare the DNA for detection by sequencing. The sequencing of ChIP processed DNA is carried out by Center for Genome Technology (CGT) of John P. Hussman Institute for Human Genomics at University of Miami Miller School of Medicine. CGT facility used DNA SMART ChIP-Seq Kit (Clontech Laboratories, Inc.) which generated Illumina-compatible sequencing libraries from 100 pg–10 ng of single-stranded (ss) or double-stranded (ds) DNA resulting from chromatin immunoprecipitation experiments. Then we performed ChIP-Seq analysis using various tools available at Galaxy, which is an open source, web-based platform (Afgan et al., 2016). It begins with

121

mapping of trimmed sequence reads to a reference genome. Next, peaks are found using peak-calling algorithms. To further analyze the data, we performed differential binding and motif analyses, which are common end points of ChIP-Seq workflows. In order to perform ChIP-Seq analysis, we used QIAGEN Company made EpiTect ChIP OneDay Kit, which is designed to isolate and purify high-quality DNA bound to our protein of interest. ChIP samples were prepared from lung EC WT, lung EC ID3$^+$, and lung EC ID3$^+$ + E2 cells as follows: 80.0 x 10^6 cells from each of all the three cell lines were cross-linked with 1% formaldehyde for 10 minutes at room temperature. Cross- linking was stopped by the addition of harvesting buffer to a final concentration of 125 mM, and cells were washed twice with PBS. Centrifuged at 2000 rpm, removed supernatant. Added 800 µL to 1 mL Lysis buffer to ppt. The cell pellet was resuspended consecutively in ChIP lysis buffers and sonicated for 6 minutes, 12 cycles (30 seconds high frequency pulsing followed by 30 seconds resting) using the Bioruptor sonicator (Diagenode, Denville, NJ). In addition to this, added 10 µL of nuclease + 20 µL reaction mixture to the sample and incubate 20 min at 37 ℃. Added 40 µL Stop buffer to produce chromatin fragments of 100 – 300 bp on average. Ran 1.2% agarose gel with sample and molecular marker to check whether the chromatin broken into appropriate sizes. Added 50 µL Protein A Beads for each IP Fraction. To prepare these beads, 100 µL of protein beads were incubated overnight with 10 µg SC-490 polyclonal ID3 antibody (SantaCruz Biotechnology), at 4 °C. The following day, the beads were rinsed and added to the sheared chromatin and incubated over- night at 4 °C. Samples were then rinsed one time each with buffer I,II,III,IV and V. The cross-linking of antibody with protein of interest-DNA complex was reversed by treating with RNAse A and Proteinase K. Procedure

controls included an input condition, obtained before DNA-protein complex sonication and further used during ChIP-Seq assays as normalization sample, and non-immunoprecipitated DNA (non-IP DNA), which was obtained just prior to ID3 immunoprecipitation. Added 400 µL Column Binding Buffer to each 200 µL IP Fraction sample. Mixed well by pipetting. Immediately transferred the entire supernatant to its own DNA Spin Column (in 2 mL Collection Tube) without wetting the rim, close the lid, and centrifuge at 11,000 × g for 1 min at room temperature. Added 600 µL Column Wash Buffer, containing ethanol, to each column, without wetting the rim, close the lid, and centrifuge at 11,000 × g for 1 min at room temperature. Placed the DNA Spin Column into a clean Elution Tube, and discarded the Collection Tube. Added 100 µL Elution Buffer to the center of the column membrane. This elution step with another 100 µL of Elution Buffer carried out for a final elution volume of 200 µL for each ChIP DNA Sample. The purified ChIP DNA Samples can be stored at -20°C. The purified DNA further processed with the help of Clean & Concentrator™ from ZYMO RESEARCH CORP for the rapid purification and concentration of high quality DNA. The ChIP DNA Clean & Concentrator™ employs a single buffer system that allows for efficient DNA adsorption to the matrix of the supplied Zymo-Spin™ Column. Then we measured the concentration of DNA in samples and inputs with the help of BioRad spectrophotometer.

Experimental protocol for Chromatin Immunoprecipitation-

Aim- To determine genome-wide mapping of ID3 target genes by using ChIP-Seq experiment. ChIP Ready Chromatin Preparation: This section describes the preparation of ChIP Ready Chromatin. It covers cell cross- linking, harvesting, lysis, and chromatin

shearing based on adherent mammalian cell lines. Certain steps may need to be optimized for experiment-dependent performance. This protocol also provides several stopping points for the flexibility to optimize those steps.

Things to do before starting-

In order to immunoprecipitate the ID3 bound chromatin, we seeded the 20.0×10^6 cells of lung EC vector, lung EC ID3$^+$, and lung EC ID3$^+$ + E2 cells on each of four 150 mm petri dishes to achieve 80.0×10^6 cells for all three cell lines.

(i) Lung Vector Control (HPMEC), (ii) Lung ID3, (iii) Lung ID3 + E2.

We treated cells with vehicle or E2 overnight. We did fixing and harvesting of chromatin from all plates - Total volume per sample = Follow the steps of Cross-Linking Cells with 1% formaldehyde. Add 1 mL cell harvesting buffer and centrifuge at 2000 rpm. Remove supernatant. Add 800 μL to 1 mL lysis buffer to ppt. We warmed IP Lysis Buffer to room temperature to prevent precipitation. We thawed Protease Inhibitor Cocktail (PIC) at room temperature. This product contains DMSO and will remain frozen below 18.4 °C.

Cross-Linking Cells- We prepared following buffers, scale up accordingly if necessary.

i. $1 \times$ PBS Buffer: 3.3 mL $10 \times$ PBS to 29.7 ml ddH$_2$O. Mixed well and placed on ice.

ii.   FRESH Fixing Buffer (1% Formaldehyde): We added 0.27 mL 37% Formaldehyde to 9.73 mL $1 \times$ PBS Buffer. Mixed and stored at room temperature in a fume hood.

iii.   Cell Harvesting Buffer: We added 15 μL Protease Inhibitor Cocktail (PIC) to 3 mL ice-cold 1 × PBS Buffer. Mix well and place on ice.

We aspirated cell culture medium from the dish. Add 10 mL FRESH Fixing Buffer to the dish. We incubated at 37 °C for 10 min. We added 1.1 mL Stop Buffer to the dish and swirl to mix thoroughly. We incubated at room temperature for 5 min. We aspirated buffer as much as possible. NOTE: From this point forward, keep the samples on ice at all times. We added 9 mL ice-cold 1 × PBS to wash the fixed cells, rocking the dish for 5 seconds. We poured off the wash buffer. Repeat once more. b. Harvesting Cells- We added 1.5 mL ice-cold Cell Harvesting Buffer to the dish. We placed dish on ice at an angle and scrape cells down to one edge of the dish with a silicone cell scraper or rubber policeman. We centrifuged the Sonication Tube at 800 × g for 10 min at 4 °C to pellet the fixed cells. Remove the supernatant. Repeat harvesting and centrifugation steps as many times as necessary to collect all of the cells. Store the cell pellet at -80 °C.

c. Cell Lysis

We added 2.2 μL Protease Inhibitor Cocktail (PIC) to 420 μL IP Lysis Buffer in a standard 1.5 mL Eppendorf tube. Mix well by pipetting. We added entire volume of IP Lysis Buffer with PIC to the cell pellet. 100 μL IP Lysis Buffer is recommended to lyse 0.5 to 1.5 million cells. We resuspended cells completely by pipetting. Incubate on ice for 10 -15 min, mixing every 5 min. Continue with next step OR store the lysate at -80 °C.

d. Shearing Chromatin- Chromatin Sonication: 800 μL total chromatin volume per sample sonicated for 6 min (12 cycles). We aliquot 800 uL volume into two 400 μL vials per sample. 800 μL = 2 x 400 μL vials per sample.

We stored one set of 400 µL samples in -80 ℃ freezer. We added 10 µL of nuclease + 20

µL reaction mixture to the sample and incubated 20 min at 37 ℃. The added 40 µL Stop

buffer. We ran 1.2% agarose gel with sample and molecular marker. Chromatin should be

broken down by sonicator into DNA pieces of size ranges from 100 bp to 300 bp as

shown in Figure 57.
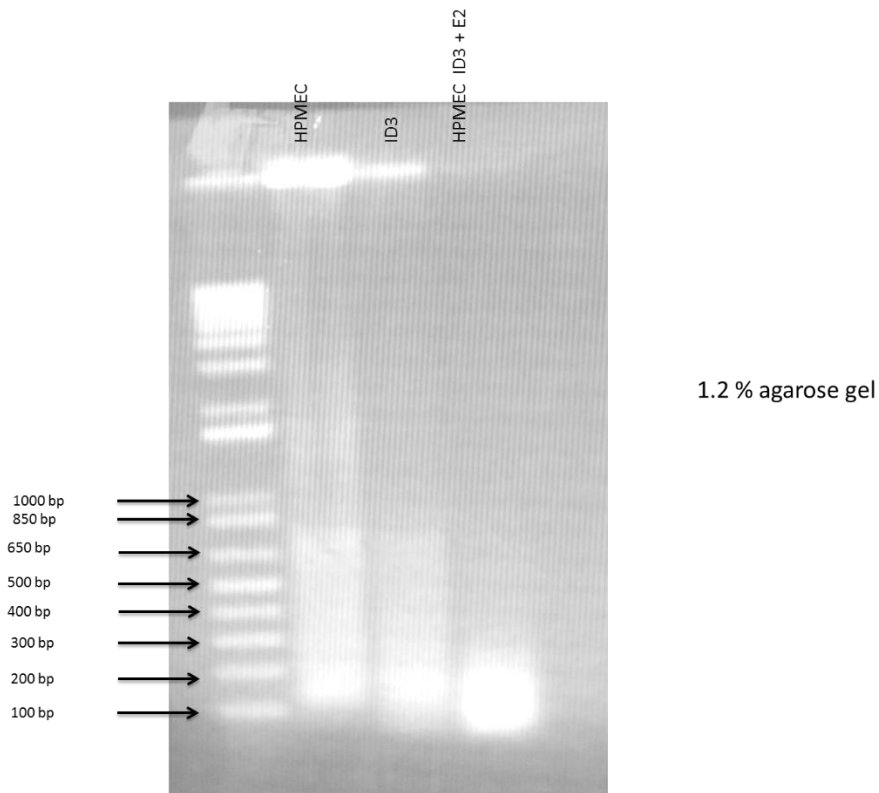
Figure 57- The picture of 1.2% agarose gel with sample and molecular marker by using UVITEC UVidoc instrument.

Chromatin is broken down by sonicator into DNA pieces of size ranges from 100 bp to 300 bp for HPMEC (Control), HPMEC ID3$^+$, and HPMEC ID3$^+$ + E2 cell samples.

Chromatin Immunoprecipitation (ChIP) Introduction-

In order to perform ChIP-Seq analysis, we used QIAGEN Company made EpiTect ChIP OneDay Kit, which is designed to isolate and purify high-quality DNA bound to our protein of interest. ChIP samples were prepared from lung EC WT, lung EC ID3$^+$, and lung EC ID3$^+$ + E2 cells as follows: 80.0 x 10^6 cells from each of all the three cell lines were cross-linked with 1% formaldehyde for 10 minutes at room temperature. Cross-linking was stopped by the addition of harvesting buffer to a final concentration of 125 mM, and cells were washed twice with PBS. Centrifuged at 2000 rpm, removed supernatant. Added 800 µL to 1 mL Lysis buffer to ppt. The cell pellet was resuspended consecutively in ChIP lysis buffers and sonicated for 6 minutes, 12 cycles (30 seconds high frequency pulsing followed by 30 seconds resting) using the Bioruptor sonicator (Diagenode, Denville, NJ). In addition to this, added 10 µL of nuclease + 20 µL reaction mixture to the sample and incubate 20 min at 37 ℃. Added 40 µL Stop buffer to produce chromatin fragments of 100 – 300 bp on average. Ran 1.2% agarose gel with sample and molecular marker to check whether the chromatin broken into appropriate sizes. Added 50 µL Protein A Beads for each IP Fraction. To prepare these beads, 100 µL of protein beads were incubated overnight with 10 µg SC-490 polyclonal ID3 antibody (SantaCruz Biotechnology), at 4 °C. The following day, the beads were rinsed and added to the sheared chromatin and incubated over- night at 4 °C. Samples were then rinsed one time each with buffer I,II,III,IV and V. The cross-linking of antibody with protein of interest-DNA complex was reversed by treating with RNAse A and Proteinase K. Procedure controls included an input condition, obtained before DNA-protein complex sonication and further used during ChIP-Seq assays as normalization sample, and non-

immunoprecipitated DNA (non-IP DNA), which was obtained just prior to ID3 immunoprecipitation. Added 400 μL Column Binding Buffer to each 200 μL IP Fraction sample. Mixed well by pipetting. Immediately transferred the entire supernatant to its own DNA Spin Column (in 2 mL Collection Tube) without wetting the rim, close the lid, and centrifuge at $11,000 \times g$ for 1 min at room temperature. Added 600 μL Column Wash Buffer, containing ethanol, to each column, without wetting the rim, close the lid, and centrifuge at $11,000 \times g$ for 1 min at room temperature. Placed the DNA Spin Column into a clean Elution Tube, and discarded the Collection Tube. Added 100 μL Elution Buffer to the center of the column membrane. This elution step with another 100 μL of Elution Buffer carried out for a final elution volume of 200 μL for each ChIP DNA Sample. The purified ChIP DNA Samples can be stored at -20 °C. The purified DNA further processed with the help of Clean & Concentrator™ from ZYMO RESEARCH CORP for the rapid purification and concentration of high quality DNA. The ChIP DNA Clean & Concentrator™ employs a single buffer system that allows for efficient DNA adsorption to the matrix of the supplied Zymo-Spin™ Column. Then we measured the concentration of DNA in samples and inputs with the help of BioRad spectrophotometer.

Illumina high-throughput Chromatin Immunoprecipitation -Sequencing (ChIP-Seq)

Chromatin-immunoprecipitation (ChIP) followed by sequencing of the immuno-precipitated DNA is an expeditiously becoming a very important method for understanding the mechanisms of Protein: DNA interactions and gene regulation by transcription factors. The sequencing of ChIP processed DNA is carried out by Center for Genome Technology (CGT) of John P. Hussman Institute for Human Genomics of Miller

School of Medicine at University of Miami. CGT facility used DNA SMART ChIP-Seq Kit (Clontech Laboratories, Inc.) which generated Illumina-compatible sequencing libraries from 100 pg–10 ng of single- stranded (ss) or double-stranded (ds) DNA resulting from chromatin immunoprecipitation (ChIP) experiments (Nos, 2012). Clontech's template switching technology, known as SMART® (Switching Mechanism at 5' End of RNA Template), has been used as a basis for cDNA synthesis for next-generation sequencing (NGS) applications (Zhu, Machleder, Chenchik, Li, & Siebert, 2001). The CGT facility generated ChIP-Seq libraries with the DNA SMART ChIP-Seq Kit, a priming site is first added to the 3' end of the DNA template using the Terminal Deoxynucleotidyl Transferase. This is followed by annealing of a proprietary DNA SMART Poly(dA) Primer, which anneals to the T-tail added by the Terminal Deoxynucleotidyl Transferase. This primer is then used by the SMARTScribe™ Reverse Transcriptase (RT) to copy the DNA strand. When the SMARTScribe RT reaches the 5' end of the DNA template, the enzyme's terminal transferase activity adds a few additional nucleotides to the 3' end of the newly synthesized DNA. The carefully designed DNA SMART Oligonucleotide base-pairs with these additional non-template nucleotides and creates an extended template, enabling the SMARTScribe RT to continue replicating to the end of the oligonucleotide. Sequencing libraries are then generated by PCR-mediated addition of Illumina adapters using primers compatible with regions on the DNA SMART Poly(dA) Primer and the DNA SMART Oligonucleotide ("Illumina Adapter Sequences Illumina Adapter Sequences Illumina Adapter Sequences," 2017).
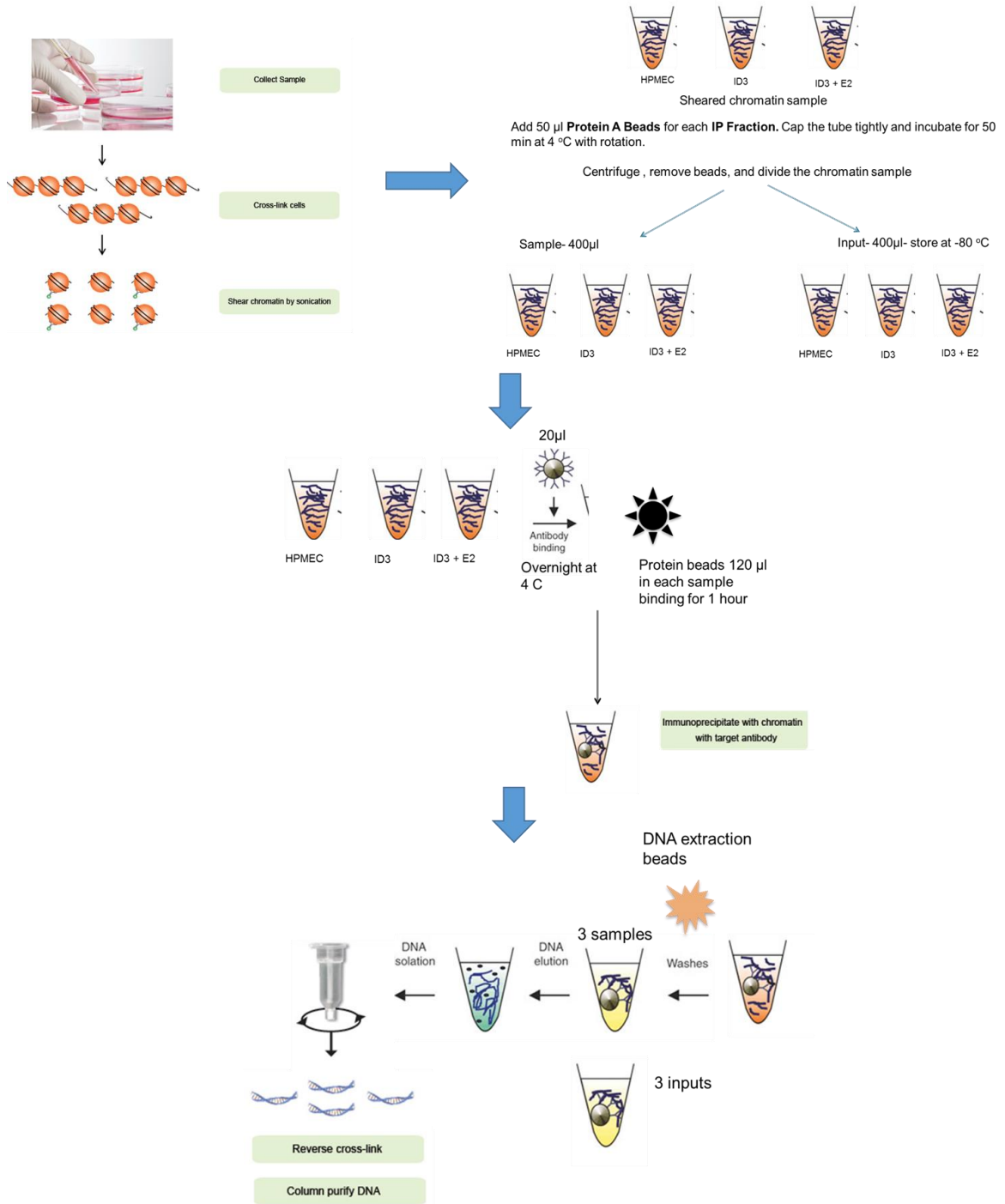
Figure 58- The diagrammatic representation of workflow of ChIP-seq experiment carried out using HPMEC (Control), HPMEC ID3+, and HPMEC ID3$^+$ + E2 cell samples.

After elution of DNA sample, we used the Chromatin Immunoprecipitation (ChIP) DNA Clean & Concentrator™ from ZYMO RESEARCH CORP. for the rapid purification and concentration of high quality DNA. The ChIP DNA Clean & Concentrator™ employs a single buffer system that allows for efficient DNA adsorption to the matrix of the supplied Zymo-Spin™ Column as shown in Figure 59.

Reagent Preparation-

Before starting, we added 24 mL 100% ethanol to the 6 mL DNA Wash Buffer concentrate to obtain the final DNA Wash Buffer solution. In 1.5 mL micro centrifuge tube, add 5 volumes of ChIP DNA Binding Buffer to each volume of sample (5:1). Mixed briefly. We transferred mixture to a provided Zymo-Spin™ Column in a Collection Tube. Centrifuged at $\geq$ 10,000 x g for 30 seconds.

We added 200 µl Wash Buffer to the column. Centrifuge at $\geq$ 10,000 x g for 30 seconds. Repeat wash step. We added 6-100 µl Elution Buffer directly to the column matrix. Transfer the column to a new 1.5 mL micro centrifuge tube and centrifuge at $\geq$ 10,000 x g for 30 seconds to elute the DNA. Ultra-pure DNA is now ready for use for DNA quantification, sequencing and other molecular applications.

Figure 59- The diagrammatic representation of elution of clean DNA sample using the Chromatin Immunoprecipitation (ChIP) DNA Clean & Concentrator™ from ZYMO RESEARCH CORP.

Then we measured the concentration of DNA in samples and inputs with the help of BioRad spectrophotometer as shown in Table 3.

Sample details	Sample Conc.	Units	Volume (uL)	260/280 Ratio
Input vector	94.6295	pg	95µl	1.8603
Input ID3 expressing	67.155	pg	95µl	1.8065
Input ID3 expressing with estrogen treatment	31.844	pg	95µl	1.8621
Sample vector	195.9	pg	95µl	1.8292
Sample ID3 expressing	76.1805	pg	95µl	1.8706
Sample ID3 expressing with estrogen treatment	53.694	pg	95µl	1.8867

Table 3- The information about the sample like concentration, volume and quality.

DNA sequencing-

We performed ChIP sample and input sequencing analysis at Center for Genome Technology (CGT) of University of Miami Miller School of Medicine. This genomic facility center used DNA SMART ChIP-Seq Kit (Clontech Laboratories, Inc. A Takara Bio Company) which generated Illumina-compatible sequencing libraries from 100 pg– 10 ng of single- stranded (ss) or double-stranded (ds) DNA resulting from chromatin immunoprecipitation (ChIP) experiments.

Clontech's template switching technology, known as SMART® (Switching Mechanism at 5' End of RNA Template), has been used as a basis for cDNA synthesis for next-generation sequencing (NGS) applications (Zhu, Machleder, Chenchik, Li, & Siebert, 2001). The CGT facility generated ChIP-Seq libraries with the DNA SMART ChIP-Seq

Kit, a priming site is first added to the 3' end of the DNA template using the Terminal Deoxynucleotidyl Transferase. This is followed by annealing of a proprietary DNA SMART Poly (dA) Primer, which anneals to the T-tail added by the Terminal Deoxynucleotidyl Transferase. This primer is then used by the SMARTScribe™ Reverse Transcriptase (RT) to copy the DNA strand. When the SMARTScribe RT reaches the 5' end of the DNA template, the enzyme's terminal transferase activity adds a few additional nucleotides to the 3' end of the newly synthesized DNA. The carefully designed DNA SMART Oligonucleotide base-pairs with these additional non-template nucleotides and creates an extended template, enabling the SMARTScribe RT to continue replicating to the end of the oligonucleotide. Sequencing libraries are then generated by PCR-mediated addition of Illumina adapters using primers compatible with regions on the DNA SMART Poly(dA) Primer and the DNA SMART Oligonucleotide as shown in Figure 60 ("Illumina Adapter Sequences Illumina Adapter Sequences Illumina Adapter Sequences," 2017).
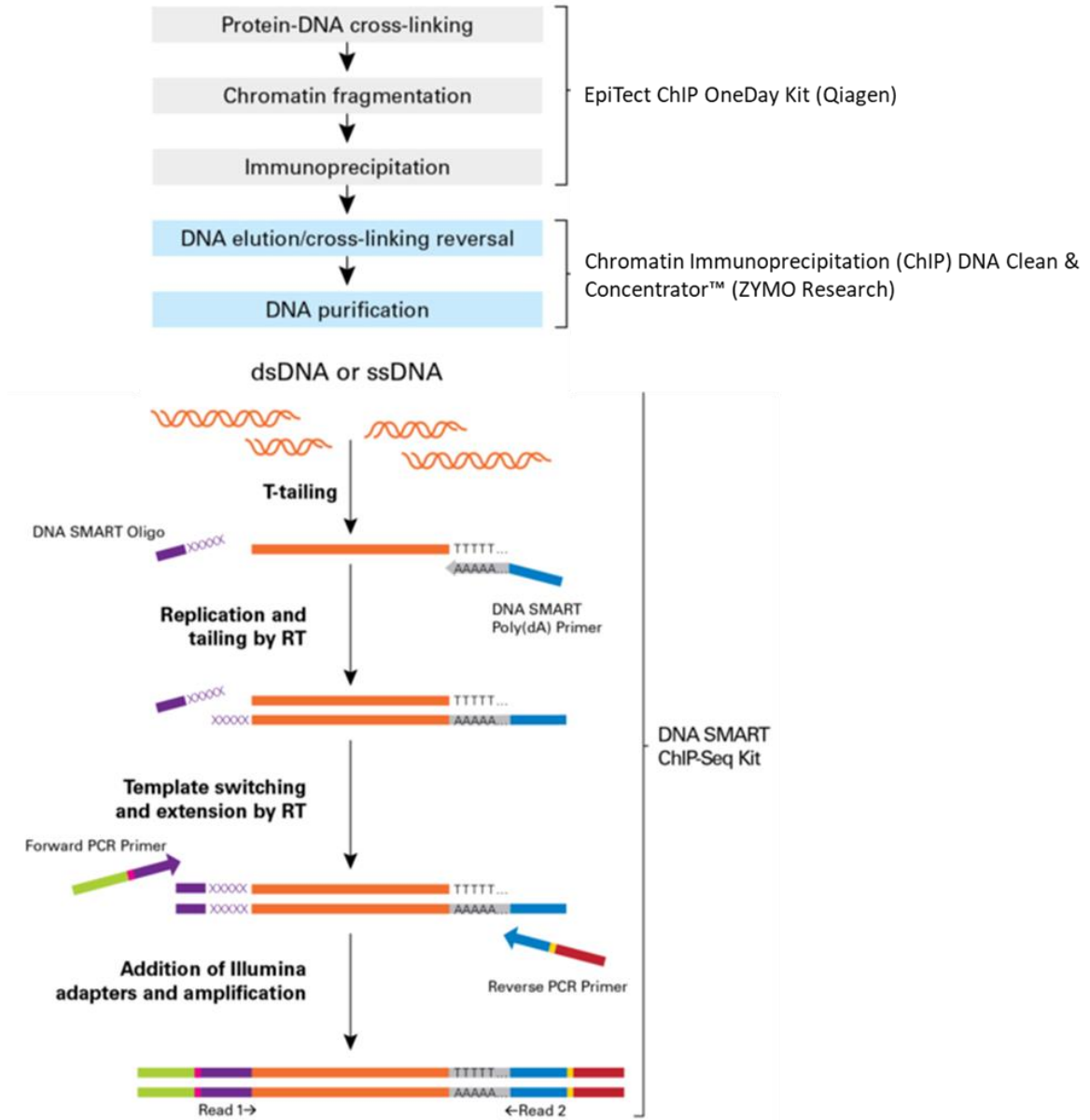
Figure 60- The diagrammatic workflow represents addition of adapter sequences at both ends of ChIP DNA using Clontech's template switching technology, known as SMART®

SAMPLE	SAMPLE ID	PF Clusters	Index Seq	Clontech DUAL Index ligated (Lot#1511520a)	TruSeq HT
Input vector	201700231-01	22,746,865	ATTACTCG-ATAGAGGC	R1-F2	D701-D502
Input ID3 expressing	201700232-01	24,084,017	TCCGGAGA-ATAGAGGC	R2-F2	D702-D502
Input ID3 expressing with estrogen treatment	201700233-01	21,705,440	CGCTCATT-ATAGAGGC	R3-F2	D703-D502
Sample vector	201700234-01	22,093,845	GAGATTCC-ATAGAGGC	R4-F2	D704-D502
Sample ID3 expressing	201700235-01	22,999,843	ATTCAGAA-ATAGAGGC	R5-F2	D705-D502
Sample ID3 expressing with estrogen treatment	201700236-01	24,738,084	GAATTCGT-ATAGAGGC	R6-F2	D706-D502

Table 4- It illustrates the list of Illumina adapters used to make sequencing libraries.

These adapters are used for labeling the samples in sequence analyzer machine

After adapter sequences addition at both ends of ChIP-Seq DNA, Illumina sequencing machine at CGT facility processed ChIP-Seq DNA information into computer files as shown in table 4. We received the raw ChIP-Seq input and sample files (.txt files) in the paired end sequenced format. Therefore, all the input and sample ChIP-Seq files have two files each (Part 1 and 2). All Illumina NGS (next-generation sequencing) systems are capable of paired-end sequencing as shown in Table 5. The paired-end sequencing reads are very crucial and significant in the case of detection of genomic rearrangements such as insertions, deletions, and inversions. It also facilitates users to detect the repetitive sequence elements, as well as gene fusions and novel transcripts. Essentially, it provides

superior alignment to a reference across DNA regions containing repetitive sequences, which improves the quality of the entire data set. It also produces longer contigs for *de novo* sequencing by filling gaps in the consensus sequence.

Field	Description	Examples
STUDY	Official HIHG project name OR PI name for external	AUT, PD, jHendrix
SPECIES	Species	h.sapiens
PREP_TYPE	protocol used for sample preparation	Agilent SS 38Mb exome, Nimblegen custom 1.5Mb, Illumina Multiplexing, whole genomic, ChIP-Seq, Illumina mRNAseq, Ribominus Whole Transcriptome, Nugene Whole Transcriptome, etc.
PLATFORM	platform	GA, Solid, HiSeq
RUN_TYPE	single-read=1x, paired-end=2x followed by read length; Full, Quad, Octet Refer to spots on SOLiD slide	1x36, 2x50, 2x75, Full 1x50, Octet 1x50
ALIGN_TO	reference used for alignment	whole genome, custom, chr3 only
HIHG_ID	HIHG generated sample number	200425179-12
CONCENTRATION	Concentration	10
CONC_UNIT	Concentration unit	ng/µL
METHOD	instrument used to measure concentration	nanodrop, bioA, qubit, picogreen, qPCR
VOLUME	DNA Volume in above concentration unit	55
VOL_UNIT	Volume unit	µL

Table 5- It illustrates the information about sequencing carried out by Illumina sequencing machine at CGT facility.

Methodology to analyze the ChIP-Sequencing data
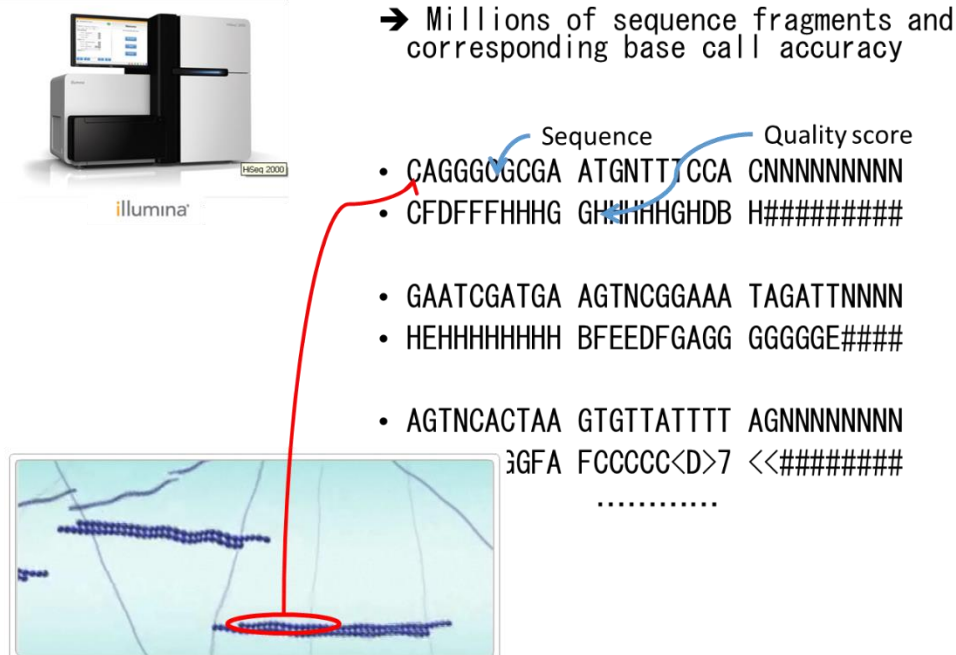
Figure 61- It shows in pictorial format that what kind of information we get in the form FASTA format from sequencing machine.

All Sanger format files saves the data in .txt files as shown in Figure 61. These files can be used as a raw data to process further analyses like peak finding, differential binding of protein of interest and comprehensive motif analysis (including motif discovery) of ChIP-Seq workflows. We analyzed the ChIP-Seq data by using tools available at Galaxy platform (Afgan et al., 2016). We uploaded the raw .txt files of input and samples in Galaxy server with the help of FileZilla which a free software, cross-platform file transfer protocol (FTP) application, consisting of FileZilla Client and FileZilla Server. We have used FileZilla to upload large sized .txt Sanger files to the Galaxy server as shown in Figure 62.
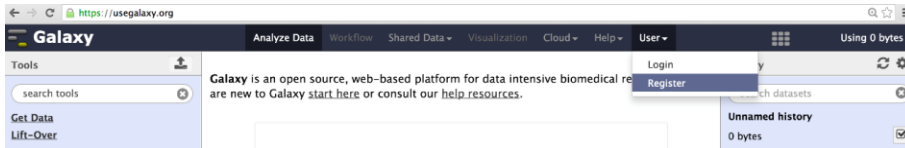
- Step 1: Go to the web server, usegalaxy.org
- Open a web browser and go to usegalaxy.org
- Let's Register
    - Click "User" ➔ "Register"

- Step 2: Upload the fastq file to Galaxy
    1. Click "Get Data" in the left column
    2. Click "Upload File from your computer
    3. Click "Paste/Fetch data"
    4. Type "fastqsanger" in the File Format box
    5. Type the "Hg38" to select the H. sapiens 2013 (Hg38)" genome
    6. In URL/Text box, type the following address
    https://goo.gl/M8p3eZ

    8. Click "Start" button
    9. Click "Close" button
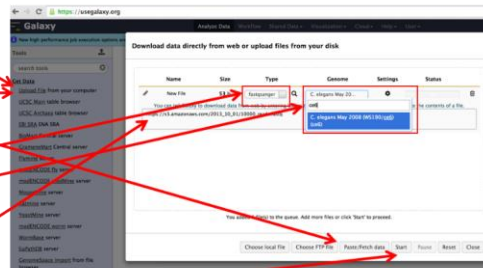
Getting Data: Uploading Large Files
Step 1: copy your file to
/nfs/galaxy/uploads/username@wi.mit.edu
using a sftp client
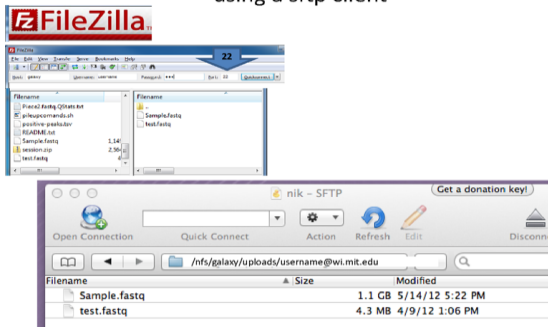
Figure 62- Diagrammatic representation of workflow of how to use Galaxy platform to upload the raw FASTA files from computer.

The FASTQ format files are saved in .txt files in local computers and contains DNA sequencing trace files. PHRED software evaluates and accredits certain quality value to the DNA sequences also known as bases (Ewing & Green, 1998). The quality value of each base is estimated in terms of probability of error. Sanger FASTQ format utilize ASCII 33–126 to encode PHRED qualities from 0 to 93 (i.e. PHRED scores with an ASCII offset of 33). In order to understand the file, Sanger FASTQ format restricted the choices to the ASCII printable characters 32–126 (decimal). The following example illustrates the application of PHRED score in Sanger FASTQ format files as shown in Figure 63.

FASTQ file:

@Read_ID
CAGGGCGCGA ATGNTTTCCA CNNNNNNNNN   ← Sequence
+
CFDFFFHHHG GHHHHHGHDB H#########   ← Quality score

- (example) We got a sequence of five bases
  Sequence: ACGTA
  The probability of error (P) is 1% for the first base A, 5% for the second and third bases, 10% for the fourth base, and 63% for the fifth base.

A	C	G	T	A
1%	5%	5%	10%	63%

P_{err} ➜ ASCII character

- Step 1: P_{err} ➜ PHRED score = $-10 \times \log_{10}(P_{err})$
  - (e.g.) P_{err} = 0.01 ➜ PHRED score = 20
- Step 2: Add +33 to the PHRED score
  - (e.g.) 20 + 33 = 53
- Step 3: Get the ASCII character that correspond to the above number
  - (e.g.) 53 ➜ 5

PHRED + 33 is called "FASTQ Sanger format"

ASCII code table

Figure 63- Diagrammatic illustrates the application of PHRED score in Sanger FASTQ format files.

140

Since the newer machines are in Sanger format 1.9, Galaxy requires 1.8. The uploaded .txt files converted using FASTQ Groomer within Galaxy in order to make Galaxy recognize the format (Blankenberg et al., 2010). We also ran FastQC analysis on these files. FastQC is a Java based quality control tool for raw high throughput sequence data (Andrews, 2016). The following tables illustrate the FastQC analysis of our pair ended raw datasets as shown in table 6. FastQC tool produces a basic text and a HTML output file that contain all of the results, including the following (Andrews, 2010). We have discussed some of the outputs from the following list:

➢ Basic Statistics
➢ Per base sequence quality
➢ Per sequence quality scores
➢ Per base sequence content
➢ Per base GC content

➢ Per sequence GC content
➢ Per base N content
➢ Sequence Length Distribution
➢ Sequence Duplication Levels
➢ Overrepresented sequences
➢ Kmer Content

➢ Basic statistics- It generates some simple composition statistics for the file analyzed. The following tables illustrate basic information of the ChIP-Seq files.

Name	Filename	Encoding	Total Sequences	Sequence length	%GC
Sample vector 1	qFelty_ID3chipseq_201700234-01_S_8_1.txt	Sanger / Illumina 1.9	22093845	35-125	40
Sample vector 2	qFelty_ID3chipseq_201700234-01_S_8_2.txt	Sanger / Illumina 1.9	22093845	35-125	36
Sample ID3 expressing	qFelty_ID3chipseq_201700235-01_S_8_1.txt	Sanger / Illumina 1.9	22999843	35-125	41
Sample ID3 expressing 2	qFelty_ID3chipseq_201700235-01_S_8_2.txt	Sanger / Illumina 1.9	22999843	35-125	40
Sample ID3 expressing with estrogen treatment 1	qFelty_ID3chipseq_201700236-01_S_8_1.txt	Sanger / Illumina 1.9	24738084	35-125	41
Sample ID3 expressing with estrogen treatment 2	qFelty_ID3chipseq_201700236-01_S_8_2.txt	Sanger / Illumina 1.9	24738084	35-125	40

Table 6- It illustrates the quality and information about sequences of ChIP Seq files

➢ Per base sequence quality – This output measures the overall range of quality values across all bases at each position in the FastQ file.

