THE ROLE OF LONG NONCODING RNAS IN CANCER AND MICROBIOTA-DERIVED ARYL HYDROCARBON RECEPTOR LIGANDS

A Dissertation

by

YATING CHENG

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Stephen H. Safe
Committee Members,	Natalie M. Johnson
	Timothy D. Phillips
	Weston W. Porter
	Yanan Tian
Intercollegiate Faculty Chair,	Ivan Rusyn

December 2016

Major Subject: Toxicology

ABSTRACT

LncRNAs are a group of non-coding RNAs containing >200 nucleotides and these RNAs have no significant protein coding potential. In the past 10-15 years the role of lncRNAs in cancers have been demonstrated, however, their function in tumors and potential for drug targeting are not well defined. We studied two lncRNAs, HOXA transcript at the distal tip (HOTTIP) and metastasis associated lung adenocarcinoma transcript 1 (MALAT1) in pancreatic cancer using the in vitro cell lines and in vivo xenograft or transgenic mouse models. Our results demonstrated that both lncRNAs are pro-oncogenic in the pancreatic cancer, supported by the observation that knockdown of HOTTIP and MALAT1 decreased cell proliferation, migration/invasion and increased apoptosis. HOTTIP might exhibit some pro-oncogenic functions, via the regulation of HOXA gene clusters in a cis-regulating manner. On the other hand, MALAT1 regulate responses of pancreatic cancer cells in part via polycomb repressive complex 2 (PRC2) dependent and independent pathways. Our transcriptomic results support the important but distinct roles of HOTTIP and MALAT1 in pancreatic cancer and we also show that MALAT1 expression can be targeted by small molecule drugs. In the second part of this dissertation, we studied several microbiota-derived aryl hydrocarbon receptor (AhR) ligands. The AhR is a ligand-activated transcription factor with an evolving role in the normal physiological development and diseases. Gut microbiota metabolites are important for mediating communication between gut microflora and the host. It has recently been shown that the gut microbiota produces several metabolites that are AhR

ligands. Microbiota-derived tryptophan metabolites and 1,4-dihydroxy-2-napthoic acid (DHNA) and related compounds are reported to be AhR ligands as evidence by their induction of cytochrome P450, family 1, subfamily A, polypeptide 1 (CYP1A1) and other AhR-responsive genes. We hypothesized the microbiota-derived AhR ligands are selective AhR modulators (sAhRMs) and their induction responses are compound, gene and cell context dependent. We have carried out extensive studies on tryptophan metabolites, DHNA and related compounds in both human and mouse colon cancer cell lines, and have observed that some of these compounds exhibited partial agonist/antagonist activities that are both gene and cell context specific.

DEDICATION

To my father who passed away when I was a sophomore. I missed him so much.

To my mother for her unconditional love and support.

To my husband for his love, patience and support.

ACKNOWLEDGEMENTS

I would like to thank my mentor Dr. Stephen H. Safe for all the support and guidance throughout my Ph.D. studies. He is a great scientist and has taught me how to be a good scientist.

Next, I would like to thank my committee members Dr. Natalie M. Johnson, Dr. Timothy D. Phillips, Dr. Weston W. Porter and Dr. Yanan Tian for their precious time and supportive advice.

I would like to thank all the people in the Safe lab. I would like to thank the current lab members, Dr. Un-ho Jin, Dr. Xi Li, Dr. Mohankumar Kumaravel, Alexandra Lacey, Erik Hedrick, and Keshav Karki. I also thank past lab members Dr. Indira Jutooru, Dr. Gayathri Chadalapaka, Dr. Sandeep Sreevalsan, and Dr. Vijayalekshmi Vasanthakumari who helped me a lot with my research. I also thank Katherine Mooney, who helped with preparing manuscripts and taught me how to use reference tools. The help from the undergraduate students is also gratefully appreciated. In the end, special thanks to Lorna Safe, who has given me some great advice in life and offered me the chance to train my managing skills in the lab.

I also thank all the faculty and staff in the Toxicology program. Special thanks to Kim Daniel, who has helped me a lot since I joined the toxicology program.

Last but not the least, I would like to thank my family and friends for their continuous support.

CONTRIBUTORS AND FUNDING SOURCES

Contributors section

Part 1, faculty committee recognition

This work was supervised by a dissertation committee consisting of Professors Stephen Safe and Yanan Tian of the Department of Veterinary Physiology and Pharmacology, Professors Western Porter and Timothy Phillips of the Department of Veterinary Integrative Biosciences, and Professor Natalie Johnson of the School of Public Health.

Part 2, student/collaborator contributions

The work presented in Section 2 is the reprinted from Cheng et al. “The long non-coding RNA HOTTIP enhances pancreatic cancer cell proliferation, survival and migration” published in May 2015 with permission from Oncotarget. All work was completed by the student, under the supervision of Dr. Stephen Safe, in collaboration with Drs. Indira Jutooru, Gayathri Chadalapaka from Department of Veterinary Physiology and Pharmacology, Dr. J. Christopher Corton of Integrated Systems Toxicology Division, US-EPA. We would like to thank the help from Drs. Susan Hester and Brian Chorley for reviewing this manuscript.

The work presented in Section 3 is a version of Cheng et al “Role of metastasis-associated lung adenocarcinoma transcript-1 (MALAT-1) in pancreatic cancer” which is under review at Carcinogenesis. All work was completed by the student, under the

supervision of Dr. Stephen Safe, in collaboration with Drs. Parisa Imanirad, Indira Jutooru, Erik Hedrick, Un-ho Jin of Department of Veterinary Physiology and Pharmacology, Dr. Aline Rodriguez-Hoffman of the Department of Veterinary Pathobiology, Drs. Ben Morpurgo, Andrei Golovko of Institute for Genomic Medicine.

The work presented in Section 4 is reprinted from Cheng et al. "Aryl hydrocarbon receptor activity of tryptophan metabolites in young adult mouse colonocytes" published in October 2015 with the permission from Drug metabolism and Disposition. All work was completed by the student, under the supervision of Dr. Stephen Safe, in collaboration with Dr. Unho Jin of Department of Veterinary Physiology and Pharmacology, Drs. Clint D. Allred, Robert S. Chapkin of Department of Nutrition and Food Science, Dr. Arul Jayraman of Department of Chemical Engineering.

The work presented in Section 5 is reprinted from Cheng et al "Microbial-derived 1,4-dihydroxy-2-naphthoic acid and related compounds as aryl hydrocarbon receptor agonists/antagonists: structure-activity relationships and receptor modeling" which was accepted in November 2016. All work was completed by the student, under the supervision of Dr. Stephen Safe, in collaboration with Dr. Un-ho Jin of Department of Veterinary Physiology and Pharmacology; Drs. Clint Allred, Laurie A. Davidson, Robert S. Chapkin of Department of Nutrition and Food Science; Drs. Arul Jayraman, Phanourios Tamamis, Mr. Asuka Orr of Department of Chemical Engineering; Evelyn

Weaver of Department of Animal Science; and Drs. Michael S. Denison, Anatoly Soshilov of Department of Environmental Toxicology at University of California.

Funding sources

The financial assistance from the National Institutes of Health (P30-ES023512) and Texas AgriLife Research is gratefully appreciated for the work in Section 2.

The financial assistance from the National Institutes of Health National Institute of Environmental Health Sciences [P30-ES023512], Texas AgriLife Research, Sid Kyle Chair Endowmentis and TAMU-CVM graduate student core facility experiential learning program funds is gratefully appreciated for the work in Section 3.

This work in Section 4 was supported by the National Institutes of Health National Institute of Environmental Health Sciences [P30-ES023512], Texas AgriLife Research, the Sid Kyle Chair Endowment and TAMU-CVM graduate student trainee research grant.

The work in Section 5 was supported by National Institutes of Health [R01-ES025713, R01-ES007685, P30-ES023512, R35-CA19707 and R01-CA202697]; Cancer Prevention Research Institute of Texas; Texas AgriLife Research; the Office of Graduate and Professional Studies and the Artie McFerrin Department of Chemical Engineering at Texas A&M University; and the Sid Kyle Endowment.

TABLE OF CONTENTS

Page

Page

LIST OF FIGURES

Page

Page

LIST OF TABLES

Page

1. INTRODUCTION AND LITERATURE REVIEW

1.1 Cancer review

1.1.1 Cancer statistics

Cancer is a major public health problem worldwide. According to the latest world cancer report, an estimated 14.1 million people were newly diagnosed with cancer and 8.2 million people died from this disease in 2012. These estimates correspond to age-standardized incidence and mortality rates of 182 and 102 per 100000, respectively. There were 32.6 million people diagnosed with cancer in the previous five years, and the predicted global burden is expected to be over 20 million new cancer cases by 2025, compare to an estimated 14.1 million new cases in 2012 (1).

The most common cancers diagnosed in the world were lung cancer (13%), breast cancer (11.9%) and colorectal cancer (9.7%). If divided by sex, among men, the most commonly diagnosed cancer was lung (16.7%), prostate (15%) and colorectal (10%) cancer, and among women, cancers of the breast (25.2%), colorectal (9.2%) and lung (8.7%) cancers were the most prevalent. The most common causes of cancer deaths were the cancer of the lung (19.4%), liver (9.1%) and stomach (8.8%). The cancer prevalence estimated for 2012 was 32.6 million in the previous 5 years. More than 60% of cancer incidence occurred in developing countries in Africa, Asia and Central and South America, and these regions also accounted for 70% of cancer deaths worldwide (1).