Green color
➔Good quality
➔sequence believable

Details on quality score: Y-axis

PHRED score = $-10 \times \log_{10}\left(P_{err}\right)$
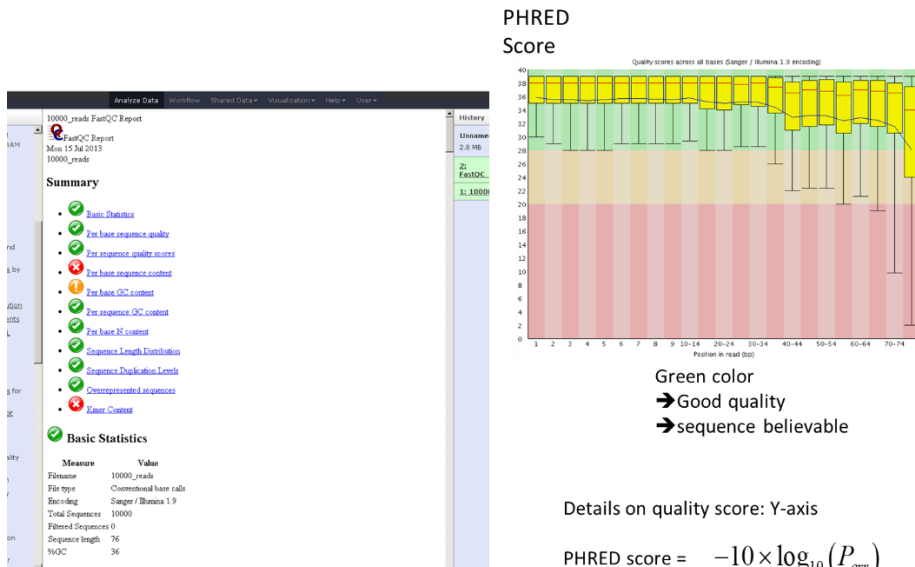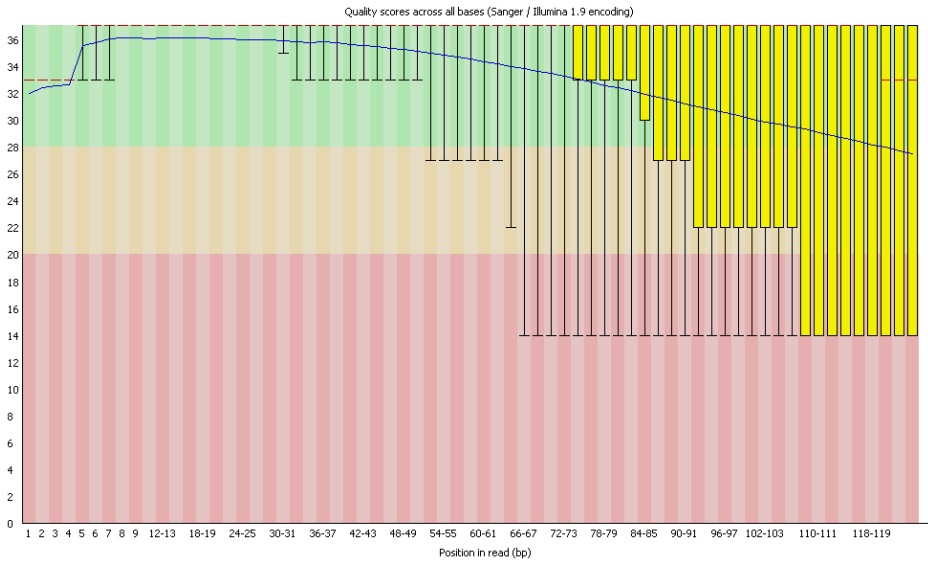
For each position a Box Whisker type plot is drawn. The elements of the plot are as follows:

- The central red line is the median value

143

- The yellow box represents the inter-quartile range (25-75%)

- The upper and lower whiskers represent the 10% and 90% points

- The blue line represents the mean quality

The y-axis on the graph shows the quality scores. The higher the score the better the base call. The background of the graph divides the y-axis into very good quality calls (green), calls of reasonable quality (orange), and calls of poor quality (red).

➢ Per sequence quality scores
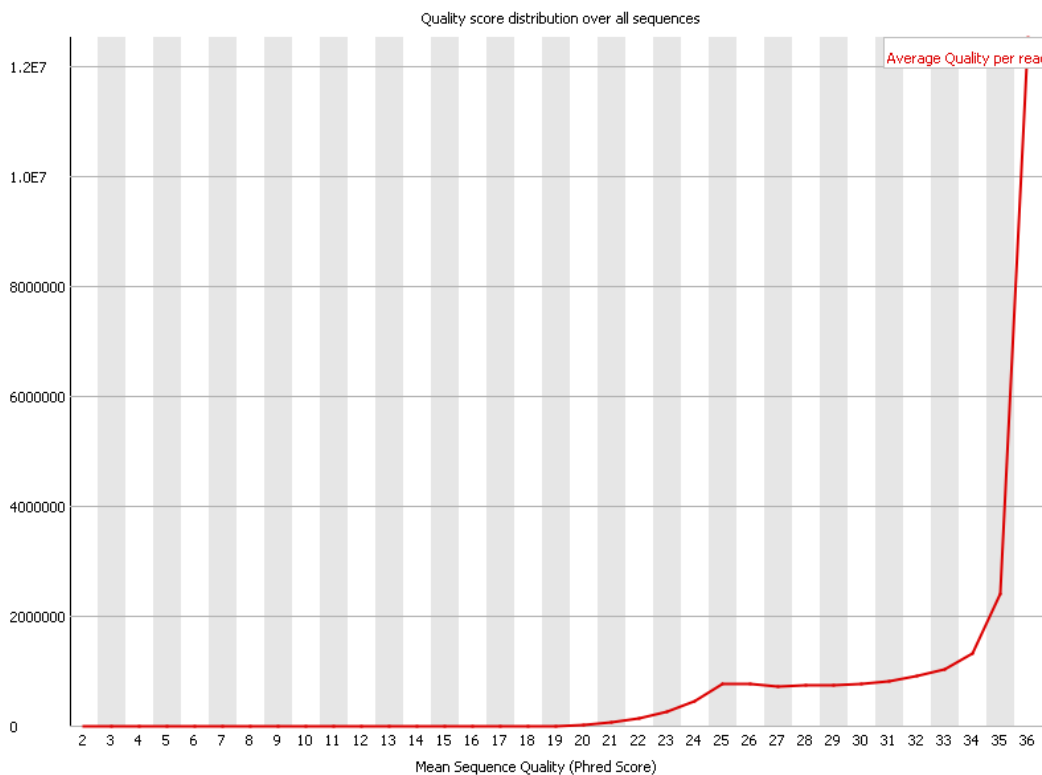
The per sequence quality score report allows to see if a subset of sequences have universally low quality values. A warning is raised if the most frequently observed mean quality is below 27 - this equates to a 0.2% error rate.

144

➢ Per Base Sequence Content- It plots out the proportion of each base position in a file for which each of the four normal DNA bases has been called.
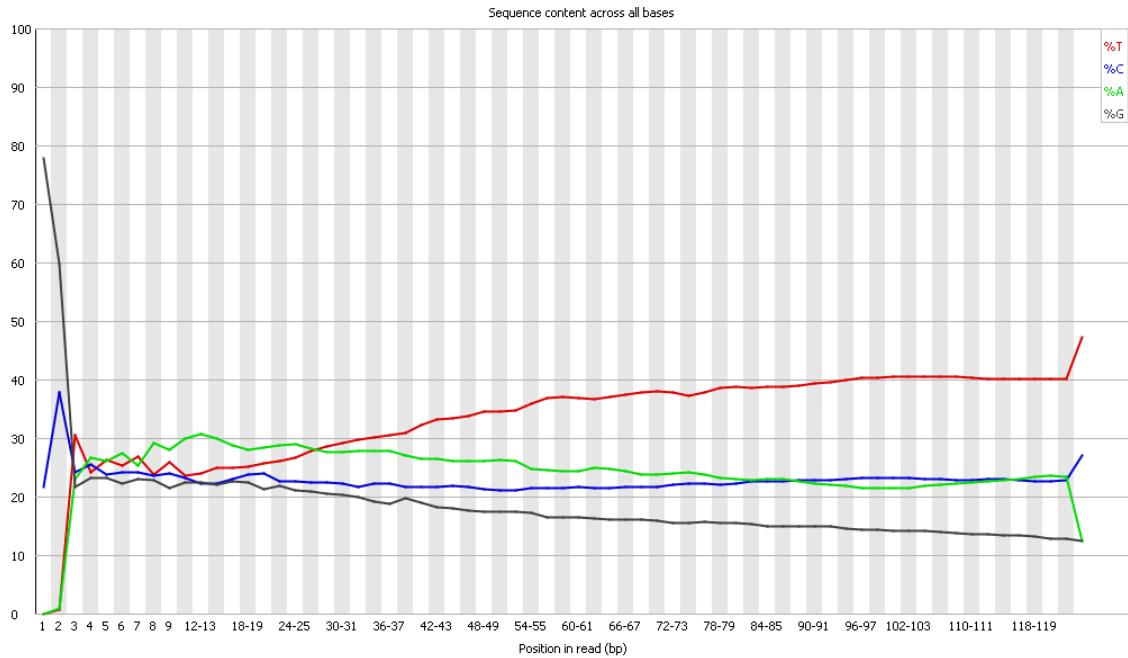
➢ Per sequence GC content- This module measures the GC content across the whole length of each sequence in a file and compares it to a modelled normal distribution of GC content.

✅ Per sequence GC content

GC distribution over all sequences
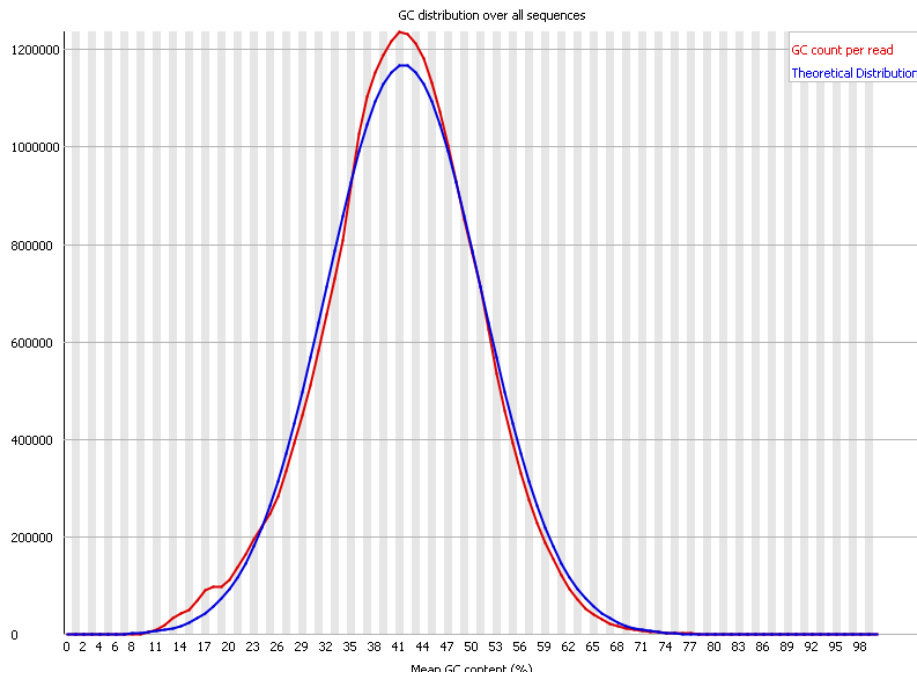
> ➤ Per base N content

✅ Per base N content

N content across all bases
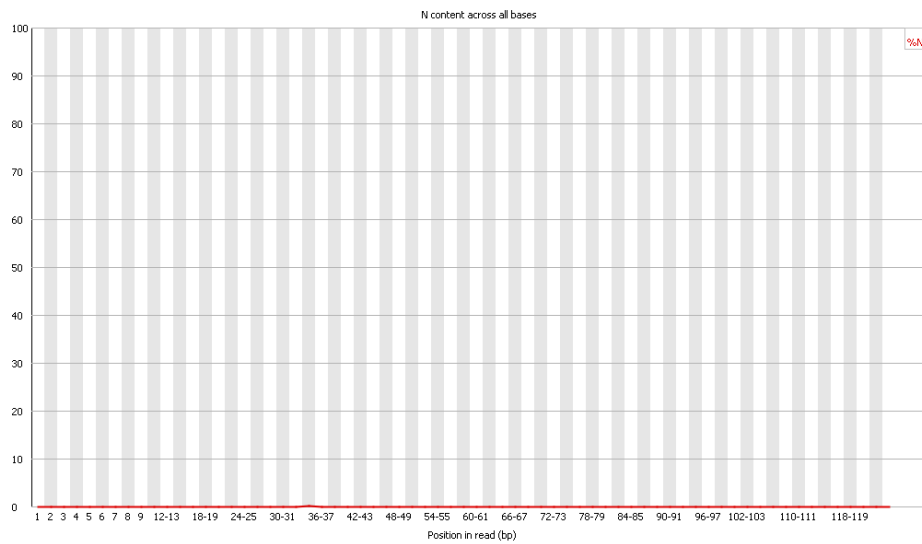
➢ Sequence Duplication Levels
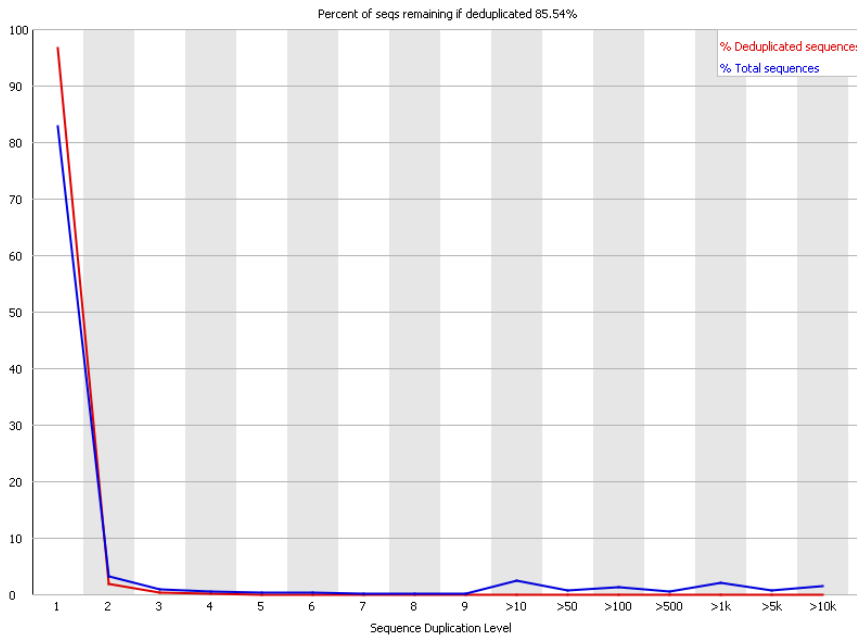
➢ Kmer content
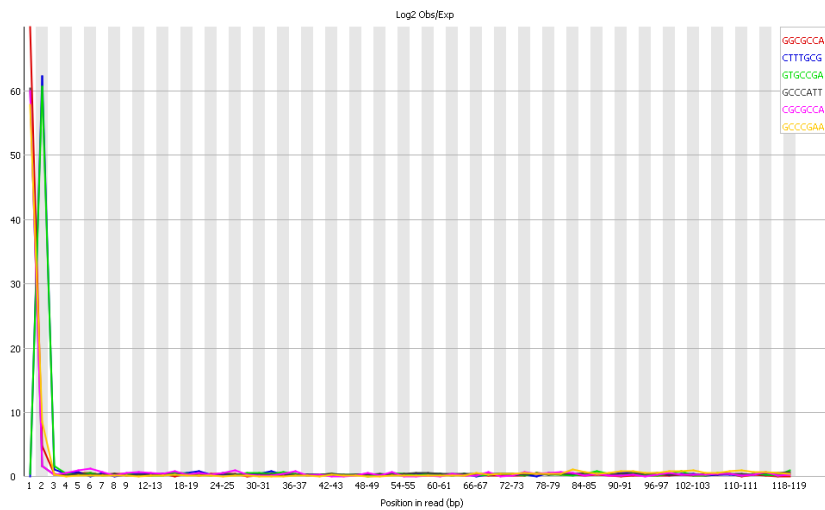
Out of all the outputs generated from FastQC analysis, per base sequence quality and sequence duplication levels are the most significant and useful outputs for further analysis. The sequence quality is calculated and measured on the basis of PHRED score (

Andrews, 2010; Cock, Fields, Goto, Heuer, & Rice, 2009). The bad quality sequences are removed from the raw Fastq files as shown in the following Figures by trimming the sequences. We used Trimmomatic tool to trim the bases in ChIP-Seq file. Trimmomatic tool is a more flexible, pair-aware and efficient preprocessing tool, optimized for Illumina NGS data (Bolger, Lohse, & Usadel, 2014). It is also equipped with the identification of short adapter sequences and quality filtering. It effectively removes the common 'adapter read-through' scenario, whereby the sequenced DNA fragment is shorter than the read length in both single and paired end data. In order to perform Trimmomatic operation on adapter sequence on our data, we selected TrySeq2 (paired-ended, for Illumina GAII) option.

The following diagrammatic process shows the Trimmomatic trimming tasks as shown in

Figure 64.
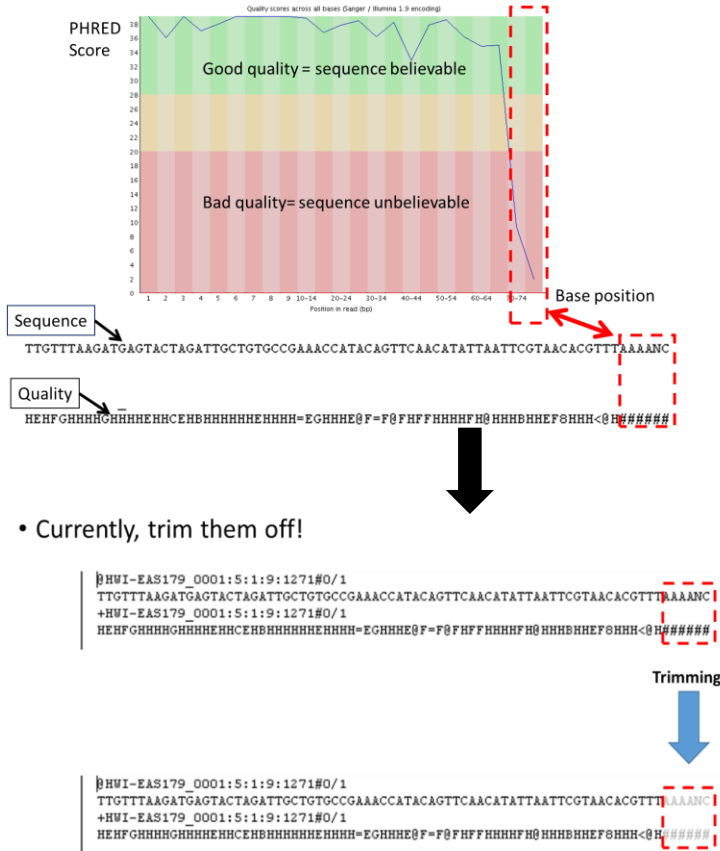
Figure 64- Diagrammatic representation of settings used in Trimmomatic software to trim the pair end adaptor sequences.

Trimmomatic process the paired end data in specific way and gives out the output of trimmed forward and reverse paired ended files. The following diagram reprenets process of Trimmomatic operation on paired end reads as shown in Figure 65 (Bolger et al., 2014; Usadellab,). We then used these trimmed forward and reverse paired ended files as sample input to align with reference human genome –Human Dec. 2013 (GRCh38/hg38) (hg38) (Miga et al., 2014a).
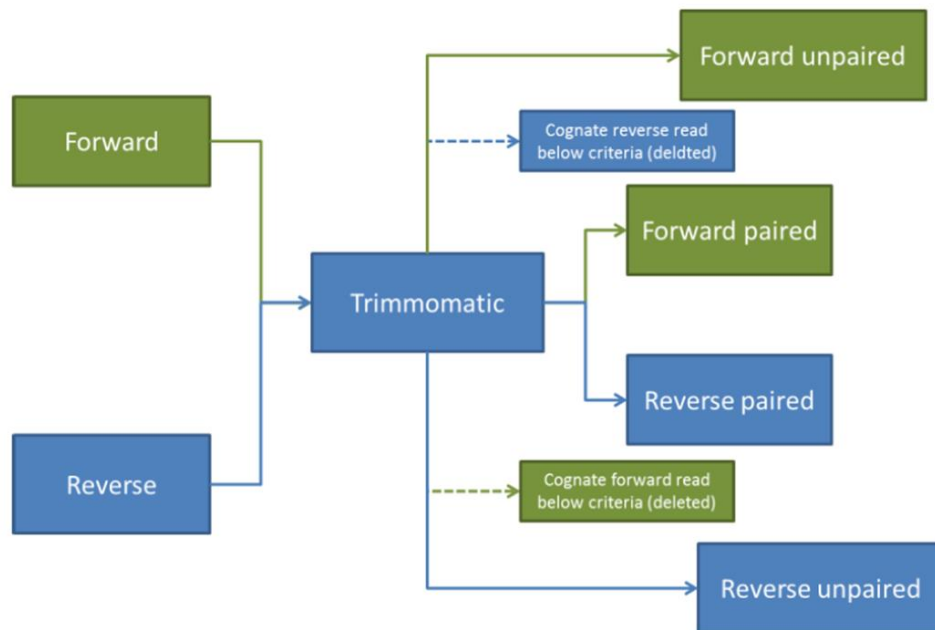
Figure 65- Flow of reads in Trimmomatic Paired End mode.

Trimmomatic process the paired end data in specific way and gives out the output of trimmed forward and reverse paired ended files

Read alignment/mapping with BOWTIE2

Bowtie2 is ultrafast read aligner and its algorithm coded by Ben Langmead. It is an ultrafast and memory-efficient tool for aligning sequencing reads to long reference sequences. Bowtie2 outputs alignments in BAM format, enabling interoperation with a

large number of other tools available at this site (Langmead and Salzberg, 2013). It

combines strengths of the full-text minute index with the flexibility and speed of

hardware-accelerated dynamic programming algorithms. We used Bowtie2 to align

trimmed paired reads with hg38 reference genome and used following settings shown in

the Figures 66 and 67.
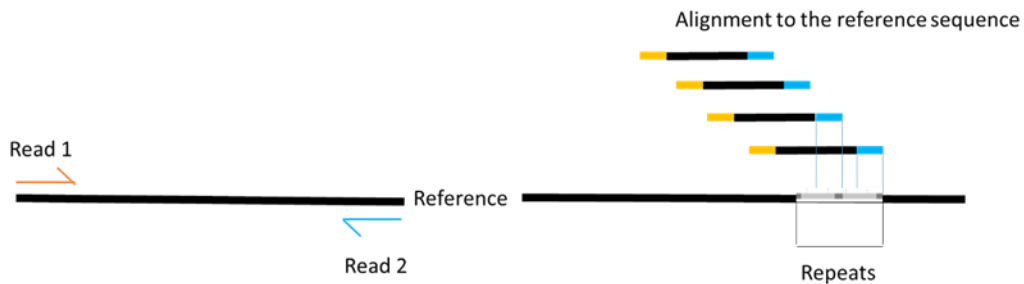
Figure 66- Paired end sequencing enabled both ends of fragments to be sequenced.
The distance between each paired read is known, alignment algorithms can use this
information to map the reads over repetitive regions more precisely.

Figure 67- Settings used in Bowtie2 for Paired end sequencing.

Bowtie2 gives output in the BAM format. The BAM format provides binary versions of store sequence data, both aligned as well as unaligned, but not in a human readable format. It is not easy to process by conventional text based processing programs, such as awk, sed, and python (Genome Analysis, 2013). Therefore, we converted BAM file into BED format with the help of bedtools, which converts sequence alignments in BAM format into BED, BED12, and/or BEDPE records as shown in Figure 68. We used following settings for bedtools.

Figure 68- Settings used in conversion BAM files from Bowtie2 software into BED format.

Peak Finding and Data Analysis
Introduction-

Peak calling is a very crucial step in ChIP-Seq workflow and study genome-wide protein-DNA interactions. Peak calling is a computational method used to identify areas in a genome that have been enriched with aligned reads as a consequence of performing a ChIP-Sequencing experiment. These areas are those where a protein interacts with DNA (Pepke, Wold, & Mortazavi, 2009).When the protein is a transcription factor, the enriched area is its transcription factor binding site (TFBS).We used the powerful novel algorithm, named Model-based Analysis of ChIP-Seq (MACS), for identifying protein interest binding sites on DNA. MACS2 is capable of capturing enriched ChIP region and improves the spatial resolution of binding sites through combining the information of both sequencing tag position and orientation in spite of genome complexity. MACS2 increases the specificity of callpeak function by comparing with the control. MACS2 estimates reads for a given location by assuming mean of the Poisson from control experiment (background reads) in 5 Kbp or 10 Kbp around the examined location. If reads in sample file comes out more than would have expected from the Poisson ($p <$ 0.00005), then MACS2 calls a peak i.e. build model for peaks. ID3 ChIP was performed using SC-490 polyclonal Anti-ID3 antibody as described above. Two Illumina TruSeq ChIP libraries were constructed from ChIP and input control DNA and sequenced using Illumina HiSeq 2500 by Center for Genome Technology (CGT) of John P. Hussman Institute for Human Genomics of Miller School of Medicine at University of Miami. We performed ChIP-Seq analysis using various tools available at Galaxy, which is an open source, web-based platform (Afgan et al., 2016). FASTQ reads were uploaded to the

155

galaxy server platform (Afgan et al., 2016) with the help of FileZilla which is a free software, cross-platform file transfer protocol (FTP) application, consisting of FileZilla Client and FileZilla Server. We also ran FastQC analysis on these files. FastQC is a Java based quality control tool for raw high throughput sequence data (Andrews, 2016). Then FASTQ file trimmed with identification of short adapter sequences and quality filtering by Trimmomatic software (Bolger et al., 2014). Then these reads mapped to the human genome –Human Dec. 2013 (GRCh38/hg38) (hg38) (Miga et al., 2014a) using Bowtie2 with default settings (Langmead & Salzberg, 2012). Next, peaks are found using peak-calling algorithms. Peak calling was performed using Model-based analysis for ChIP-Seq (MACS2 (2.1.0) options gsize 2451960000 --bw 300 --ratio 1.0 --slocal 1000 --llocal 10000 --keep-dup 1 --bdg --qvalue 0.05 ''--tsize=35 --, lower mfold=5, lower mfold= 50, --pvalue=1e-5'') on each ChIP-Seq file against the matching input file (Feng, Liu, Qin, Zhang, & Liu, 2012; Zhang et al., 2008). For the analysis, peaks with a false discovery rate (FDR) of <5% were selected. MACS2 is very conducive and it yields detailed information for each peak, such as genome coordinates, p-value, FDR, fold enrichment, and summit (peak center). For further processing, we converted tabular format file of MACS2 call peak into interval type by modulating settings. We filtered the data file by setting cut off value for the fold enrichment of the peaks above 5. The summit of the peak represent maximum binding probability of transcription factor or protein of interest on the particular position of DNA. Therefore, we decided to select 200 base pairs region around summit location which is also called as summit region. In order to get summit region we subtracted and added 100 base pairs from summit location. Then cut down the three columns containing chromosome number, start site of summit region of peak and

156

end site of summit region of peak. This process resulted into 200 base pairs region of peak, which signifies most predictable binding sites of our protein of interest ID3. These peaks with start and end position associated with chromosome number were copied and pasted in tested region in the Genomic Regions Enrichment of Annotations Tool (GREAT). GREAT uses ChIP-Seq identified binding events as the TFs binding region on DNA. GREAT associates the proximal and distal transcription start site (TSS) of target gene to the genomic region and then annotates these genomic regions with gene ontologies. GREAT uses statistical algorithm to study enrichment analysis for the associations between genomic regions with annotations. GREAT requires statistical significance from binomial test, but also filters on fold enrichment to identify enrichments of large effect (McLean et al., 2010). GREAT works perfectly to annotate the peaks (genomic regions) with known genes but it uses Human: GRCh37 (UCSC hg19, Feb/2009) as a reference genome, which is earlier version of Human reference genome as shown in Figure 69. In order to annotate the peak regions with genes of updated version of human genome- Human Dec. 2013 (GRCh38/hg38) (hg38), we used RnaChipIntegrator. RnaChipIntegrator integrates analyses of genes (or other genomic feature data) against a set of peaks of ChIP data (Additional file 5), identifying the nearest peaks to each feature and vice versa (Briggs, Donaldson, 2010).
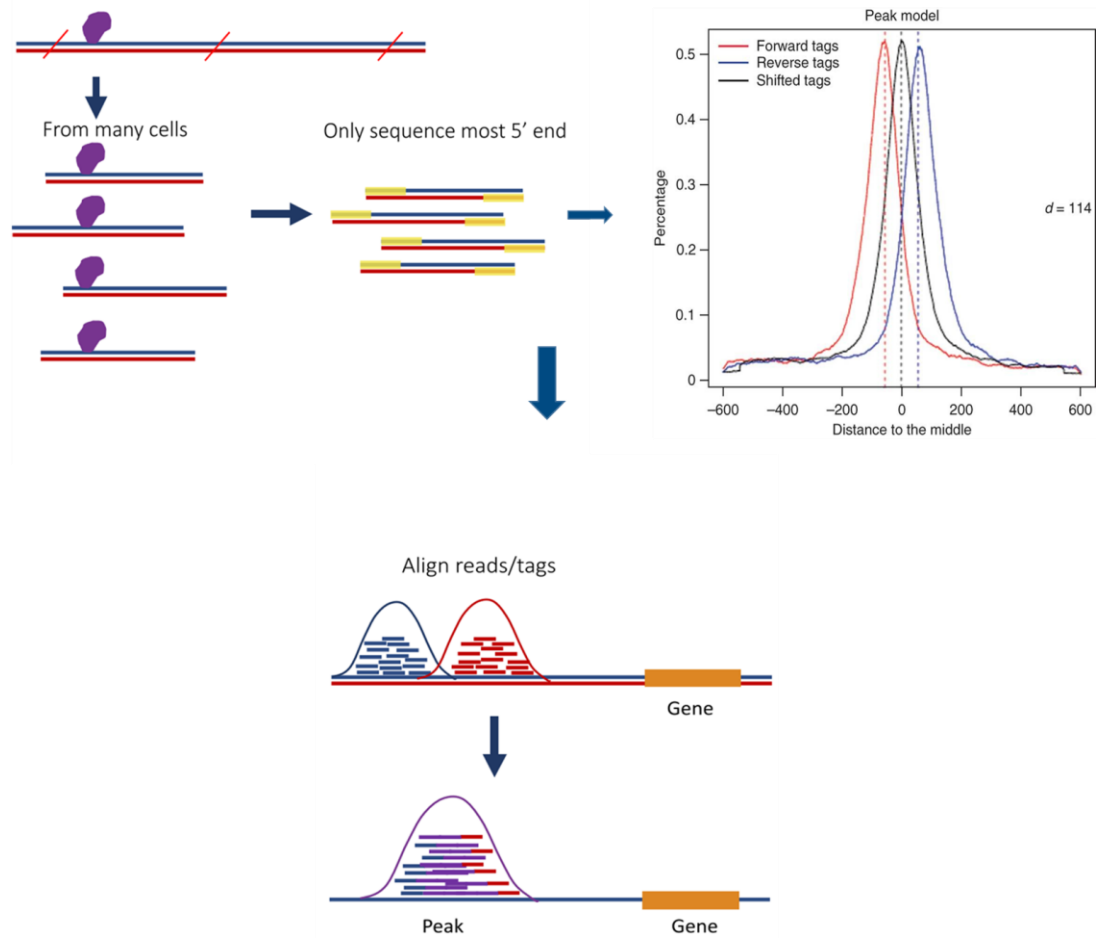
Figure 69- Diagrammatic representation of Peak finding analysis process of MACS2 software. (from Zhang et al. Model-based Analysis of ChIP-Seq (MACS). *Genome Biology* **(2008) vol. 9 (9) pp. R137)**

MACS2 empirically models the shift size of ChIP-Seq tags, and uses it to improve the spatial resolution of predicted binding sites of protein of interest in this case ID3.

The collections of aligned reads (three ChIP-Seq and three input controls) were normalized to the same sequencing depth by randomly removing aligned reads. Duplicated reads were removed before normalization. Peak finding was carried out by running MACS2 (Zhang et al. 2008; options ''--tsize=35 --bw=110, lower mfold=5, lower mfold= 50,   --pvalue=1e-5'') on each ChIP-Seq file against the matching input

file. For the analysis, peaks with a false discovery rate (FDR) of <5% were selected.

MACS2 is very conducive and yield detailed information for each peak, such as genome coordinates, p-value, FDR, fold enrichment, and summit (peak center) as shown in Figure 70.

1	2	3	4	5	6	7	8	9	10
chr	start	end	length	abs_summit	pileup	-log10(pvalue)	fold_enrichment	-log10(qvalue)	name
chr1	9904	10536	633	10139	104.00	96.61920	17.44187	90.50296	MACS2_peak_1
chr1	180760	181053	294	180900	18.00	29.52513	15.52369	24.58319	MACS2_peak_2
chr1	1008957	1009127	171	1009064	4.00	6.89398	4.50072	3.19484	MACS2_peak_3
chr1	1443542	1443715	174	1443633	21.00	15.41446	7.42154	11.06625	MACS2_peak_4
chr1	1595241	1595489	249	1595345	7.00	9.89087	6.53629	5.93337	MACS2_peak_5
chr1	3239929	3240119	191	3240062	11.00	13.17505	8.35351	8.96295	MACS2_peak_6
chr1	3485936	3486107	172	3485958	9.00	7.49291	5.33891	3.75320	MACS2_peak_7
chr1	3908292	3908559	268	3908420	11.00	11.14918	7.25171	7.07982	MACS2_peak_8
chr1	4356995	4357254	260	4357106	9.00	10.33176	6.96126	6.33323	MACS2_peak_9
chr1	4778566	4778740	175	4778610	10.00	14.95925	9.02927	10.63023	MACS2_peak_10
chr1	5225231	5225405	175	5225263	6.00	10.42914	6.30101	6.39365	MACS2_peak_11
chr1	5942420	5942590	171	5942449	3.00	5.23845	3.60058	1.96461	MACS2_peak_12
chr1	6106776	6107242	467	6106931	7.00	12.28785	7.20116	8.12237	MACS2_peak_13
chr1	6422712	6422882	171	6422733	5.00	8.62824	5.40087	4.74977	MACS2_peak_14
chr1	6458008	6458180	173	6458029	9.00	10.33176	6.96126	6.33323	MACS2_peak_15
chr1	8037535	8037709	175	8037565	16.00	19.55753	11.16573	14.98898	MACS2_peak_16
chr1	8843995	8844258	264	8844156	19.00	18.07804	9.56341	13.58942	MACS2_peak_17
chr1	9899460	9899637	178	9899583	5.00	7.45811	5.10458	3.73098	MACS2_peak_18
chr1	10330509	10330726	218	10330623	5.00	7.86864	5.22061	4.09762	MACS2_peak_19

Figure 70- Settings used in MACS2 software for peak finding analysis.

For further processing, we converted tabular format file of MACS2 call peak into interval type by modulating settings. We filtered the data file by setting cut off value for the fold enrichment of the peaks above 5 as shown in Figure 71.

Filter data on any column using simple expressions (Galaxy Version 1.1.0) ▾ Options

Filter

70: MACS2 callpeak on data 67 and data 68 (narrow Peaks) ▾

Dataset missing? See TIP below.

With following condition

c8>=5

Double equal signs, ==, must be used as shown above. To filter for an arbitrary string, use the Select tool.

Number of header lines to skip

0

✔ Execute

⚠ Double equal signs, ==, must be used as *"equal to"* (e.g., **c1** == **'chr22'**)

ℹ **TIP:** Attempting to apply a filtering condition may throw exceptions if the data type (e.g., string, integer) in every line of the columns being filtered is not appro on strings). If an exception is thrown when applying the condition to a line, that line is skipped as invalid for the filter condition. The number of invalid skipped li issue".

ℹ **TIP:** If your data is not TAB delimited, use *Text Manipulation->Convert*

1	2	3	4	5	6	7	8	9	10
chr	start	end	length	abs_summit	pileup	-log10(pvalue)	fold_enrichment	-log10(qvalue)	name
chr1	9904	10536	633	10139	104.00	96.61920	17.44187	90.50296	MACS2_peak_1
chr1	180760	181053	294	180900	18.00	29.52513	15.52369	24.58319	MACS2_peak_2
chr1	1008957	1009127	171	1009064	4.00	6.89398	4.50072	3.19484	MACS2_peak_3
chr1	1443542	1443715	174	1443633	21.00	15.41446	7.42154	11.06625	MACS2_peak_4
chr1	1595241	1595489	249	1595345	7.00	9.89087	6.53629	5.93337	MACS2_peak_5
chr1	3239929	3240119	191	3240062	11.00	13.17505	8.35351	8.96295	MACS2_peak_6
chr1	3485936	3486107	172	3485958	9.00	7.49291	5.33891	3.75320	MACS2_peak_7
chr1	3908292	3908559	268	3908420	11.00	11.14918	7.25171	7.07982	MACS2_peak_8
chr1	4356995	4357254	260	4357106	9.00	10.33176	6.96126	6.33323	MACS2_peak_9
chr1	4778566	4778740	175	4778610	10.00	14.95925	9.02927	10.63023	MACS2_peak_10
chr1	5225231	5225405	175	5225263	6.00	10.42914	6.30101	6.39365	MACS2_peak_11
chr1	5942420	5942590	171	5942449	3.00	5.23845	3.60058	1.96461	MACS2_peak_12
chr1	6106776	6107242	467	6106931	7.00	12.28785	7.20116	8.12237	MACS2_peak_13
chr1	6422712	6422882	171	6422733	5.00	8.62824	5.40087	4.74977	MACS2_peak_14
chr1	6458008	6458180	173	6458029	9.00	10.33176	6.96126	6.33323	MACS2_peak_15
chr1	8037535	8037709	175	8037565	16.00	19.55753	11.16573	14.98898	MACS2_peak_16
chr1	8843995	8844258	264	8844156	19.00	18.07804	9.56341	13.58942	MACS2_peak_17
chr1	9899460	9899637	178	9899583	5.00	7.45811	5.10458	3.73098	MACS2_peak_18
chr1	10330509	10330726	218	10330623	5.00	7.86864	5.22061	4.09762	MACS2_peak_19

Figure 71- Setting used for sorting the peak data.