The American Cancer Society updates cancer statistics in the United States (US) every year (2). Cancer is the second most common cause of death in the US, following after heart disease and in 2016, an estimated 1,685,210 new cases of cancer will be diagnosed and there will be 595,690 estimated deaths. The most abundant tumor types among new cases are cancers of the prostate (21%), lung (14%) and colon (8%) in men and breast (29%), lung (13%) and colon (8%) in women. The most common causes of cancer deaths in the US are cancers of the lung (27%), prostate (8%) and colon (8%) in men, and lung (26%), breast (14%) and colon (8%) in women. Overall, the cancer mortality rate in this country has declined due to early detection and improved therapy, however, the rates from some cancers such as pancreatic cancer remain the same or have even increased. The battle against cancer will last for a long time and more research and development is required to decrease incidence and improve treatment.

1.1.2 Theory of carcinogenesis

In 1971, President Nixon signed the National Cancer Act into law and declared the "war on cancer". More than 45 years have passed by and significant improvements in our knowledge on cancer have been achieved, however, the etiologies of this disease and identification of early biomarkers of disease have not kept pace. In this section, two different theories of carcinogenesis, namely the somatic mutation theory (SMT) and tissue organization field theory (TOFT) are discussed. The importance of theories of carcinogenesis is best supported by Einstein who once said, "It is the theory which decides what we can observe".

The essence of SMT is that cancer originates from a single cell and progress with an accumulation of genomic somatic mutations (3). In 1914, Boveri published his monograph in Germany, and until 1929, when his monograph concerning the origin of malignant tumors, was translated, his theory of somatic mutations drew attention worldwide. An oncogene is a gene whose mutation causes a gain of function leading to cancer and a tumor suppressor gene is the gene whose mutation causes a loss of function (4). The SMT theory have been supported by the development of a genetic model for colorectal tumorigenesis and in this genetic model, it is proposed that different gene mutations drive the progression of colorectal tumorigenesis (5). This model has also been used to describe the stepwise development of other cancers (6-8). Later on, the driver mutations and passenger mutations were identified, and it was postulated that small numbers of driver mutations are key mutations leading to cancer and these mutated genes are oncogenes and tumor suppressor genes. The remaining mutations are passenger mutations that do not have a causal role in carcinogenesis (9).

Hanahan et al summarized six hallmarks of cancer in 2000 (3) and ten years later two emerging hallmarks and two enabling characteristics were added, in order to rationalize the complexity of the pathogenesis of human cancer (Figure 1). Those hallmarks helped to rationalize the biological traits underlying different cancer types and these two publications have been among the most highly cited papers in the cancer research field.

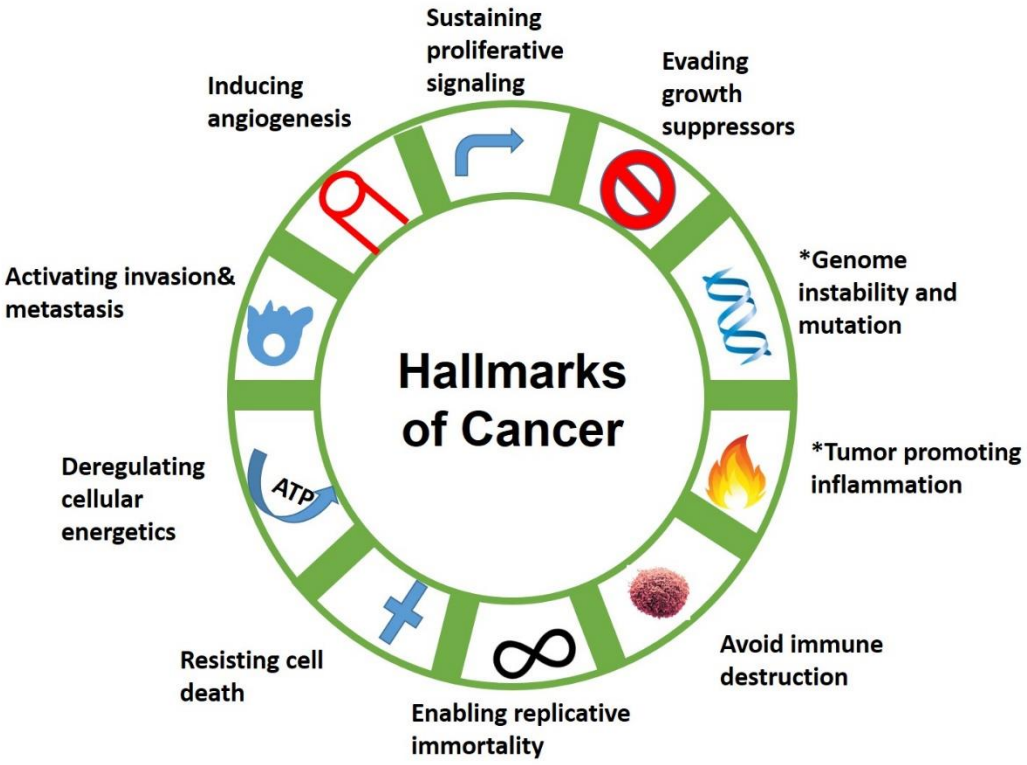


Figure 1. Illustration of the hallmarks of cancer. *enabling characteristics. Modified from (10).

1.1.2.1 Sustaining proliferative signaling

Normal cells strictly control their proliferation by regulating the production and release of growth-promoting signals and the normal proliferation is essential for cell homeostasis and function. Cancer cells sustain the proliferative signaling via different approaches. They can produce the growth factor by themselves and thereby stimulate proliferation by autocrine pathways. Stroma cells also provide difference growth factors for enhancing cancer cell growth (11,12). Cancer cells can either increase the expression of growth factor receptor on the cell surface or change the conformation of the receptor

to enhance growth factors binding which subsequently leads to uncontrolled growth. Cancer cells can also sustain proliferation signaling in a growth-factor independent pathway, such as activation of additional downstream pathways and disruption of negative feedback that attenuates proliferative signaling. For example, the B-RAF proto-oncogene, serine/threonine kinase (BRAF) is an oncogenic serine/threonine kinase which is downstream of RAS-RAF pathway, and an activation mutation of B-RAF protein lead to constitutive activation of Raf-mitogen protein (MAP)-kinase pathway (13). Phosphatase and tensin homolog (PTEN) was identified as a negative regulator of the phosphoinositide 3-kinase (PI3K) pathway by degrading phosphatidylinositol (3,4,5) trisphosphate (PIP3). Therefore inactivation of PTEN could promotes PI3K signaling which accelerates cell growth (14).

1.1.2.2 Evading growth suppressors

In addition to sustaining the proliferative signal, cancer cells have evolved to evade growth suppressors, mainly through loss of function of tumor suppressor genes. P53 is a tumor suppressor and a well-known checkpoint inhibitor necessary for repairing DNA damage in the cells (15,16). Almost 40% of cancers possess p53 mutation (17). P53 senses cellular stress, such as abnormal growth-promoting signals, genome damage such as double strand breaks and induces apoptosis or senescence which is necessary prior to DNA repair. In this way p53 functions as growth suppressors. Another well-known tumor suppressor gene is retinoblastoma (RB), which is a gate-keeper for cell cycle progression. RB protein binds and inactivated E2F transcription factors which promote

production of S stage cyclins and cyclin-dependent kinases, and inhibition by Rb blocks cells entering into the S phase and cell growth (18,19). Loss of function of RB is often observed in cancers and the inactivation of RB leads to cancer cells evading growth suppression by p53 (20).

1.1.2.3 Resisting cell death

This hallmark is complicated and controversial. There are several types of cells deaths including apoptosis, autophagy, mitotic catastrophe, necrosis, and senescence, which have been observed in different cancer types. Cancer death counters the uncontrolled cell proliferation of cancer cells, which prevents cancer progression. However, the functions of cell death in cancer are bidirectional and can either promote or inhibit cancer growth.

1.1.2.3.1 Apoptosis

Apoptosis is a natural barrier to cancer formation and is disabled in several cancers. Cancer cells have developed different strategies to evade apoptosis. P53, “the guardian of the genome” can induce apoptosis in cells with DNA damage by increasing expression of phorbol-12-myristate-13-acetate-induced protein 1 (Pmaip1 or also known as noxa) and bcl2 binding component 3 (BBC3 or known as Puma), which are pro-apoptotic proteins involved in the apoptotic pathways (21). Inactivation of p53 results in cells that escape from apoptosis. Tumor cells can also prevent cell death by upregulating anti-apoptotic factors such as bcl2 and bcl2 like 1 (bcl2l1) also known as bcl-xl). The anti-apoptotic factors function by localizing pro-apoptotic proteins in the mitochondrial

membrane to inhibit disruption of membrane integrity. The disruption of mitochondrial integrity is essential for the release of cytochrome-c and other pro-apoptotic factors into the cytoplasm (22). On the other hand, there are some controversial discoveries by the pathologist suggesting that evading apoptosis not always be considered a hallmark since increased apoptosis is usually observed simultaneously in cancer lesions and increased apoptosis in breast cancer (23,24) is associated with a poor clinical outcome. In addition, overexpression of anti-apoptotic protein BCL2 is linked to the better overall survival of cancer patients, and conversely, caspase3 and other apoptosis-inducing proteins have been found to promote tumorigenesis (25).

1.1.2.3.2 Autophagy

Autophagy is a physiological reaction of the cell in response to cellular stress such as hypoxia and insufficient nutrition. During autophagy, cellular structures and organelles are broken down and degraded and the catabolic products are then used for biosynthesis and energy metabolism to support cellular needs during stressful conditions (26,27). Autophagy plays a role in both tumor cell growth and death. For example, autophagy suppresses tumor growth in the early stages of tumorigenesis whereas treatment with radiation therapy and cytotoxic drugs often induces autophagy, which can protect cancer cells, and thereby plays a role in radiation and drug resistance. In addition, tumor dormancy can be induced by autophagy, and the growth-arrested tumor can become functional upon more favorable and this results in tumor relapses (28-30).

1.1.2.3.3 Senescence

Senescence was first described in 1961 as a state of cells that exit cell cycle and senescence is important not only for normal development but also tumorigenesis. Activation of senescence typically depends on the induction of P53 or RB pathways. Senescence also plays a dual role in tumorigenesis. Oncogene-induced senescence is a barrier for tumor progression (31), whereas the senescence-associated secretory phenotype (SASP), the dark side of tumor progression, transforms senescent fibroblasts in the tumor microenvironment into pro-inflammatory cells capable of promoting tumor progression (32).

1.1.2.3.4 Mitotic catastrophe

Mitotic catastrophe is an intrinsic oncosuppressive mechanism, which can sense the mitotic failure, DNA damage, and cytokinesis failure and drive the defective cells into apoptosis or senescence. Therefore, failed mitotic catastrophe promotes the sustained proliferation of defective cells and promotes tumor cell growth (33).

1.1.2.3.5 Necrosis

Necrotic cell death releases pro-inflammatory signals to the surrounding environment and recruits inflammatory cells to eliminate necrotic debris. It has been shown that necrosis promotes tumor growth by releasing pro-survival inflammatory cells and pro-proliferation regulatory factors, such as interleukin 1 alpha (IL1A). Dying cells also release a number of factors such as high mobility group B1 protein (HMGB1). HMGB1

which is a ligand for the receptor for advanced glycation end products (RAGE), and this can lead to increased cancer cell growth, motility, invasion, angiogenesis and metastasis (34-37).

1.1.2.4 Enabling replicative immortality

The unlimited replicative potential of cancer cells is associated with telomeres that are hexanucleotide repeats in tandem on the ends of every chromosome. After a limited number of cell divisions, normal cells undergo senescence and enter the quiescent state. Each cell division progressively erodes and shortens the telomeres on the chromosome. Complete loss of telomeres leaves the ends of chromosomes unprotected and predisposes the DNA to form end-to-end fusions with other chromosomes leading to genomic instability and cell death. In contrast, cancer cells are immortalized by preventing the shortening of telomerase by upregulating telomerase (hTERT) enzyme. Cancer cells express high levels of telomerase adding telomere repeat segments to the ends of DNA after each replicative cycle, supporting the immortality of cancer cells (10,38).

1.1.2.5 Inducing angiogenesis

Angiogenesis is a term describing the alteration in blood vessel growth associated with tumors following hypoxic stress. The formation of blood vessel provides tumor cells with nutrients and oxygen and is crucial for tumor growth and progression. Folkman and coworkers observed that tumors without vasculature can only grow up to ~4mm, and once reaching that size, the nutrients and oxygen provided by the surrounding peripheral

blood vessels will not be sufficient to support further tumor growth (39). In cancer cells, the angiogenic switch is turned on and the balance of angiogenesis inducers and inhibitors shifts into the pro-angiogenesis status. Factors regulating angiogenesis can be stimulatory or inhibitory. For example, the overexpression of vascular endothelial growth factor (VEGF) and/or fibroblast growth factors (FGF) is observed in many cancers compared to normal tissues, whereas the angiogenic inhibitors such as thrombospondin-1 or interferon beta are decreased in cancer cells compared to normal cells (40,41).

1.1.2.6 Activating invasion and metastasis

The mortality of cancer patients is usually caused by the metastasis of the primary tumor into lymph nodes or other organs. This process is complex and includes activation of extracellular proteases and cell to cell adhesion molecules such as cadherins and integrins which allow cancer cells the potential to become invasive and metastasis. Epithelial-mesenchymal-transition (EMT) converts polarized, immotile epithelial cells to migratory mesenchymal cells and has been shown to promote cancer metastasis, drug resistance and increased tumor recurrency (42,43). There have been debates with regard to the necessity of EMT in metastasis after the appearance of publications stating that EMT is not required for lung and pancreatic cancer metastasis (44,45) and the heterogeneity and plasticity of EMT should be taken into consideration (46,47). Cells undergoing EMT express low levels of epithelial genes (such as E-cadherin) and increased expression of mesenchymal genes (such as vimentin and N-cadherin). Loss of

E-cadherin is considered to be a hallmark of EMT (43,48). The crosstalk between the neoplastic cells and stromal cells is necessary for invasion and metastasis. Cancer cells release stimulatory signals to the surrounding tumor stroma and in response, the cancer-associated fibroblast located in the stroma secrete chemokines such as C-C motif chemokine ligand 5 (CCL5) that stimulate invasive behavior in cancer cells (49,50).

1.1.2.7 Reprogramming energy metabolism

Besides the deregulated control of cell proliferation, adjustment of energy metabolism is also an enabling characteristic of the tumor. Warburg effect is considered as a metabolic hallmark for most cancer cells. It was observed that cancer cells convert excessive amounts of glucose into lactate even under the aerobic conditions. Although the efficiency of adenosine triphosphate (ATP) production by glycolysis is low compared to mitochondrial oxidative phosphorylation, cancer cells find ways to maximize energy generation, partially by increasing glucose transporters, which in turn upregulates glucose import into the cytoplasm (10,34-37). Oncogenes, such as RAS, MYC, and mutated tumor suppressors are closely associated with glycolytic fueling (51-53). In 2016, Thompson and coworkers recently discussed six hallmarks of cancer-associated metabolic changes, namely, deregulated uptake of glucose and amino acids, use of opportunistic models of nutrient acquisition, use of glycolysis/ tricarboxylic acid (TCA) cycle intermediating biosynthesis and NADPH production, increased demand for nitrogen, alterations of metabolite-driven gene regulation, and metabolic interactions

with the microenvironment. While few tumors display all six hallmarks, most tumors display several of them (54).

1.1.2.8 Evading the immune system

The immune system is a natural barrier against tumorigenesis. Both innate and adaptive immune systems have a substantial role in tumor inhibition. The mice with combined deficiencies of natural killer (NK) cells and CD8+ cytotoxic T lymphocytes (CTLs), CD4+ Th1 helper T cells exhibit higher tumor incidence compared with mice devoid of T cells or NK cells individually (55,56). Colon and ovarian cancer patients with tumors heavily infiltrated with CTLs and NK cells exhibit better prognosis than patients expressing lower levels of these lymphocytes (57).

1.1.2.9 Genome instability and mutation

Genome instability and mutations are the enabling characteristics of cancer which help maintain the hallmarks of cancer. There is a proposed model indicating that genes driving cells proliferation induce DNA replication stress, which in turn generates genomic stability and selects cells for escape from apoptosis (58). Cancer cells also increase rates of mutation by increasing the sensitivity of cells to mutagenic agents, through breaking down one or more components of the machinery that maintains the genomic stability (10).

1.1.2.10 Tumor promoting inflammation

Another enabling characteristic of cancer is inflammation, which plays a crucial role in promoting tumorigenesis (59). Inflammation supports the sustained proliferation of tumor cells by enriching the tumor microenvironment with growth factors. In addition, inflammation can provide tumor cells with proangiogenic factors and matrix metalloproteinases which facilitate EMT, tumor invasion and metastasis (60). Moreover, the activated inflammatory cells generate reactive oxygen species (ROS) and reactive nitrogen intermediates (RNI) that inducing mutations contributing to genome instability (61).

In summary, 8 hallmarks and 2 enabling characteristics of cancers were discussed by Hanahan and Weinberg (10). Many studies have been carried and ongoing discover more hallmarks of cancer. West and the coworkers identified cancer system hallmarks, namely, network entropy is increased in cancer compared to the normal phenotype and differentiated network entropy is anti-correlated with differential gene expression. In addition, cancer cell dormancy is also proposed to be another hallmark. The ability of cancer cells to become quiescent may be critical for evolving malignancies with implications for understanding cancer initiation, progression and treatment resistance. Cancer cell dormancy is a dynamic cells state conferring a fitness advantage to an evolving malignancy under stress (62). Alternative splicing changes proteins encoded by mRNAs and these proteins may have distinct properties. Over 94% of human genes are alternatively spliced (63) and alternative splicing emerges as an essential element in

gene regulation in almost every cellular functions and aberrant alternative splicing is increasingly linked to cancer. The adenomatous polyposis coli (APC) gene is critical for the change of epithelium to the hyperplastic epithelium, which is involved in the regulation of cell cycle, apoptosis, cell adhesion and cytoskeletal architecture (64), and several alternative splice isoforms of APC have been discovered. One of the APC transcripts contains an additional exon termed exon 1A and this exon 1A which expressed at 3.5 fold higher levels in colon cancer compared to normal mucosa. The inclusion of exon 1A results in less production of protein because exon 1A leads to an extra stop codon in exon 2 (65).

DNA mutations, altered epigenetic profiles and deregulated signaling pathways can be interwoven among different hallmarks. For example, telomerase is primarily discussed under the hallmark “unlimited replication capacity”. It is proposed that telomerase is the central regulator of all the hallmarks of cancer (38). The concept of hallmarks is a good reference for organizing the vast array of cancer data. However, it is important to remember that cancer is a dynamic evolving system and these proposed hallmarks are not applicable to all cancer cells (66).