The cut off value for the fold enrichment of the peaks above 5 used in MACS2 software.

From the above table, we selected summit location of each peak region. The summit location is considered to be the center location of the peak. Each peak region generally consists of 500-600 base pairs. The above table shows the start and end location of the peak correspond to chromosome number. The summit of the peak represents maximum binding probability of transcription factor or protein of interest on the particular position of DNA. Therefore, we decided to select 200 base pairs region around summit location, which is also called as summit region. In order to get summit region we subtracted and added 100 base pairs from summit location. Then cut down the three columns containing chromosome number, start site of summit region of peak and end site of summit region of peak. This process resulted into 200 base pairs region of peak, which signifies most predictable binding sites of our protein of interest ID3. The process is shown in the following Figure 72.

Figure 72- **Diagrammatic representation of workflow of setting use for selecting only 200 bp region around summit location of peaks and chromosome number of peaks.**

These peaks with start and end position associated with chromosome number copied and pasted in tested region in the Genomic Regions Enrichment of Annotations Tool (GREAT). GREAT uses ChIP-Seq identified binding events as the TFs binding region on DNA. GREAT associates the proximal and distal transcription start site (TSS) of target gene to the genomic region and then annotates these genomic regions with gene ontologies. GREAT uses statistical algorithm to study enrichment analysis for the associations between genomic regions with annotations as shown in Figure 73. GREAT requires statistical significance from binomial test, but also filters on fold enrichment to identify enrichments of large effect (McLean et al., 2010).

Figure 73- The Diagrammatic representation of workflow of setting used for peak data in GREAT software.

Motif discovery

We used MEME-ChIP 4.3.0 (Bailey, 2011; Ma, Noble, & Bailey, 2014) to discover motifs. We selected 10% of the strongest peaks for each experiment (reads under peak). The location of the peak was extracted and extended by 100 bp on either side. MEME was instructed to report the top 10 motifs between 5 and 30 bases in length. We accepted all motifs that showed significant similarity (S. Gupta, Stamatoyannopoulos, Bailey, & Noble, 2007) to the reported motif (Mathelier et al., 2014; Sandelin, 2004), MA0074.1, P-value $<1 \times 10^{-5}$). MEME was instructed to report the top 10 motifs between 5 and 30 bases in length. We accepted all motifs that showed significant similarity motifs present in HOmo sapiens COmprehensive MOdel COllection (HOCOMOCO) v11 database (Kulakovskiy et al., 2017). The options were set to perform MEME-ChIP (DREME E-value ≤0.05, CentriMO score ≥5).

Data Visualization

We used Integrated Genome Browser (IGB) which is an open-source tool for visualizing of ChIP Seq datasets (Freese, Norris, & Loraine, 2016). To visualize the data, we selected human genome –Human Dec. 2013 (GRCh38/hg38) (hg38) as a refrence genome and then uploaded the Bowtie2 mapped aligned sequenced .BAM file and MACS2 generated ID3 peaks .*BED* file in the IGB. In order to visualize the peaks, we followed the IGB user guide instructions. We viewed genomic sequence of peaks with highest peak score and peaks which fall into the promoter region of gene with the help of IGB Genomic Sequence Viewer. In order to check the peaks in the promoter region of the gene, we used RnaChipIntegrator (Briggs, Donaldson, 2010) which shows the information of overlapping peaks into the promoter region of the gene.

Results

Global identification of ID3 binding sites in HPMEC and HPMEC cells by ChIP-Seq analysis

In order to understand the global binding patterns of ID3, we applied ChIP-Seq technology, which is a novel approach for identifying transcription factor binding sites genome-wide (Barski et al., 2007; Johnson, Mortazavi, Myers, & Wold, 2007). We performed experiment in replicates of HPMEC ID3$^+$ ChIP and HPMEC ChIP samples. After sequencing analysis, we obtained a total of 45,999,686 and 44,187,690 sequence tags respectively for HPMEC ID3$^+$ ChIP and HPMEC ChIP that were mapped uniquely using Bowtie2 to the reference human genome GRCh38 (hg38) by allowing two mismatches (Miga et al., 2014b; Univeristy of California Santa Cruz, 2013b) (Langmead & Salzberg, 2012) . Peak calling was performed on HPMEC ID3$^+$ ChIP sample using Model-based analysis for ChIP-Seq (MACS2 (2.1.0) on each ChIP-Seq file against the matching input file (Feng et al., 2012; Zhang et al., 2008). The q-value =0.01 (minimum FDR) cutoff to call significant regions. Using MACS2, we identified a total of 2493 ID3 peaks in HPMEC ID3$^+$ ChIP using control ChIP-Seq data as the negative control. These mapped regions then associated with gene ontologies with the help of GREAT software which uses GRCh37 (hg19) as a reference genome (McLean et al., 2010). GREAT associated 2336 ID3 peak-binding regions (200 base pairs bp) around summit location of the peak) of HPMEC ID3$^+$ ChIP sample to transcription start site (TSS) of 2834 known genes (Fig.68B) (Additional file 1). Out of 2834 known genes, 470 genes were associated with only one binding region, 1589 genes were associated with two regions and 1 gene was associated with three regions (Fig.75B). Results additionally showed 276 genes were not associated with any region. We calculated the distance of the ID3 binding regions to

TSS (transcription start sites) and then tabulated the frequency across the distance intervals before TSS and after TSS against region-gene association (Fig. 75A). Figure 75A shows that distance of peak around the TSS sites. We found that about 4.19% ID3 (153 of 3651) binding regions were mapped within ±5 Kb of TSS (Figure 68B), and about 24.92% (910 binding regions) were mapped within ± 50 Kb of TSS. Similarly, MACS2 identified 3603 ID3 binding regions i.e. peaks in HPMEC ChIP sample using control ChIP-Seq data as the negative control. GREAT associated 3516 ID3 peaks binding regions (200 bp around summit location of the peak) that were mapped to TSS (transcription start site) of 3935 known genes (Fig.76B) in HPMEC ChIP sample (Additional file 2). Out of 3935 known genes, 782 genes were associated with only one binding region, 2524 genes were associated with two regions per gene. While 208 genes were not associated with any region. We calculated the distance of the ID3 binding regions to TSS (transcription start sites) in HPMEC ChIP sample and then tabulated the frequency across the distance intervals before TSS and after TSS against region-gene association (Fig. 3A). Figure 76A shows that the distance of peak around the TSS sites. We found that about 3.86% ID3 (226 of 5841) binding regions were mapped within ±5 kb of TSS (Figure 76B), and about 22.80% (1332 binding regions) were mapped within ± 50 kb of TSS. Furthermore, we also checked the genome-wide protein-DNA interaction patterns for HPMEC ID3$^+$ ChIP and HPMEC ChIP data with the help of CEAS (cis-regulatory element annotation system) tool which provides summary statistics on ChIP peaks enrichment in important genomic regions such as individual chromosomes, promoters, gene bodies or exon (Shin, Liu, Manrai, & Liu, 2009) (Fig.77). Association of ID3 bound regions with annotated genomic features indicated that ID3 tends to localize

within intronic regions. A total of 37.4% and 41% peaks were found in HPMEC ID3[+] ChIP and HPMEC ChIP samples respectively. Very small percentage of peaks were found in the coding exons which comprised of 0.6% and 0.5% in HPMEC ID3+ ChIP and HPMEC ChIP samples respectively. 5'-UTR regulatory regions showed 1.3% and 0.3% of peaks in HPMEC ID3[+] ChIP and HPMEC ChIP samples respectively (Figure 77). In addition, ID3 binding sites were found to be localized in the promoter region of the gene in HPMEC ID3[+] ChIP (3.7%) and HPMEC ChIP (3.6%) samples (Figure 77). We identified ID3 target genes by GREAT and compared the genes between HPMEC ID3[+] ChIP and HPMEC ChIP samples in Fig. 78. We found total of 1240 ID3 gene targets between both ChIP Seq samples. We found only 22.4% similarity in the genes associated with binding regions of ID3 in both ChIP Seq samples. These findings may suggest that ID3-depedent binding to DNA is different in the HPMEC (wild type) and HPMEC ID3[+] cell lines.
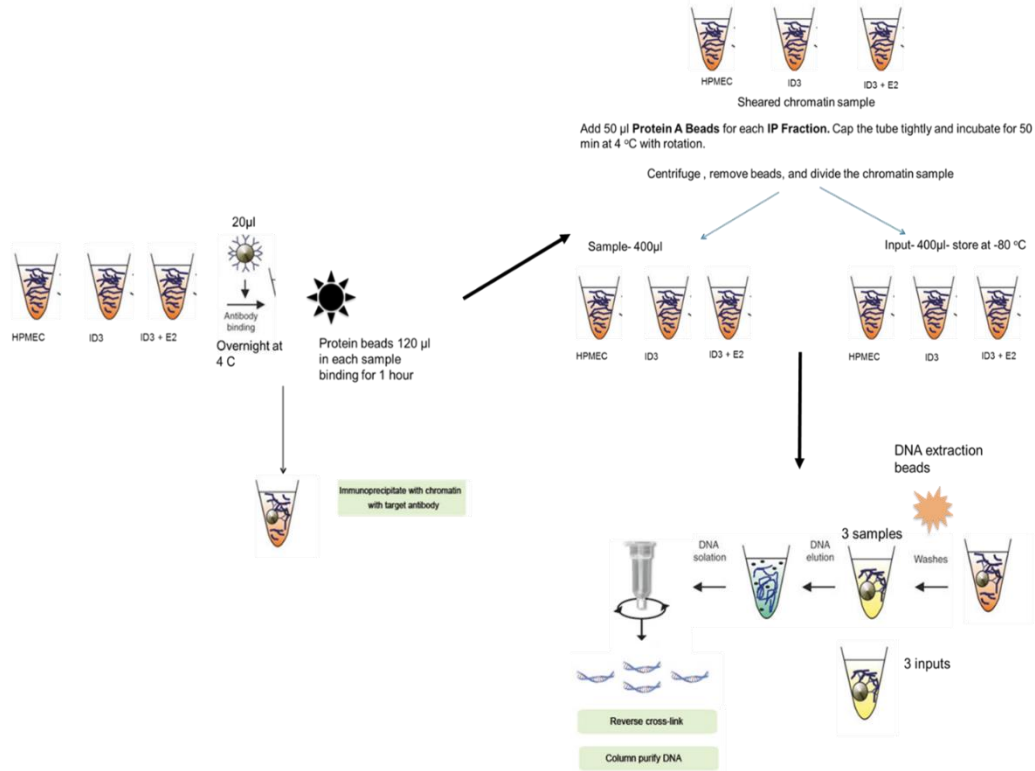
Figure 74- The Diagrammatic representation of workflow of ChIP-Seq experiment carried out using HPMEC (Control), HPMEC ID3+, and HPMEC ID3 + E2 cell samples.
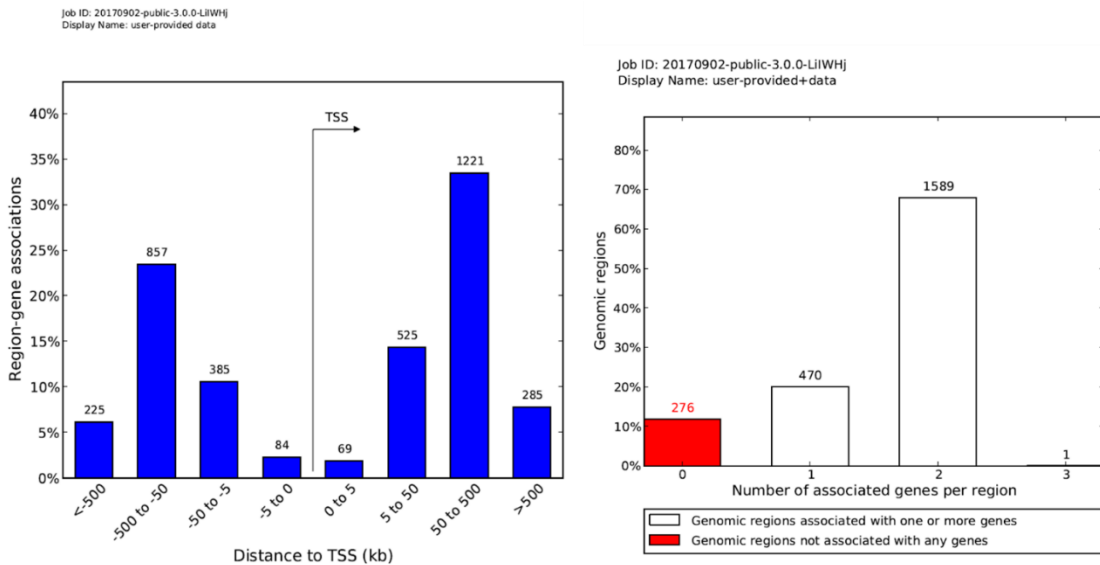
Figure 75 A- The graphical representation of MACS2 generated peaks for HPMEC ID3+ cell sample binned by orientation and distance to TSS. B) The graphical representation of number of gene associated per ID3 binding peak regions by GREAT tool.
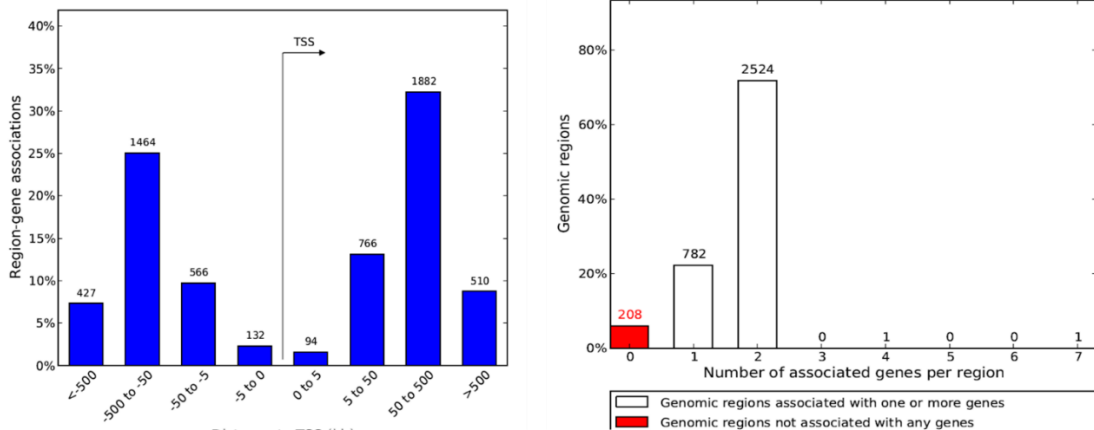
Figure 76A) The graphical representation of MACS2 generated peaks for HPMEC cell sample binned by orientation and distance to TSS. B) The graphical representation of number of gene associated per ID3 binding peak regions by GREAT tool.
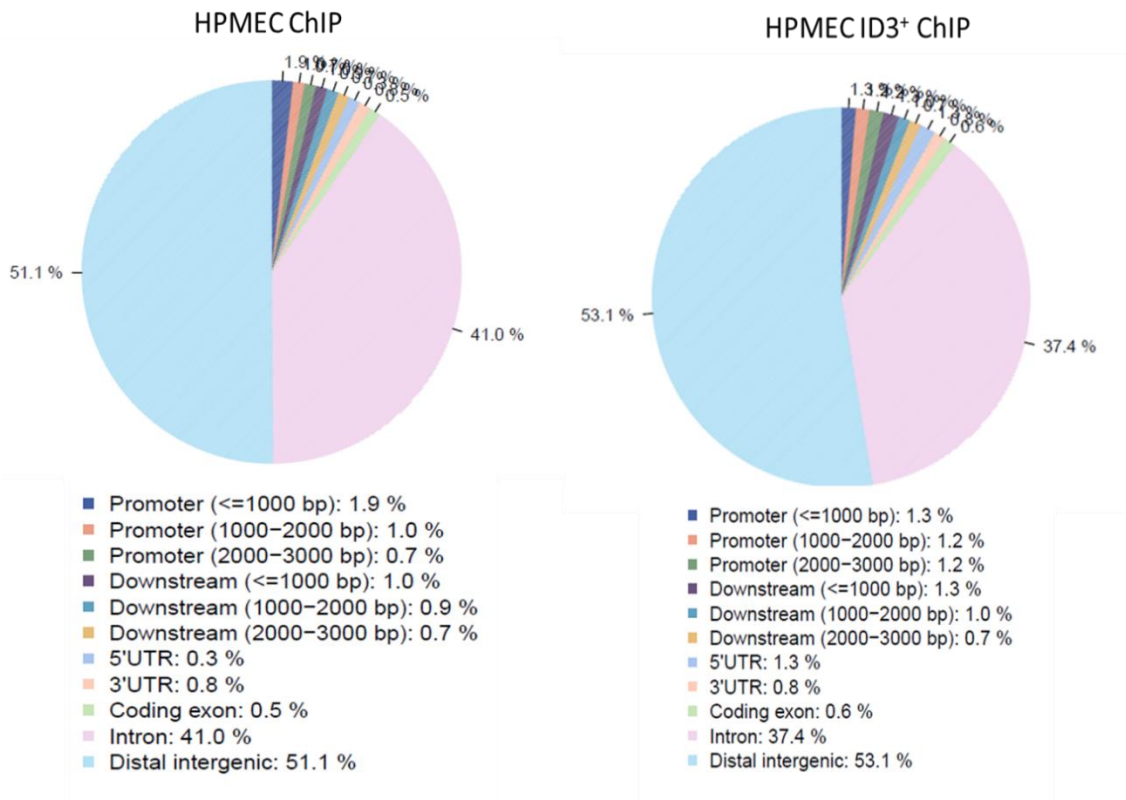
Figure 77- Diagrammatic representation of MACS2 identified ChIP peaks characteristics created with the help of CEAS tool. Diagrams show percentage of peaks overlapping with gene features for ChIP-Seq HPMEC and HPMEC ID3+ cell samples.
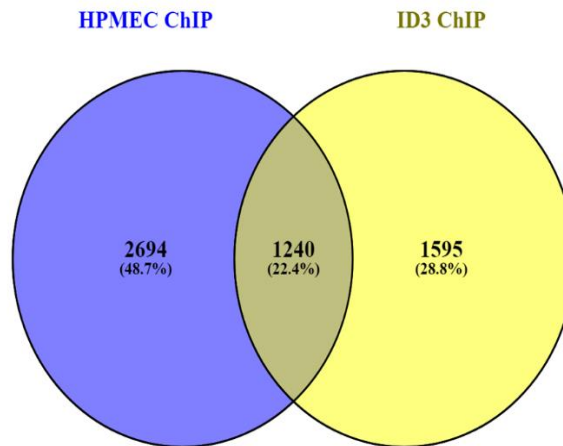
Figure 78- A Venn diagram showing the overlap of ID3 target genes among HPMEC and HPMEC ID3+ cell samples.

Snyder et al. recently used ChIP Seq to map binding sites of eGFP-ID3 and other key TFs on human K562 cell line (ENCSR005NMT) and made data available at ENCODE. We used GEO accession: GSE91408 data from ENCODE to check ID3 binding regions (Dunham et al., 2012; Sloan et al., 2016). We processed GSE91408_ENCFF468VTJ_optimal_idr_thresholded_peaks_GRCh38.bigBed file to annotate the peak or binding regions of ID3 to gene annotation using GREAT. We found total of 9717 genes were associated with 24336 binding regions of ID3 in ENCODE data. When we compared the genes generated from HPMEC ID3$^+$ ChIP Seq data with ENCODE ID3 – K562 ChIP data, 64.60% (1831 out of 2834 genes) of genes of HPMEC ID3$^+$ ChIP seq data were overlapped with ENCODE ID3 – K562 ChIP data (Fig.79) (Additional file 3). The difference between numbers of genes associated with ID3 binding region of HPMEC ID3$^+$ ChIP-Seq data and ENCODE ID3 – K562 ChIP data could be due to the use of different cell line (HPMEC VS K562), different antibodies,

170

Snyder et al. used the anti-GFP antibody (11814460001, Roche) or differences in the analysis pipeline and downstream analysis procedures.
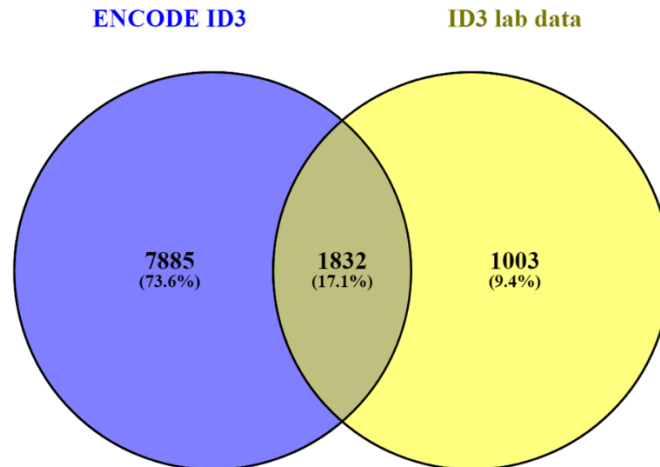
Figure 79- A Venn diagram showing the overlap of ID3 target genes among ENCODE ID3 – K562 and HPMEC ID3+ ChIP data.

This suggests that there are common sets of genes regulated by ID3 in humans K562 cells and in HPMEC cells. However, we identified many ID3 binding sites that are only present in the HPMEC cell line, suggesting that ID3 targets different genes in the context of lung EC cells. These findings clearly validated ChIP-Seq as an efficient and powerful technique for mapping ID3 binding sites in HPMEC cells.

Identification of the DNA binding consensus sequences of ID3 and other known TF binding sites in the ID3 bound regions of the ChIP data. The recognition of interaction of ID3 protein with TFs found from ChIP-Seq data.

To find out the *de novo* ID3 binding sequence in the ID3 binding regions (peak regions) of HPMEC and HPMEC ID3[+] cell ChIP samples, we used the MEME-ChIP suite

171

program (Bailey, 2011; Ma et al., 2014). We also checked whether other TFs have their own unique and enriched binding motifs in the ID3 binding regions of HPMEC ID3$^+$ and HPMEC ChIP sample that we identified. We found a repeated sequence of GAGAGAGAGAGAGAGAGAGAGAGAGAGAGA with a very high E-value of 2E-62 Enriched motifs (E-value $\leq$ 10 using the binomial test) in HPMEC ID3$^+$ ChIP sample. There are 348 instances at which this motif found in 2493 ID3 binding regions. We also manually checked GAGAGAGAGAGA repeat sequences in peaks with highest peak score and peaks which fall into the promoter region of gene identified by RnaChipIntegrator (Briggs, Donaldson, 2010). The data is visualized with the help of Integrated genome browser (IGB) which is a visual analytics platform for ChIP seq data analysis (Freese et al., 2016). In addition, we also processed ID3 bound peak regions of HPMEC ChIP sample and found most significant, unknown motif TCCCAGCWMYTYGGGAGGCTGAGGCRGGAG sequence with a very high E-value Enriched motifs (E-value $\leq$ 10 using the binomial test) of 1.30E-05. There are 45 instances of this motif found in 3603 ID3 binding regions. To identify possible interactions between ID3 and other TFs, MEME-ChIP suite containing DREME and Centrimo screening was performed to identify the other most over-represented motifs or consensus sequences of TFs in the ID3 binding regions identified by MACS2 (Bailey, 2011; Ma et al., 2014). We found IRF1, BC11A, IRF4, PRDM1, FOXJ3, SMAD4, ZBTB6, GATA1, and STAT2 TFs binding motifs in the binding region of HPMEC ID3$^+$ ChIP sample (Table 7). The consensus sequences with their associated TFs for HPMEC ID3$^+$ ChIP sample are shown in Table 7. In addtion to this, the consensus sequences with their associated TFs ZBT7B, NKX28, TGIF1, RFX2, MYOG, ZBT18, MXI1, SOX2,

MAF, MEIS1, SOX3, SOX9 ,and TBX3 for HPMEC ChIP sample are shown in Table 8. We have shown all three samples MEME-ChIP generated output in additional file 4. Since ID3 lacks DNA binding domain, it regulates the transcription machinery in a dominant negative manner. ID3 binds to basic Helix-loop-helix I (bHLH I) proteins via the helix loop helix (HLH) motif. ID3 protein interactions block the DNA binding activity of basic HLH (bHLH) transcription factors (Doke et al., 2018). The molecular mechanism for the role of ID3 proteins in the transcriptional control was initially thought to be only dependent on its ability to dimerize with bHLH factors acting as dominant-negative. Nevertheless, in addition to bHLH proteins, researchers have shown the interaction of ID3 with non-bHLH proteins. Researchers have shown that ID3 interaction with Pax-5 and ternary complex factor (TCF) subfamily of ETS-domain proteins (Roberts, Deed, Inoue, Norton, & Sharrocks, 2001; Yates, Atherton, Deed, Norton, & Sharrocks, 1999). Therefore, it is important to know that TFs could bind to the ID3 binding regions that we identified and act as ID3 cooperators for the regulation of gene expression. We compared known TFs which were generated using MEME-ChIP suite analysis from HPMEC ID3$^+$ and HPMEC ChIP samples. We did not find any TFs overlapped between both HPMEC ID3$^+$ and HPMEC ChIP samples. Although we did find only one (GATA1) overlap of TF when we compared TFs generated from HPMEC ID3$^+$ and ENCODE ID3 – K562 ChIP data. Researchers have already shown that interaction of GATA1 with E47 protein acts as the master regulators of hematopoiesis and binds to β-globin promoter in Id cDKO bone marrow cells (Zhao et al., 2016). This data may suggests that ID3 indirectly interacts with GATA1. It is interesting to see that there is no overlap of any TFs between HPMEC ID3$^+$ and HPMEC ChIP samples. This

finding corroborates the idea that ID3 gene regulation with copartner TFs is different in ID3 overexpressed HPMEC than wild type HPMEC. Moreover, we tried to understand the connection between ID3 and 8 TFs generated from HPMEC ID3^{+} ChIP data using STRING database in the Fig. 80 and 81. The STRING netwok showed how this proteins are interconnected with various pathways (Szklarczyk et al., 2015b; Jensen et al., 2015). In addition to GATA1, other proteins like STAT2, SMAD4 and IRF1 did not show any direct connection with ID3 in Figure 82A. But Figure 82B showed indirect interaction of ID3 with SMAD4 throgh BMP2, as SMAD4 is central component of the BMP signalling and ID3 is direct target of BMP2 in a variety of cell types (Lim et al., 2015; O'Shaughnessy, Christiano, & Jahoda, 2004; Voorneveld et al., 2014). In case of STAT2-JAK-IRF4 did not show any type of interaction with ID3. But there is evidence of cytokine-mediated STAT activation directly modulate Id2 expression. Morover. researchers also suggested possibilty of cytokines indirect regulatation of Id3 expression through modulation of Id2 expression in memory T cells of mice (Yang et al., 2011). However, in order to rule out any nonspecific effect of the ID3 antibody on the recognition of these transcription factors not related to ID3, we compared their amino acid sequences with the ID3 protein by means of the DNAStar alignment program. No significant similarities were observed among the considered proteins, thus ruling out any unspecific binding of the ID3 antibody. Although ID3 ChIP-Seq data demonstrated an enrichment of genomic regions with overrepresentation of general transcriptional regulatory TFs, functional experiments should be performed to demonstrate an interaction of ID3 with these TFs. Our data underscore the importance of synergistic interactions between ID3 and tissue-specific TFs to ultimately modulate gene expression.

Gene ID for optimal peaks	Species	Quality	Consensus	E-value Enriched motifs (E-value ≤ 10 using the binomial test)	Region Width	Region Matches
GAGAGAGAGAGAGAGAGAGAGAGAGAGAGA				2E-62	47	348
GAGAGAGAGAGAGAGAGAGAG				8.8E-57	54	475
GAGAGAGAGAGAGAG				1.60E-45	54	425
IRF1	HUMAN	A	RAAASNGAAAGTGAAASTRR	4E-12	81	716
GGGAGAGA				1.9E-11	71	172
BC11A	HUMAN	C	RAAAGAGGAAGTGARAV	1.2E-09	72	1014
IRF4	HUMAN	C	RRRGRGGAASTGARAS	8.7E-09	81	1061
PRDM1	HUMAN	C	RAAAGTGAAAGTGA	0.000000031	53	515
FOXJ3	HUMAN	A	TTGTTTATKKTTW	0.0035	70	340
SMAD4	HUMAN	C	STGTCTSDCACCT	0.0047	90	891
ZBTB6	HUMAN	D	VGRTGATRGAGCC	0.03	106	594
STAT2	HUMAN	B	TYASTTTCNBTTYCH	0.005	12	390
GATA1	HUMAN	A	CTGBNNVVBRSAGATAAGR	7.50E-03	76	251

Table 7- Main TFs binding motifs overrepresented in the ID3 binding regions of HPMEC ID3+ ChIP data.

ID	Alt ID		Consensus	E-value	Region Width	Region Matches
TCCCAGCWMYTYGGGAGGCTGAGGCRGGAG		MEME-3	TCCCAGCWMYTYGGGAGGCTGAGGCRGGAG	1.30E-05	39	45
ZBT7B	HUMAN		SSCSGTTGCCATGGVRACSGS	4.20E-03	12	34
NKX28	HUMAN		TTCAAGKRC	7.40E-03	88	326
TGIF1	HUMAN		TGACAGS	8.40E-03	84	620
RFX2	HUMAN		CKGTTGCCATGGCRACVS	1.40E-02	5	15
AGTRGCTG		DREME-6	AGTAGCTG	1.60E-02	9	15
MYOG	HUMAN		CHSCAGCTGYYYB	2.20E-02	94	215
ACAGGCRTGMGCCACCRYGCC		MEME-5	ACAGGCRTGMGCCACCRYGCC	7.00E-02	48	45
AYTACAGG		DREME-3	AYTACAGG	1.50E-01	21	21
ZBT18	HUMAN		KCCAGMTGTKB	2.10E-01	84	416
MXI1	HUMAN		SCCACGTGSYB	2.90E-01	100	457
SOX2	HUMAN		BBCCTTTGTYYYB	4.00E-01	46	321
SOX2	HUMAN		BCCWTTGTYATGYWRA	6.20E-01	55	360
MAF	HUMAN		VAAWBTGCTGASTHWGCMD	6.50E-01	86	483
MEIS1	HUMAN		GMTGACAGV	6.60E-01	82	532
SOX3	HUMAN		BCCWTTGTBYB	7.10E-01	48	293
SOX9	HUMAN		YCWTTGKYY	8.40E-01	38	317
TBX3	HUMAN		DRAGGTGBSAR	9.00E-01	90	609

Table 8- Main TFs binding motifs overrepresented in the ID3 binding regions of HPMEC ChIP data.

175
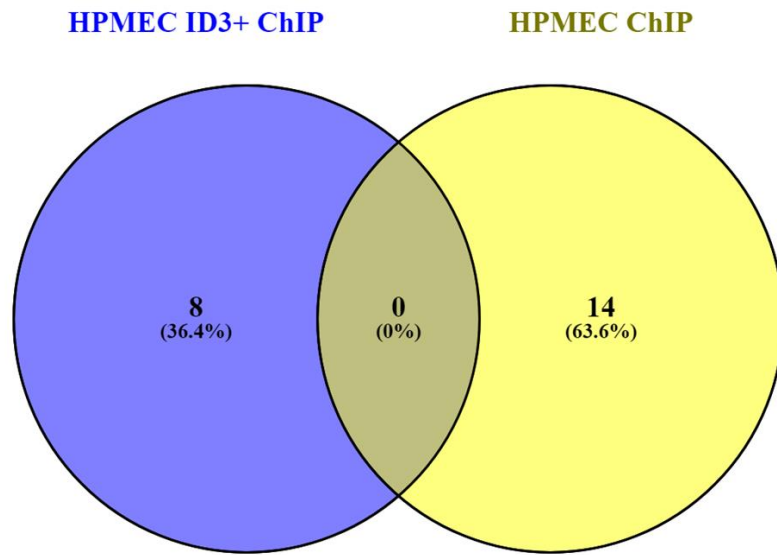
Figure 80- A Venn diagram showing there is no overlap of TFs copartners of ID3 in the ID3 binding regions among HPMEC ChIP and HPMEC ID3+ ChIP data.
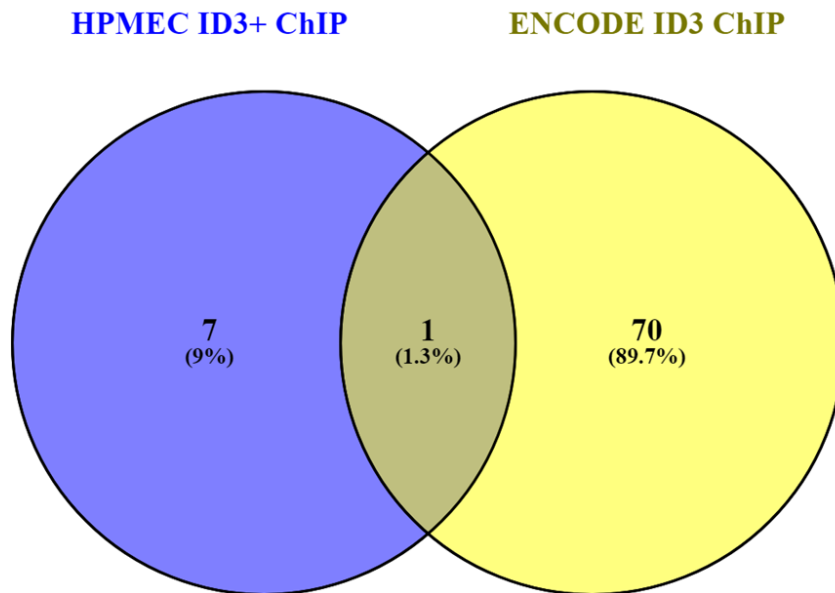
Figure 81- A Venn diagram showing the one overlap of TF copartner of ID3 in the ID3 binding regions among ENCODE ID3 – K562 and HPMEC ID3[+] ChIP data.