SMT has been the dominant view in cancer research field, however, with the emergence of the “omics” era, SMT is facing a crisis, and many scientists have questioned that adequacy of SMT and has been highlighted by Baker and coworkers (67), and these include non-genotoxic-dependent carcinogenesis and the lack of mutations in some

tumors. Some scientists have suggested that TOFT provides a better explanation of these observations. The TOFT suggests that the hallmarks of cancers are primarily due to altered tissue organization (68). Tumors are more complex and less organized than normal tissues and TOFT is centered on the tissue level of biological organization. The default state of the cell is proliferation and motility, and the target of carcinogens is tissue instead of individual genes and thereby carcinogens disrupt the reciprocal interactions among cells and tissues. The TOFT focuses on the reciprocal relations among cells and tissues and the key point of TOFT is that tissue/cell becomes uncontrolled. The view of TOFT is that mutations are not necessary for carcinogenesis and gene instabilities causing mutations are the byproducts of carcinogenesis and cancer can arise in tissues not exposed to carcinogens (67). In contrast, in the SMT model, the default state of the cell is quiescence and abnormal proliferation is promoted by inactivation of tumor suppressor genes and/or activation of oncogenes. Another premise is that cells are in a quiescent state, and motility should also be actively triggered (3). Figure 2 illustrates the difference between SMT and TOFT. In the SMT, carcinogens lead to mutations in the epithelial cells and this subsequently results in the excessive proliferation of epithelial cells. Factors secreted by the growing tumor will alter the stroma and altered stroma will, in turn, promote and support tumor progression. On the other hand, the TOFT model states that carcinogens will hit stroma and lead to altered cell-ECM and cell-cell interactions in the epithelium and in the end result in tumor formation in the epithelium (69).

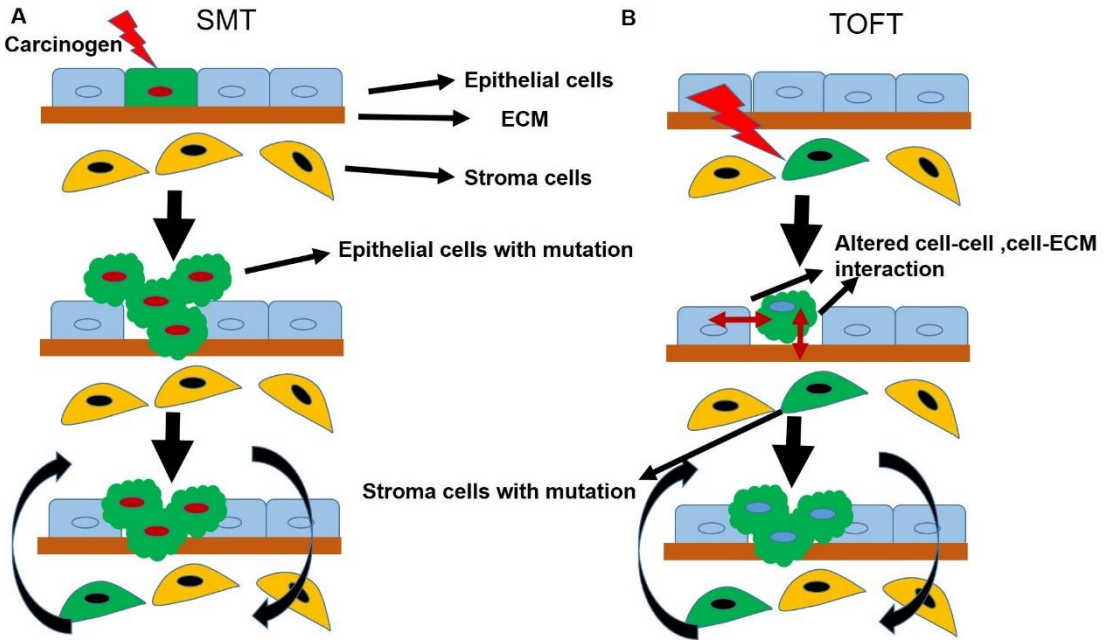


Figure 2. Illustrations of SMT and TOFT models. A) In the SMT, carcinogens lead to the mutation of epithelial cells. The epithelial cells with advantageous mutations propagate and form tumor. The tumor secretes paracrine factors leading to the alteration of stroma and the altered stroma in turn support tumor progression. B) In the TOFT model, carcinogens affect the stroma and lead to the altered cell-ECM and cell-cell interactions in the epithelium, resulting in tumor formation in the epithelium. Modified from (69).

It has been debated whether SMT and TOFT are compatible and if the two theories should be combined (68,70-73). SMT is still the dominant cancer theory and hallmarks of cancer have been used as a basis for developing new therapies for cancer treatment and these include genes /pathways as therapeutic targets. However, the minimal overall benefits of targeted therapies suggest that current strategies need to be improved. The increasing observations and accumulation of omics data seriously question the validity of SMT indicating that SMT cannot be used to explain everything about carcinogenesis, but this does not mean that SMT should be abandoned.

Both SMT and TOFT are theories, not facts, and the debate and discussion of the two theories will be a motivation for further exploration of the origins of cancer with the ultimate goal of curing this disease. The war on cancer is still on.

1.1.3 Pancreatic cancer

In the ten leading cancer types, pancreatic cancer is the fourth-leading cause of death in the US for both males and females. Pancreatic cancer is also the 8th among 10 cancer types for the estimated development of new cancer cases. It is estimated that 53,070 new cases (27,670 men and 25,400 women) will be diagnosed with pancreatic cancer and 41,780 people (21,450 men and 20,330 women) will die of this disease in 2016 (2). Pancreatic cancer is considered as a disease of the elderly since very few cases are diagnosed before the age of 40 and the median age of diagnosis is between 60-75 year of age (74).

The pancreas, a gland situated behind the stomach and in front of the spine, consists of endocrine and exocrine compartments. The endocrine pancreas makes and secretes insulin, glucagon, somatostatin and pancreatic polypeptide into the blood, and this portion of pancreas accounts for 1-2% of the pancreatic mass. The exocrine pancreas, comprising more than 95% of the pancreas, includes acinar and duct cells with associated connective tissue, vessels, and nerves. The exocrine pancreas primarily makes and secretes digestive enzymes (75).

Pancreatic adenocarcinoma (PDAC) accounts for 90% of pancreatic cancer with a five-year survival rate of 5%. Pancreatic neuroendocrine tumors are less common (1-2%) with a much improved 5-year survival rate of 65%. Other cancer types include cystic neoplasms (intraductal papillary, mucinous neoplasms, mucinous cystic neoplasms, serous cystadenoma, solid-pseudopapillary neoplasms) and acinar carcinoma (74).

1.1.3.1 Pancreatic cancer progression model and stages classification

A proposed model for pancreatic tumor formation is based on pathology and altered gene expression data derived from analysis of human pancreatic tumor specimens (76,77). The pervasive lesions, termed as pancreatic intraepithelial neoplastic lesions (PanINs) progress through different stages (from PanIN1 to PanIN2 to PanIN3) and the difference stages of PanINs harbor distinct nuclear and structural differences (76). The normal ductal pancreatic epithelium is a single layer of low columnar epithelium with amphophilic cytoplasm, and the characteristics of progressing PanINs are depicted in Table 1. Accompanied by the differences in histology changes, the underlying genetic changes have also been investigated (Figure 3). In the proposed genetic progression model, KRAS mutations are followed by subsequent inactivation of the tumor suppressor gene cyclin-dependent kinase inhibitor 2A (CDKN2A) (or INK4A/ARF, which encodes p14/Arf and p16/Ink4A), followed by inactivation of p53 (TP53) and SMAD4 (78).

Table 1. Different stages of PanIN development.

Stages	Histology characteristics
PanIN-1A	Columnar, mucin-producing ductal epithelium, elongated
PanIN-1B	Development of papillary architecture
PanIN-2	Nuclear atypia: nuclear enlargement, partially loss of polarity, crowding
PanIN-3	Marked cytologic atypia, complete loss of polarity

The mutant KRAS oncogene is expressed in 90% of PDAC (79,80) and KRAS mutations are considered an early event in the progression from normal pancreatic cells to invasive carcinoma (81,82) and the mutation is found in 36, 44.5 and 87% of PanIN-1a,1b, and 2-3 lesions, respectively (81). KRAS is located on chromosome 12 and encodes a small GTPase, which can promote cell growth and proliferation. The most common KRAS mutation is at codon12 and its mutation affects the GTPase activity by blocking it from going into the inactive state. The constitutively activated KRAS interacts to GTP and produces uncontrolled growth signals and continuous KRAS signaling is essential for tumor initiation, progression and maintenance (83,84).

Another frequently mutated gene is p16/CDKN2A, which is mutated in 82% of PDAC (85,86) and the mutation usually, occurs in the early or intermediate grade PanIN. P16/CDKN2A is a tumor suppressor gene that prevents cell progression through G1 cell cycle checkpoint by preventing phosphorylation of the retinoblastoma protein, which in turn affects the downstream inhibition of E2F transcription factor-mediated gene expression (87).

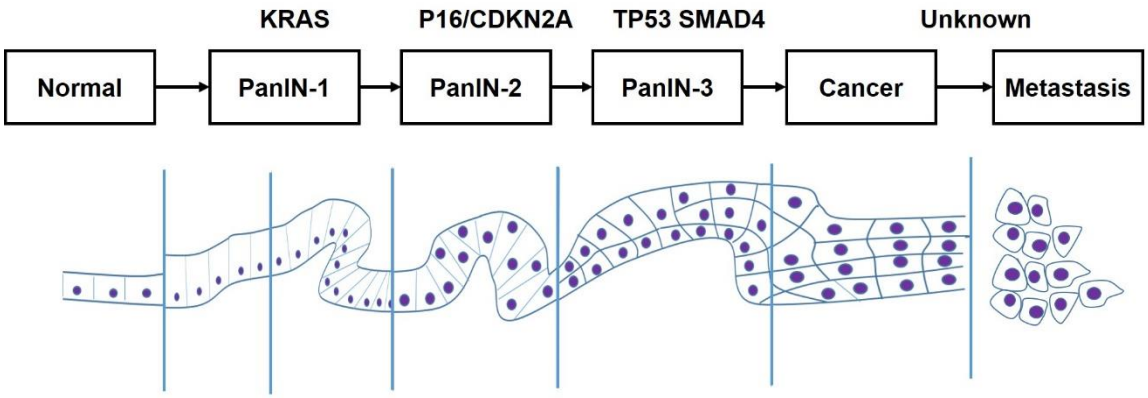


Figure 3. The schematic model of pancreatic tumor progression.

P53 is the most common somatic mutation found consistently in various human cancers and p53 mutations are detected in over 50% of PDACs. P53 modulates the cellular response to cytotoxic stress, DNA damage, and/or hypoxia via either cell cycle arrest or induction of apoptosis. In the pancreatic cancer progression model, p53 occurs at the PanIN3 stage and loss of P53 has been associated with abnormal mitoses and severe nuclear abnormalities (88). P53 is considered as a progression factor rather than an inductive agent. SMAD4 is another tumor suppressor inactivated in more than 50% of pancreatic cancer cases (89) and is associated with poor diagnosis (90,91). SMAD4 is downstream of transforming growth factor β (TGFβ) superfamily-induced signals and is important for gene transcription and growth arrest. It was reported that loss of SMAD4 could accelerate the process of the activated KRAS-induced neoplasms into malignancy (92).

Table 2. Categories of TNM system (93).

T		N		M	
TX	The main tumor cannot be assessed	NX	The regional lymph nodes cannot be assessed	M0	The tumor has not spread to other parts of the body
T0	No evidence of a primary tumor	N0	The tumor cannot be found in the nearby lymph nodes	M1	The tumor has spread to other parts of the body
Tis	Carcinoma in situ	N1	Cancer has spread to nearby lymph nodes		
T1	The tumor growth is limited in the pancreas and the diameter is 2 cm or less				
T2	The tumor growth is limited in the pancreas and the diameter is more than 2 cm				
T3	The tumor has grown beyond the pancreas but not spread into major blood vessels and nerves				
T4	The tumor has grown beyond the pancreas and spread into major blood vessels and nerves				

Defining the stages of pancreatic cancer is important for choosing treatment options and predicting the efficacy of treatment. The most often used staging system for pancreatic cancer is the American Joint Committee on Cancer (AJCC) TNM system, where T represents the size of the primary tumor and whether it has grown beyond the pancreas and spread into other sites of the body. N indicates the spread of tumor to nearby lymph nodes and M delineates the metastasis of tumor to other parts of the body. Table 2 provides the different categories of T, N, and M respectively. The higher the number, the

more severe the condition is. Based on the different categories of individual T, N, M, the stages of pancreatic cancer can be determined (Table 3) (93). Besides the TNM system, the clinicians often use a simpler staging system, which divides cancers into resectable, borderline resectable and unresectable. Resectable tumors often include stage IA, IB and IIA, and the entire tumor can be surgically removed. The borderline resectable tumor is the one which has just reached the blood vessels and there is a possibility that the entire tumor can be removed. Lastly, the unresectable tumors are the ones either locally advanced or metastatic, and this category often includes the stage IIB and most III in the TNM system). At this stage, the entire tumor is unlikely to be removed by surgery. Understanding these two staging systems is important to determine the most suitable treatment for pancreatic cancer patients.

Table 3. Stages of pancreatic cancer using TNM system (93).

Stage	Stage grouping	Stage description
0	Tis, N0, M0	The tumor is limited to the top layers of cells in the pancreas and has not spread beyond the pancreas.
IA	T1, N0, M0	The tumor is limited in the pancreas and is 2cm across or less. The tumor has not spread beyond the pancreas.
IB	T2, N0, M0	The tumor is limited in the pancreas with size larger than 2 cm across. The tumor has not spread beyond the pancreas.
IIA	T3, N0, M0	The tumor has grown beyond the pancreas but not into major blood vessels or nerves. The tumor has spread to nearby lymph nodes.
IIB	T1-T3, N1, M0	The tumor has spread to nearby lymph nodes and is either confined to the pancreas or growing outside the pancreas.
III	T4, Any N, M0	The tumor has invaded into the major blood vessels or nerves but has not metastasized.
IV	Any T, Any N, M1	The tumor has spread to distant organs.

1.1.3.2 Experimental models for studying pancreatic cancer

Research on cancer extensively utilizes different cell culture and in vivo models and for pancreatic cancer, this includes xenograft & orthotopic immunodeficient mice and several transgenic mouse models.

Cancer cell lines and primary cells in culture are excellent models for mechanistic studies (94). The phenotype and genotype of different pancreatic cancer cell lines have been thoroughly characterized. There are studies which linked the differential expression of specific proteins in the cell lines to the tumor growth, migration and invasion (95,96), as well as drug resistance (97). For example, CXCR4 expression was evaluated in 11 different PDAC cell lines and six of 11 cell lines expressed CXCR4 and among them, three cell lines (Aspc1, Capan1, and Hs766T) had high expression of CXCR4 transcripts. They discovered that CXCR4 expression was higher in cell lines driven from metastatic lesions compared to those derived from primary tumors. They concluded that CXCR4 receptor is frequently expressed in metastatic pancreatic tumor cells, and promote survival and migration/invasion (96). However, there are limitations of using cell line for study pancreatic cancer. The long-term culturing of cell lines in the laboratory may cause cross-contamination, mycoplasma growth, and outgrowth of subtypes, and this may help explain the discrepancy of reports from different labs using the same cell lines. Primary cell cultures are also used but not as frequently as cell lines because of the short maintenance of primary cells in vitro. Primary cells maintain many important markers and functions close to in vivo when compared to cell lines (98,99).

Organoids culture is the advanced form of primary cell culture. The organoids are entirely made up of pancreatic ductal cells, without the contamination of surrounding cell types. The organoids grow within special medium and can be transplanted subsequently into mice, where they fully resemble pancreatic cancer. The organoids emerge as a good model for studying early stages of tumor progression. The researchers can grow cells from normal pancreas and diseased pancreatic tissues in vitro. By comparing the organoids from normal pancreas and pancreatic cancers, we gain more insight into the molecular pathways correlating with disease progression. In addition, organoids can be generated rapidly from tiny needle biopsies from pancreatic cancer patients and those patient-derived organoids will be a good platform to investigated new therapeutic regimens (100). Xenograft mouse models include heterotopic and orthotopic xenograft models (101). Studies with subcutaneous heterotopic mouse models can be readily carried out and are low cost, however, the results of various therapeutic regimens in the mouse model may not necessarily predict effects on human patients. For example, 9-AC is drug showing curative potential in human colon cancer xenografts, and subsequently in other tumor types (102-104), however, 9-AC only showed a modest effect in ovarian cancer and malignant lymphoma, and was not active against lung cancer or colon cancer in human trials (105). Orthotopic xenograft mouse models inject or implant human cancer cell lines directly into the same organ from which the tumor originated. Thus the orthotopic xenograft models may closely resemble the human patient's tumors, however, the immunodeficient mouse models of orthotopic implantation do not completely mimic the cancer patient due to differences in the mouse

vs human tumor microenvironment, which is an important factor in tumor progression and metastasis. Tumor microenvironment or stroma consists of extracellular matrix components, cancer-associated fibroblast, vascular and immune cells. Stroma supports the tumor growth and promotes metastasis and can also serve as a barrier against drug delivery (106). The lack of stroma in the orthotopic mouse model somewhat distracts from the relevance of this model for mechanistic and drug efficacy studies.

1.1.3.2.1 Genetic modified mouse models for pancreatic cancer

Genetically modified mouse models (GEMMs) have been widely used for studies various cancers because of the high degree conservation in genes between the mouse and human. In addition, the mouse genome has been fully sequenced and the genome is relative easy to be manipulated. Mouse models can help us understand the genetic alterations in cancer cells behave during tumor formation and progression. Moreover, understanding the interactions of environmental factors with the genome, and finally the study of GEMMs can help develop strategies for detecting cancers at early stages and thereby improve the outcomes of various therapies (107).

The first GEMMs for cancer were generated by overexpressing viral and cellular oncogenes in specific tissues. The first GEMM for breast cancer is generated by replacing large portions of the Myc gene promoter with a hormonally inducible mouse mammary tumor virus promoter which leads to mammary tumor formation in the pregnant mice (108). Another GEMM generated in the same year using plasmids

containing the SV40 early region genes and a metallothionein fusion gene which were injected into mouse eggs, and a high percentage of the offsprings developed the brain tumors (109). An autosomal dominant mutation, multiple intestinal neoplasia (Min), is a nonsense mutation in the murine Apc gene. The C57BL/6L Apc min strain is a widely used mouse model for spontaneous intestinal adenomas. C57BL/6L Apc min heterozygous mice fed a 15% fat diet generally develop more than 30 adenomas throughout the small and large intestinal tract and subsequently die of anemia or obstruction at approximately four months of age (110). The transgenic adenocarcinoma of the mouse prostate (TRAMP) model is widely used mouse model for studying prostate cancer and Male TRAMP mouse spontaneously develop prostate cancer after the puberty. In TRAMP mice the expression of oncoprotein, SV40T antigen (TAg) is transcriptionally controlled by the rat probasin promoter and the specific probasin promoter ensures the TAg is only expressed in prostate epithelial cells (111,112).