Approved symbol	Approved name	HGNC ID	Location
BCL11A	B cell CLL/lymphoma 11A	HGNC:13221	2p16.1
IRF4	interferon regulatory factor 4	HGNC:6119	6p25.3
PRDM1	PR/SET domain 1	HGNC:9346	6q21
FOXJ3	forkhead box J3	HGNC:29178	1p34.2
SMAD4	SMAD family member 4	HGNC:6770	18q21.2
ZBTB6	zinc finger and BTB domain containing 6	HGNC:16764	9q33.2
STAT2	signal transducer and activator of transcription 2	HGNC:11363	12q13.2
GATA1	GATA binding protein 1	HGNC:4170	Xp11.23

Table 9- A detailed list of TFs binding in the peak regions of ID3 from HPMEC ID3$^+$ ChIP sample.
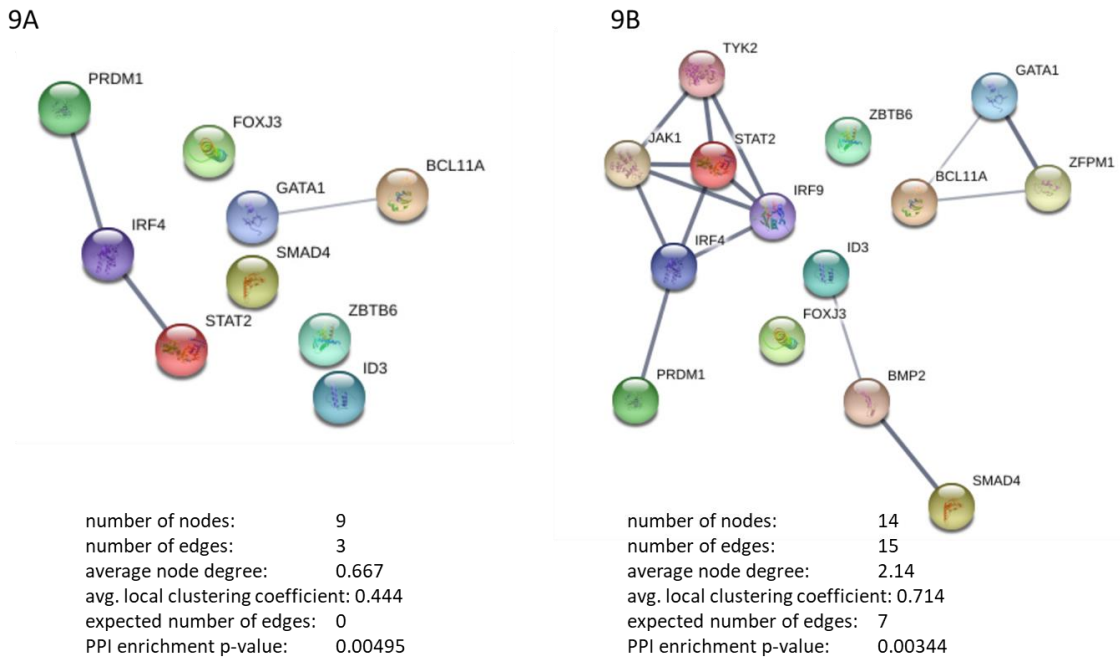
Figure 82- A) STRING protein network illustrates 8 proteins interaction with ID3 protein and p= 0.00495 shows that interactions among the protein is significant. B) STRING has added a total of 5 proteins to network to show a network around our earlier 9 proteins

References-

Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Čech, M., … Goecks, J. (2016). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Research*, *44*(W1), W3–W10. http://doi.org/10.1093/nar/gkw343

Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. *Babraham Bioinformatics*, http://www.bioinformatics.babraham.ac.uk/projects/. http://doi.org/citeulike-article-id:11583827

Andrews, S. (2016). Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. Retrieved November 20, 2017, from https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Bailey, T. L. (2011). DREME: Motif discovery in transcription factor ChIP-Seq data. *Bioinformatics*, *27*(12), 1653–1659. http://doi.org/10.1093/bioinformatics/btr261

Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., … Zhao, K. (2007). High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell*, *129*(4), 823–837. http://doi.org/10.1016/j.cell.2007.05.009

Benezra, R., Davis, R. L., Lockshon, D., Turner, D. L., & Weintraub, H. (1990). The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. *Cell*, *61*(1), 49–59. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2156629

Blankenberg, D., Gordon, A., Von Kuster, G., Coraor, N., Taylor, J., Nekrutenko, A., & Team, G. (2010). Manipulation of FASTQ data with galaxy. *Bioinformatics*, *26*(14), 1783–1785. http://doi.org/10.1093/bioinformatics/btq281

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. http://doi.org/10.1093/bioinformatics/btu170

Briggs PJ, Donaldson IJ, Z. L. (). RnaChipIntegrator — RnaChipIntegrator 1.0.3 documentation. Retrieved February 13, 2018, from https://github.com/fls-bioinformatics-core/RnaChipIntegrator

Cock, P. J. A., Fields, C. J., Goto, N., Heuer, M. L., & Rice, P. M. (2009). The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Research*, *38*(6), 1767–1771. http://doi.org/10.1093/nar/gkp1137

Das, J. K., & Felty, Q. (2014a). Microvascular Lesions by Estrogen-Induced ID3: Its Implications in Cerebral and Cardiorenal Vascular Disease. *Journal of Molecular Neuroscience*, 618–631. http://doi.org/10.1007/s12031-014-0401-9

Das, J. K., & Felty, Q. (2014b). PCB153-induced overexpression of ID3 contributes to the development of microvascular lesions. *PloS One*, *9*(8), e104159. http://doi.org/10.1371/journal.pone.0104159

Doke, M., Avecilla, V., & Felty, Q. (2018). Inhibitor of Differentiation-3 and Estrogenic Endocrine Disruptors: Implications for Susceptibility to Obesity and Metabolic Disorders. *BioMed Research International*, *2018*, 1–16. http://doi.org/10.1155/2018/6821601

Dunham, I., Kundaje, A., Aldred, S. F., Collins, P. J., Davis, C. A., Doyle, F., … Birney, E. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, *489*(7414), 57–74. http://doi.org/10.1038/nature11247

Ewing, B., & Green, P. (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Research*, *8*(3), 186–94. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9521922

Felty, Q. (2006). Estrogen-induced DNA synthesis in vascular endothelial cells is mediated by ROS signaling. *BMC Cardiovascular Disorders*, *6*, 16. http://doi.org/10.1186/1471-2261-6-16

Felty, Q., & Porther, N. (2008). Estrogen-induced redox sensitive Id3 signaling controls the growth of vascular cells. *Atherosclerosis*, *198*(1), 12–21. http://doi.org/10.1016/j.atherosclerosis.2007.12.048

Feng, J., Liu, T., Qin, B., Zhang, Y., & Liu, X. S. (2012). Identifying ChIP-Seq enrichment using MACS. *Nature Protocols*, *7*(9), 1728–1740. http://doi.org/10.1038/nprot.2012.101

Freese, N. H., Norris, D. C., & Loraine, A. E. (2016). Integrated genome browser: Visual analytics platform for genomics. *Bioinformatics*, *32*(14), 2089–2095. http://doi.org/10.1093/bioinformatics/btw069

Gupta, G. P., Perk, J., Acharyya, S., de Candia, P., Mittal, V., Todorova-Manova, K., … Massague, J. (2007). ID genes mediate tumor reinitiation during breast cancer lung metastasis. *Proceedings of the National Academy of Sciences*, *104*(49), 19506–19511. http://doi.org/10.1073/pnas.0709185104

Gupta, S., Stamatoyannopoulos, J. A., Bailey, T. L., & Noble, W. (2007). Quantifying similarity between motifs. *Genome Biology*, *8*(2), R24. http://doi.org/10.1186/gb-2007-8-2-r24

Illumina Adapter Sequences Illumina Adapter Sequences Illumina Adapter Sequences. (2017). Retrieved from https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/experiment-design/illumina-adapter-sequences-1000000002694-03.pdf

Johnson, D. S., Mortazavi, A., Myers, R. M., & Wold, B. (2007). Genome-wide mapping of *in vivo* protein-DNA interactions. *Science*, *316*(5830), 1497–1502. http://doi.org/10.1126/science.1141319

Kulakovskiy, I. V, Vorontsov, I. E., Yevshin, I. S., Sharipov, R. N., Fedorova, A. D., Rumynskiy, E. I., … Makeev, V. J. (2017). HOCOMOCO: towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. *Nucleic Acids Research*, *46*(D1), D252–D259. http://doi.org/10.1093/nar/gkx1106

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. http://doi.org/10.1038/nmeth.1923

Langmead and Steven L Salzberg. (2013). Bowtie2. *Nature Methods*, *9*(4), 357–359. http://doi.org/10.1038/nmeth.1923.Fast

Lim, J., Tu, X., Choi, K., Akiyama, H., Mishina, Y., & Long, F. (2015). BMP–Smad4 signaling is required for precartilaginous mesenchymal condensation independent of Sox9 in the mouse. *Developmental Biology*, *400*(1), 132–138. http://doi.org/10.1016/j.ydbio.2015.01.022

Lyden, D., Young, A. Z., Zagzag, D., Yan, W., Gerald, W., O'Reilly, R., … Benezra, R. (1999). Id1 and Id3 are required for neurogenesis, angiogenesis and vascularization of tumour xenografts. *Nature*, *401*(6754), 670–677. http://doi.org/10.1038/44334

Lyden, D., Young, a Z., Zagzag, D., Yan, W., Gerald, W., O'Reilly, R., … Benezra, R. (1999). Id1 and Id3 are required for neurogenesis, angiogenesis and vascularization of tumour xenografts. *Nature*, *401*(October), 670–677. http://doi.org/10.1038/44334

Ma, W., Noble, W. S., & Bailey, T. L. (2014). Motif-based analysis of large nucleotide data sets using MEME-ChIP. *Nature Protocols*, *9*(6), 1428–1450. http://doi.org/10.1038/nprot.2014.083

Mathelier, A., Zhao, X., Zhang, A. W., Parcy, F., Worsley-Hunt, R., Arenillas, D. J., … Wasserman, W. W. (2014). JASPAR 2014: An extensively expanded and updated open-access database of transcription factor binding profiles. *Nucleic Acids Research*, *42*(D1), D142-7. http://doi.org/10.1093/nar/gkt997

McLean, C. Y., Bristor, D., Hiller, M., Clarke, S. L., Schaar, B. T., Lowe, C. B., … Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. *Nature Biotechnology*, *28*(5), 495–501. http://doi.org/10.1038/nbt.1630

Miga, K. H., Newton, Y., Jain, M., Altemose, N., Willard, H. F., & Kent, E. J. (2014a). Centromere reference models for human chromosomes X and y satellite arrays. *Genome Research*, *24*(4), 697–707. http://doi.org/10.1101/gr.159624.113

Miga, K. H., Newton, Y., Jain, M., Altemose, N., Willard, H. F., & Kent, E. J. (2014b). Centromere reference models for human chromosomes X and y satellite arrays. *Genome Research*, *24*(4), 697–707. http://doi.org/10.1101/gr.159624.113

Mueller, C., Baudler, S., Welzel, H., Böhm, M., & Nickenig, G. (2002). Identification of a novel redox-sensitive gene, Id3, which mediates angiotensin II-induced cell growth. *Circulation*, *105*(20), 2423–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12021231

Norton, J. D. (2000). ID helix-loop-helix proteins in cell growth, differentiation and tumorigenesis. *J Cell Sci*, *113 (Pt 2*, 3897–3905. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11058077

Nos, C. (). DNA SMART ᵀᴹ ChIP-, *634867*(21115), 1–23.

O'Shaughnessy, R. F. L., Christiano, A. M., & Jahoda, C. A. B. (2004). The role of BMP signalling in the control of ID3 expression in the hair follicle. *Experimental Dermatology*, *13*(10), 621–629. http://doi.org/10.1111/j.0906-6705.2004.00206.x

Pepke, S., Wold, B., & Mortazavi, A. (2009). Computation for ChIP-Seq and RNA-Seq studies. *Nature Methods*, *6*(11S), S22. http://doi.org/10.1038/nmeth.1371

Roberts, E. C., Deed, R. W., Inoue, T., Norton, J. D., & Sharrocks, A. D. (2001). Id Helix-Loop-Helix Proteins Antagonize Pax Transcription Factor Activity by Inhibiting DNA Binding, *21*(2), 524–533. http://doi.org/10.1128/MCB.21.2.524–533.2001

Robertson, G., Hirst, M., Bainbridge, M., Bilenky, M., Zhao, Y., Zeng, T., … Jones, S. (2007). Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nature Methods*, *4*(8), 651–657. http://doi.org/10.1038/nmeth1068

Ruzinova, M. B., & Benezra, R. (2003). Id proteins in development, cell cycle and cancer. *Trends in Cell Biology*, *13*(8), 410–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12888293

Rvtests - Genome Analysis Wiki. (). Retrieved from https://genome.sph.umich.edu/wiki/BAM

Sandelin, A. (2004). JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Research*, *32*(90001), 91D–94. http://doi.org/10.1093/nar/gkh012

Sharma, P., Patel, D., & Chaudhary, J. (2012). Id1 and Id3 expression is associated with increasing grade of prostate cancer: Id3 preferentially regulates CDKN1B. *Cancer Medicine*, *1*(2), 187–97. http://doi.org/10.1002/cam4.19

Shin, H., Liu, T., Manrai, A. K., & Liu, S. X. (2009). CEAS: Cis-regulatory element annotation system. *Bioinformatics*, *25*(19), 2605–2606. http://doi.org/10.1093/bioinformatics/btp479

Sloan, C. A., Chan, E. T., Davidson, J. M., Malladi, V. S., Strattan, J. S., Hitz, B. C., … Cherry, J. M. (2016). ENCODE data at the ENCODE portal. *Nucleic Acids Research*, *44*(D1), D726–D732. http://doi.org/10.1093/nar/gkv1160

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., … Von Mering, C. (2015). STRING v10: Protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Research*, *43*(D1), D447–D452. http://doi.org/10.1093/nar/gku1003

Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M, Bork P, Jensen LJ, von M. C. (2015). STRING: functional protein association networks. *STRING v10: Protein-Protein Interaction Networks, Integrated over the Tree of Life.* http://doi.org/43:D447-52

Tanaka, F., Oyanagi, H., Takenaka, K., Ishikawa, S., Yanagihara, K., Miyahara, R., … Wada, H. (2003). Glomeruloid Microvascular Proliferation Is Superior to Intratumoral Microvessel Density as a Prognostic Marker in Non-Small Cell Lung Cancer. *Cancer Research*, *63*(20), 6791–6794. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14583475

Tuder, R. M., Groves, B., Badesch, D. B., & Voelkel, N. F. (1994). Exuberant endothelial cell growth and elements of inflammation are present in plexiform lesions of pulmonary hypertension. *The American Journal of Pathology*, *144*(2), 275–85. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7508683

Univeristy of California Santa Cruz. (-a). UCSC Genome Browser Gateway. Retrieved November 21, 2017, from http://genome.ucsc.edu/cgi-bin/hgGateway?db=hg38

Univeristy of California Santa Cruz. (-b). UCSC Genome Browser Gateway. Retrieved February 10, 2018, from https://genome.ucsc.edu/cgi-bin/hgGateway?db=hg38

Usadellab. (2011). Trimmomatic Manual: V0.32. http://doi.org/10.1093/bioinformatics/btu170

Voorneveld, P. W., Kodach, L. L., Jacobs, R. J., Liv, N., Zonnevylle, A. C., Hoogenboom, J. P., … Hardwick, J. C. H. (2014). Loss of SMAD4 Alters BMP Signaling to Promote Colorectal Cancer Cell Metastasis via Activation of Rho and ROCK. *Gastroenterology*, *147*(1), 196–208.e13. http://doi.org/10.1053/j.gastro.2014.03.052

Wassermann, M., Wassermann, D., Cucos, S., & Miller, H. J. (1979). World Pcbs Map: Storage And Effects In Man And His Biologic Environment In The 1970s. *Annals of the New York Academy of Sciences*, *320*(1 Health Effect), 69–124. http://doi.org/10.1111/j.1749-6632.1979.tb56594.x

Yang, C. Y., Best, A., Knell, J., Yang, E., Sheridan, A. D., Jesionek, A. K., … Goldrath, A. W. (2011). The transcriptional regulators Id2 and Id3 control the formation of distinct memory CD8+ T cell subsets. *Nature Immunology*, *12*(12), 1221–9. http://doi.org/10.1038/nature09421.Oxidative

Yates, P. R., Atherton, G. T., Deed, R. W., Norton, J. D., & Sharrocks, A. D. (1999). Id helix – loop – helix proteins inhibit nucleoprotein complex formation by the TCF ETS-domain transcription factors. *The EMBO Journal*, *18*(4), 968–976. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1171189/pdf/000968.pdf

Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., … Liu, X. S. (2008). Model-based Analysis of ChIP-Seq (MACS). *Genome Biology*, *9*(9),

Zhao, Q., Chang, C., Gonzalez, J. P., Alzahrani, K., Button, J. L., & Fraidenraich, D. (2016). Combined Id1 and Id3 deletion leads to severe erythropoietic disturbances. *PLoS ONE*, *11*(4). http://doi.org/10.1371/journal.pone.0154480

Zhu, Y. Y., Machleder, E. M., Chenchik, A., Li, R., & Siebert, P. D. (2001). Reverse transcriptase template switching: a SMART approach for full-length cDNA library construction. *BioTechniques*, *30*(4), 892–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11314272

CHAPTER 4

The whole transcriptome shotgun sequencing (RNA Sequencing) of ID3 target genes reveals the role of ID3 in the proliferation and dysregulation of lung endothelial cells.

Introduction-

In recent years, next-generation sequencing methods, for instance, ChIP-Seq and RNA-Seq, due to their distinct advantages in increasing specificity and sensitivity of gene expression, they have become a popular choice in biological studies. Specifically, RNA-Seq has become an attractive alternative to microarrays in the inference of differential expression between several conditions or tissues, for it gives more accurate detection and measure of gene expression. The power of sequencing RNA lies in the fact that the twin aspects of discovery and quantification can be combined in a single high-throughput sequencing assay called RNA-sequencing (RNA-Seq). Every RNA-Seq experimental scenario could potentially have different optimal methods for transcript quantification, normalization, and ultimately differential expression analysis.

Materials and Methods
Cell culture and treatment conditions

Primary cells, Human pulmonary microvascular endothelial cells or Human lung endothelial cells (HPMEC-ST1.6R). HPMEC cells were maintained in Dulbecco's modified Eagle's medium (DMEM-F12) supplemented with 5% (v/v) fetal bovine serum (FBS). Cells were cultured at 37 ºC in a humidified atmosphere with 5% CO_2.

Estrogenic chemical treatments

Stock solutions of PCB153 and 17β-estradiol were prepared in dimethyl sulfoxide (DMSO). The same amount of DMSO as in PCB and estradiol-treated cells were added to control cultures. The level of DMSO in experimental media was less than 0.1%. PCB

blood levels have been reported to reach approximately 1000 ng/mL (∼3 μM) in occupationally exposed individuals (Wassermann et al., 1979). Our unpublished data showed a significant increase in PCB- induced vascularization with PCB153 concentrations of 10–100 ng/mL. Based on known PCB blood levels from occupational exposure and our preliminary results; we chose a PCB dose of 100 ng/mL (∼0.3 μM) to expose endothelial cells. Our previous studies also showed endothelial cell proliferation and vascular tube formation at physiological doses of 17β-estradiol (Das & Felty, 2014a; Felty, 2006; Felty & Porther, 2008b) therefore, HPMEC and HPMEC ID3⁺ were exposed to estradiol at 100 ng/mL (∼3.6 nM). PCB congeners 2,2′,4,4′,5,5′-hexachlorobiphenyl (PCB153) were purchased from AccuStandard (NewHaven, CT) and dissolved in dimethyl sulfoxide (DMSO). All other chemicals and reagents were purchased from Sigma (St. Louis, MO).

ID3 overexpression

The HPMEC cells were stably transfected with either Precision LentiORF for ID3 (Thermo Scientific Open Biosystems) or empty vector lentiviral pLEX-JRED/TurboGFP by the trans-lentiviral packaging kit with Express-in transfection reagent according to the manufacturer's instructions. We used the MOI (multiplicity of infection) of 25 and selected cells that overexpressed ID3 with blasticidin S (5 mg/ml) as per manufacturer's instructions. Cells expressing TurboGFP were identified by fluorescence microscopy.

The whole transcriptome shotgun sequencing (RNA-Sequencing)
We extracted total RNA from the HPMEC (control), HPMEC ID3$^+$ and HPMEC ID3$^+$ +

PCB153 cells with the help of Invitrogen™ TRIzol™ Reagent RNA extraction method

Pub. No. MAN0001271 (Fisher Scientific Rio, Ares, Hannon, & Nilsen, 2010). The

protocol required using TRIzol™ Reagent. TRIzol™ Reagent maintains the integrity of

the RNA due to highly effective inhibition of RNase activity while disrupting cells and

dissolving cell components during sample homogenization (Chomczynski P, 1983). After

dislodging of growing monolayer of cells from petri-dishes with TRIzol™ Reagent,

chloroform is added to cell sample, and is allowed to separate into a clear upper aqueous

layer (containing RNA), an interphase, and a red lower organic layer (containing the

DNA and proteins). RNA is precipitated from the aqueous layer with isopropanol.

Isolated RNA is then used to prepare as a sample for RNA sequencing analysis. We

performed RNA sequencing analysis at Center for Genome Technology (CGT) of

University of Miami Miller School of Medicine. The CGT facility used Illumina: TruSeq

Stranded mRNA Sample Preparation Guide for RNA sequencing analysis (TruSeq,

2013).

Experiment procedure for RNA Sequencing (RNA-Seq)
Introduction-

To begin with RNA sequencing, we extracted total RNA from the HPMEC (control),

HPMEC ID3$^+$ and HPMEC ID3$^+$ + PCB153 cells with the help of Invitrogen™ TRIzol™

Reagent RNA extraction method Pub. No. MAN0001271 (Fisher Scientific Rio et al.,

2010). The protocol required using TRIzol™ Reagent, which is a monophasic solution of

phenol, guanidine isothiocyanate, and other proprietary components helps the isolation of

a variety of RNA species of large or small molecular size. TRIzol™ Reagent maintains the integrity of the RNA due to highly effective inhibition of RNase activity while disrupting cells and dissolving cell components during sample homogenization (Chomczynski P, 1983). After dislodging of growing monolayer of cells from petri-dishes with TRIzol™ Reagent, chloroform is added to cell sample, and is allowed to separate into a clear upper aqueous layer (containing RNA), an interphase, and a red lower organic layer (containing the DNA and proteins). RNA is precipitated from the aqueous layer with isopropanol. Isolated RNA is then used to prepare as a sample for RNA sequencing analysis.

Method-

1×10^5–1×10^7 cells grown in monolayer in 3.5 cm culture dish (10 cm^2). We added 0.3–0.4 mL of TRIzol™ Reagent per 1×10^5—10^7 cells directly to the culture dish to lyse the cells. We did pipet the lysate up and down several times to homogenize. We centrifuged the lysate for 5 minutes at $12,000 \times$ g at 4–10 °C, then transfer the clear supernatant to a new tube. We incubated for 5 minutes to permit complete dissociation of the nucleoproteins complex. We added 0.2 mL of chloroform per 1 mL of TRIzol™ Reagent used for lysis, then securely cap the tube. We centrifuged the sample for 15 minutes at $12,000 \times$ g at 4 °C. The mixture separates into a lower red phenol-chloroform, and interphase, and a colorless upper aqueous phase. We transferred the aqueous phase containing the RNA to a new tube by angling the tube at 45° and pipetting the solution out.

We separated phases. We then added 0.5 mL of isopropanol to the aqueous phase, per 1 mL of TRIzol™ Reagent used for lysis. We centrifuged for 10 minutes at 12,000 × g at 4°C.Total RNA precipitate forms a white gel-like pellet at the bottom of the tube. We discarded the supernatant with a micropipette. We resuspended the pellet in 1 mL of 75% ethanol per 1 mL of TRIzol™ Reagent used for lysis. Vortex the sample briefly, then centrifuge for 5 minutes at 7500 × g at 4°C. Vacuum or air dry the RNA pellet for 5–10 minutes. We resuspended the pellet in 20–50 μL of RNase-free water, 0.1 mM EDTA, or 0.5% SDS solution by pipetting up and down. We incubated in a water bath or heat block set at 55–60 °C for 10–15 minutes. We determined the RNA yield using absorbance method-Absorbance at 260 nm provides total nucleic acid content, while absorbance at 280 nm determines sample purity. Since free nucleotides, RNA, ssDNA, and dsDNS absorb at 260 nm, they all contribute to the total absorbance of the sample. We diluted sample in RNase-free water, then measure absorbance at 260 nm and 280 nm. We calculated the RNA concentration using the formula A260 × dilution × 40 = μg RNA/mL. Calculate the A260/A280 ratio. A ratio of ~2 is considered pure.

RNA sequencing-

We performed RNA sequencing analysis at Center for Genome Technology (CGT) of University of Miami Miller School of Medicine. The CGT facility analyzed the quality of RNA samples by absorption method. The concentration of three RNA samples prior to use for sequencing are given as follows in table 10-

Project	Prep	volume (ul)	bioA Conc (ng/ul)	Total RNA input (ng)	input volume (ul)
		219.46	84.20	500	5.94
qRNASeq1		245.80	102.00	500	4.90
	TruSeq	232.01	91.1	500	5.49

Table 10- It gives the information about the sample like concentration, volume, and quality.

The CGT facility used Illumina: TruSeq Stranded mRNA Sample Preparation Guide for RNA sequencing analysis (TruSeq, 2013). They have used adapter sequence specific to each samples. The adapter sequences are listed in the following table 11.

Lane	Study_Project	Sample	Barcode sequence	PF Clusters	total PF Reads	FC_ID
1	qFelty_qRNASeq1	201603252-01	ATCACG	4,551,510		160314_NS500167_0017_AHYF3HBGXX
2	qFelty_qRNASeq1	201603252-01	ATCACG	4,523,244		160314_NS500167_0017_AHYF3HBGXX
3	qFelty_qRNASeq1	201603252-01	ATCACG	4,654,605		160314_NS500167_0017_AHYF3HBGXX
4	qFelty_qRNASeq1	201603252-01	ATCACG	4,582,159		160314_NS500167_0017_AHYF3HBGXX
1	qFelty_qRNASeq1	201603252-01_	ATCACG	8,249,593		160325_NS500167_0018_AHY7TMBGXX
2	qFelty_qRNASeq1	201603252-01_	ATCACG	8,137,621		160325_NS500167_0018_AHY7TMBGXX
3	qFelty_qRNASeq1	201603252-01_	ATCACG	8,342,614		160325_NS500167_0018_AHY7TMBGXX
4	qFelty_qRNASeq1	201603252-01_	ATCACG	8,172,774	51,214,120	160325_NS500167_0018_AHY7TMBGXX
1	qFelty_qRNASeq1	201603253-01	TTAGGC	5,029,343		160314_NS500167_0017_AHYF3HBGXX
2	qFelty_qRNASeq1	201603253-01	TTAGGC	4,959,964		160314_NS500167_0017_AHYF3HBGXX
3	qFelty_qRNASeq1	201603253-01	TTAGGC	5,138,406		160314_NS500167_0017_AHYF3HBGXX
4	qFelty_qRNASeq1	201603253-01	TTAGGC	5,047,644		160314_NS500167_0017_AHYF3HBGXX
1	qFelty_qRNASeq1	201603253-01_	TTAGGC	9,095,323		160325_NS500167_0018_AHY7TMBGXX
2	qFelty_qRNASeq1	201603253-01_	TTAGGC	8,928,221		160325_NS500167_0018_AHY7TMBGXX
3	qFelty_qRNASeq1	201603253-01_	TTAGGC	9,219,859		160325_NS500167_0018_AHY7TMBGXX
4	qFelty_qRNASeq1	201603253-01_	TTAGGC	9,025,336	56,444,096	160325_NS500167_0018_AHY7TMBGXX
1	qFelty_qRNASeq1	201603254-01	ACTTGA	5,162,332		160314_NS500167_0017_AHYF3HBGXX
2	qFelty_qRNASeq1	201603254-01	ACTTGA	5,110,305		160314_NS500167_0017_AHYF3HBGXX
3	qFelty_qRNASeq1	201603254-01	ACTTGA	5,272,969		160314_NS500167_0017_AHYF3HBGXX
4	qFelty_qRNASeq1	201603254-01	ACTTGA	5,177,994		160314_NS500167_0017_AHYF3HBGXX
1	qFelty_qRNASeq1	201603254-01	ACTTGA	7,523,473		160325_NS500167_0018_AHY7TMBGXX
2	qFelty_qRNASeq1	201603254-01	ACTTGA	7,408,121		160325_NS500167_0018_AHY7TMBGXX
3	qFelty_qRNASeq1	201603254-01	ACTTGA	7,603,456		160325_NS500167_0018_AHY7TMBGXX
4	qFelty_qRNASeq1	201603254-01	ACTTGA	7,437,656	50,696,306	160325_NS500167_0018_AHY7TMBGXX

Table 11- It gives the information about sequencing reads obtained per the sample.

Methodology to analyze the RNA-sequencing data

Illumina TruSeq RNA libraries were constructed and sequenced using Illumina HiSeq

2000 at CGT. The CGT facility at UM sent us the data in the FASTA format single ended

files. We received FASTA files for HPMEC (control), HPMEC ID3$^+$, and HPMEC ID3$^+$

+ PCB153 samples. The FASTA files were constructed for strand-specific Single-end

cDNA sequencing on the Truseq sequencer (Illumina, UM, Miami). Low-quality

sequences analyzed by FASTQC (Andrew, 2010) and adapter sequences were trimmed

by Trimmomatic, v.0.33 (Bolger et al., 2014; Usadellab, 2013), and aligned to the

reference human genome (GRCh38, GENCODE release 23 primary assembly) by the

sequence alignment software Spliced Transcripts Alignment to a Reference, RNA STAR

(Dobin et al., 2013), Bowtie2 (Langmead & Salzberg, 2012), and HISAT (Kim,

Langmead, & Salzberg, 2015). We mapped the sample FASTQ files separately by three

different mapping software to check the mapping efficiency of these various software and

establish optimal pipeline strategy to analyze RNA Seq data to achieve correct

interpretation. Bowtie2 and HISAT are ungapped mapper. In other hand RNA STAR is a

gapped mapper which able detect non-canonical splice junctions (Conesa et al., 2016).

After all filtering and mapping, a high percentage of poor or unaligned reads were

rejected and removed from further analysis. Overall, the mean number of raw reads was

50 million per sample. Of the mean 5.6 million high-quality reads, 24.2% mapped to a

unique location in the reference human genome. RNA STAR, Bowtie2, and HISAT

mapping software aligned millions of short reads to the reference genome and outputs

either SAM file (.sam) is a tab-delimited text file that contains sequence alignment data

or a BAM file (.bam) is the binary version of a SAM file. SAM or BAM file can be used

as input for a read summarization program suitable for counting reads generated from RNA. We used featureCounts program which implements highly efficient chromosome hashing and feature blocking techniques. It is considerably faster than existing methods by an order of magnitude for gene-level summarization (Liao, Smyth, & Shi, 2014). FeatureCounts produces a table containing counted reads, per gene, per row. Optionally the last column can be set to be the effective gene-length. These tables are compatible with the DESeq2 Galaxy wrapper by IUC. Column names are added as metadata object. In order to make the expression levels of genes comparable and further conduct differential expression analysis, normalization is a crucial step in data processing. The normalization step aims to adjust the systematic technical effects and reduce the noise on the data as well. Mostly, to compare the transcriptome data, the null hypothesis is being tested on the logarithmic fold change (LFC) between treatment and control for a gene's expression. The null hypothesis for the logarithmic fold change between treatment and control for a gene's expression is considered to be exactly zero, i.e., that the gene is not at all affected by the treatment. The one way to normalize the raw RNA-Seq data by reads per kilo base of exon model per million reads (RPKM) (Mortazavi, Williams, McCue, Schaeffer, & Wold, 2008) which is a within-sample normalization method that will remove the feature-length and library-size effects. RPKM and its subsequent derivatives fragments per kilo base of exon model per million mapped reads (FPKM), a within-sample normalized transcript expression measure analogous to RPKs acts as a RNA-Seq gene expression values. Normalization with RPKM method performs poorly when samples have highly and differentially expressed features (Bullard, Purdom, Hansen, & Dudoit, 2009; Hansen, Brenner, & Dudoit, 2010). Ranking by fold change, on the other

hand, is complicated by the noisiness of LFC estimates for genes with low counts. Furthermore, the number of genes called significantly differentially expressed depends as much on the sample size and other aspects of experimental design (Boer et al., 2001). Inferential method is another way that treat each gene separately suffer from lack of power, due to the high uncertainty of within-group variance estimates and the small number of samples in typical high-throughput assay experiments – often as few as two or three replicates per condition. In high-throughput assay analysis, this limitation can be overcome by pooling information across genes, specifically, by estimating variance-mean dependence in count data and test for differential expression based on a model using the negative binomial distribution (Conesa et al., 2016). DEseq2 is software which uses negative binomial distribution statistics, which gives framework to facilitate gene ranking and visualization based on stable estimation of effect sizes, as well as testing of differential expression with respect biological significance. DESeq2, is a method for differential analysis of count data, using shrinkage estimation for dispersions and fold changes to improve stability and interpretability of estimates (Anders & Huber, 2010; Love, Huber, & Anders, 2014) as shown in Figure 83 and 84. This enables more quantitative analysis focused on the strength rather than the mere presence of differential expression. DESeq2 takes count tables that generated from the featureCounts as input. Count tables must be generated for each sample individually. DESeq2 internally corrects for library size, so it is important to provide un-normalized raw read counts as input. We used variance-stabilizing transformation to account for differences in sequencing depth. P-values were adjusted for multiple testing using the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995). A false discovery rate adjusted p-value (i.e., q-value) <

0.05 was set for the selection of DE genes. A nominal adjusted P value for each algorithm, which resulted in a median actual precision of 0.9 (FDR = 0.1).

Inputs- The following table gives some examples of factors and their levels:

Factor	Factor level 1	Factor level 2
Treatment	Treated	Untreated
Condition	Overexpression	Wildtype
Gender	Female	Male

Output- DESeq2 generates a tabular file containing the different columns and optional visualized results as PDF.

Column	Description
1	Gene Identifiers
2	Mean normalized counts, averaged over all samples from both conditions
3	Logarithm (to basis 2) of the fold change (See the note in inputs section)
4	Standard error estimate for the log2 fold change estimate
5	Wald statistic
6	p value for the statistical significance of this change
7	p value adjusted for multiple testing with the Benjamini-Hochberg procedure which controls false discovery rate (FDR)
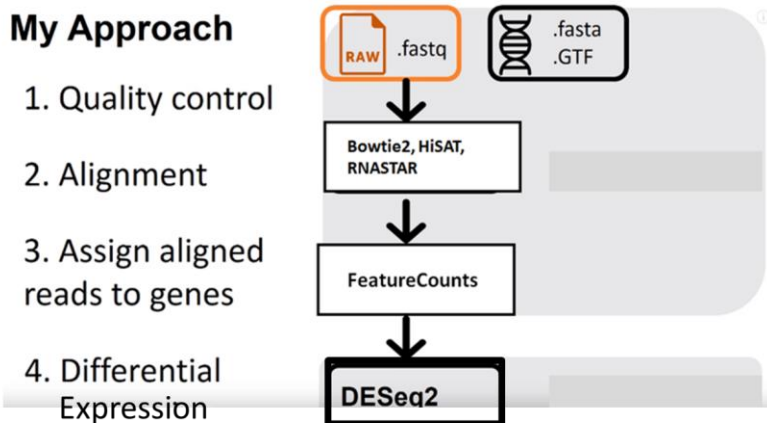
Figure 83- It illustrates the information about the approach of RNA Seq analysis used in this study.

194

Figure 84- It illustrates the setting used for DESeq2 analysis of differentially expressed genes in ID3 overexpressed sample VS vector control.

Integrative analysis of ChIP- and RNA- Seq data with Binding and expression target analysis (BETA)

BETA is a software package that integrates ChIP-Seq data of TF with transcriptome quantification and then differential gene expression data to infer direct target genes. We used available open source at http:// cistrome.org/BETA. We used BETA-minus software and input only binding data i.e. peak regions bed file generated from MACS2 output which predicted target genes. The purpose of this protocol is to predict genes that are the direct targets of TFs or chromatin regulators. BT minus defined the targets as the genes with a high regulatory potential, derived only from TF binding within a specific region (100 kb default). BETA calculates regulatory potential for each gene using the equation $S_g = \sum_{i=1}^{k} e^{-(0.5+4\Delta_i)}$. BETA minus is only equipped to predict TF target genes. It cannot predict whether the factor binding is activating or repressing gene expression (S. Wang et al., 2013). We checked the potential target genes output from BETA minus with DESeq2 differential expressed gene list and then sorted them into upregulated or downregulated genes list on the basis of fold change of ID3 overexpressed count data over HPMEC control count data.

Results-

To examine the global changes in the ID3 gene overexpression HPMEC cells by using whole transcriptome shotgun sequencing.

We performed global transcriptome analysis on HPMEC and HPMEC ID3$^+$ cells. We isolated RNA from HPMEC and HPMEC ID3$^+$ cells and performed RNA Seq analysis comparing gene expression. We mapped FASTAQ files to a reference genome an

ungapped aligner Bowtie2. Mapping with gapped mapper allows for the identification of novel genes or transcripts. However, we did not require any novel transcript discovery, therefore we used Bowtie2. Bowtie2 mapped reads, then used for transcript identification and quantification. We found total of 1061 genes (significantly change over control with $P < 0.05$) were differentially expressed in HPMEC ID3$^+$ than HPMEC cells (Additional file 8A). We then selected only selected differentially expressed genes with at least one fold change in expression. Therefore, out of of 1061 genes, we finally found 43 genes that were downregulated and 22 upregulated genes. DESeq2 normalized the gene expression counts taken from RNA Seq data of HPMEC ID3$^+$ and HPMEC cells. An MA-plot is used to visualize differences between measurements taken in HPMEC ID3$^+$ and HPMEC cells. An MA-plot is a plot of log-fold change (M-values, i.e., the log of the ratio of level counts for each gene between two samples) against the log-average (A-values, i.e., the average level counts for each gene across the two samples) (Love et al., 2014). The MA-plot is shown in Figure 85. We used volcano plots to show significance versus fold-change on the y and x axes, respectively. A volcano plot combines a measure of statistical significance from a statistical test (e.g., a p value from an ANOVA model) with the magnitude of the change, enabling quick visual identification of genes that display large magnitude changes that are also statistically significant.The green colour marked genes detected as differentially expressed and named at Padj<0.001 when Bonferroni correction testing adjustment is used (Figure 86). The dispersion estimates for normaized data is given in additional file 7. A representative heat map is shown in Figure 87 (complete image of heat map is given in additional file). The complete list of raw data of expressed genes is given in additional file 8B**.** To further examine the biological

significance of the differentially expressed genes, we used FunRich which is an open access standalone functional enrichment and interaction network analysis tool (Pathan et al., 2015, 2017). G-protein NF-Kβ signaling pathway, immune response, DNA replication, cell proliferation, and MAPK cascade are topmost significant biological processes network were expressed after ID3 overexpression which shown in Figure 88 (Additional file 9). The top 20 canonical pathways modulated by ID3 are shown in Figure 82. We found that the topmost pathways modulated by ID3 were regulation of cell cycle checkpoints, oxidative stress response, EGFR dependent signaling network, biological oxidation and VEGF and VEGFR signaling network (Additional file 10). Further, analysis of molecular and cellular functions of the differentially expressed genes revealed that they are involved in protein-metal ion binding, enzyme binding, cadherin binding, sequence specific DNA binding, and transcription regulator activity (Figure 90) (Additional file 11). KEGG pathways were significantly enriched in these ID3-regulated genes, which included DNA/RNA events, cell cycle, cell-cell interactions (focal adhesions, adherent junctions), cancer-related pathways (MAPK, LKB1, mTOR and VEGF) signaling events.
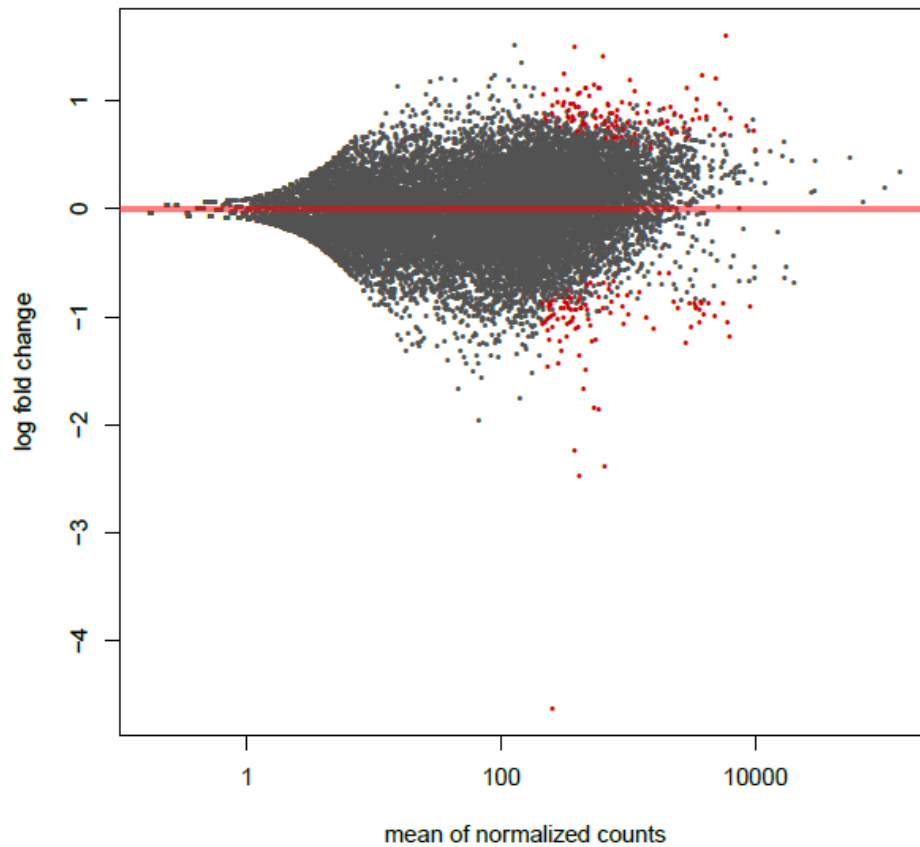
Figure 85- The MA-plot

It visualizes the differences between gene expression in HPMEC ID3$^+$ and HPMEC (control) cells, by transforming the data onto M (log2 fold changes attributable to a gene expression in HPMEC ID3$^+$ over the mean of normalized counts for both HPMEC ID3$^+$ and HPMEC (control) cells. Points have been colored red whose adjusted p value is less than 0.1.
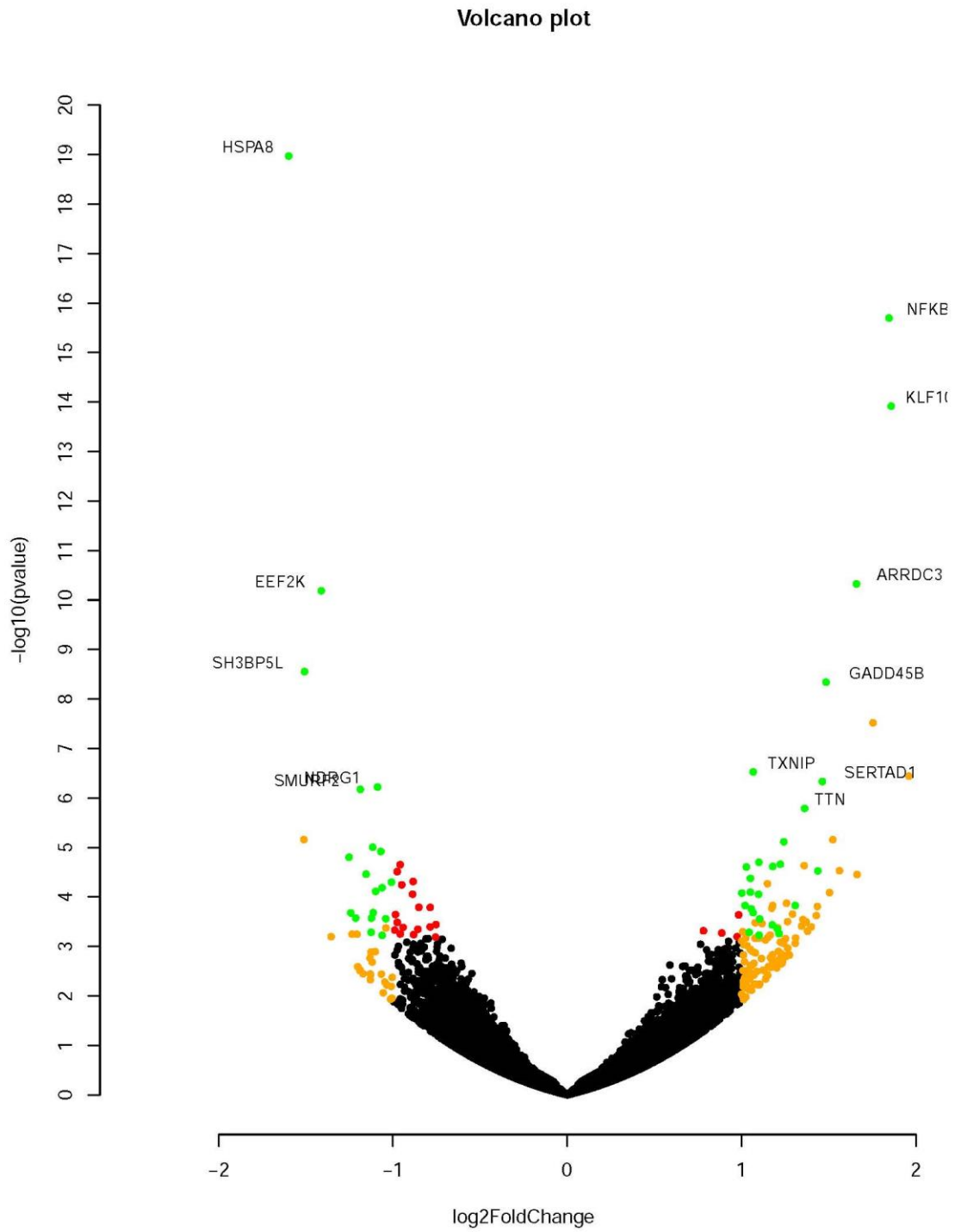
Figure 86- Volcano plot

It illustrates fold differences in individual gene expression in HPMEC ID3$^+$ vs HPMEC (control) cells, and associated P values (negative log10), in RNA Seq experiment.
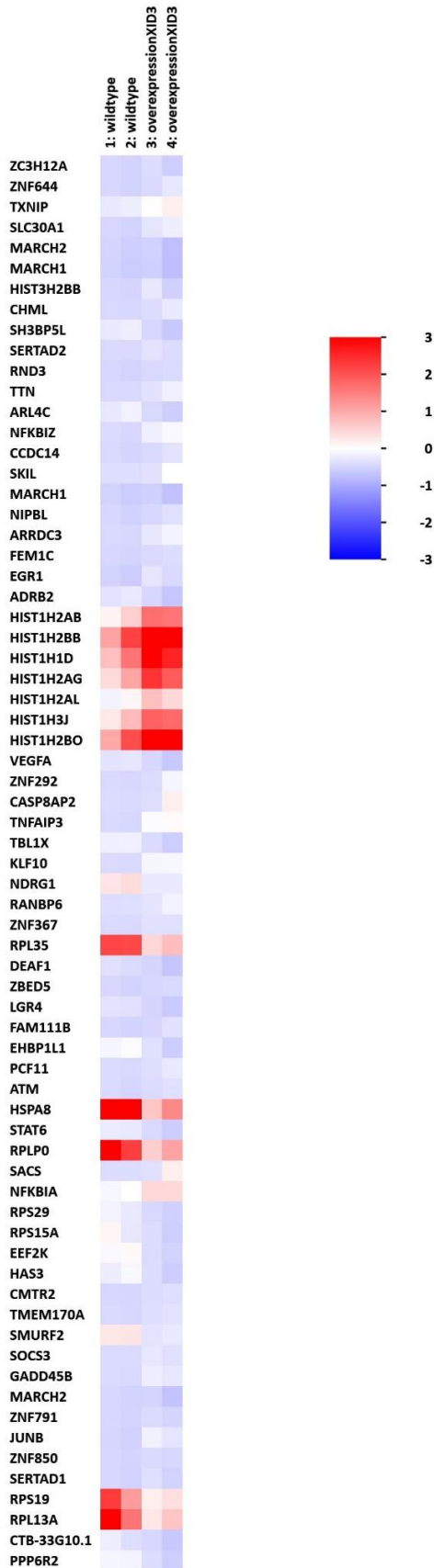
Figure 87- HeatMap

Gene expression profiles of HPMEC ID3$^+$ and HPMEC (control) cells are illustrated by heat map. Heat map of differentially expressed genes between control and ID3 overexpressed cells is shown. Increased or decreased mRNA expression is represented by red or blue, respectively.
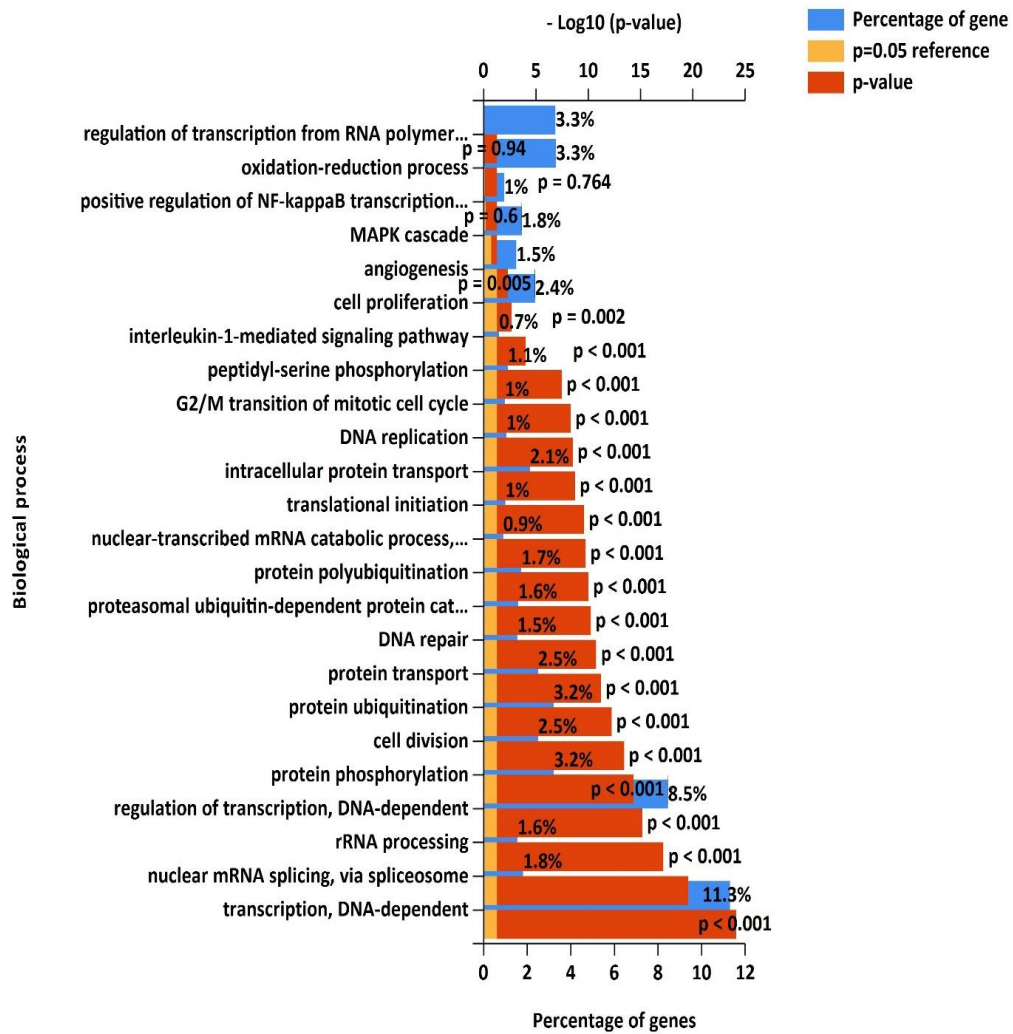
Figure 88- FunRich GEO Biological Process

Differentially expressed genes were subjected to FunRich GO enrichment analysis to elucidate the biological processes and the top 20 biological processes are shown.
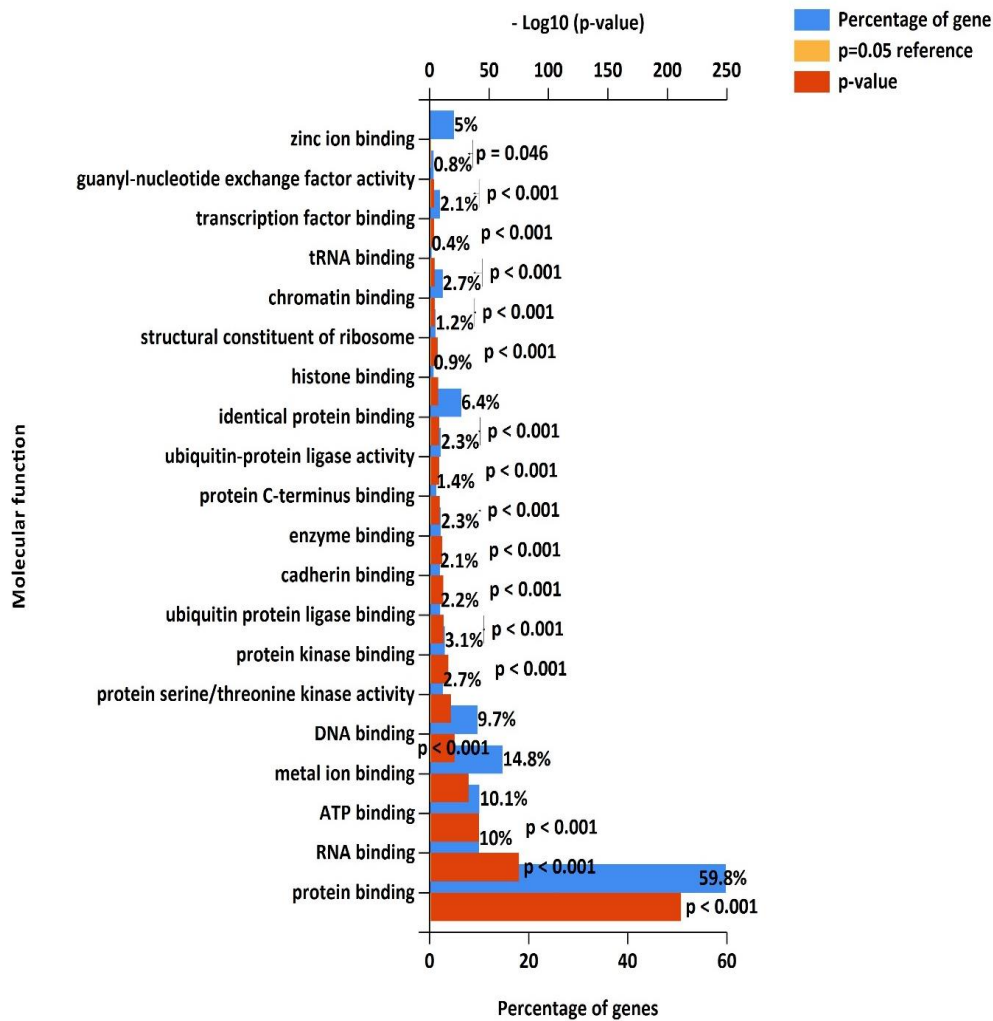
Figure 89- FunRich Molecular Function

Differentially expressed genes were subjected to FunRich GO enrichment analysis to elucidate the molecular function of DE genes. The top 20 molecular functions in which DE genes involved are shown.
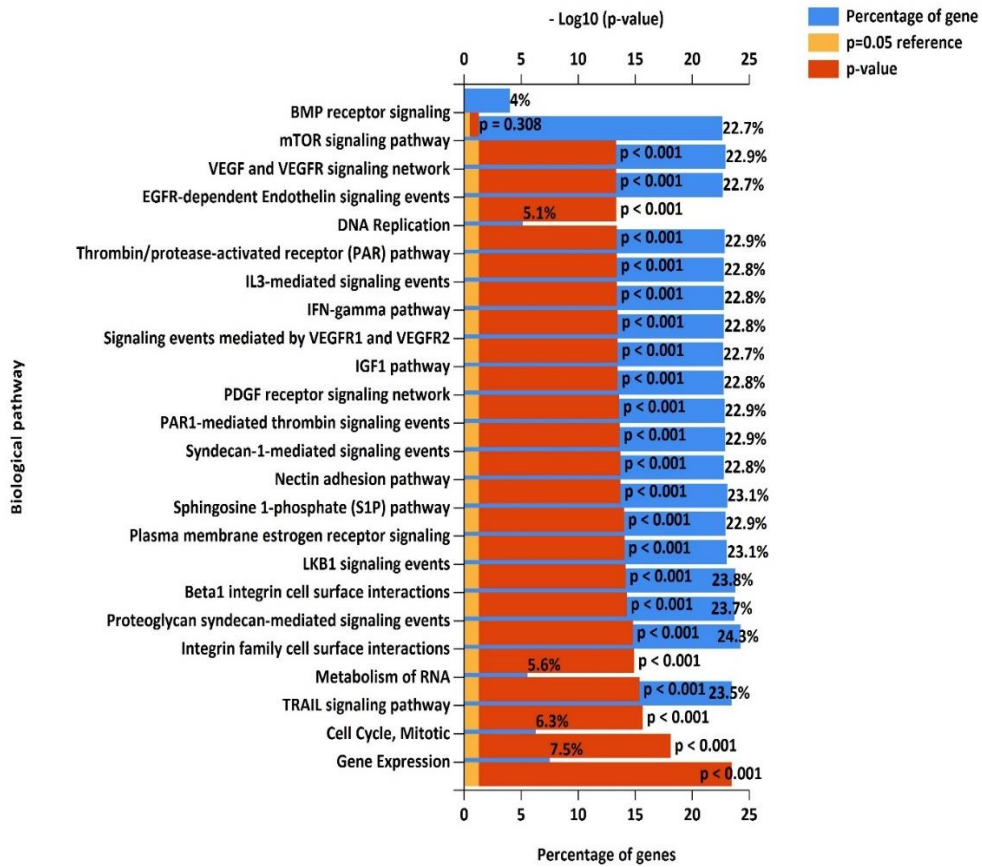
Figure 90- FunRich Biological Pathways

Differentially expressed genes were subjected to pathway analysis using FunRich enrichment analysis and the top 20 biological pathways are shown.

Understanding ID3 gene regulatory mechanisms by integrating ChIP-Seq and RNA-Seq data in HPMEC cells.

To answer the question whether ID3 binding to the genomic regions has functional consequences through changes in the expression level of target genes, the ChIP-Seq data were integrated with the RNA Seq data. High-throughput technologies like, RNA-Seq and ChIP-Seq are widely used to measure gene expression and to obtain genome-wide maps of transcription factor occupancies (Angelini & Costa, 2014; Park, 2009; Wang, Gerstein, & Snyder, 2009). The interplay between transcriptomics and epigenomics is very important to understand the chromatin accessibility to the transcription machinery which regulates gene expression and, vice versa, some non-coding RNAs can affect local chromatin states (Wang et al., 2011) In the pipeline process of ChIP Seq data, MACS2 and SICER10 are some of the peak caller tools uses to identify precise TF-binding sites called peak regions. These peak regions then associated with nearest gene. This peak-based method to identify targets is to assign the proximal nearest gene or the gene containing peaks in its promoter region (Boyer et al., 2005; Rougemont & Naef, 2012). Researchers have shown that most of the TF ChIP-Seq data contain only a small percentage of binding at the promoter region. In addition the use of nearest peaks to assign target genes is very unreliable (Hua et al., 2008). Therefore, it is very important to integrate ChIP-Seq and RNA-Seq data to predict gene expression during various cellular and metabolic pathways (Angelini & Costa, 2014). We used BETA to predict precise gene targets of ID3 in HPMEC cells. We used BETA minus on Cistrome Galaxy platform, which predicted ID3 direct gene targets by using .BED peaks file of ID3 binding regions. We gave two ranks (R) to every gene target predicted by BETA minus (g): one is based on decreasing regulatory potential (Rgb), such that is Rgb = 1 for the

gene with the largest regulatory potential score, and the other is based on the increasing

of the P value (Rge), that is, Rge = 1 for the most strongly differentially expressed gene.

Then the rank product of the gene (g), RPg = (Rgb/n)*(Rge/n). The RP can be interpreted

as a P value; because it shows the probability that the gene has a regulatory potential rank

≤Rgb and a differential expression changed rank ≤Rge. On the basis of RP, we judged the

targets with an RP less than 10^{-3} will be more likely to be the true target genes of ID3

(Breitling, Armengaud, Amtmann, & Herzyk, 2004). We found total of 27 most

significant genes (RPg <0.001), which is supposed to be regulated by ID3 (Additional file

13). We then further analyzed by using gene ontology–based tools such as STRING and

FunRich to link functions to this set shown table 12, Figures 91 and 92. If we decrease

the significance level up to RPg <0.05 then list increased to 493 genes. We also identified

19 genes whose promoter region bound by ID3 and also differentially expressed in RNA

Seq data. These data shows that 19 gene were directly regulated by ID3 (additional file

14) (Thomas et al., 2003).

GO biological process complete	Homo sapiens - REFLIST (21042)	(fold Enrichment)	(raw P-value)	(FDR)
cellular response to zinc ion (GO:0071294)	19	47.13	1.25E-08	1.93E-04
cellular response to cadmium ion (GO:0071276)	31	24.07	3.89E-06	1.51E-02
response to zinc ion (GO:0010043)	54	16.58	2.91E-06	1.51E-02
cellular response to metal ion (GO:0071248)	154	8.72	1.49E-06	1.16E-02
cellular response to inorganic substance (GO:0071241)	178	7.55	4.62E-06	1.43E-02
positive regulation of RNA metabolic process (GO:0051254)	1504	2.48	2.46E-05	4.24E-02
regulation of transcription, DNA-templated (GO:0006355)	3680	1.87	1.74E-05	4.51E-02
regulation of nucleic acid-templated transcription (GO:1903506)	3704	1.85	1.85E-05	4.11E-02
regulation of RNA biosynthetic process (GO:2001141)	3712	1.85	1.90E-05	3.69E-02
regulation of RNA metabolic process (GO:0051252)	3909	1.79	2.98E-05	4.62E-02

Table 12- A detailed list of GO biological processes associated with most significant genes identified by BETA software.
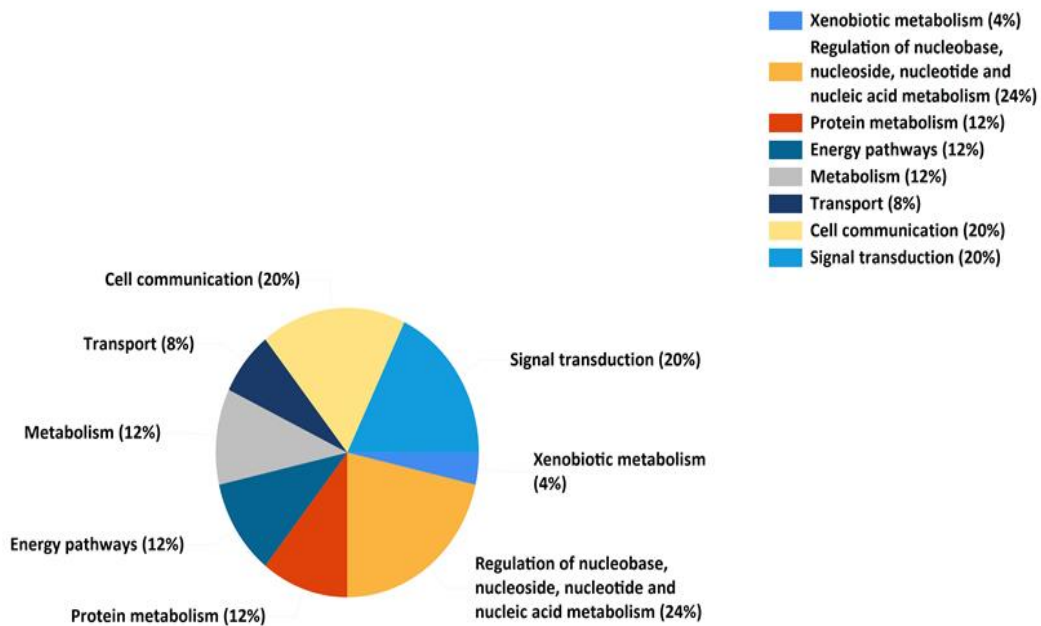
Figure 91- FunRich GO enrichment analysis of ID3 regulated significant genes involved in molecular functions

ID3 regulated significant genes were subjected to FunRich GO enrichment analysis to elucidate the molecular function. The top 10 molecular functions in which these genes involved are shown.

Figure 92- FunRich GO enrichment analysis of ID3 regulated significant genes involved in biological pathways

ID3 regulated significant genes were subjected to FunRich GO enrichment analysis to elucidate the biological pathway. The top biological processes in which these genes involved are shown.

A)

EGFR TYROSINE KINASE INHIBITOR RESISTANCE

Non-small cell lung cancer | Pancreatic cancer | Glioma

ERBB signaling pathway

EGFR mutation

EGF | TGFα | HRG | HGF | IGF | VEGF | PDGF | FGF | GAS6 | IL-6

EGFR tyrosine kinase inhibitors
○ Gefitinib
○ Erlotinib

Bypass pathway

Her2 mutation
Her2 overexpression
Her3 overexpression

c-Met amplification
HGF overexpression

IGFR abnormality
EGFRvIII
VEGF/VEGFR abnormality
PDGF/PDGFR abnormality
FGF/FGFR abnormality
AXL overexpression
IL-6 abnormality

JAK | STAT3 | Src | STAT3

Abnormal dimerization of STAT3

Jak-STAT signaling pathway

Downstream pathway

mTOR signaling pathway

mTOR | p70S6K | S6 → Protein synthesis | eIF-4BP | eIF-4E

Loss of PTEN | PTEN

PIK3CA mutation

PI3K | PIP3 | PKB/Akt | Bad | GSK-3 | Foxo3a

PI3K-Akt signaling pathway

GAB1

PLCγ | DG | PKC

Shc | Grb2 | Sos | Ras | Raf | MEK | ERK

MAPK signaling pathway

Low expression of NF1 | NF1 | K-RAS mutation | BRAF mutation

Impairment of apoptotic pathway

Bim | Bcl-XL | Bcl-2 | Bax | Apoptosis

Bim BH3 deletion

Mitochondrion

Epithelial mesenchymal transition (EMT)

Differentiation | Growth | Proliferation | Survival | Motility | Angiogenesis

DNA | Bim

01521 6/21/16
(c) Kanehisa Laboratories

B)

mTOR SIGNALING PATHWAY

Lysosome

Amino acids | SLC38A9 | V-ATPase | Regulator | RNF152 | Skp2 | GATOR1 | GATOR2 | SESN2 | Leu

SLC7A5 | SLC3A2

Amino acids

STRAD | LKB1 | MO25 | FNIP | FLCN

RagA/B | RagC/D

Energy deficit | AMP | AMPK

AMPK signaling pathway

Hypoxia | REDD1

Wnt | Frizzled LRP5/6 | Dvl | GSK3B

TNFα | TNFR | IKKβ

TSC1/2 TBC1D7 | Rheb

RSK | ERK1/2 | MEK | Raf | Ras

MAPK signaling pathway

Ins/IGF | R | Grb2 | SOS | IRS1 | PI3K | PTEN | PIP3 | PDK1 | Akt | IKKα

Grb10 | S6K

Insulin signaling pathway | PI3K-Akt signaling pathway

Raptor | Deptor | mTOR | mLST8 | Tel2 | Tti1 | PRAS40

mTORC1

CLIP-170 → Microtubule organization
Grb10 → Lipolysis
Lipin-1 → Lipid biosynthesis | Glycerolipid metabolism
ATG1 → Autophagy | Regulation of autophagy
4E-BP | eIF4E → Cap-dependent translation | Protein synthesis
S6K | eIF4B → Ribosome biogenesis
S6

mSin1 | Deptor | mTOR | mLST8 | Tel2 | Tti1 | Rictor | Protor

mTORC2

Rho → Cytoskeletal organization | Regulation of actin cytoskeleton
PKC
SGK1 → Cell survival

04150 7/4/16
(c) Kanehisa Laboratories

210

C)

Figure 93- A,B,C ID3 regulated significant genes involved in EGFR, mTOR and Focal adhesion signaling biological pathway are shown using KEGG pathway database.

Figure 94- The cell signaling biological pathways shown in diagram are involved in cell proliferation.

Discussion

We applied ChIP-Seq technology to identify global ID3 binding regions in HPMEC and HPMEC ID3$^+$ cells. To our best knowledge, this is the first global analysis of ID3's binding regions in HPMEC cells. ID3 belongs to a family of HLH proteins identified and named over two decades ago for its dual role as both inhibitors of the differentiation process and inhibitors of DNA binding. We show here for the first time that ID3 is part of a more general mechanism of transcriptional regulation than was at first thought. Our ChIP Seq data show that ID3 binds to a subset of 1240 target genes shared by both HPMEC ChIP and HPMEC ID3$^+$ ChIP samples (Figure 78). However, only a minor fraction (3.5-3.7%) of the ID3 binding sites identified which are close to TSSs. This moderate percentage may be explained by the ID3 binding site distribution, where most of the binding sites are related with orphan intergenic regions. In this regard, our study demonstrates ID3 binding sites in regions distal to TSSs, preferentially in distal intergenic regions, which highlights a potential role as a distal or alternative transcriptional regulator, although this does not rule out indirect regulation.

MEME-ChIP showed that ID3 binds to statistically significant GAGAGAGAGA motif sequence on the genomic DNA. Researchers have not been confirmed yet any ID3 binding sequences. However, Zaragoza et al. showed in genome wide mapping experiment that Id2 binds to individually as well as along with E2F4 on c-myc promoter of mouse liver cells (Ferrer-Vicens et al., 2014). Id2 is one of the members of the ID family proteins, which lacks DNA binding domain similar to ID3. In addition, genome wide mapping (ChIP on chip) and then MEME analysis showed overrepresentation of similar GAGAGAGA motifs also known as GAF motif for polycomb/trithorax response

elements in Drosophila melanogaster (Ringrose, 2017; Ringrose, Rehmsmeier, Dura, & Paro, 2003). These evidences support the hypothesis of ID3 binding to the DNA. There is another possibility that ID3 may be regulating gene expression by interacting with other known TFs and then binding to DNA as a protein complex. This type of protein complex binding is also known as cofactor based targeting binding. Targeting of non-DNA-binding cofactors to DNA was earlier viewed solely via protein–protein interactions with DNA-bound TFs. This is known as indirect interaction with DNA through the TFs. However, several studies have described a provocative alternative mechanism in which recruited cofactors interact directly with DNA when part of a larger, multi-protein complex (Babb, Huang, Aufiero, & Herr, 2001; Siggers, Duyzend, Reddy, Khan, & Bulyk, 2011). The non-DNA-binding cofactors mediate sequence-specific interactions that preferentially stabilize the binding to composite DNA sites containing specific auxiliary motifs. This represents yet another mechanism to enhance the DNA sequence specificity of multi-protein regulatory complexes. The regulation of sulfur metabolism genes in yeast involves a multi-protein complex composed of a sequence-specific TF Cbf1 and two non-DNA-binding cofactors Met28 and Met4 (Blaiseau & Thomas, 1998; Kuras, Cherest, Surdin-Kerjan, & Thomas, 1996). Cbf1 is a bHLH protein and binds as a homodimer to E-box sites with a consensus 5′-CACGTG-3′ core (Kuras et al., 1996). Therefore, we predict it may be possible that ID3 regulating gene expression by interacting with TFs like BCL11A, IRF4, PRDM1, FOXJ3, SMAD4, ZBTB6, STAT2, and GATA1. Researchers have already shown that loss of SMAD4 and overexpression of IRF4 increases the cell proliferation in Non-small cell lung cancer (NSCLC) (Haeger et al., 2016; Qian et al., 2017). In addition, our lab and other researchers have shown ID3

involvement in EC proliferation acts as a disease phonotype for NSCLC (Castañon et al., 2013). The regulation of ID3 through interaction with other TFs proteins enhances the cooperation in the modulation of transcription of other genes. Taking into account this complete transcriptional scenario, our data describe potential interactions of ID3 with both common TFs and core elements, which could cooperate in chromatin remodeling for transcriptional regulation in lung endothelial cells.

Identification of biological processes controlled by overexpressed ID3 in HPMEC has been mainly associated to lung EC proliferation and checking at cell cycle phases through its transcriptional role in key EC proliferation-related genes (Al-husseini et al., 2015; Das & Felty, 2014b; Felty, Yoo, & Kennedy, 2010). At this regard, we observed a downregulation of Cell division protein kinase 6 (CDK6), which is an enzyme encoded by the CDK6 gene. It is important for the G1 phase progression and G1/S transition of the cell cycle (Meyerson et al., 1992). Researchers have shown that deregulation of CDK6 has been shown to be important in lymphoid malignancies by increasing angiogenesis, a hallmark of cancer (Diaz-Moralli, Tarrado-Castellarnau, Miranda, & Cascante, 2013). Interestingly, we observed four fold downregulation of EGR1. EGR1 acts as a tumor suppressor in NSCLC (Zhang et al., 2013). Our data showed that ID3 significantly downregulated EGR1 which could result into HPMEC hyper proliferation. Intriguingly, ID3 modulates the expression of several genes involved in carcinogenesis and lung malignancies in Figures 93 A, B, C and 87 (insulin signaling, mTOR signaling, EGFR signaling and MAPK pathways), inflammatory response and angiogenesis (VEGFR signaling). We observed VEGFA overexpression in RNA Seq data, which is angiogenesis marker. Moreover, we found upregulation of gene HSPA8 that is a heat

214

shock family protein 70. Its functions contribute to biological processes including signal transduction, apoptosis, autophagy, protein homeostasis, and cell growth. It codes for the protein HSP72 which is shown to be over expressed in gastric cancers, colon cancers, breast cancers, and lung cancers (Wang et al., 2013a) (Mayer & Bukau, 2005; Wang et al., 2013b). We also observed EEF2K gene upregulation which codes for calcium/calmodulin-dependent eukaryotic elongation factor 2 kinase enzyme (Ryazanov et al., 1997). Researchers have shown upregulation EEF2K gene in various cancers (Leprivier et al., 2013).

Moreover, MyCancerGenome has compiled database on molecular profiling of lung cancer which is comprised of oncogenes associated with two main histologic subtypes: non-small cell lung cancer and small cell lung cancer (SCLC) (Lovly, Horn, 2016). The database has identified the total of 23 oncogenes known to be important for lung cancer pathogenesis. We compared this list of genes with DE gene list from RNA Seq data that resulted into 11 common oncogenes. Similarly, comparison has been done between the lists of lung cancer oncogenes with BETA integrative analysis list, we found MAP2K1, and MET oncogenes were common. We also used pulmonary arterial hypertension knowledgebase (PAHKB), which provides 341 human PH-related genes (293 protein coding and 48 non-coding genes) to compare with DE genes and BETA integrative analysis data (Zhao, Austin, Hemnes, Loyd, & Zhao, 2014). We found common 69 genes and 30 genes in DE genes and BETA integrative analysis data respectively. This evidence shows that potential involvement of ID3 in regulation of genes, which are responsible in lung endothelial cells proliferation. This endothelial cell hyper proliferation is disease phenotype for both PAH and lung cancer (additional file 15).

Conclusion

State-of-the-art cis-regulatory sequencing studies have been combined with RNA sequencing to further characterize the functional relevance of TF- interacting DNA regions and thus to define their transcriptional output. In our study, we describe ID3 as a master regulator of key cellular processes for lung cells biology, including cell cycle regulation, DNA repair, replication and metabolism, and cell polarity, and define a large set of genes whose expression is modulated by ID3.

References-

Al-husseini, A. A., Kraskauskas, D., Mezzaroma, E., Nordio, A., Farkas, D., Drake, J. I., … Nordio, A. (2015). Vascular endothelial growth factor receptor 3 signaling contributes to angioobliterative pulmonary hypertension. *Pulmonary Circulation*, *5*(1), 101–116. http://doi.org/10.1086/679704.I

Anders, S., & Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biology*, *11*(10). http://doi.org/10.1186/gb-2010-11-10-r106

Andrew, S. (2010). Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. Retrieved February 1, 2018, from https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Angelini, C., & Costa, V. (2014). Understanding gene regulatory mechanisms by integrating ChIP-Seq and RNA-Seq data: statistical solutions to biological problems. *Frontiers in Cell and Developmental Biology*, *2*, 51. http://doi.org/10.3389/fcell.2014.00051

Babb, R., Huang, C. C., Aufiero, D. J., & Herr, W. (2001). DNA Recognition by the Herpes Simplex Virus Transactivator VP16: a Novel DNA-Binding Structure. *Molecular and Cellular Biology*, *21*(14), 4700–4712. http://doi.org/10.1128/MCB.21.14.4700-4712.2001

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, *57*(1), 289–300. http://doi.org/10.2307/2346101

Blaiseau, P.-L., & Thomas, D. (1998). Multiple transcriptional activation complexes tether the yeast activator Met4 to DNA. *The EMBO Journal*, *17*(21), 6327–6336. http://doi.org/10.1093/emboj/17.21.6327

Boer, J. M., Huber, W. K., Sültmann, H., Wilmer, F., Von Heydebreck, A., Haas, S., … Poustka, A. (2001). Identification and classification of differentially expressed genes in renal cell carcinoma by expression profiling on a global human 31,500-element cDNA array. *Genome Research*, *11*(11), 1861–1870. http://doi.org/10.1101/gr.184501

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. http://doi.org/10.1093/bioinformatics/btu170

Boyer, L. A., Tong, I. L., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., … Young, R. A. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell*, *122*(6), 947–956. http://doi.org/10.1016/j.cell.2005.08.020

Breitling, R., Armengaud, P., Amtmann, A., & Herzyk, P. (2004). Rank products: A simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Letters*, *573*(1–3), 83–92. http://doi.org/10.1016/j.febslet.2004.07.055

Bullard, J. H., Purdom, E., Hansen, K. D., & Dudoit, S. (2009). Evaluation of Statistical Methods for Normalization and Differential Expression in mRNA-Seq Experiments Evaluation of Statistical Methods for Normalization and Differential Expression in mRNA-Seq Experiments. *U.C. Berkeley Div. Biostat. Pap. Ser.*, *11*(1), 94. http://doi.org/10.1186/1471-2105-11-94

Castañon, E., Bosch-Barrera, J., López, I., Collado, V., Moreno, M., López-Picazo, J. M., … Gil-Bazo, I. (2013). Id1 and Id3 co-expression correlates with clinical outcome in stage III-N2 non-small cell lung cancer patients treated with definitive chemoradiotherapy. *Journal of Translational Medicine*, *11*, 13. http://doi.org/10.1186/1479-5876-11-13

Chomczynski P, S. N. (1983). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem*, *162*(June 1989), 1993. http://doi.org/10.1006/abio.1987.9999

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., … Mortazavi, A. (2016). A survey of best practices for RNA-Seq data analysis. *Genome Biology*. http://doi.org/10.1186/s13059-016-0881-8

Das, J. K., & Felty, Q. (2014a). Microvascular Lesions by Estrogen-Induced ID3: Its Implications in Cerebral and Cardiorenal Vascular Disease. *Journal of Molecular Neuroscience*, 618–631. http://doi.org/10.1007/s12031-014-0401-9

Das, J. K., & Felty, Q. (2014b). Microvascular Lesions by Estrogen-Induced ID3: Its Implications in Cerebral and Cardiorenal Vascular Disease. *Journal of Molecular Neuroscience*, *55*(3), 618–631. http://doi.org/10.1007/s12031-014-0401-9

Diaz-Moralli, S., Tarrado-Castellarnau, M., Miranda, A., & Cascante, M. (2013). Targeting cell cycle regulation in cancer therapy. *Pharmacology & Therapeutics*, *138*(2), 255–271. http://doi.org/10.1016/J.PHARMTHERA.2013.01.011

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., … Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-Seq aligner. *Bioinformatics*, *29*(1), 15–21. http://doi.org/10.1093/bioinformatics/bts635

Felty, Q. (2006). Estrogen-induced DNA synthesis in vascular endothelial cells is mediated by ROS signaling. *BMC Cardiovascular Disorders*, *6*, 16. http://doi.org/10.1186/1471-2261-6-16

Felty, Q., & Porther, N. (2008). Estrogen-induced redox sensitive Id3 signaling controls the growth of vascular cells. *Atherosclerosis*, *198*(1), 12–21. http://doi.org/10.1016/j.atherosclerosis.2007.12.048

Felty, Q., Yoo, C., & Kennedy, A. (2010). Gene expression profile of endothelial cells exposed to estrogenic environmental compounds: Implications to pulmonary vascular lesions. *Life Sciences*, *86*(25–26), 919–927. http://doi.org/10.1016/j.lfs.2010.04.007

Ferrer-Vicens, I., Riffo-Campos, Á. L., Zaragozá, R., García, C., López-Rodas, G., Viña, J. R., … García-Trevijano, E. R. (2014). *In vivo* genome-wide binding of Id2 to E2F4 target genes as part of a reversible program in mice liver. *Cellular and Molecular Life Sciences*. http://doi.org/10.1007/s00018-014-1588-1

Fisher Scientific -, T. (). TRIzol Reagent User Guide - Pub. no. MAN0001271 - Rev. A.0. Retrieved from https://tools.thermofisher.com/content/sfs/manuals/trizol_reagent.pdf

Haeger, S. M., Thompson, J. J., Kalra, S., Cleaver, T. G., Merrick, D., Wang, X. J., & Malkoski, S. P. (2016). Smad4 loss promotes lung cancer formation but increases sensitivity to DNA topoisomerase inhibitors. *Oncogene*, *35*(5), 577–586. http://doi.org/10.1038/onc.2015.112

Hansen, K. D., Brenner, S. E., & Dudoit, S. (2010). Biases in Illumina transcriptome sequencing caused by random hexamer priming. *Nucleic Acids Research*, *38*(12), e131–e131. http://doi.org/10.1093/nar/gkq224