Conditional gene expression in the mouse allows for the expression of a gene in a time and space-dependent manner and the most common strategy used is the Cre-loxP system. Cre is a recombinase mediating DNA recombination between two loxP sites and its expression can be driven by promoter of genes that express in certain types of cells. The loxP sequences are introduced into the specific gene locus in the mouse. The loxP mice can be crossed with transgenic mice expressing Cre recombinase to generate tissue-specific conditional knockout mice. One advantage of the Cre-loxP mouse model is to prevent the embryo lethality induced by total knockout or knockdown a specific gene in

all tissues. For example, Brca1 knockout (KO) mice die during embryonic development, however when Brca1 is conditionally knocked out only in mammary gland epithelial cells, they are viable but develop mammary tumors (113). The Cre-loxP system can also be used to conditional activate an oncogene by using lox-stop-lox sequence. For example, when the loxP-stop-loxP-krasG12D is introduced into the endogenous kras locus, oncogenic kras only expressed upon the Cre-mediated removal of STOP sequence (107,114).

Table 4 summarizes the current mouse models for PDAC and those models have been widely used in the study of pancreatic cancers. Active mutated KRAS is detected in over 90% of PDAC and early research on generating PDAC mouse models focused on the manipulation of KRAS (115). The first model faithfully reproducing PDAC observed in humans was developed in 2003, and expresses KRASG12D which is activated by the Pdx1-Cre or Pft1a/P48-Cre (116). Pdx1 is homeodomain protein and both exocrine and endocrine cells of pancreas develop from Pdx1 expressing progenitor cells and the PDX1 expression persists postnatally in the exocrine component of the pancreas (117). PFT1A/P48 is expressed later than PDX1 and is essential to commit cells to pancreatic fate (118). The Pdx1-Cre; LSL-KRASG12D and Ptf1a-Cre; LSL-KRASG12D mice develop the PanIN lesions that are indistinguishable from those in human patients (116). Addition mutations in tumor suppressors can be applied to accelerate the progression of those PanIN lesions to invasive PDAC. P53 knockout mice are viable indicating that P53 is dispensable for mouse embryonic development (119). The LSL-KRASG12D

mice develop invasive PDAC after 35-70 weeks, and when crossed with homozygous p53 knockout alleles, the mice develop PDAC between 11 and 25 weeks of age (median survival 15 weeks), and when crossed with heterozygous p53 knockout alleles, the mice develop PDAC more rapidly with a median survival of 7.9 weeks (83). In addition to P53 knockout mouse models, the knockin mouse models of p53 R172H and R270H have been generated by different groups. There is evidence showing that accumulation of P53 mutants R172H and R270H increases the incidence of metastasis of osteosarcomas and epithelial carcinomas (120,121), whereas p53 knockout mice rarely develop metastases. For example, the cooperation of P53R172H and KRASG12D promotes chromosomal instability and induces widely metastatic PDAC in mice (122).

There are also other mouse models generated on the basis of the KRAS derived mouse model. A novel immunocompetent mouse model with robust stroma has been generated to serve as a valuable tool for preclinical evaluation of new therapies (123). This mouse model recapitulates the tumor stroma and immune microenvironment, moreover, a bioluminescent mouse model can be used to monitor the proliferation of early stages of pancreatic cancer. The MITO-Luc mouse expresses luciferase in active proliferating cells and proliferation events can be visualized non-invasively in mice by bioluminescence imaging. The MITO-Luc mouse crossed with Pdx1-Cre; LSL-KRASG12D (KC) and Pdx1-Cre; LSL-KRASG12D, TRP53R172H/+ (KPC). The results obtained with these mice show that abnormal proliferation events take place in the early stages of pancreatic carcinogenesis prior to tumor formation (124).

Table 4. Genetic engineered models for pancreatic cancer.

Mouse model	Survival (median)	References
$KRas^{LSLG12D/+}$; $Ink4a/Arf^{lox/lox}$; Pdx1-Cre	2 months	(125)
$KRas^{LSLG12D/+}$; $Ink4a/Arf^{-/-}$; Pdx1-Cre	18.3 weeks	(126)
$KRas^{LSLG12D/+}$; $Trp53^{lox/lox}$; Pdx1-Cre	6.2 weeks	(126)
$KRas^{LSLG12D/+}$; $Ink4a/Arf^{-/-}$; $Trp53^{lox/lox}$; Pdx1-Cre	7.2 weeks	(126)
$KRas^{LSLG12D/+}$; $Trp53^{R172H/+}$; Pdx1-Cre	5 months	(122)
$KRas^{LSLG12D/+}$; $Smad4^{lox/lox}$; Pdx1-Cre	13.1 weeks	(127,128)
$KRas^{LSLG12D/+}$; $Ink4a/Arf^{lox/lox}$; $Smad4^{lox/lox}$; Pdx1-Cre	7.4 weeks	(127)
$KRas^{LSLG12D/+}$; $Ink4a/Arf^{lox/lox}$; $Smad4^{lox/lox}$; Ptf1a-Cre	8.8 weeks	(127)
$KRas^{LSLG12D/+}$; $Tgfbr2^{lox/lox}$; Ptf1a-Cre	59 days	(129)
$KRas^{LSLG12V/+}$; Elastase-tTA/tetO-Cre	>12m	(130)
$KRas^{LSLG12V/+}$; $Trp53^{+/-}$; Elastase-tTA/tetO-Cre	6m	(130)
$KRas^{LSLG12D/+}$; Elastase-TGFa; Ptf1a-Cre	7m	(131)
$KRas^{LSLG12D/+}$; $Smad4^{lox/lox}$; Ptf1a-Cre	15.7 weeks	(92,127)
$KRas^{LSLG12D/+}$; $Trp53^{R270H/+}$; Pdx1-Cre	168 days	(132)
$KRas^{LSLG12D/+}$; $Trp53^{R270H/+}$; $Brca2^{Tr/+}$; Pdx1-Cre	143 days	(132)
$KRas^{LSLG12D/+}$; $Trp53^{R270H/+}$; $Brca2^{Tr/lox}$; Pdx1-Cre	84 days	(132)
$KRas^{LSLG12D/+}$; Pdx1-Cre	NA	(132)
$KRas^{LSLG12D/+}$; $Brca2^{Tr/+}$; Pdx1-Cre	NA	(132)
$KRas^{LSLG12D/+}$; $Brca2^{Tr/lox}$; Pdx1-Cre	NA	(132)
$KRas^{LSLG12D/+}$; $Brca2^{lox/lox}$; Pdx1-Cre	300 days	(133)
$Lkb1^{lox/lox}$; Pdx1-Cre	68 days	(134)
$KRas^{LSLG12D/+}$; $Lkb1^{lox/+}$; Pdx1-Cre	141 days	(134)
$KRas^{LSLG12D/+}$; $p21^{+/-}$; Pdx1-Cre	75 days	(134)
$KRas^{LSLG12D/+}$; $Notch1^{lox/lox}$; Ptf1a-Cre	336 days	(135)
$KRas^{LSLG12D/+}$; $Notch2^{lox/lox}$; Ptf1a-Cre	521 days	(135)
$KRas^{LSLG12V/+}$; $Trp53^{lox/lox}$; Elastase-tTA/tetO-Cre	NA	(136)
$KRas^{LSLG12V/+}$; $Ink4a/Arf^{lox/lox}$; Elastase-tTA/tetO-Cre	~8m	(136)
$KRas^{LSLG12D/+}$; $Usp9x^{lox/+}$; Pdx1-Cre	NA	(137)
$KRas^{LSLG12D/+}$; $Trp53^{lox/lox}$; $Brca1^{lox/lox}$; Pdx1-Cre	40 days	(138)
$KRas^{LSLG12D/+}$; $Rb^{lox/lox}$; Pdx1-Cre	10 weeks	(139)
$KRas^{LSLG12D/+}$; Ptf1a-Cre	NA	(116)
$KRas^{LSLG12D/+}$; MUC1.Tg; Ptf1a-Cre	NA	(140)
$iKRas^{LSLG12D/+}$; $Trp53^{lox/+}$; Ptf1a-Cre	15 weeks	(83)
$iKRas^{LSLG12D/+}$; $Trp53^{lox/lox}$; Ptf1a-Cre	7.9 weeks	(83)

1.1.3.3 Biomarkers for pancreatic cancer

Most pancreatic cancers (96%) originate in the exocrine cells with no obvious signs of tumor initiation and this makes it difficult for early detection of this disease. Identification of early biomarker that will predict future cancers will be invaluable for the early diagnosis, prognosis, and development of effect therapies (141,142). Because of the asymptomatic nature of pancreatic caners in the early stages, less than 20% of PDAC patients are diagnosed with localized resectable cancer that is amenable to therapy and surgical resection.