Hua, S., Kallen, C. B., Dhar, R., Baquero, M. T., Mason, C. E., Russell, B. A., … White, K. P. (2008). Genomic analysis of estrogen cascade reveals histone variant H2A.Z associated with breast cancer progression. *Molecular Systems Biology*, *4*, 188. http://doi.org/10.1038/msb.2008.25

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: A fast spliced aligner with low memory requirements. *Nature Methods*, *12*(4), 357–360. http://doi.org/10.1038/nmeth.3317

Kuras, L., Cherest, H., Surdin-Kerjan, Y., & Thomas, D. (1996). A heteromeric complex containing the centromere binding factor 1 and two basic leucine zipper factors, Met4 and Met28, mediates the transcription activation of yeast sulfur metabolism. *The EMBO Journal*, *15*(10), 2519–2529. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8665859

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. http://doi.org/10.1038/nmeth.1923

Leprivier, G., Remke, M., Rotblat, B., Dubuc, A., Mateo, A.-R. F., Kool, M., … Sorensen, P. H. (2013). The eEF2 kinase confers resistance to nutrient deprivation by blocking translation elongation. *Cell*, *153*(5), 1064–79. http://doi.org/10.1016/j.cell.2013.04.055

Liao, Y., Smyth, G. K., & Shi, W. (2014). FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30*(7), 923–930. http://doi.org/10.1093/bioinformatics/btt656

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. *Genome Biology*, *15*(12), 550. http://doi.org/10.1186/s13059-014-0550-8

Lovly, C., L. Horn, W. P. (2016). Molecular Profiling of Lung Cancer - My Cancer Genome. Retrieved February 20, 2018, from https://www.mycancergenome.org/content/disease/lung-cancer/

Mayer, M. P., & Bukau, B. (2005). Hsp70 chaperones: Cellular functions and molecular mechanism. *Cellular and Molecular Life Sciences*, *62*(6), 670–684. http://doi.org/10.1007/s00018-004-4464-6

Meyerson, M., Enders, G. H., Wu, C. L., Su, L. K., Gorka, C., Nelson, C., … Tsai, L. H. (1992). A family of human cdc2-related protein kinases. *The EMBO Journal*, *11*(8), 2909–17. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1639063

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods*, *5*(7), 621–628. http://doi.org/10.1038/nmeth.1226

Park, P. J. (2009, October 8). ChIP-Seq: Advantages and challenges of a maturing technology. *Nature Reviews Genetics*. http://doi.org/10.1038/nrg2641

Pathan, M., Keerthikumar, S., Ang, C. S., Gangoda, L., Quek, C. Y. J., Williamson, N. A., … Mathivanan, S. (2015). FunRich: An open access standalone functional enrichment and interaction network analysis tool. *Proteomics*, *15*(15), 2597–2601. http://doi.org/10.1002/pmic.201400515

Pathan, M., Keerthikumar, S., Chisanga, D., Alessandro, R., Ang, C. S., Askenase, P., … Mathivanan, S. (2017). A novel community driven software for functional enrichment analysis of extracellular vesicles data. *Journal of Extracellular Vesicles*, *6*(1). http://doi.org/10.1080/20013078.2017.1321455

Pulmonary Arterial Hypertension KnowledgeBase (PAHKB) Home. (). Retrieved February 22, 2018, from https://bioinfo.uth.edu/PAHKB/index.html

Qian, Y., Du, Z., Xing, Y., Zhou, T., Chen, T., & Shi, M. (2017). Interferon regulatory factor 4 (IRF4) is overexpressed in human non-small cell lung cancer (NSCLC) and activates the Notch signaling pathway. *Molecular Medicine Reports*, *16*(5), 6034–6040. http://doi.org/10.3892/mmr.2017.7319

Ringrose, L. (2017). *Epigenetics and Systems Biology*. Academic Press. Retrieved from https://books.google.com/books

Ringrose, L., Rehmsmeier, M., Dura, J.-M., & Paro, R. (2003). Genome-wide prediction of Polycomb/Trithorax response elements in Drosophila melanogaster. *Developmental Cell*, *5*(5), 759–71. http://doi.org/10.1016/S1534-5807(03)00337-X

Rio, D. C., Ares, M., Hannon, G. J., & Nilsen, T. W. (2010). Purification of RNA Using TRIzol (TRI Reagent). *Cold Spring Harbor Protocols*, *2010*(6), pdb.prot5439-prot5439. http://doi.org/10.1101/pdb.prot5439

Rougemont, J., & Naef, F. (2012). Computational analysis of protein-DNA interactions from ChIP-Seq data. *Methods in Molecular Biology*, *786*, 263–273. http://doi.org/10.1007/978-1-61779-292-2_16

Ryazanov, A. G., Ward, M. D., Mendola, C. E., Pavur, K. S., Dorovkov, M. V, Wiedmann, M., … Hait, W. N. (1997). Identification of a new class of protein kinases represented by eukaryotic elongation factor-2 kinase. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(10), 4884–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9144159

Siggers, T., Duyzend, M. H., Reddy, J., Khan, S., & Bulyk, M. L. (2011). Non-DNA-binding cofactors enhance DNA-binding specificity of a transcriptional regulatory complex. *Molecular Systems Biology*, *7*(1), 555–555. http://doi.org/10.1038/msb.2011.89

Thomas, P. D., Campbell, M. J., Kejariwal, A., Mi, H., Karlak, B., Daverman, R., … Narechania, A. (2003). PANTHER: A library of protein families and subfamilies indexed by function. *Genome Research*, *13*(9), 2129–2141. http://doi.org/10.1101/gr.772403

TruSeq, I. (2013). Illumina: TruSeq Stranded mRNA Sample Preparation Guide. *Manual*, (October), 1–4. http://doi.org/# RS-122-9004DOC

Usadellab. (). Trimmomatic Manual: V0.32. http://doi.org/10.1093/bioinformatics/btu170

Wang, L. Y., Wang, P., Li, M. J., Qin, J., Wang, X., Zhang, M. Q., & Wang, J. (2011). EpiRegNet: Constructing epigenetic regulatory network from high throughput gene expression data for humans. *Epigenetics*, *6*(12), 1505–1512. http://doi.org/10.4161/epi.6.12.18176

Wang, S., Sun, H., Ma, J., Zang, C., Wang, C., Wang, J., … Liu, X. S. (2013). Target analysis by integration of transcriptome and ChIP-Seq data with BETA. *Nature Protocols*, *8 VN-re*(12), 25022515. http://doi.org/10.1038/nprot.2013.150

Wang, X., Wang, Q., Lin, H., Li, S., Sun, L., & Yang, Y. (2013a). HSP72 and gp96 in gastroenterological cancers. *Clinica Chimica Acta*, *417*, 73–79. http://doi.org/10.1016/j.cca.2012.12.017

Wang, X., Wang, Q., Lin, H., Li, S., Sun, L., & Yang, Y. (2013b). HSP72 and gp96 in gastroenterological cancers. *Clinica Chimica Acta*, *417*, 73–79. http://doi.org/10.1016/j.cca.2012.12.017

Wang, Z., Gerstein, M., & Snyder, M. (2009, January 1). RNA-Seq: A revolutionary tool for transcriptomics. *Nature Reviews Genetics*. http://doi.org/10.1038/nrg2484

Wassermann, M., Wassermann, D., Cucos, S., & Miller, H. J. (1979). World PCBs Map: Storage And Effects In Man And His Biologic Environment In The 1970s. *Annals of the New York Academy of Sciences*, *320*(1 Health Effect), 69–124. http://doi.org/10.1111/j.1749-6632.1979.tb56594.x

Zhang, H., Chen, X., Wang, J., Guang, W., Han, W., Zhang, H., … Gu, Y. (). EGR1 decreases the malignancy of human non-small cell lung carcinoma by regulating KRT18 expression. http://doi.org/10.1038/srep05416

Zhao, M., Austin, E. D., Hemnes, A. R., Loyd, J. E., & Zhao, Z. (2014). An evidence-based knowledgebase of pulmonary arterial hypertension to identify genes and pathways relevant to pathogenesis. *Mol. BioSyst.*, *10*(4), 732–740. http://doi.org/10.1039/C3MB70496C

CHAPTER 5

Determination of ID3 targeted genes resulted from the integrative analysis involved

in endothelial cell proliferation

Abstract

PAH is an incurable end-stage lung disease that is often fatal and characterized by the presence of plexiform vascular lesions in pulmonary arterioles. ECs in plexiform vascular lesions (a pathological hallmark of PAH) have been characterized as hyper-proliferative and apoptosis-resistant, which is consistent with a neoplastic hypothesis. Epidemiological studies have reported that women are up to four times as likely to present with severe PAH as men while animal studies have implicated endogenous estrogen as a contributing factor for PVL. The impact of exogenous estrogen or estrogenic PCBs is growing as an environmental risk factor, particularly exposure to PCBs, in the development of PAH that clearly justify the need for further studies to determine the contribution of the environment to PVLs. The integrative analysis of ChIP- and RNA-Seq data from our lab resulted in ID3 specific targeted genes, which are associated with vascular reprogramming. Out of these ID3 target genes, some of the genes have been shown involved in vascular reprogramming and EC proliferation. The objective of the research is to investigate the interaction of ID3 with these genes and elucidate molecular pathogenesis of a highly disordered or neoplastic-like endothelial vascular lesion to more accurately address the clinical reality of short survival and no cure for IPAH patients with PVLs. Our aim is to generate mechanistic data of this interaction in our cell model with ID3 in order to identify target genes involved in promoting the growth of a tumor-like vascular lesion.

Introduction-

The integrative ChIP and RNA Seq data of ID3 targeted gene demonstrated that DNAJC2 and MALAT1 genes were significantly expressed. Therefore, we decided to check the DNAJC2 and MALAT1 genes epigenetically regulate ID3 and PCB-induced vascular lesions *in vitro*. DNAJC2 is a gene in *Homo sapiens* codes for the protein which is an officially named as DnaJ heat shock protein family (Hsp40) member C2 (*Entrez Gene: ZRF1 zuotin related factor 1*, ; Matsumoto-Taniura, Pirollet, Monroe, Gerace, & Westendorf, 1996). DNAJC2 gene is also known as MPHOSPH11, MPP11, MIDA1, and Zuotin-related factor 1, ZRF1 (Gray, Yates, Seal, Wright, & Bruford, 2015). DNAJC2 gene belongs to the family of the M-phase phosphoprotein (MPP), which encodes a phosphoprotein with a J domain and a Myb DNA-binding domain, which localizes to both the nucleus and the cytosol. The protein acts as a molecular chaperone for nascent polypeptide chains as they exit the ribosome (Matsumoto-Taniura et al., 1996). It has been shown that DNAJC2 protein is related to pathways like a cellular response to heat stress and cellular Senescence (Gray et al., 2015). Researchers have also shown the evidence of involvement of DNAJC2 as leukemia-associated antigen and in primary head and neck squamous cell tumors (Demajo et al., 2014; Resto et al., 2000). The DnaJ domain of DNAJC2 protein, which is part of the amino-terminal part of the protein is conserved ubiquitously from yeast to mammals. This amino part mediates interactions with heat shock proteins and is mainly implicated in protein-chaperone functions associated with ribosomes (Yan et al., 1998). In addition to this, DNAJC2 can bind to nucleic acids through its DnaJ domain (Wilhelm, Reinbolt, Gangloff, Dirheimer, & Wilhelm, 1994; Zhang, Lockshin, Herbert, Winter, & Rich, 1992) and has been suggested

to have a role in transcriptional regulation (Eisenman & Craig, 2004; Yoshida, Inoue, Shoji, Ikawa, & Obinata, 2004). Inou et al. have showed that MIDA1 i.e. DNAJC2 protein is associated with ID1 protein which regulates cell growth. They have also shown that DNAJC2 can directly bind to DNA through amino acid sequence contains tryptophan-mediated repeats (Tryp-med repeats) similar to the DNA binding region of the c-Myb oncoprotein (Inoue, Shoji, & Obinata, 2000; Shoji, Obinata, 1999). Recently, researchers have shown that DNAJC2 protein is a regulator of the expression of polycomb targets - polycomb repressive complexes 1 and 2 (PRC1 and PRC2), which play a crucial role in several processes including cell proliferation, and senescence in stem cells and cancer (Sauvageau & Sauvageau, 2010). In humans, the protein structure of DNAJC2 contains an additional C-terminal extension with 2 SANT domains. SANT domains are similar to the Myb-DNA binding motif, (Boyer, Latek, & Peterson, 2004) which is required for Myb proto-oncogene to bind DNA as shown in Figure 95. Indeed, the researchers have shown that DNAJC2 can directly bind to DNA. Researchers have shown the evidence of binding of DNAJC2 to monoubiquitinated histone H2A (H2Aub1) protein through its ubiquitin-binding domain (UBD). In proliferating conditions of cells, polycomb proteins occupy H2Aub1 mark. *In vitro* studies have shown that DNAJC2 displaces PRC1 from chromatin during differentiation stage by competing for the H2Aub1 mark. Once DNAJC2 protein attached to the chromatin by DnaJ binding domain, it induces transcriptional activation of differentiation genes (Richly et al., 2010). In addition to this, DNAJC2 protein is also responsible for self-renewing, the pluripotent state of endothelial stem cells (ESC). Researchers have shown that DNAJC2 protein directly interacts with ID1 protein to maintain a self-renewing state of ESCs. Similar to

H2Aub1 protein, DNAJC2 protein interacts with ID1 through its UBD domain (Aloia, Gutierrez, Caballero, & Di Croce, 2015; Ruzinova & Benezra, 2003). ID1 protein regulates the cell differentiation through DNAJC2 protein since ID1 protein belongs to the family of ID proteins- ID1, ID2, ID3, and ID4. ID proteins are helix-loop-helix (HLH) proteins also known as inhibitors of differentiation function as dominant negative regulators of basic HLH (bHLH) transcriptional regulators that drive cell lineage commitment and differentiation of cells. Earlier researchers and our lab have shown that ID3 regulates cell differentiation, stemness to the cells and angiogenesis (Das et al., 2015; Morisaki et al., 2016). With this background information, we checked whether proteins ID3 and DNAJC2 also interact with each other.
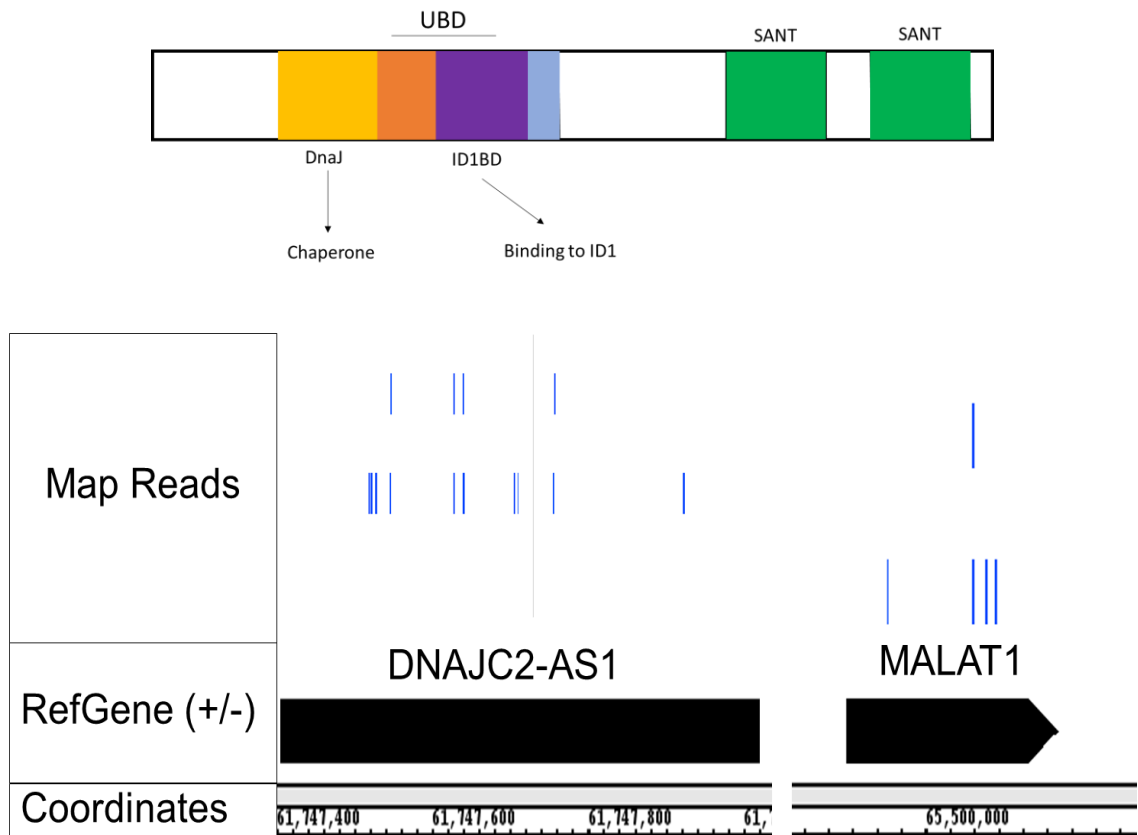
Figure 95- Representative Map Reads from ID3 ChIP-Seq shows immunoprecipitated chromatin containing DNA regions corresponding to genes DNAJC2-AS1 and MALAT1

226

Experiment –

Western Blot-

In order to determine the interaction between ID3 and DNAJC2 protein, we first checked the expression of DNAJC2 protein in HPMEC ID3$^+$ cell line with control, E2, PCB153 and co-treatment E2 +PCB153 treatments. We seeded the HPMEC ID3$^+$ cells in four 100mm petri dishes each. We added 0.1% DMSO (control), E2 (100 ng/mL), PCB153 (100 ng/mL) and combination treatment of E2 +PCB153 to separate petri dishes. We analyzed that combination treatment of E2 +PCB153 increase 2-3-fold change in DNAJC2 protein expression in HPMEC ID3$^+$ cells compared to control. E2 (100 ng/mL), PCB153 (100 ng/mL) treatment showed significant (P<0.01) increased than control as shown in Figure 96.
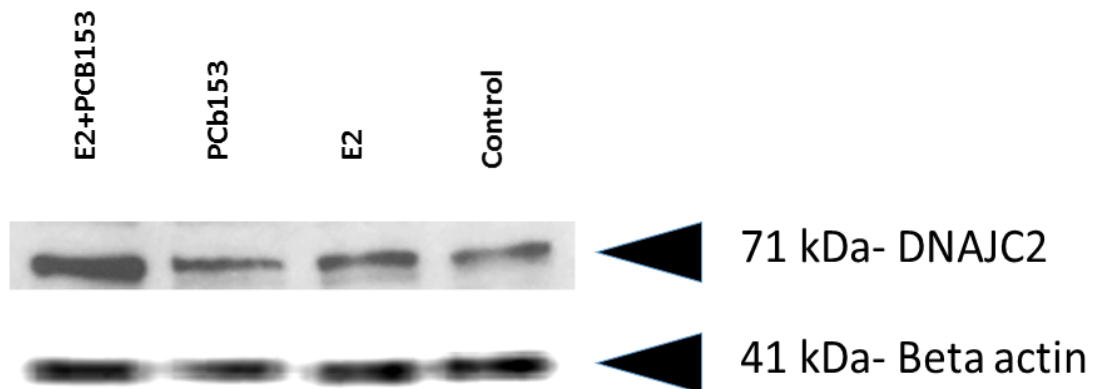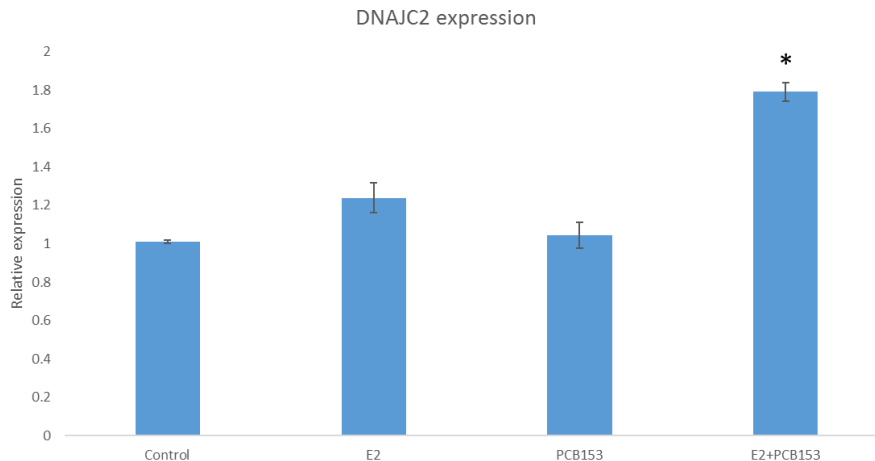
Figure 96- Western Blot analysis of DNAJC2 protein

Detection of DNAJC2 protein 71 KDa and beta actin as a control 41 KDa in Lung EC ID3$^+$ cell line. Gel percentage 7.5%. Dilution of DNAJC2 antibody (1: 2000) i.e. 5 µL in 10 mL of blocking buffer- Primary antibody is rabbit Ab (polyclonal). Incubation time overnight at 4 ºC for Primary antibody. Secondary Ab is anti-rabbit Ab. Dilution of Secondary antibody (1: 50000). E2 & PCB153 = 100ng/mL

Immunoprecipitation
ID3 co-immunoprecipitation of DNAJC2-

Background- *In vivo*, the mouse homolog to DNAJC2 known as MIDA1 (mouse Id associated 1) has been shown to bind to ID1 and positively regulate cell growth. Hence, our aim is to generate mechanistic data of this interaction in our cell model with ID3 in order to identify target genes involved in promoting the growth of a tumor-like vascular lesion. We postulate that ID3 is a novel co-activator of DNAJC2 and ID3-DNAJC2 protein-protein interaction is critical in driving the growth of tumor-like vascular lesions.

Procedure- For immunoblots of immunoprecipitated (IP) lysate the following method was used. ID3 protein was pulled down using magnetic beads (Dynabeads protein G). In

228

Brief, the lysate was incubated with 10 mg of antibody and 50 mL (1.5 mg) of Dynabeads (Invitrogen) in microcentrifuge tubes with rotation for 10 minutes at room temperature. Samples were then placed on a magnetic particle concentrator (DynaMagTM-2, Invitrogen) and the supernatant was discarded. Dynabeads–Ab–Ag complex was washed 3 times with washing buffer and tubes were placed on magnetic particle concentrator to remove the supernatant. The magnetic Dynabeads–Ab–Ag complex was suspended with 20 mL of Elution Buffer and boiled in SDS sample buffer for 5 min. Samples were placed on the magnetic particle concentrator and the supernatant containing protein were separated by SDS-PAGE and immunoblotted as described above as shown in Figure 97.

Experimental Design: ID3 and DNAJC2 cross-linked to DNA (chromatin) will be measured by Western blot using lung EC ID3[+] and/or lung EC WT.

Using approved ChIP protocol:

Step 1: Fix cells to cross-link DNA to protein.

Step 2: Sonicate to lyse cells and shear chromatin. Note: run a DNA gel to determine optimum conditions for shearing DNA.

Step 3: Immunoprecipitated DNA bound by ID3 and/or DNAJC2.

Experiment 1 (ID3 bound DNA)

Sample 1: IP chromatin cross-linked proteins with polyclonal ID3 Ab.

Sample 2: IP chromatin cross-linked proteins with IgG Ab (negative control).

Sample 3: Whole cell lysate (Lung EC ID3[+] positive control)

We loaded the immunoprecipitated lysate in SDS-PAGE and detected ID3 using antibody. We ran three samples: (i) ID3 IP lysate; (ii) IgG IP lysate (negative control); and (iii) whole cell lysate 100 µg (positive control).

Experiment 2 (DNAJC2 bound DNA)

Sample 1: IP chromatin cross-linked proteins with polyclonal DNAJC2 Ab

Sample 2: IP chromatin cross-linked proteins with IgG Ab (negative control).

Sample 3: Whole cell lysate (Lung EC ID3$^+$ positive control)

We loaded the immunoprecipitated lysate in SDS-PAGE and detected DNAJC2 using antibody. We ran three samples: (i) DNAJC2 IP lysate (ii) IgG IP lysate (negative control) and (iii) whole cell lysate 100 µg (positive control).

Procedure-

Experiment-1 - Dilution of ID3 antibody (1: 1000) i.e 10 µL in 10 mL of blocking buffer-BSA (0.5g +10 mL TBST). Primary antibody is rabbit Ab (polyclonal). Incubation time overnight at 4 degrees Celsius for Primary antibody. Secondary Ab is anti-rabbit Ab. Dilution of Secondary antibody (1: 5000) i.e. 2 µL in 10 mL of blocking buffer. Molecular weight of ID3 is 17 KDa. Separating Gel percentage- 15% DNAJC2 (abcam) ab134572 Rabbit polyclonal used for IP. ID3 (CalBioReagent) monoclonal used for western blot.

Experiment-2 - Dilution of DNAJC2 antibody (1: 2000) i.e 5 µL in 10 mL of blocking buffer. Primary antibody is rabbit Ab (polyclonal). Incubation time overnight at 4 ºC for Primary antibody. Secondary Ab is anti-rabbit Ab. Dilution of Secondary antibody (1:

50000). Molecular weight of DNAJC2 is 71 KDa. Separating Gel percentage- 7.5%. DNAJC2 (abcam) ab134572 Rabbit polyclonal used for western blot. ID3 (Santacruz Biotechnology sc- 490) monoclonal antibody.

Results-
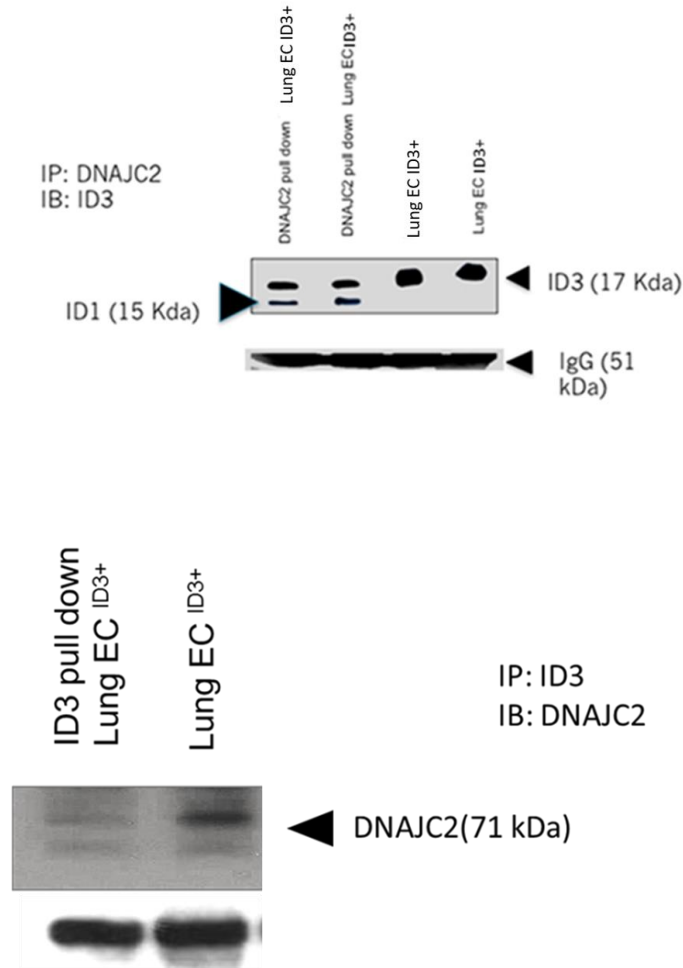
Figure 97- Immunoprecipitation of ID3 and DNAJC2

ID3 protein was isolated by immunoprecipitation (IP) technique using magnetic beads and detected by immunoblot. ID3 protein of immunoprecipitated DNAJC2 protein was determined with anti-ID3 antibody. ID3 bands detected from immunoprecipitated cell lysate corresponded with the molecular weight of 17 KDa for ID3. Electrochemiluminescence (ECL) band intensity for phosphorylated ID3 was imaged with a Bio-Rad Versa Doc instrument. Similarly, DNAJC2 protein detected from
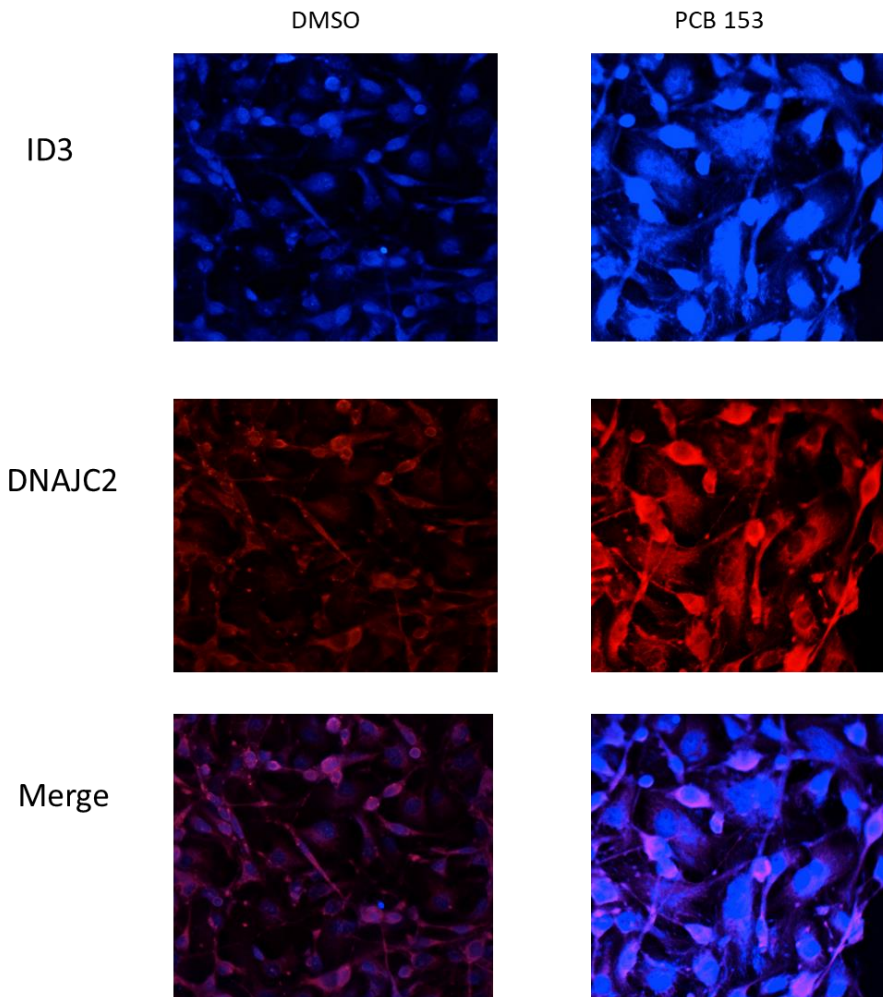
Figure 98- Co-localization of ID3 and DNAJC2

Confocal microscopy of the co-localization of ID3 (blue) with DNAJC2 (Red) in lung EC ID3$^+$ cells. Cells were treated with DMSO (Left) or PCB153 (0.3 µM) (Right) for 24h. ID3 was determined by immunofluorescent labelling with rabbit anti-ID3 polyclonal antibody (Alexa Fluor 566-conjugated secondary antibody). DNAJC2 was determined by immunofluorescent labelling with mouse monoclonal IgG (Alexa Fluor 633-conjugated secondary antibody).

Immunofluorescence and immunoprecipitation experiments showed that ID3 and DNAJC2 interact with each other as shown in Figure 97 and 98.

CRISPR/Cas9 experiment

Hypothesis: The CRISPRCas9 MALAT1/DNAJC2 (ZRF1) epigenetically regulate the phenotypic expression like spheroid size and numbers in lung EC ID3[+].

Background: Our aim is to generate mechanistic data of this interaction in our cell model with ID3 in order to identify target genes involved in promoting the growth of a tumor-like vascular lesion. We postulated that ID3 is a novel co-activator of DNAJC2 and ID3-DNAJC2 protein-protein interaction is critical in driving the growth of tumor-like vascular lesions. In order to determine the interaction between ID3 and DNAJC2 proteins, we did a co-immunoprecipitation experiment in two ways. In the first experiment, we used an ID3 polyclonal antibody to pull own ID3 bound proteins from Lung EC ID3[+] protein sample and then immunoblotted with the monoclonal DNAJC2 antibody. We found the expression of DNAJC2 protein at 71 KDa, which is confirmed by using standard protein ladder for molecular weight determination on western blots. In another similar experiment, we used the DNAJC2 antibody for immunoprecipitation from the sample and then immunoblotted with the monoclonal ID3 antibody. We found the expression of ID3 protein at 17 KDa, which is confirmed by using standard protein ladder for molecular weight determination on western blots. Moreover, we did mRNA sequencing experiment with lung EC ID3[+] cells and analyzed the data. We found out that Metastasis Associated Lung Adenocarcinoma Transcript1 (MALAT1) mRNA was expressed higher folds in lung EC ID3[+] cells compared to the control wild type HPMEC cells.

Therefore, to check whether MALAT1 and DNAJC2 proteins co-regulate with ID3 protein, we decided to silence the expression of MALAT1 and DNAJC2 proteins and study the effect of silencing on the stem like property of lung EC ID3$^+$ cells.

MALAT1 gene is non-protein coding gene which is also known as Nuclear Paraspeckle Assembly Transcript 2 (Ji, Diederichs, Wang, Böing, et al., 2003; Kryger, Fan, Wilce, & Jaquet, 2012; Zhang et al., 2017). Noncoding RNAs (ncRNAs) are defined as untranslated regulatory RNA molecules. In humans, at least 98% of the total genome does not get translated into protein. Only 2% of total DNA sequences responsible for non-protein coding regions (Yin, Hamblin, & Chen, 2014). So, the non-protein coding part of genome consists of Noncoding RNAs which include microRNAs, and long noncoding RNAs (Bartel & Chen, 2004). *MALAT1* gene produces non-long coding RNA. In addition to this, More than 8000bp sequence present in MALAT1 gene is highly conserved in 33 mammalian species and the highest expression of MALAT1 mRNA is found to be in pancreas and lung (Chen, Li, Lodish, & Bartel, 2004; Ji, Diederichs, Wang, Boïng, et al., 2003). *MALAT1* gene transcribes into a precursor transcript. This transcript resides in the nucleus, which is thought to form molecular scaffolds for ribonucleoprotein complexes. Then, RNase P cleaves of a tRNA-like small ncRNA (known as mascRNA) from the 3' end of transcript and extract a long non-coding RNA (Ma et al., 2015; Rinn & Chang, 2012). Primarily, *MALAT1* gene is identified by subtractive hybridization in stage I NSCLC (non-small cell lung cancer) tumors (Hutchinson et al., 2007; Ji, Diederichs, Wang, Boïng, et al., 2003).

At the beginning, many researchers have been originally identified lncRNA MALAT1 as a prognostic marker for non-small cell lung cancer, in the cerebrovascular pathogenesis of ischemic stroke. Researchers have done several experiments to prove the elevated level of lncRNA MALAT1 responsible for NSCLC. In one of the experiments, scientist tested increasing levels of MALAT1 expression in cultured NSCLC cell lines (A549 and HTB58) via qPCR analysis (Ma et al., 2015). Gutschner et al. 2013 showed *in vivo* and *in vitro* experiments that NSCLC cell lines transfected with MALAT1- shRNA proliferate (Gutschner et al., 2013). Moreover, another group of researchers in the literature showed that inhibition of MALAT1 forces cells to switch its gears from anti-proliferative to impaired endothelial cell proliferation *in vitro* and *in vivo* and reduced retinal vessel growth. They have also showed that when siRNA transfected for MALAT1 treated with HUVECs reduces the expression of MALAT1 induces angiogenic sprouting and migration but represses proliferation. These researchers have also found out that when HUVECs exposed to hypoxia environment, the expression MALAT1 increased substantially (Michalik et al., 2014). Similarly, when we treated with lung EC ID3[+] cells to SU5416 which is a selective synthetic inhibitor of the Flk-1/KDR vascular endothelial growth factor (VEGF) receptor tyrosine kinase in presence of hypoxia environment in our lab, we found the expression of stem cell markers Oct-4, Sox-2, and Nanog by cell cytometry and immunohistochemistry analysis. At the same time, surprisingly we found the elevated expression of VEGFR3 markers on lung EC ID3[+] cells (Doke & Felty, 2016). We proposed these stem likes cells, which are pluripotent cell that gives rise to a clonal pulmonary vascular lesion that supports the aggressive growth of NSCLC.

Looking at the close relation of lncRNA MALAT1 and ID3 protein to metastatic growth of NSCLC, therefore, we studied the effect of lncRNA MALAT1 on the lung EC ID3$^+$.

Experiments-

Spheroid Assays-

Question: Does the long noncoding RNA (lncRNA) MALAT1 and DNAJC2 regulate the growth of ID3 expressing lung endothelial spheroids (lung EC ID3$^+$)?

Purpose: Small noncoding RNAs (i.e. microRNAs), >200-nucleotide long transcripts, namely lncRNAs, can interfere with gene expressions and signaling pathways at various stages. ID3 may be epigenetically regulated by the long-noncoding RNA MALAT1 based on the following evidence: MALAT1 expression is augmented by hypoxia in endothelial cells (a contributing factor in the rodent model of PAH) and MALAT1 supports the proliferative response of endothelial cells both phenomena are correlated with ID3. Thus, the purpose of this experiment is to determine whether lncRNA MALAT1 regulates the formation of ID3 expressing endothelial spheroids. It is unclear whether endothelial- or smooth muscle-based lesions have a similar impact on the course of PAH.

Experimental Design:

We followed the protocol for endothelial spheroid and 3-D cell culture as described in Das JK et al. Microvascular Research (2015). In brief, cells were suspended in serum-free stem cell medium containing DMEM/F12 (1:1) supplemented with B27®. For EC spheroid generation, approximately 100-150 cells per well were seeded in an ultra-low attachment 96-well plate (Corning Inc, Lowell, MA).

2. Experiment 2 Design:

- Transfect (i) Lung EC ID3$^+$ and (ii) Lung EC vector with MALAT1 sgRNA CRISPR/Cas9 All-in-One Lentivector set (Human) Applied Biological Materials (ABM) K1258505.

- Transfect (i) Lung EC ID3$^+$ and (ii) Lung EC vector with DNAJC2 sgRNA CRISPR/Cas9 All-in-One Lentivector set (Human) Applied Biological Materials (ABM) K0614705.

- Transfect (i) Lung EC ID3$^+$ and (ii) Lung EC vector with Scrambled sgRNA CRISPR/Cas9.

After 24-48 h transfection, seed cells in triplicate (100-150 cells per well) in ultra-low attachment 96-well plates as described previously and recorded phenotype changes.

Results: We found that knockdown of lncRNA MALAT1 resulted in the following phenotype changes- decreased endothelial spheroid number, decreased endothelial spheroid size when compared to the control endothelial spheroids transfected with Scrambled sgRNA CRISPR/Cas9. The nuclear transcript MALAT1 has been functionally associated with gene regulation and alternative splicing and its regulation has been shown to impact proliferation, apoptosis, migration and invasion. Therefore, these results showed ID3 may be epigenetically regulated by the long-noncoding RNA MALAT1 (metastasis-associated lung adenocarcinoma transcript 1) as shown in Figure 99, 100 and 101.
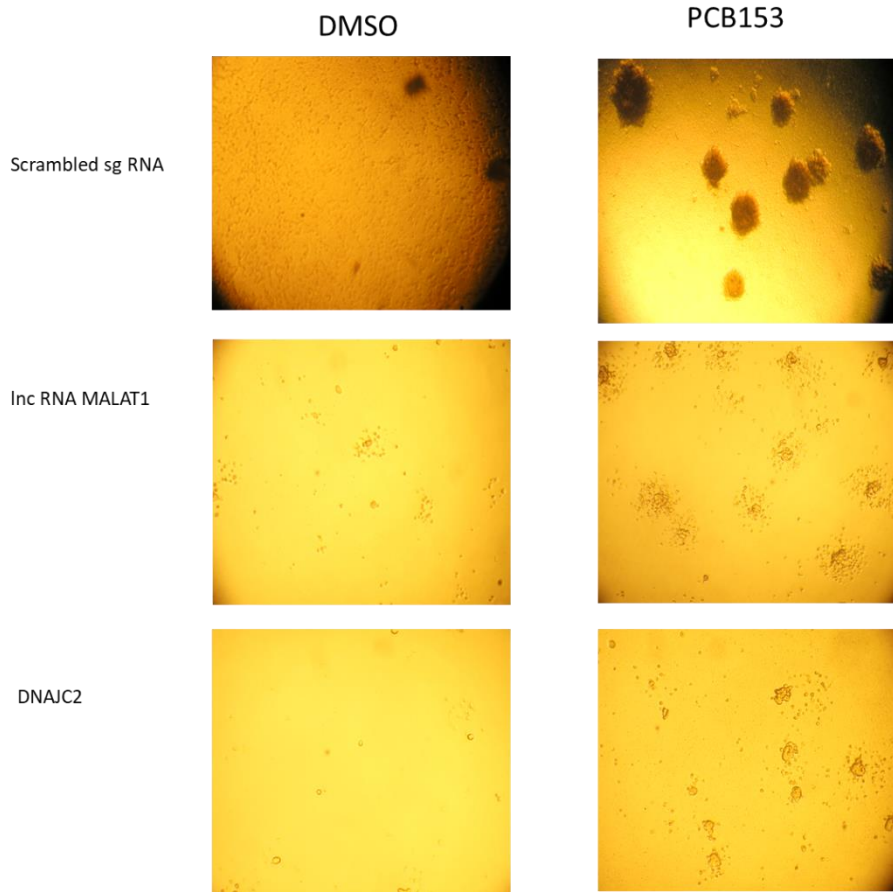
Figure 99- It shows pictures of spheroids with CRISPR/CAS9 MALAT1 /ZRF1 treatment.

CRISPR/CAS9 MALAT1 /ZRF1 treatment decreases spheroid size and number in the PCB153 treated wells. The decrese in the size and number of spheroid in CRISPR/CAS9 MALAT1 /ZRF1 treated wells are significantly lower compared to scrambled sgRNA control. Scale bar= 10µm.

Spheroid number Lung EC ID3 at 10 day
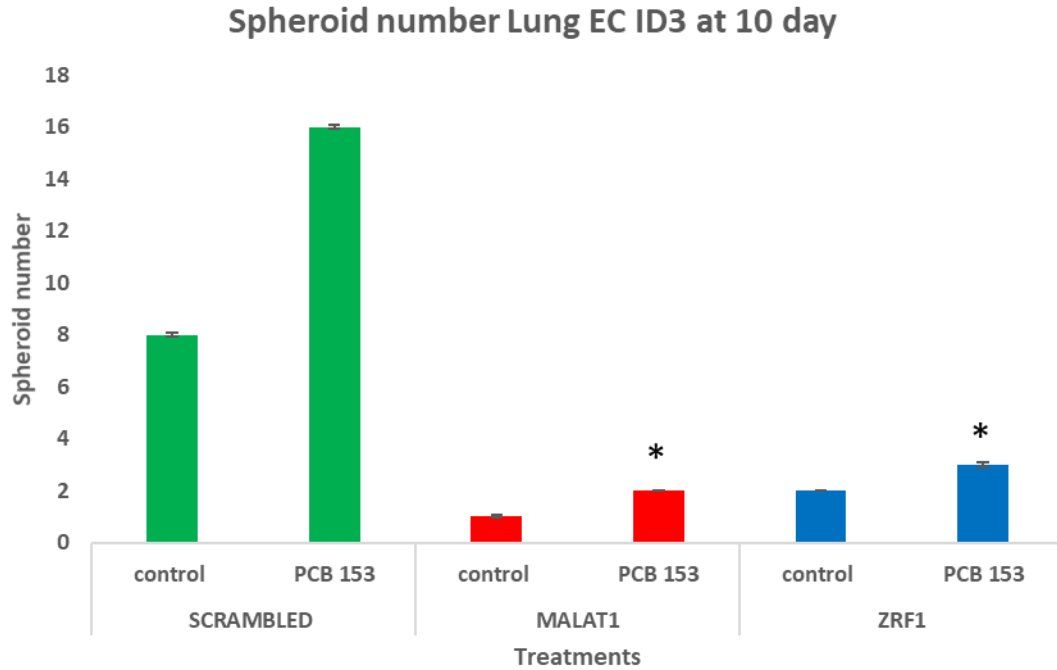
Figure 100- CRISPR/CAS9 MALAT1 /ZRF1 knockdown reduced spheroid number

CRISPR/CAS9 MALAT1 /ZRF1 knockdown ID3 expression in lung EC ID3[+] and result in a decrease in spheroid number compared to spheroid with scrambled sgRNA CRISPR/Cas9. (*)PCB153 treatment in CRISPR/CAS9 MALAT1 /ZRF1 significantly different from PCB153 treatment in scrambled sgRNA (P<0.01) at day 10. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.
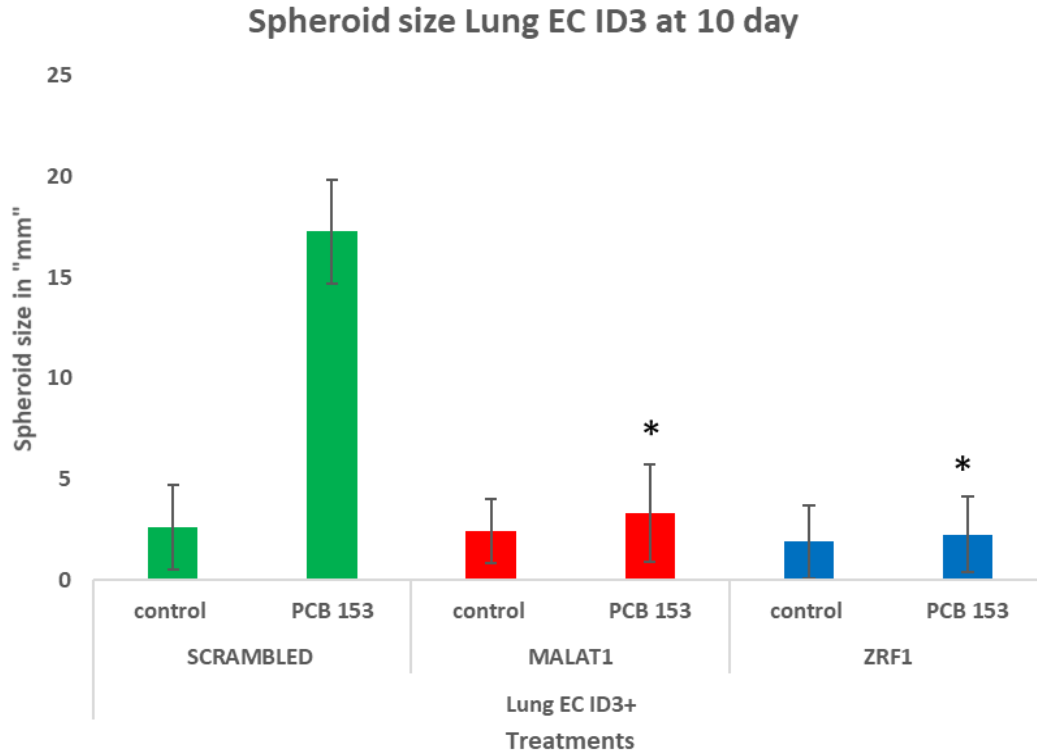
Figure 101- CRISPR/CAS9 MALAT1 /ZRF1 knockdown reduced spheroid size

CRISPR/CAS9 MALAT1 /ZRF1 knockdown ID3 expression in lung EC ID3$^+$ and result in a decrease in spheroid diameter compared to spheroid with scrambled sgRNA CRISPR/Cas9. (*)PCB153 treatment in CRISPR/CAS9 MALAT1 /ZRF1 significantly different from PCB153 treatment in scrambled sgRNA (P<0.01) at day 10 Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.

PCR experiment

Purpose: To investigate how non-coding RNAs regulate ID3 when exposed to PCBs.

RNA Isolation and Quantitative real-time PCR (qRT-PCR)

Total RNA was isolated and purified using the RNeasy Mini Kit from Qiagen (Valencia, CA). RNA sample quality was verified by electrophoresis on a denaturing agarose gel and then reverse transcribed into cDNA using the RT² First Strand Kit from SuperArray Bioscience Corporation (Frederick, MD) according to the manufacturer's protocol. The PCR reactions using cDNA were performed in an Applied Biosystems 7300 Real-Time PCR System using RT2 SYBR Green/ROX qPCR Master Mix and the manufacturer's thermal cycler protocol (95 °C for 10 min, followed by 40 cycles at 95 °C for 15 s, 60 °C for 1 min). The threshold cycle (CT) from each well was determined using ABI 7300 SDS software.

MALAT1-set siRNA/shRNA/RNAi Lentivector (Human)

The siRNA specifically targeting MALAT1 was constructed by abm Company (USA), HPMEC and HPMEC ID3⁺ cells were transfected with siRNA targeting MALAT1 using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. Cells transfected with Scramble was used as the negative control, the Scramble siRNA sequence was as follows: sense, AACAUCUCUUGGGCUUUUAUU, antisense, GAGAAUUAAGUAAAUGCCCUU. Cells were collected after transfection for RNA extraction

siRNA Vectors Workflow

Viral and non-viral mediated gene delivery using abm products
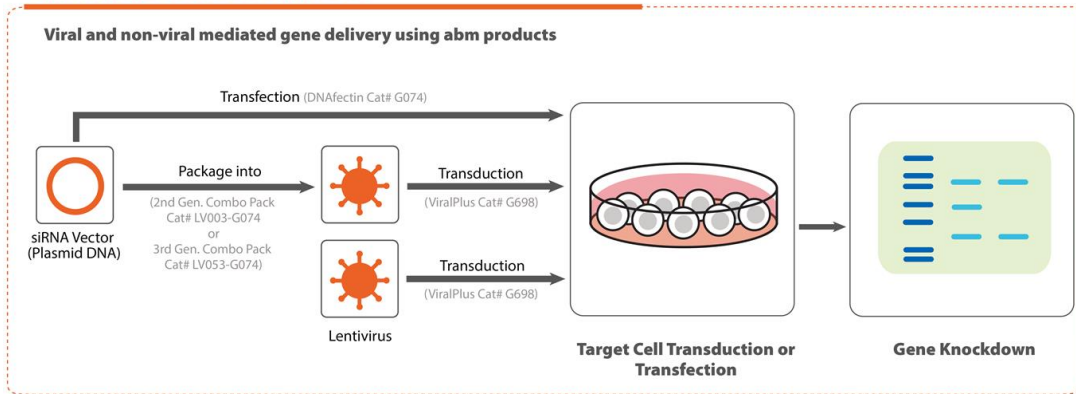
Target Sequence   1762: GACCTTGAAATCCATGACG

5374: TCAGAAGAGTTGCTTCATTTCATCTGGGA

5427: ACAGATAAGTTTAACTTGCATCTGCAGTA

5605: TGGGCTTCAGTGATGGGATAGTACACTTC

Sequencing Primers   TACGTCCAAGGTCGGGCAGGAAGA

Mammalian Selection - Puromycin. Direct non-viral plasmid transfection for immediate expression. Package into Lentiviral particles for high efficiency transduction and stably integrated expression. RNA interference lentiviral vectors contain siRNAs. We employed a dual convergent promoter system where the sense and antisense strands of the siRNA are expressed by two different promoters rather than in a hairpin loop to avoid any possible recombination events that can occur.

LncRNA MALAT1 polymerase chain reaction (PCR) array

PCR analysis of cDNA from RNA isolated from HPMEC ID3$^+$ treated with vehicle control (0.1% DMSO), PCB153, or 17β-estradiol for 24h was performed in triplicate

using RT2. RT² lncRNA qPCR Assay for Human MALAT1 (LPH18065A) and RT² lncRNA qPCR Assay for Human HIF1A-AS2 (LPH00171A) from Qiagen Inc. were used to compare the relative expression of genes. Relative changes in gene expression between the treated and untreated HPMEC ID3$^+$ RNA were calculated by the ΔΔCT method as previously described (Livak & Schmittgen, 2001). First, target genes in were normalized to the endogenous control or housekeeping genes: RPL13A, GAPDH, or ACTB. Next, each treatment group was normalized to the untreated group (vehicle control, 0.1% DMSO). Using the SA Biosciences PCR Array Data Analysis Web Portal, fold changes in gene expression were analyzed between groups using the equation-

ΔΔCT= [CT target gene−Avg.CT housekeeping genes] treated group−CT target gene− Avg. CT housekeeping genes] control group. Each value is a repeat of three independent experiments. ANOVA and Fisher's least significant difference test were used for multi-group comparisons, with a P value of less than 0.05 considered significant.

Results and Conclusion-

To validate our results from RNA Seq studies, we did RT-qPCR analysis. We validated the expression of lncRNA MALAT1 in HPMEC ID3$^+$ cells. When we treated the cells with E2 and PCB153, the MALAT1 showed 4-5 fold higher expression compared to control as shown in Figure 102. These results signify the interaction between ID3 overexpression and MALAT1 is significant. Researchers have already shown a critical role for MALAT1 in NSCLC cancer metastasis, which is the major cause of cancer recurrence and tumor-related death. Therefore, we hypothesized that MALAT1 and ID3

might be a potential therapeutic target to inhibit endothelial cell proliferation due to exposure to PCB.
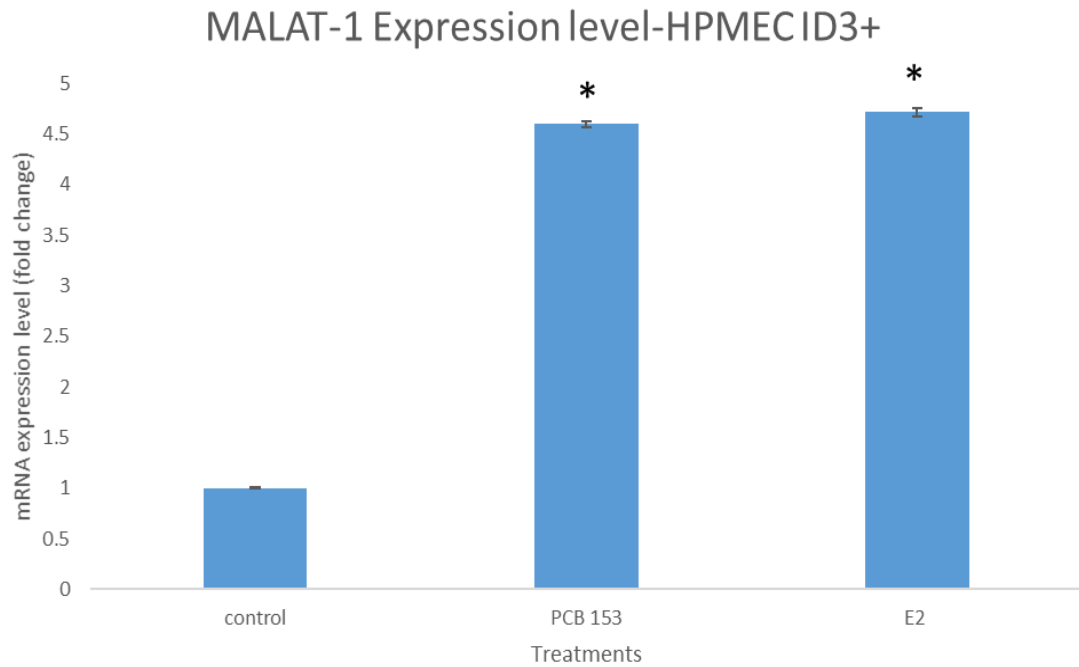
Figure 102- MALAT1 expression in HPMEC ID3+ cells.

Upregulation of lncRNA MALAT1 mRNA expression when treated with E2 and PCB153. HPMEC ID3[+] cells were treated with either vehicle or E2 or PCB153 for 24 h, and the mRNA expression of MALAT1 was validated by using RT-qPCR. Data are represented as mean ± SE. * p<0.01.
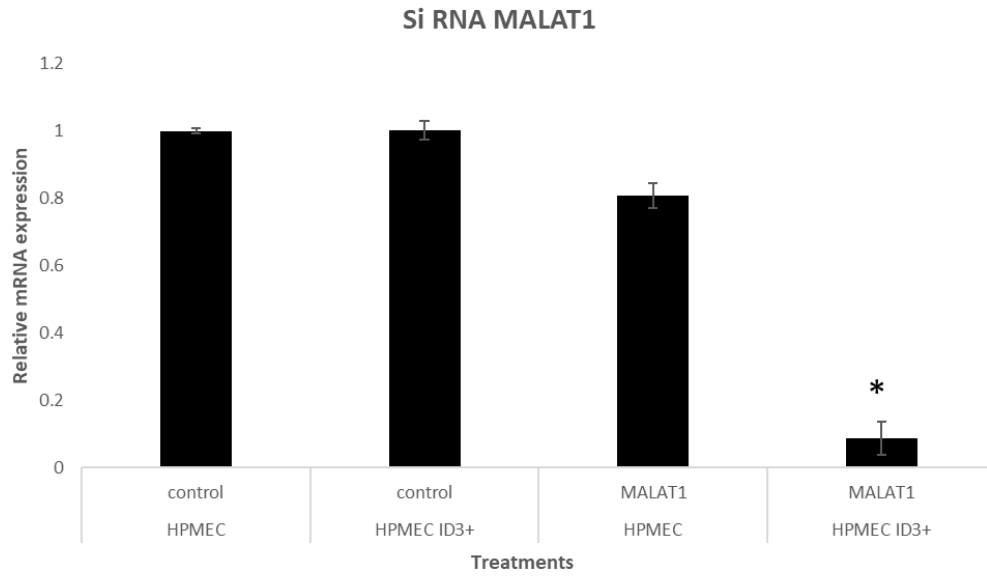
Figure 103- Knockdown of MALAT1 in HPMEC ID3⁺ cells measured by qRT-PCR

HPMEC and HPMEC ID3⁺ cells were transfected with siRNA targeting MALAT1, QRT-PCR assay was employed to detect the endogenous expression of MALAT1 to confirm the knockdown efficiency. The HPMEC cells transfected with Scramble was used as the negative control.*P < 0.01 vs. the control.

References-

Aloia, L., Gutierrez, A., Caballero, J. M., & Di Croce, L. (2015). Direct interaction between Id1 and Zrf1 controls neural differentiation of embryonic stem cells. *EMBO Reports*, *16*, 63–70. http://doi.org/10.15252/embr

Bartel, D. P., & Chen, C.-Z. (2004). Opinion: Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. *Nature Reviews Genetics*, *5*(5), 396–400. http://doi.org/10.1038/nrg1328

Boyer, L. A., Latek, R. R., & Peterson, C. L. (2004). Opinion: The SANT domain: a unique histone-tail-binding module? *Nature Reviews Molecular Cell Biology*, *5*(2), 158–163. http://doi.org/10.1038/nrm1314

Chen, C., Li, L., Lodish, H. F., & Bartel, D. P. (2004). Lineage Differentiation. *Science*, *303*(5654), 83–6. http://doi.org/10.1126/science.1091903

Das, J. K., Voelkel, N. F., & Felty, Q. (2015). ID3 contributes to the acquisition of molecular stem cell-like signature in microvascular endothelial cells: Its implication for understanding microvascular diseases. *Microvascular Research*, *98*, 126–138. http://doi.org/10.1016/j.mvr.2015.01.006

Doke, M., & Felty, Q. H. (2016). Abstract 3274: VEGF─receptor antagonist, Sugen 5416, sensitizes pulmonary endothelial stem-like cells to estrogens: A microvascular model for the progression of lung cancer. *Cancer Research*, *76*(14 Supplement). Retrieved from http://cancerres.aacrjournals.org/content/76/14_Supplement/3274

Eisenman, H. C., & Craig, E. A. (2004). Activation of pleiotropic drug resistance by the J-protein and Hsp70-related proteins, Zuo1 and Ssz1. *Molecular Microbiology*, *53*(1), 335–344. http://doi.org/10.1111/j.1365-2958.2004.04134.x

Entrez Gene: ZRF1 zuotin related factor 1. (). Retrieved from http://www.ncbi.nlm.nih.gov/sites/entrez?Db=gene&Cmd=ShowDetailView&TermToSearch=27000

Gray, K. A., Yates, B., Seal, R. L., Wright, M. W., & Bruford, E. A. (2015). Genenames.org: the HGNC resources in 2015. *Nucleic Acids Research*, *43*(Database issue), D1079-85. http://doi.org/10.1093/nar/gku1071

Gutschner, T., Hämmerle, M., Eißmann, M., Hsu, J., Kim, Y., Hung, G., … Diederichs, S. (2013). The non-coding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Res. February*, *1*(733), 1180–1189. http://doi.org/10.1158/0008-5472.CAN-12-2850

Hutchinson, J. N., Ensminger, A. W., Clemson, C. M., Lynch, C. R., Lawrence, J. B., & Chess, A. (2007). A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics*, *8*(1), 39. http://doi.org/10.1186/1471-2164-8-39

Ji, P., Diederichs, S., Wang, W., Boïng, S., Metzger, R., Schneider, P. M., … Muïler-Tidow, C. (2003). MALAT-1, a novel noncoding RNA, and thymosin b4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene*, *22*, 8031–8041. http://doi.org/10.1038/sj.onc.1206928

Ji, P., Diederichs, S., Wang, W., Böing, S., Metzger, R., Schneider, P. M., … Müller-Tidow, C. (2003). MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene*, *22*(39), 8031–41. http://doi.org/10.1038/sj.onc.1206928

Kryger, R., Fan, L., Wilce, P. A., & Jaquet, V. (2012). MALAT-1, a non protein-coding RNA is upregulated in the cerebellum, hippocampus and brain stem of human alcoholics. *Alcohol (Fayetteville, N.Y.)*, *46*(7), 629–34. http://doi.org/10.1016/j.alcohol.2012.04.002

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. *Methods*, *25*(4), 402–408. http://doi.org/10.1006/meth.2001.1262

Ma, X. Y., Wang, J. H., Wang, J. L., Ma, C. X., Wang, X. C., & Liu, F. S. (2015). Malat1 as an evolutionarily conserved lncRNA, plays a positive role in regulating proliferation and maintaining undifferentiated status of early-stage hematopoietic cells. *BMC Genomics*, *16*, 676. http://doi.org/10.1186/s12864-015-1881-x

Matsumoto-Taniura, N., Pirollet, F., Monroe, R., Gerace, L., & Westendorf, J. M. (1996). Identification of Novel M Phase Phosphoproteins by Expression Cloning. *Molecular Biology of the Cell Nurse*, *7*, 1455–1469. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC275994/pdf/mbc00016-0138.pdf

Michalik, K. M., You, X., Manavski, Y., Doddaballapur, A., Zörnig, M., Braun, T., … Dimmeler, S. (2014). Long noncoding RNA MALAT1 regulates endothelial cell function and vessel growth. *Circulation Research*, *114*(9), 1389–1397. http://doi.org/10.1161/CIRCRESAHA.114.303265

Morisaki, Y., Niikura, M., Watanabe, M., Onishi, K., Tanabe, S., Moriwaki, Y., … Misawa, H. (2016). Selective Expression of Osteopontin in ALS-resistant Motor Neurons is a Critical Determinant of Late Phase Neurodegeneration Mediated by Matrix Metalloproteinase-9. *Scientific Reports*, *6*(January), 27354. http://doi.org/10.1038/srep27354

Richly, H., Rocha-Viegas, L., Ribeiro, J. D., Demajo, S., Gundem, G., Lopez-Bigas, N., … Di Croce, L. (2010). Transcriptional activation of polycomb-repressed genes by ZRF1. *Nature*, *468*(7327), 1124–1128. http://doi.org/10.1038/nature09574

Rinn, J. L., & Chang, H. Y. (2012). Genome Regulation by Long Noncoding RNAs. *Annual Review of Biochemistry*, *81*(1), 145–166. http://doi.org/10.1146/annurev-biochem-051410-092902

Ruzinova, M. B., & Benezra, R. (2003). Id proteins in development, cell cycle and cancer. *Trends in Cell Biology*, *13*(8), 410–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12888293

Sauvageau, M., & Sauvageau, G. (2010). Polycomb Group Proteins: Multi-Faceted Regulators of Somatic Stem Cells and Cancer. *Cell Stem Cell*, *7*(3), 299–313. http://doi.org/10.1016/j.stem.2010.08.002

Wilhelm, M. L., Reinbolt, J., Gangloff, J., Dirheimer, G., & Wilhelm, F. X. (1994). Transfer RNA binding protein in the nucleus of Saccharomyces cerevisiae. *FEBS Letters*, *349*, 260–264. Retrieved from http://ac.els-cdn.com/0014579394006830/1-s2.0-0014579394006830-main.pdf?_tid=b4a7df8a-6fe0-11e7-803e-00000aacb35e&acdnat=1500839836_dfa5266605a2cda5b7e0558f94ab383a

Yan, W., Schilke, B., Pfund, C., Walter, W., Kim, S., & Craig, E. A. (1998). Zuotin, a ribosome-associated DnaJmolecular chaperone. *The EMBO Journal*, *17*(16), 4809–4817. http://doi.org/10.1093/emboj/17.16.4809

Yin, K.-J., Hamblin, M., & Chen, Y. E. (2014). Non-coding RNAs in cerebral endothelial pathophysiology: Emerging roles in stroke. *Neurochemistry International*, *77*, 9–16. http://doi.org/10.1016/j.neuint.2014.03.013

Yoshida, M., Inoue, T., Shoji, W., Ikawa, S., & Obinata, M. (2004). Reporter gene stimulation by MIDA1 through its DnaJ homology region. *Biochemical and Biophysical Research Communications*, *324*(1), 326–332. http://doi.org/10.1016/j.bbrc.2004.09.059

Zhang, L., Niyazi, H. E. X. D., Zhao, H. R., Cao, X. P., Abudula, M. N. S., Ye, W. J., … Bao, Y. X. (2017). Effects of miRNA-143 and the non-coding RNA MALAT1 on the pathogenesis and metastasis of HeLa cells. *Genetics and Molecular Research : GMR*, *16*(1). http://doi.org/10.4238/gmr16019269

Zhang, S., Lockshin, C., Herbert, A., Winter, E., & Rich, A. (1992). Zuotin, a putative Z-DNA binding protein in Saccharomyces cerevisiae. *The EMBO Journal*, *11*(10), 3787–96. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1396572

CHAPTER 6

The ID3 Protein: A Novel, Potential Target for Therapy to Inhibit the Proliferative Plexiform Lesion in PAH and NSCLC.

Introduction

The proliferative vascular lesion is a major cause of high morbidity and mortality in severe end stage IPAH patients because the pulmonary vascular lesion progressively obliterate pulmonary arterioles and impinge upon the unobstructed vessels resulting in heart failure. The prognosis of these patients remains poor because the current treatment modalities rely on vasodilators, which are not designed to control the growth of PVL, which has become a major clinical barrier for the treatment of end stage PAH. We have earlier shown that PCB-induced ROS upregulated ID3 levels and this re-programmed adult human ECs into iPSCs. Furthermore, we have demonstrated that ID3 signals VEGFR3 expression *in vitro* (Das et al., 2015). Hence, our findings support the idea that ID3 re-programs the adult ECs through inducing VEGFR3 to induce pluripotent endothelial stem cells. This is consistent with the fact that VEGFR3 knockout mice are embryonically lethal and display defects in remodeling the primary vascular plexus. Recently, the expression of VEGFR3 has been shown in non-malignant orbital vascular lesions (Atchison et al., 2016). Based on these evidences, we hypothesize that ID3-VEGFR3 signaling act to drive stemness by enhancing proliferation and inhibiting differentiation. Therefore, we propose that blocking this signaling pathway will suppress stem cell–related cell proliferation. We postulate the following signaling pathway:

PCBs → ID3 → VEGFR3 and/or PCB → ROS → ID3→ VEGFR3

Our preliminary data have shown that PCB-induced ROS contribute to cell proliferation and endothelial spheroid growth. Using the previously published data, we decided to use BNC420 drug to check whether epigenetic regulation of ID3 is affected in the presence and absence of PCBs.

250

Materials and Methods
Cell culture and treatment conditions

Primary cells, Human pulmonary microvascular endothelial cells or Human lung endothelial cells (HPMEC-ST1.6R.). HPMEC cells were maintained in Dulbecco's modified Eagle's medium (DMEM-F12) supplemented with 5% (v/v) fetal bovine serum (FBS). Cells were cultured at 37 ℃ in a humidified atmosphere with 5% CO_2.

Estrogenic chemical treatments

Stock solutions of PCB153 was prepared in DMSO. The same amount of DMSO as in PCB and estradiol-treated cells were added to control cultures. The level of DMSO in experimental media was less than 0.1%. PCB blood levels have been reported to reach approximately 1000 ng/mL (∼3 μM) in occupationally exposed individuals (Wassermann et al., 1979). Our unpublished data showed a significant increase in PCB- induced vascularization with PCB153 concentrations of 10–100 ng/mL. Based on known PCB blood levels from occupational exposure and our preliminary results; we chose a PCB dose of 100 ng/mL (∼0.3 μM) to expose endothelial cells. Our previous studies also showed endothelial cell proliferation and vascular tube formation at physiological doses of 17β-estradiol (Das & Felty, 2014a; Felty, 2006; Felty & Porther, 2008b) therefore, HPMEC and HPMEC ID3[+] were exposed to estradiol at 100 ng/mL (∼3.6 nM). PCB congeners 2,2′,4,4′,5,5′-hexachlorobiphenyl (PCB153) were purchased from AccuStandard (NewHaven, CT) and dissolved in dimethyl sulfoxide (DMSO). All other chemicals and reagents were purchased from Sigma (St. Louis, MO).

BNC420 Treatment

Stock solutions of BNC420 was prepared in dimethyl sulfoxide (DMSO). The same amount of DMSO as in PCB and estradiol-treated cells were added to control cultures. The level of DMSO in experimental media was less than 0.1%. BNC420 is a novel VEGFR3 selective inhibitor provided by Bionomics Limited via approved Material Transfer Agreement (Leske et al., 2015). The single oral dose of BNC420 used in mice is at 75 mg/kg, which is well below the IC50 value. This dose is equivalent to 0.15 mM (Mol. Wt. = 501.55). We treated the cells with lower dose than 0.15 mM. Therefore, BNC420 is dissolved in DMSO at a stock conc. of 10 mM and further to achieve working concentration of 0.1 mM stock solution was diluted in culture medium for *in vitro* use as per Bionomics suggestions.

ID3 overexpression

The HPMEC cells were stably transfected with either Precision LentiORF for ID3 (Thermo Scientific Open Biosystems) or empty vector lentiviral pLEX-JRED/TurboGFP by the trans-lentiviral packaging kit with Express-in transfection reagent according to the manufacturer's instructions. We used the MOI (multiplicity of infection) of 25 and selected cells that overexpressed ID3 with blasticidin S (5 mg/ml) as per manufacturer's instructions. Cells expressing TurboGFP were identified by fluorescence microscopy.

Western blot

Whole cell lysates were prepared with lysis buffer containing 25 mM Tris-HCl buffer (pH 8.0), 150 mM NaCl, 0.2% NP-40, 10% glycerol, 8 mM b-glycerophosphate, 2.5 mM

sodium pyro- phosphate, 10 mM NaF, 0.2 mM Na3VO4, 1 mM DTT and 10 ml/ml protease inhibitor cocktail (Sigma-Aldrich) followed the published protocol of our laboratory (Das & Felty, 2014a). Proteins were quantified using the Bradford Assay Reagent (Bio-Rad) according to the manufacturer's instructions. Proteins (35–75 mg) were separated by 15% SDS-PAGE and transferred to polyvinylidene fluoride (PVDF) membranes (Millipore). Membranes were blocked with 5% nonfat milk and incubated with the following antibodies: ID3 (Cal BioReagent) and b-actin (Cell Signaling). Anti-body dilutions used were according to manufacturer's recommendations for detection by immunoblot. Gel percentage 15%. Primary antibody ID3 monoclonal Ab- 1:2500 and Secondary Ab- 1:10000. Membranes were then incubated with horseradish peroxidase-conjugated secondary IgG antibodies and visualized with ECL Plus Western blot reagents (GE Healthcare, Amersham). The membranes were re-probed for b-actin as a loading control. Electrochemiluminescence (ECL) intensity of detected target proteins was imaged and quantified with a Bio-Rad Versa Doc instrument. All immunoblots have completed a minimum of three times for each experiment.

Immunofluorescence

The cells will be fixed with methanol and incubated at −20 °C for 15 min. Immunofluorescence staining was performed according to standard procedures. Briefly, the cells were blocked with 3% normal goat serum at 4 °C for 1 h and then incubated with antibodies: anti-ID3 overnight at 4 °C. After washing with 1× phosphate buffered saline (PBS), cells were incubated with anti-mouse-IgG Alexa Fluor® 633 for ID3 (Santa Cruz Biotechnology Inc.). Cells were washed with PBS mounted with Fluoromount-

GTM re- agent. The spheroids with GFP or RFP were photographed and analyzed using a Nikon C1 laser SCANNING confocal microscope.

Endothelial spheroid assay

HPMEC ID3$^+$ cells were suspended in serum-free DMEM/F12 (1:1) culture medium supplemented with B27. For EC spheroid formation, approximately 100–150 HPMEC ID3$^+$ cells were seeded separate well in an ultra-low attachment 96-well plate (Corning Inc. Lowell, MA) followed the published protocol of our laboratory (Das & Felty, 2014a). The effect of estrogenic chemical PCB153 was determined by pre-treating the cells with 100 ng/mL E2 2h prior to exposure with 100 ng/mL (~0.3 µM) PCB153 on the day of seeding cells. Endothelial spheroids were grown for 10 days in liquid culture in the absence or presence of PCB153. A total of 15 endothelial spheroids with a minimum diameter of 50 µm were counted in each experimental group. Pictures were taken on 10d after seeding. We counted number and size of spheroids on 10d. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons

MTT assay

To determine the effect of extra estrogen environmental pollutant PCB153 on the metabolic activity of lung endothelial cells, we performed some cell viability and cell proliferation assays. For cell viability measurement, we performed MTT assay which is colorimetric assay and measures the metabolic activity of cells. Actually, MTT is a tetrazolium dye known as 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide. NAD(P)H-dependent oxidoreductase enzymes present in mitochondria of metabolically active and dividing cells. This enzyme catalyses tetrazolium dye and converts into

insoluble formazan, which gives a purple colour (Berridge & Tan, 1993; M. V. Berridge et al., 2005). The intensity of the purple colour which can be measured at an optical density (OD) of 590 nm and its proportional to a number of metabolic active cells. We seeded HPMEC and HPMEC ID3[+] cells with 50% confluence in 96 well clear-walled plate. We allowed the cells to grow for 24h in 5% FBS culture medium. After 24 hours, we gave treatments, 0.1% DMSO (control), PCB153 (100 ng/mL) to make total volume 100 µL in each well. We measured the metabolic activity of cells after 48h with the addition of MTT dye and evaluating the purple colour intensity at 590nm.

SRB assay

The sulforhodamine B (SRB) assay is used for measurement of the total cellular protein content of adherent cells in a 96-well format. This assay is useful in the determination of cellular proliferation and growth. We seeded HPMEC and HPMEC ID3[+] cells with 50% confluence in 96 well clear-walled plate. We allowed the cells to grow for 24h in 5% FBS culture medium. After 24 hours, we gave treatments, 0.1% DMSO (control), PCB153 (100 ng/mL) to make total volume 100 µL in each well. After an incubation of 48h, cell monolayers were fixed with 10% (wt/vol) trichloroacetic acid and stained for 30 min, after which the excess dye is removed by washing repeatedly with 1% (vol/vol) acetic acid. The protein-bound dye is dissolved in 10 mM Tris base solution for OD determination at 590 nm using a microplate reader (Vichai & Kirtikara, 2006).

BrdU assay

BrdU assay is a direct method for evaluation of cell proliferation. In this rapid method, Bromodeoxyuridine (5-bromo-2'-deoxyuridine, BrdU) is a synthetic nucleoside and analogue of thymidine gets incorporated into DNA synthesis of proliferating cells. It is a suitable method for detection and quantification of the stimulatory effect of environmental chemicals (Russo et al., 1984). For BrdU assay, we used Roche's kit and followed the manual instructions.

We seeded HPMEC and HPMEC ID3$^+$ cells in 5% FBS DMEM F-12 media with 50% confluence and cultured with addition of BrdU in 96 well clear-walled plate. After 24 hours, we gave treatments, 0.1% DMSO (control), PCB153 (100 ng/mL) to make total volume 100 µL in each well. After an incubation of 48h, cell monolayers fixed with FixDenat. Followed by fixation we added anti- BrdU-POD, which binds to BrdU in newly synthesized DNA. This immune complex detected by a plate reader at 450 nm.

STATISTICAL ANALYSIS

All experiments were performed at least three times and data are expressed as mean ± standard deviation Graphpad prism version 5.01 was used to perform graphics and the two-tailed student's t-test was used to compare statistical significance. P<0.05 was set as a statistical significance. Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.