A significant amount of research on identification of non-invasive and highly sensitive and specific biomarkers in blood, urine, stool, saliva or pancreatic juice has been carried out. Some potential biomarkers include altered gene, protein expression, and DNA methylation patterns. Carbohydrate antigen 19-9 (CA19-9) is a well-established FDA-approved biomarkers and early detection of CA19-9 is a prognostic marker for survival after resection surgery (143,144). However, CA19-9 is not a particularly sensitive biomarker for diagnosis of early pancreatic cancer (145), and one limitation of CA19-9 is that it may be elevated in nonmalignant conditions such as pancreatitis and acute cholangitis. Christer and coworker identified a 25-serum biomarker signature, composing of different cytokines and complement proteins (C1, C3, C5, CD40, eotaxin, GM-CSF, IgM, IL-11, IL-12, IL-16, IL-1a, 1L-1ra, IL-2, IL-3, IL-4, IL-7, Integrin a-10, MCP-1, MCP-3, Mucin-1, properdin, TGFB1 and TNF-A, TNF-B, and VEGF). They screened 148 patients with pancreatic cancer, chronic pancreatitis, autoimmune

pancreatitis (AIP) and healthy people. Using this 25-serum biomarker, they can differentiate pancreatic cancer patients from people who are healthy or have benign pancreatic diseases with a high diagnostic potential (AUC of 0.88)(146). Victoria and coworkers have also identified a unique panel of cytokines that can be used along with CA19-9 to enhance the ability to distinguish between pancreatic cancer and benign pancreatic disease (147).

Non-coding RNAs (ncRNAs) have been shown to play a role in pancreatic cancer progression and abnormal level of ncRNAs in serum have been observed in PC patients (148,149). For example, salivary miR-3679-5p and miR-940 are reliable biomarkers for distinguishing resectable pancreatic cancer patients from health control and patients with benign pancreatic diseases with sensitivity and specificity of 70% (150). Long noncoding RNAs also have the potential to be non-invasive biomarkers. For example, the expression of HOTTIP-005 and RP11-567G11 and their fragments in the plasma is significantly higher in patients with PC compared to healthy people ($p<0.0001$) (151).

Pancreatic cancer often metastasizes to liver, lung and skeletal system, indicating that pancreatic tumor cells are able to circulate to distant organs. Circulating tumors cells (CTCs) are cells that detached from the original tumor site and enter the blood circulation either through passive shedding of tumor cells (152) or active transport involving the epithelial-to-mesenchymal transition (153). There is evidence that the presence of CTCs correlates with a poor survival rate in pancreatic cancer patients

(154,155). In a study in Tokyo, the median survival times (MSTs) of CTC-positive patients were 110.5 days compare to 375.8 days for the CTC-negative patients ($p<0.001$), and when they limited the analysis only for the stage-IVb patients, the MSTs of the CTC-positive and negative patients were 52.5 and 308.3 days ($p<0.01$), respectively (155).

Changes in epigenetics can also be used as a biomarker for detecting pancreatic cancer. Many studies suggest that cancer-specific DNA methylation patterns can be detected in circulating tumor cells in body fluids and biopsies or in free-floating DNA (156,157). Joko et al showed a high frequency (97%) of methylation of BNC1 and ADAMTS1 in CTCs from patients with invasive pancreatic cancers, and the methylation on beads (MOB) technology can be used with serum from pancreatic cancer patients to detect methylation changes with high sensitivity and specificity. Kiesiel and coworkers examined six methylation biomarkers and KRAS mutation in the pancreatic juice from 61 PC patients, 22 chronic pancreatitis patients, and 19 healthy people. CD1D is the most discriminant marker for detection of the PC versus normal pancreas or chronic pancreatitis (AUC value for PC compared to the normal pancreas or chronic pancreatitis is 0.92), suggesting that methylated CD1D is a sensitive and specific biomarker for PC (158).

Exosomes are extracellular vesicles produced during endosomes formation and are sources of circulating DNA (159). Glypican-1 (GPC1), a membrane-anchored protein, is

overexpressed in a variety type of cancers such as glioma, breast and pancreatic cancer (160-162). Melo and the coworkers isolated the exosomes from the serum of PC patients and healthy controls. The levels of GPC1 positive exosomes in PC patients are significantly higher that levels healthy patients and those with benign pancreatic disease. GPC1 positive exosomes showed 100% sensitivity and specificity in each stage of PC (Carcinoma in situ, stage I and stages II-IV). In addition, GPC1 positive exosomes can also predict inform the metastatic tumor burden of PC patients. An average of 58.5% of GPC1 positive exosomes were observed in the PC patients with distant metastasis, and this was significantly higher in patients with restricted metastasis in lymph nodes (~50.5%) and with no metastases (~39.9%) (162).

Biomarkers can also help to predict the optimal therapy for PC patients. Actinin 4 (ACTN4) is an actin-binding protein, closely associated with cancer invasion and metastasis (163) and is overexpressed in pancreatic cancer patients (164). Clinical trials have compared chemotherapy alone and chemoradiotherapy (CRT) for treating locally advanced pancreatic cancer (LAPC) treatment and the increase in ACTN4 copy number can be used as a biomarker for CRT for treating LAPC. Patients with an increased copy number of ACTN4 had a worse prognosis of overall survival than patients with a lower (normal) copy number (165). Another example is that patients with advanced pancreatic cancer exhibiting a normal baseline of serum albumin ($\geq$ 3.4 g/dL) have increased benefit from treatment with bevacizumab compared to those with decreased levels of serum albumin ($<$ 3.4 g/dL) (166). Biomarkers can also be used for evaluating the

therapeutic efficacy of drugs, for examples, the HER3 expression is a predictive biomarker for pertuzumab efficacy in HER3-low expressing pancreatic cancer (167). Another challenge is to identify biomarkers in high-risk patients. PC-594, a novel circulating 36-carbon long-chain polyunsaturated fatty acid, is significantly reduced in the serum of pancreatic cancer patients at different stages (168,169). Because of the lack of stage effect, it is possible that the PC-594 levels may be decreased prior to the development of PC and it was concluded that decreased PC-594 may be a significant prognostic marker for PC (169).

Biomarkers are continually being identified, however, future studies are needed to validate whether newly discovered biomarkers can be used for early detection of pancreatic cancer. It is likely that a combination of different biomarkers may be a useful approach for increasing the sensitivity and specific of biomarkers for PC. Using the combination of MMP7 and CA19-9 expression in the plasma of 92 PC patients and 31 patients with chronic pancreatitis gave a predictive value of 100% on discriminating between patients with carcinoma (high in MMP7 and CA19-9) and those with benign pancreatic disease (low in MMP7 and CA19-9) (170).

As discussed above, current research has been successful for identifying pancreatic cancer biomarkers, but there is still a lack of validation for the expression pattern of genes/proteins biomarkers can be used for designing clinical trials. The combination of newly discovered biomarkers along with CA19-9 has shown some success in increasing

sensitivity and specificity, but this combination also required validation with improved technology and deeper understanding of the molecular mechanisms of pancreatic cancer development.

1.1.3.4 Therapy for pancreatic cancer

Cancer therapies have evolved and improved and this is due, in part to progress in understanding the biology of this complex disease. There are different types of therapy for treating pancreatic cancer, including surgical management, chemotherapy, radiation therapy and targeted therapies. The types of treatment used for treating this disease will depend on the stage, critical gene biomarkers, and tumor classification. However, for those PC patients that are inoperable, the overall success of current therapies is low and the 5-year survival rate is 5%.

Chemotherapy and radiation therapy are still the primary treatments for advanced pancreatic cancer. Surgery is usually considered when the tumor is still in the pancreas and does not extend far beyond the pancreas. The primary goal of pancreatic cancer surgery is to completely remove of the tumor and clear the regional lymph nodes (171,172). Systemic chemotherapy, radiation therapy or combinations (chemoradiotherapy) have been applied after surgical resection (adjuvant therapy) or before (neoadjuvant therapy) in the effort to reduce mortality rates (173). Surgical removal of the tumor followed by six months of gemcitabine treatment increased the median patient survival to 22.8 months with a one-year survival rate of 70% (174).

In contrast to cytotoxic chemotherapy or radiation therapy which target all rapidly dividing cells, targeted therapy interacts with key factors/pathways that are differentially expressed in cancer cells compared to normal cells. In this way, targeted therapies minimize the side effects encountered with systemic cytotoxic chemotherapies. Overexpression of vascular endothelial growth factor (VEGF) has been correlated with increased rates of PC recurrence and metastasis (175,176). Bevacizumab is a humanized anti-VEGF monoclonal antibody which has been extensively studied as an anti-angiogenic agent for treating pancreatic cancer patients. In a phase III trial of patients with metastatic pancreatic cancer, the addition of bevacizumab to bevacizumab gemcitabine and erlotinib is well tolerated and improves the progression-free survival but not overall survival (177). A recent meta-analysis of targeted therapies for treating advanced pancreatic cancer examined the effectiveness of anti-EGFR, anti-angiogenesis, anti-PSA, FTASE inhibitor, anti-IGFR1, anti-apoptosis, anti-MEK, immunotherapy, anti-PI3K, c-KIT inhibitor therapies and showed that there was no overall survival benefit from these targeted therapies (178). The clinical trials examined in this study were primarily those using conventional agents and this may have masked the efficacy of targeted therapies. Therefore, there is a need not only to identify predictive factors that will allow for selection of patients that will benefit from targeted therapies, but also to develop new and more effective treatment for pancreatic cancer patients.

1.2 Long non-coding RNAs

In 1970, Francis Crick proposed the developed central dogma of molecular biology, which describes the flow of genes into proteins: DNA to RNA to protein (179). However, the human genome project showed that there are only <25,000 protein-coding genes in our genome and this only accounts for two percent of the total DNA (180,181). The majority of gene transcripts are not translated into protein (182). Initially, those RNAs with no protein-coding potential were called the "dark matter of human genome", or junk RNA. During the last 10-20 years, the importance of non-coding RNAs (ncRNAs) has been emerging and it is now believed that ncRNAs strongly contribute to the diverse structure and function of eukaryotic organisms. (182,183).