Results-

Selection of optimum BNC420 Treatment dose

BNC420 is a novel VEGFR3 selective inhibitor made by company Biogenomics. It has been shown in mouse model that angiogenesis specially inhibited by BNC420. In every protein targeted molecular drug treatment research revolved around ensuring cellular protein homeostasis. The efficacy of drug treatment is based on assessment of how it affects cellular viability, that is, their cytotoxic effect. Investigations of cytotoxicity often require testing several different concentrations and drug exposure times using cells in culture. Therefore, we used colorimetric tetrazolium salt assay to test putative cytotoxic effects associated with exposure to a BNC420. MTT assay monitors metabolic activity of cultured cells, was adapted to analyze the viability of cells exposed to BNC420. It was found that 24h exposure to increasing concentration of BNC420 reduced the viability of lung EC ID3$^+$ cells, with the most severe effect observed at 10 µM dose of 24 h exposure with MTT reductions approaching 30% of non-exposed cells as shown in Figure 104. For a series of incubations for 24 h, it was found that concentration 0.1 µM did not affect the viability, and BNC420 concentrations of 0.5 µM reduced the MTT reduction rate to approximately 20% the level displayed by cells receiving vehicle alone. Generally, less than 10 µM concentration of a drug in plasma is accepted because higher than 10 µM of a drug inhibit necessary enzymes. The most of the drugs on the market have less than 10 µM therapeutic plasma concentrations. In most cases, the IC50 of the best candidate compound should be lower to 10 µM, which is acceptable for NIH to screen the NCI60 program. Therefore, lower dose the IC50 of candidate drug was, the higher potentially it become the new drug for any diseases. (Hansen & Bross, 2010; Hughes, Rees,

Kalindjian, & Philpott, 2011; Wang, Newton, Liu, Atkins, & Lu, 2000; Williams et al., 2002). We decided to use the dose of 0.1 µM to treat the lung cell and measure the changes at molecular level in the expression level of ID3 protein as well as phenotypic changes. The single oral dose of BNC420 used in mice is at 75 mg/kg by Biogenomics Company, which is well below the IC50 value. This dose is equivalent to 0.15 mM (Mol. Wt. = 501.55). We tried to mimic similar drug concentration in Lung EC ID3$^+$ cells with lower dose than LC50 value. We treated the cells with lower dose than 0.15 mM. BNC420 inhibited Lung EC ID3$^+$ cells with 50% survival at 15 µM.
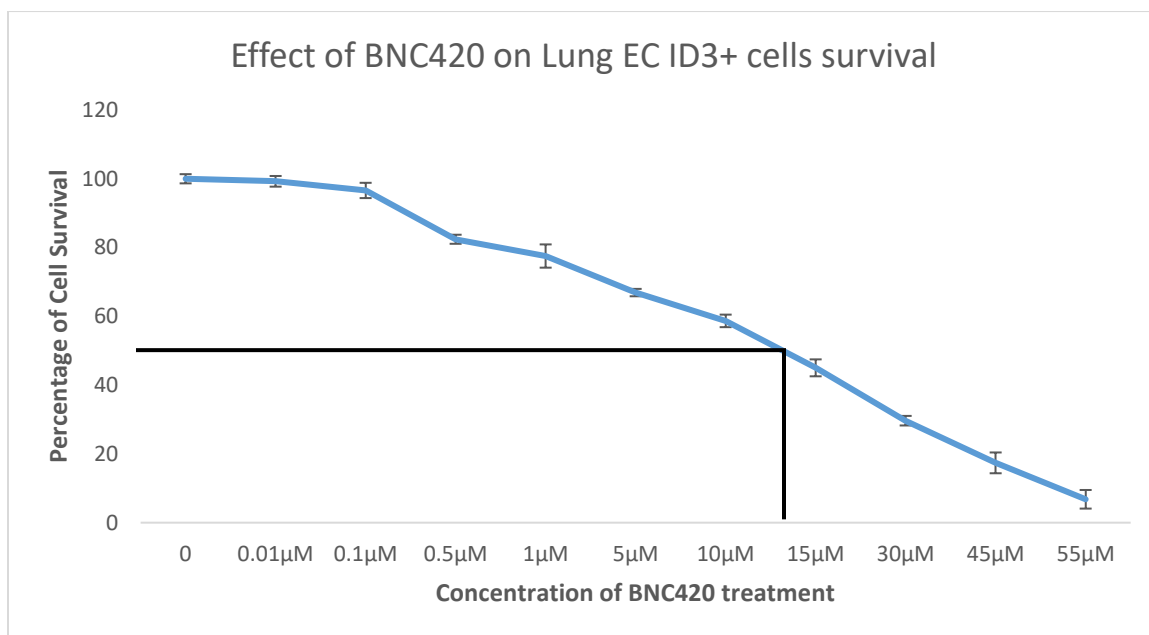

Figure 104- Graph of cell cytotoxity assay for BNC420 drug

Graph of cell survival was determined 24 h after exposure to BNC420. LC50~15 µM by using MTT assay. Results shown are averages of triplicates ± SD (<10%).

BNC420 inhibits PCB153-induced lung endothelial cell proliferation

We have earlier tested the lung endothelial cell proliferation in presence of estrogenic chemicals by MTT, SRB and BrdU assay. In this experiment, we tested the effect of BNC420 drug on the cell proliferation of HPMEC ID3$^+$ cells in presence or absence of PCB153. The PCB153 induced cell proliferation measured by MTT, SRB and BrdU assay. We tested range of BNC420 concentration starting from 10 µM to 0.1 µM (Figure 105 A,B,C). Our data revealed that in presence of treatment PCB153 (100 ng/mL) with addition of different concentration of BNC420 drug ranges from 10 µM to 0.1 µM showed a significant gradual inhibition in cell proliferation. We measured the growth in a number of cells in terms increase in OD value. Our MTT assay results were corroborated by SRB and BrdU assay. The data from SRB and BrdU assays reported that treatment PCB153 (100 ng/mL) showed a significant gradual inhibition in cell proliferation. In terms of SRB assay, the total protein content of treated cells was inhibited. The BrdU assay data results showed the synthesis of new DNA in proliferating cells was significantly inhibited in treatments than control. These results prove that PCB153 (100 ng/mL) is an estrogenic chemical drives the endothelial cells proliferation. These data shows that BNC420 with dose of 0.1 µM and 1 µM effectively inhibits PCBindcued cellulr proliferation without having any cellular toxicity effects.

Lung EC ID3 MTT assay
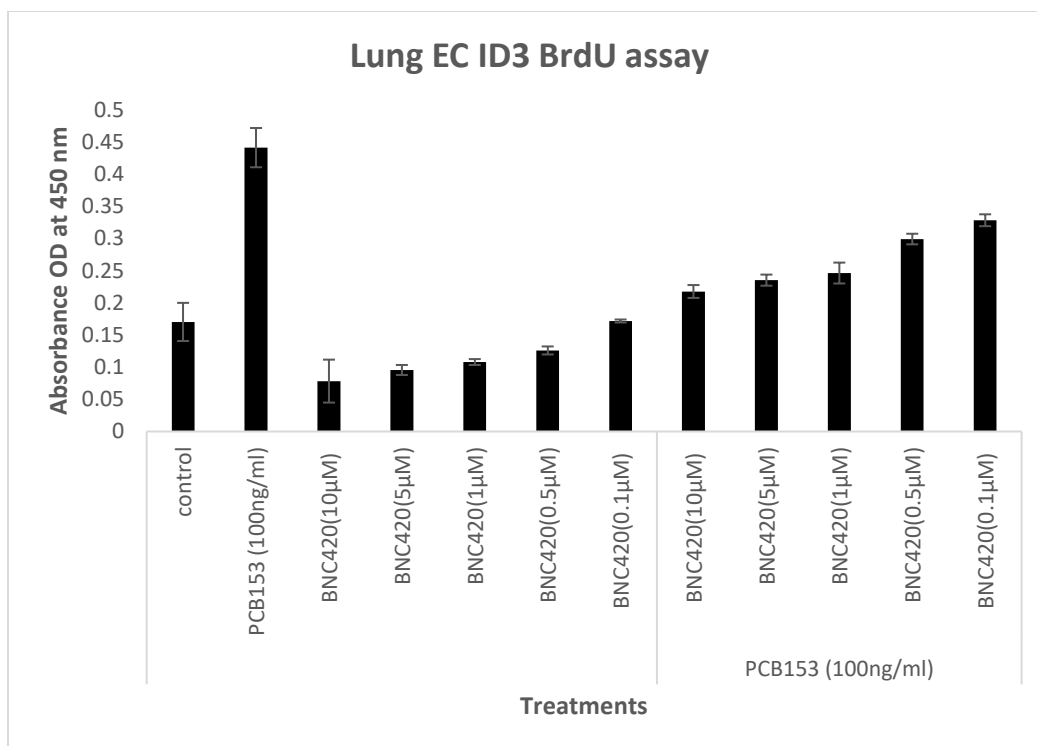

Lung EC ID3 SRB assay
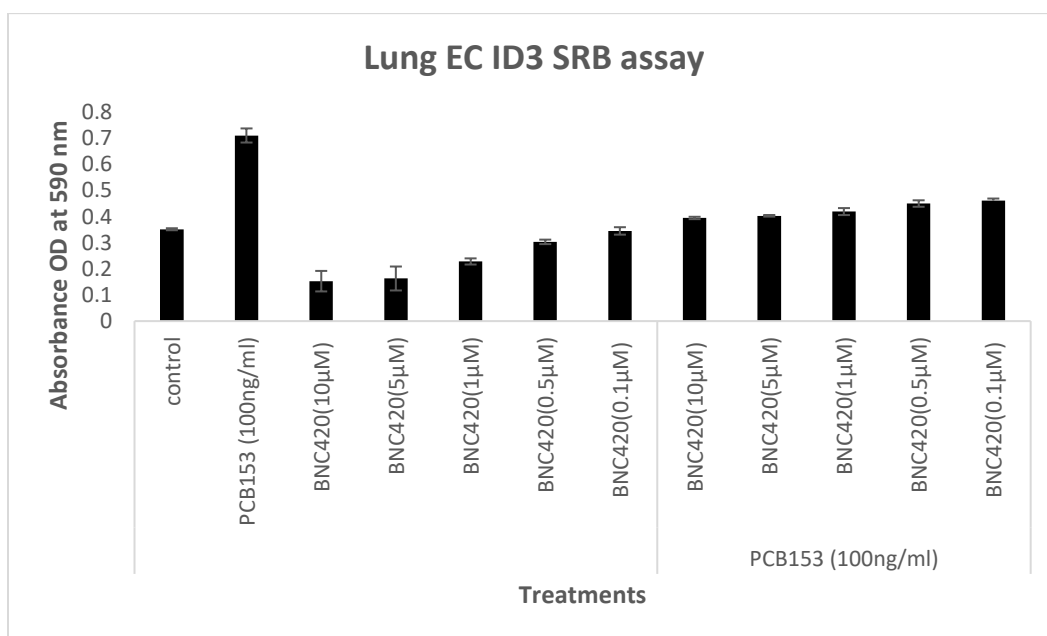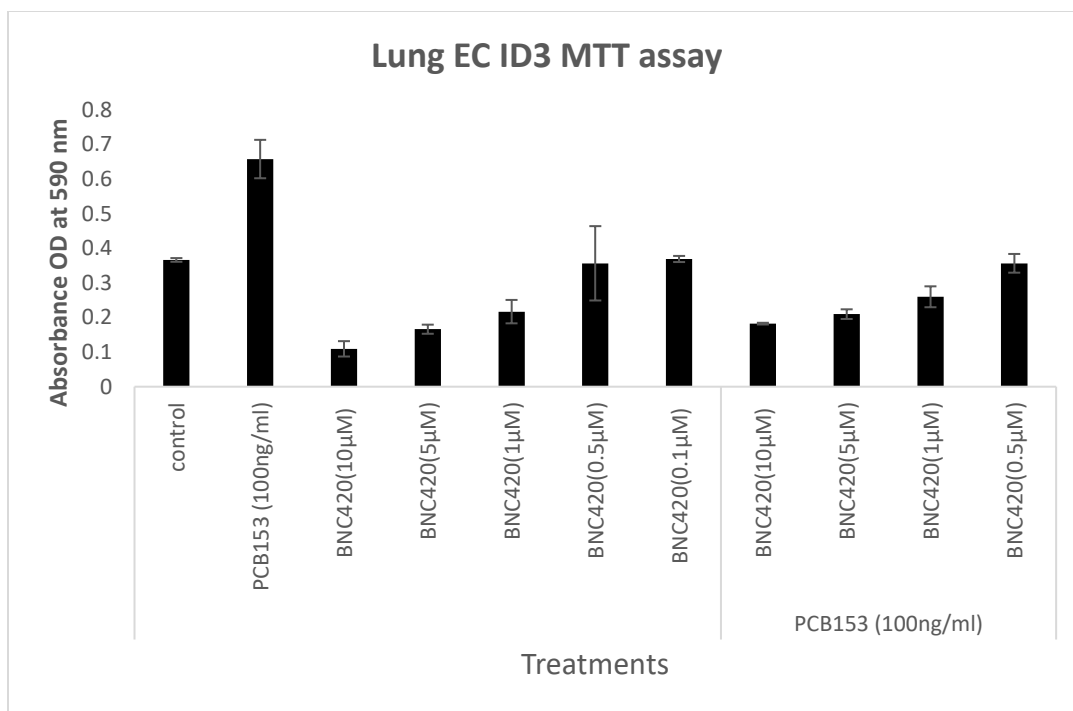
Figure 105 A, B, C- BNC420 inhibits the PCB153 induced cell proliferation

A, B, C- BNC420 inhibits the PCB153 induced cell proliferation measured using MTT, SRB and BrdU assay 24 h after exposure to BNC420. Results shown are averages of triplicates ± SD (<10%).

BNC420 inhibits PCB153-induced lung EC spheroid formation

The endothelial cells lines the innermost layer of artery maintains the balance between vasoconstriction and vasodilation of pulmonary artery. When endothelial dysfucntion happens, the cells undergo proliferation and form plexiform lesions, which grow in the three-dimension structure resembles spheroid. These cells do not follow characteristics of the cells, which grow in a monolayer. The cells derived from tumour tissue alter their both genotype and phenotype characteristics once they cultured in monolayer (Lee et al., 2006). Therefore, we tried to imitate the conditions of tumour *in vitro*. In addition to this,

recent data from our laboratory showed that ID3 induced a molecular stem cell-like signature—CD133[+] VEGFR3[+] CD34[+] in microvascular ECs. These spheroid also form angiogenic sprouting network. We grew the cells in no FBS DMEM F-12 medium supplemented with B-27 reduced growth supplement. We used a stem cell sphere-forming assay to determine the effect of BNC420 drug on HPMEC ID3[+] cells spheres. We treated spheroid-containing wells with 0.1 µM BNC420 in presence or absence of estrogenic chemicals like E2 and PCB153 as shown in Figure 106. We checked whether BNC420 inhibits spheroid growth of HPMEC ID3[+] cells. We found that BNC420 significantly inhibited the growth of spheroids in terms of spheroid number and spheroid size compare to E2 and PCB153 treated spheroid-containing wells as shown in Figures 107 and 108.
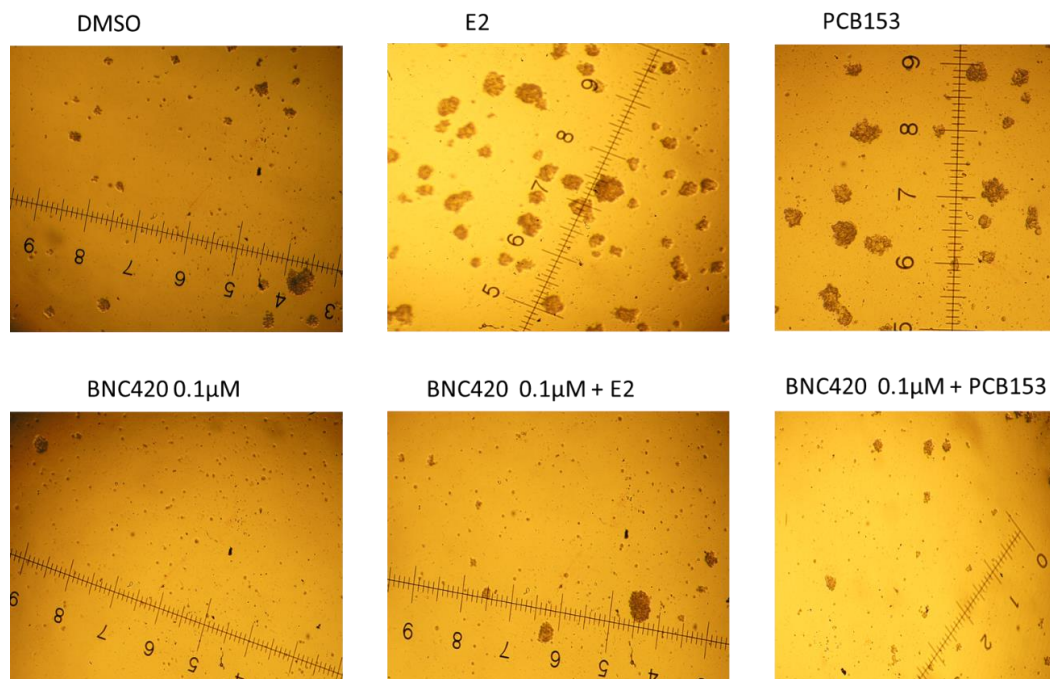

Figure 106- shows pictures of spheroids with BNC420 treatment.

BNC420 treatment decreases Spheroid size and number in the PCB153 and E2 treated wells. The decrease in the size and number of spheroid in BNC420 treated wells are significantly lower compared to respective control. Scale bar= 10µm.
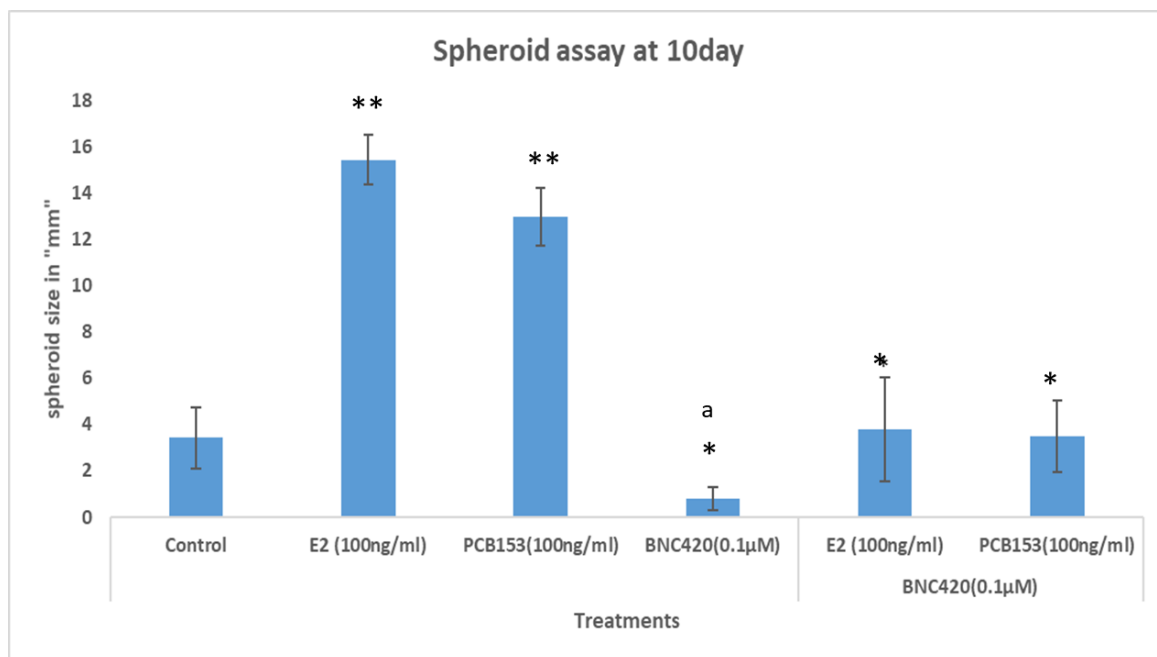
Spheroid assay at 10day

Figure 107- BNC420 treatments decreased the size of spheroids compared to control.

Graph shows quantification of spheroid diameter. Error bars represent the mean sphere size ± SD in 5 spheres. E2 & PCB153= 100 ng/ mL. Representative microphotographs of treatment groups are shown inset. E2 & PCB153= 100 ng/mL. BNC420= 0.1 μM Measurements were performed from a representative experiment repeated three times in triplicate. (*a) BNC420 treatment significantly different from vehicle control at (P<0.01). (**) PCB & E2 treatment significantly different from vehicle control at (P<0.01). (*) BNC420 treatment significantly different from PCB153 + BNC420 & E2 + BNC420 (P<0.05) at day 10 Data; Tukey HSD test for multiple comparisons. Magnification X200.
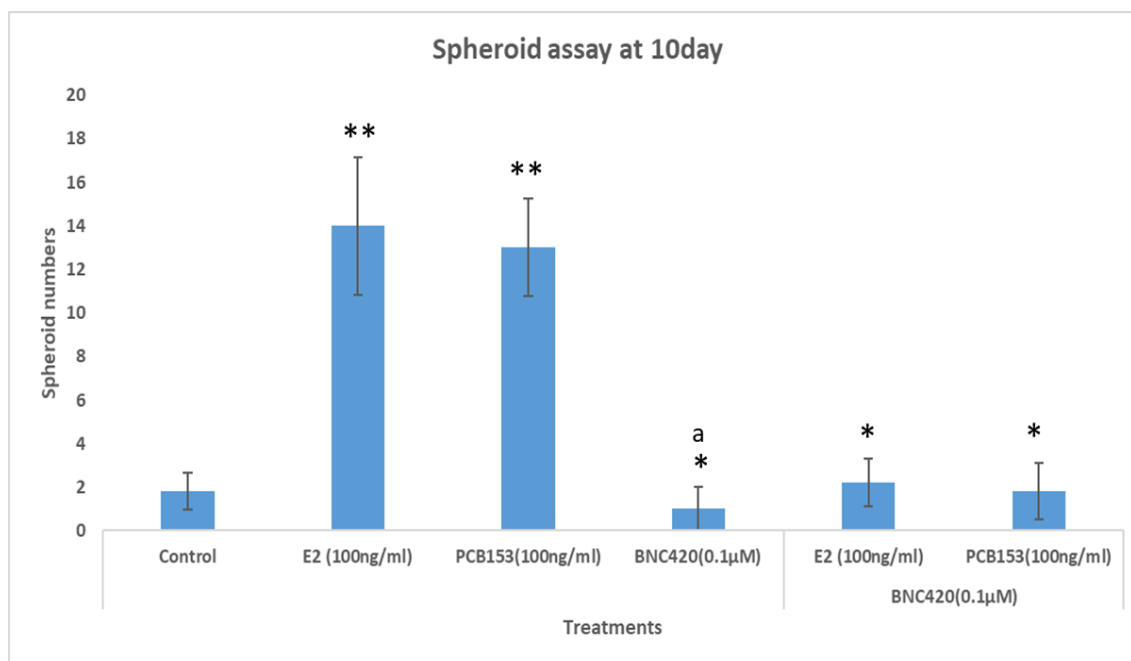
Figure 108- BNC420 treatments decreased the number of spheroids compared to control.

Graph shows quantification of spheroid number. Error bars represent the mean sphere number ± SD in 5 wells. E2 & PCB153= 100 ng/mL. BNC420= 0.1 μM Measurements were performed from a representative experiment repeated three times in triplicate. (*a) BNC420 treatment significantly different from vehicle control at (P<0.01). (**) PCB & E2 treatment significantly different from vehicle control at (P<0.01). (*) BNC420 treatment significantly different from PCB153 + BNC420 & E2 + BNC420 (P<0.05) at day 10 Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons. Magnification X200.

BNC420 inhibits PCB153-induced ID3 expression

We performed immunoblot and fluorescent microscopy experiment after detecting change in spheroid and cell viability assay due to the treatment of BNC420. The immunoblot and fluorescent microscopy results revealed that BNC420 significantly down regulated ID3 protein expression in lung EC ID3[+] cells compared to control as shown in Figure 109, 110 and 111.
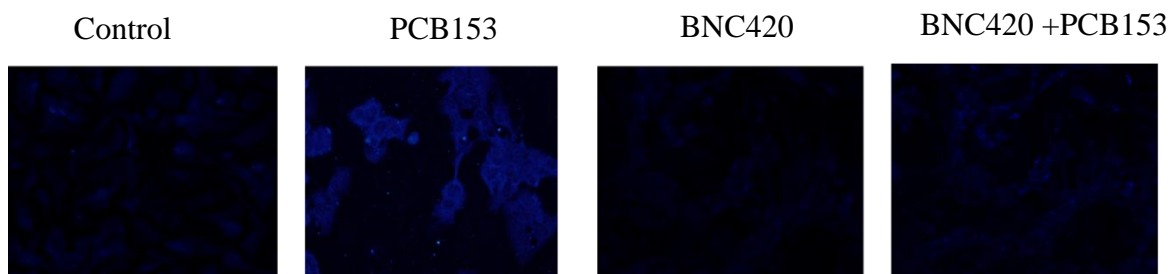
Figure 109- Representative images of fluorescent microscopy to study the effect of BNC420 drug on ID3 expression

Fluorescent microscopy experiment was used to compare the effect of BNC420 0.1 µM on PCB153 treated lung EC ID3$^+$ cells. The experiment showed that BNC420 treatment decreased the expression of ID3. BNC420= 0.1 µM, PCB153 = 100 ng/mL (~0.3 µM). Scale bar= 50 µm

Figure 110- Western blot analysis of inhibition of PCB-induced ID3 expression by BNC420

The experiment shows inhibition of ID3 protein by 0.1 µM BNC420 dose with or without exposure to PCB153. BNC420= 0.1 µM, PCB153 = 100 ng/ mL (~0.3 µM).

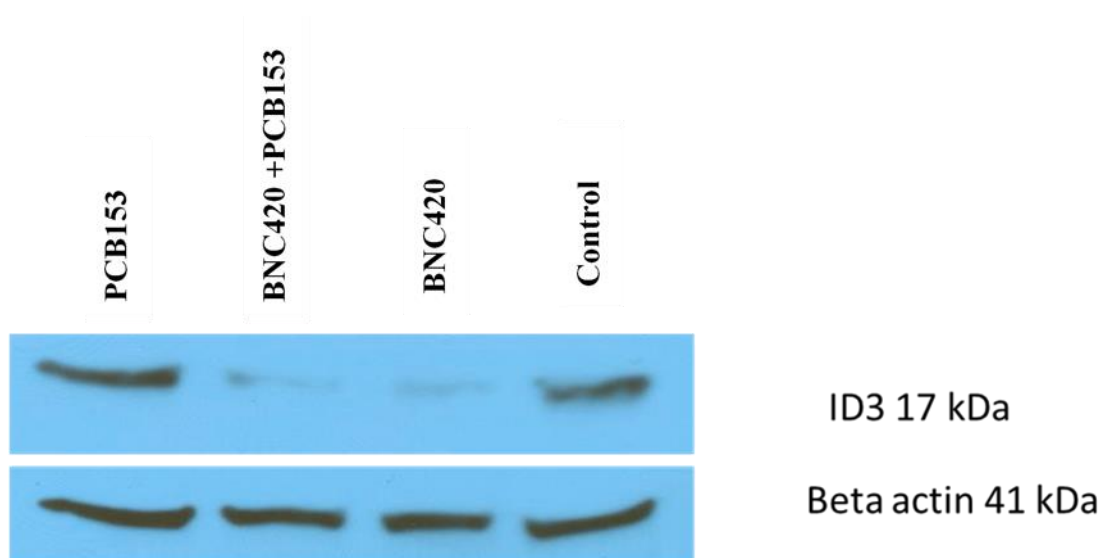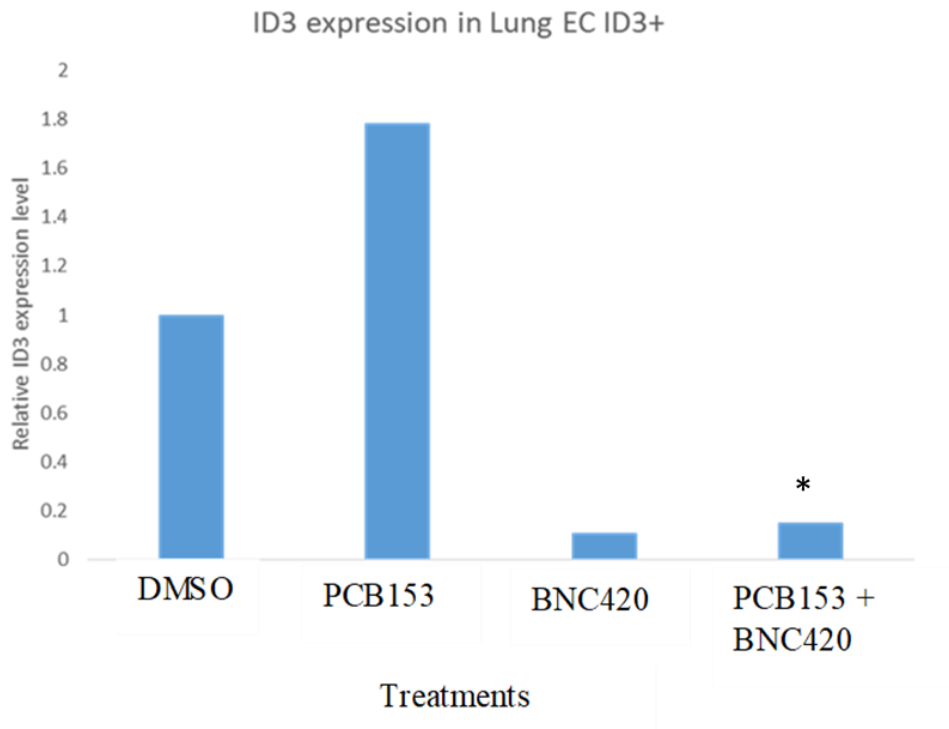**Figure 111- The relative expression level of ID3 protein in BNC420 drug treated lung EC ID3+ cells**

Shows the relative expression level of ID3 protein in lung EC ID3[+] cell line when treated with PCB153, BNC 0.1 μM and PCB153 + BNC 0.1 μM. PCB153 = 100 ng/ mL (~0.3 μM). (*) PCB513 treatment significantly different from PCB153 + BNC420 ($P<0.05$) at day 10 Data were analyzed by ANOVA; Tukey HSD test for multiple comparisons.

Conclusion

Using a novel VEGFR3 selective inhibitor BNC420 (provided by Bionomics Limited via approved Material Transfer Agreement), we determined BNC420 treatment of hiPSCs controls the epigenetic regulation of ID3 and regression of PCB-induced vascular lesions *in vitro*. These results are very promising can give impetus for further development of BNC420 drug and can used as a molecular target against ID3 protein. Molecular targeting overexpreesed ID3 expression can regress angiogenesis and lung EC cell proliferation which are markers of NSCLC and PAH.

References

Atchison, E. A., Garrity, J. A., Castillo, F., Engman, S. J., Couch, S. M., & Salomaõ, D. R. (2016). Expression of Vascular Endothelial Growth Factor Receptors in Benign Vascular Lesions of the Orbit: A Case Series. *Ophthalmology*, *123*(1), 209–213. http://doi.org/10.1016/j.ophtha.2015.09.009

Berridge, M. V., & Tan, A. S. (1993). Characterization of the Cellular Reduction of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT): Subcellular Localization, Substrate Dependence, and Involvement of Mitochondrial Electron Transport in MTT Reduction. *Archives of Biochemistry and Biophysics*, *303*(2), 474–482. http://doi.org/10.1006/abbi.1993.1311

Berridge, M. V., Herst, P. M., & Tan, A. S. (2005). Tetrazolium dyes as tools in cell biology: New insights into their cellular reduction (pp. 127–152). http://doi.org/10.1016/S1387-2656(05)11004-7

Das, J. K., & Felty, Q. (2014). Microvascular Lesions by Estrogen-Induced ID3: Its Implications in Cerebral and Cardiorenal Vascular Disease. *Journal of Molecular Neuroscience*, 618–631. http://doi.org/10.1007/s12031-014-0401-9

Das, J. K., Voelkel, N. F., & Felty, Q. (2015). ID3 contributes to the acquisition of molecular stem cell-like signature in microvascular endothelial cells: Its implication for understanding microvascular diseases. *Microvascular Research*, *98*, 126–138. http://doi.org/10.1016/j.mvr.2015.01.006

Felty, Q. (2006). Estrogen-induced DNA synthesis in vascular endothelial cells is mediated by ROS signaling. *BMC Cardiovascular Disorders*, *6*, 16. http://doi.org/10.1186/1471-2261-6-16

Felty, Q., & Porther, N. (2008). Estrogen-induced redox sensitive Id3 signaling controls the growth of vascular cells. *Atherosclerosis*, *198*(1), 12–21. http://doi.org/10.1016/j.atherosclerosis.2007.12.048

Hansen, J., & Bross, P. (2010). A Cellular Viability Assay to Monitor Drug Toxicity. In *Protein Misfolding and Cellular Stress in Disease and Aging: Concepts and Protocols. Methods in Molecular Biology.* (Vol. 648, pp. 303–311). Humana Press, Totowa, NJ. http://doi.org/10.1007/978-1-60761-756-3_21

Hughes, J. P., Rees, S., Kalindjian, S. B., & Philpott, K. L. (2011). Principles of early drug discovery. *British Journal of Pharmacology*, *162*(6), 1239–49. http://doi.org/10.1111/j.1476-5381.2010.01127.x

Lee, J., Kotliarova, S., Kotliarov, Y., Li, A., Su, Q., Donin, N. M., … Fine, H. A. (2006). Tumor stem cells derived from glioblastomas cultured in bFGF and EGF more closely mirror the phenotype and genotype of primary tumors than do serum-cultured cell lines. *Cancer Cell*, *9*(5), 391–403. http://doi.org/10.1016/j.ccr.2006.03.030

Leske, A., Lavranos, T., Beaumont, D., Brown, C., Paul, D., Inglis, D., … Kremmidiotis, G. (2015). BNC420 is a novel VEGFR3 selective inhibitor, which unlike the pan-VEGFR inhibitor Sunitinib, suppresses lymphatic metastasis in a model of metastatic melanoma. *Cancer Research*, *75*(15). http://doi.org/10.1158/1538-7445.AM2015-4115

Russo, A., Gianni, L., Kinsella, T. J., Klecker, R. W., Jenkins, J., Rowland, J., … Myers, C. (1984). Pharmacological evaluation of intravenous delivery of 5-bromodeoxyuridine to patients with brain tumors. *Cancer Research*, *44*(4), 1702–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6704976

Vichai, V., & Kirtikara, K. (2006). Sulforhodamine B colorimetric assay for cytotoxicity screening. *Nature Protocols*, *1*(3), 1112–1116. http://doi.org/10.1038/nprot.2006.179

Wang, R. W., Newton, D. J., Liu, N., Atkins, W. M., & Lu, A. Y. (2000). Human cytochrome P-450 3A4: *in vitro* drug-drug interaction patterns are substrate-dependent. *Drug Metabolism and Disposition: The Biological Fate of Chemicals*, *28*(3), 360–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10681383

Wassermann, M., Wassermann, D., Cucos, S., & Miller, H. J. (1979). World Pcbs Map: Storage And Effects In Man And His Biologic Environment In The 1970s. *Annals of the New York Academy of Sciences*, *320*(1 Health Effect), 69–124. http://doi.org/10.1111/j.1749-6632.1979.tb56594.x

Williams, J. A., Ring, B. J., Cantrell, V. E., Jones, D. R., Eckstein, J., Ruterbories, K., … Wrighton, S. A. (2002). Comparative Metabolic Capabilities of CYP3A4, CYP3A5, and CYP3A7. *Drug Metabolism and Disposition*, *30*(8), 883–891. http://doi.org/10.1124/dmd.30.8.883

Overall summary

Stable ectopic expression of ID3 in lung endothelial cells contributed to endothelial-mesenchymal transition (EndMT) and cell proliferation. ID3 overexpressing cells showed the loss of VE-cadherin and gain of N-Cadherin and vimentin, which are markers of EndMT. Using an endothelial spheroid assay, an established method to measure aberrant hyper-proliferation of endothelial cells in PAH patients, we show that stable ectopic expression of ID3 increased the number and size of vascular spheres.

We investigated ID3 transcriptional reprogramming using ChIP-seq and RNA-seq technology. We show here for the first time that ID3 is part of a more general mechanism of transcriptional regulation. Our ChIP-Seq data show that ID3 binds to a subset of approximately 1200 target genes. Comprehensive motif analysis of ChIP-seq data using the MEME Suite software toolkit revealed that ID3 bound to the GAGAGAGAGA motif sequence on genomic DNA. We also show a significant preference of ID3 binding to motifs associated with transcription factors IRF1, ZRF1, BC11A, IRF4, PRDM1, FOXJ3, SMAD4, ZBTB6, GATA1, and STAT2. Using an integrative approach of ChIP-Seq and RNA-Seq data, we identified 27 genes was bound by ID3 and RNA was differentially expressed in ID3 overexpressing cells.

Our data demonstrated that PCB153 and/or ID3 induces proliferation of lung endothelial cells via transcriptional reprogramming. Discoveries from these findings will lay the necessary groundbreaking work for testing the efficacy of ID3 antagonists for the

prevention and treatment of pathological vascular remodeling as well as provide a new paradigm by which PCBs may contribute to lung vascular toxicity.

VITA

MAYUR DOKE

Born, Pune, India

2006-2009   B.Sc., Biotechnology

Mumbai University, India

2009-2011   M.Sc., Environmental Science

Mumbai University, India

2013-2015   Graduate Research Assistant

Florida International University, Miami, Florida, USA

2015-2018   Doctoral Candidate

Florida International University, Miami, Florida, USA

PUBLICATIONS AND PRESENTATIONS

Avecilla, V., Doke, M., & Felty, Q. (2017). Contribution of Inhibitor of DNA Binding/Differentiation-3 and Endocrine Disrupting Chemicals to Pathophysiological Aspects of Chronic Disease. *BioMed Research International*, *2017*, 1–22. https://doi.org/10.1155/2017/6307109

Doke, M., Avecilla, V., & Felty, Q. (2018). Inhibitor of Differentiation-3 and Estrogenic Endocrine Disruptors: Implications for Susceptibility to Obesity and Metabolic Disorders. *BioMed Research International*, *2018*, 1–16. https://doi.org/10.1155/2018/6821601

Doke, M., Das, J., & Felty, Q. (2015). ID3 mediated vascular reprogramming of PCB exposed endothelial cells and its potential contribution to lung tumorigenicity, *198*(1), 84827.

Doke, M., Felty, Q., & Das, J. (2017). The role of environmental chemical PCB153 in aggressive growth of plexiform lesions associated with lung cancer. *European Journal of Surgical Oncology*, *43*(11), 2218. https://doi.org/10.1016/j.ejso.2017.10.119

Doke, M., & Felty, Q. H. (2016). Abstract 3274: VEGF─receptor antagonist, Sugen 5416, sensitizes pulmonary endothelial stem-like cells to estrogens: A microvascular model for the progression of lung cancer. *Cancer Research*, *76*(14 Supplement). Retrieved from http://cancerres.aacrjournals.org/content/76/14_Supplement/3274

Roy, D., Morgan, M., Yoo, C., Deoraj, A., Roy, S., Yadav, V. K., … Doke, M. (2015). Integrated bioinformatics, environmental epidemiologic and genomic approaches to identify environmental and molecular links between endometriosis and breast cancer. *International Journal of Molecular Sciences*, *16*(10), 25285–25322.