Based on their size, ncRNAs can be classified into small, medium and long noncoding RNAs (184). The size of small ncRNAs is 18-31 nucleotides (nt), and these include small interfering RNAs (siRNAs), microRNAs (miRNAs), medium ncRNAs contain ~31 to 200 nt, including transfer RNA (tRNA), small nuclear (snRNAs) and small nucleolar RNAs (snoRNAs). The third category is long non-coding RNAs (lncRNAs) which have a sequence length > 200 nt. The pie chart represents the major categories of the 16,592 non-coding RNAs (185) (Figure 4). Based on their function, ncRNAs can also be divided into housekeeping ncRNAs (ribosomal, transfer, snRNAs), regulatory ncRNAs (miRNAs, siRNAs and lncRNAs), promoter-associated RNAs (PARs) and enhancer RNAs (eRNAs) (186-188).

LncRNAs are the largest group in the ncRNA family. LncRNAs are highly heterogeneous and complex and it is estimated that >80% of the transcribed mammalian genome is associated with the generation of lncRNAs (189,190).

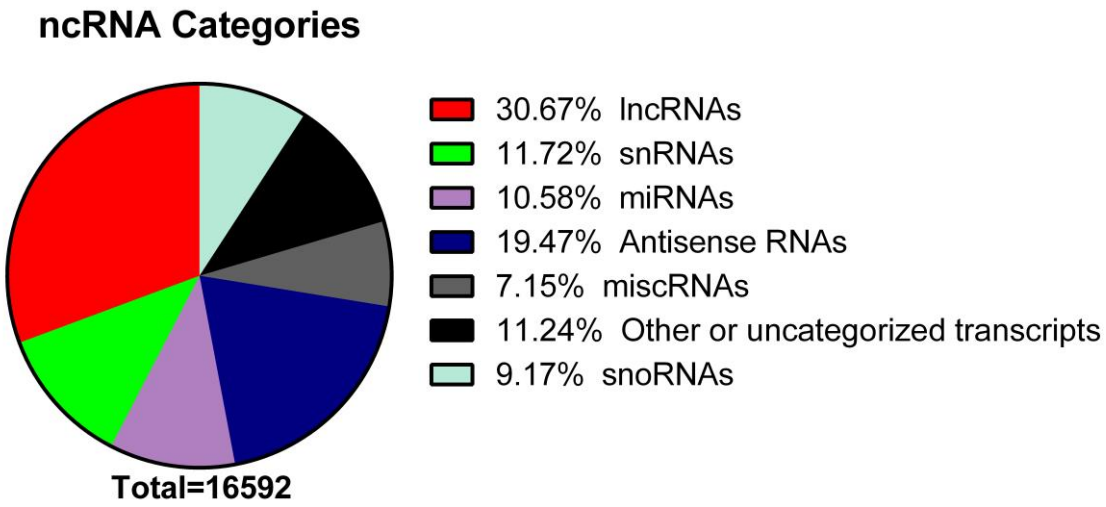


Figure 4. The classification of ncRNA categories. Modified from (184).

Initially lncRNAs were classified based on their relationship to the corresponding mRNAs or localization and direction with respect to a specific mRNA/RNA and this can be sense, antisense, intergenic, intronic and bidirectional (191). Sense lncRNAs are transcribed from the sense strand of protein-coding genes and may partially overlap with the protein-coding gene or cover the entire sequence. For example, cold-assisted intronic noncoding RNA (COLDAIR) is transcribed from an intron of flowering locus C (FLC) protein (192). Antisense lncRNAs, in contrast, are transcribed from the antisense strand of protein-coding genes. Based on the GENCODE, antisense lncRNAs can overlap with

an exon of sense gene, transcript from the intron of a sense gene but do not have exon-exon overlap with this sense gene, or cover the entire sequence of a sense gene through an intron (193). LncRNAs transcribed from intergenic regions are named intergenic lncRNAs, while, those transcribed from introns of protein-coding genes are named as intronic lncRNAs. Bidirectional lncRNAs are oriented head to head with a protein-coding gene within 1kb. In 2016, Sanbonmatsu and coworkers proposed to classify lncRNAs based on their structure. According to the secondary structure, lncRNAs can be categorized into three different types, highly structured RNAs with subdomains and complex structural motifs; loosely structured RNAs with multiple stem-loops, but lacking hierarchical domain structure and complex motifs; unstructured, disordered RNAs, which lack secondary structure (194). Based on the effects on the DNA sequences, lncRNAs can also be classified into cis-lncRNAs, and trans-lncRNAs. Cis-lncRNAs regulate the expression of genes in close genomic proximity and an example is the HOXA transcript at the distal tip (HOTTIP) which regulates HOXA gene cluster genes via interaction with histone modification complexes (195). Trans-lncRNAs target distant gene loci; for example, HOX transcript antisense RNA (HOTAIR) is transcribed from the HOXC gene but can affect target genes at the more distal HOXD gene locus. (196).

1.2.1 Molecular mechanisms of lncRNAs

LncRNAs have important regulatory roles at the transcriptional, translational and epigenetic level, however, the molecular mechanisms through which lncRNAs act are

still poorly understood. Based on the molecular mechanisms of action, lncRNAs can be classified into several categories 1) signal: respond to diverse stimuli, 2) decoy: bind and titrate away a protein target, 3) guide: bind proteins and then direct the complex to specific targets to affect gene transcription, and post-transcriptional pathways, 4) effect: form RNA dimers with mRNA sequences to block transcription-associated sites, regulate the stability, cleavage, and translation of protein-coding genes, 5) scaffold: serve as platforms for different proteins to form ribonucleoprotein complexes (197). The classification is not fixed and sometimes a lncRNA can function via different mechanisms in different cells and tissues. Regulatory mechanisms of lncRNAs are summarized below.

1.2.1.1 Transcription regulation mediated by lncRNAs

Many lncRNAs affect transcription, for instance, lncRNA-p21 is a p53 target and negatively regulates p53 downstream genes through the physical interactions with heterogeneous nuclear ribonucleoprotein K (hnRNAP-K). This interaction is required for genomic localization of hnRNAP-K at the repressed gene sites and regulation of p53 mediated apoptosis (198). Another p53 induced lncRNA, p21 associated ncRNA DNA damage activated (PANDA) is derived from the CDK1A promoter and it function as a decoy for nuclear transcription factor-Y alpha (NF-YA) to limit expression of pro-apoptotic genes leading to cell-cycle arrest (199).

1.2.1.2 Regulation of mRNA processing

According to central dogma, DNA is transcribed into premature mRNA, which is subsequently spliced into different mature isoforms (200). Serine/arginine (SR) splicing factors regulate alternative splicing in the nucleus. Metastasis associated lung adenocarcinoma transcript 1 (MALAT1) is an lncRNA that interacts with SR proteins and affects the location of splicing factors in nuclear speckle domains. In addition, MALAT1 regulates the phosphorylation of SR proteins, suggesting that MALAT1 can regulate alternative splicing by modulating splicing factors (201). Besides regulating the mRNA splicing, lncRNAs can also directly affect mRNA editing (202).

1.2.1.3 Post-transcriptional or translational regulation

The last stop for center dogma flow is mRNA to protein. After mRNA is transported into the cytoplasm, lncRNAs can regulate protein expression by affecting translation efficiency or regulating mRNA stability.

Some lncRNAs are transported into the cytoplasm and many studies have shown that lncRNAs modulate mRNA translation into protein and also mRNA stability. One example of a lncRNA that affects translation efficiency is GAS5, which binds to cMyc mRNA and this interaction enhances the binding of eukaryotic translation initiation factor-4E (elF4E) with cMyc mRNA (203). Many lncRNAs have been shown to bind to 3’ UTR of mRNA and regulate mRNA stability. PCNA Antisense RNA 1 (PCNA-AS1),

a lncRNA antisense to PCNA mRNA, is upregulated in hepatocellular carcinoma and affects PCNA mRNA stability by RNA hybridization (204).

Besides those categorized general function of lncRNAs, there are several emerging function of lncRNAs and their molecular mechanisms of action.

1.2.1.4 LncRNAs in epigenetics

Transcription of some lncRNAs resides within promoters of protein-coding genes and this can lead to deposition of methylation markers and recruitment of repressor complexes (205). One example is the lncRNA IRT1 which resides within the promoter region of gene IME1, a master regulator of meiosis. IRT1 transcription recruits the Set2 histone methyltransferase and the Set3 histone deacetylase complex to form repressive chromatin at the IME1 promoter. By inhibiting Ime1, IRT1 plays an essential role in the mating-type control of gametogenesis in yeast (206).

LncRNAs can also recruit histone-modifying enzymes in cis or trans configuration. X-inactive specific transcript (Xist), one of the first lncRNAs to be identified, is located on the Xist inactivation center (XIC) of the genome (207), and interacts with the polycomb repressive complex 2 (PRC2) to target the X chromosome and is responsible for the epigenetic regulation of H3K27 trimethylation (208). HOTAIR, is expressed within the HOXC gene cluster and represses transcription of HOXD genes in trans by regulating PRC2 recruitment. The 5’ end of HOTAIR binds to enhancer of zeste 2 polycomb

repressive complex 2 subunit (EZH2) and mediates its recruitment to HOXD cluster (209).

LncRNAs also serve as scaffolds for recruiting multiple histone modification complexes. For example, the 3' end of HOTAIR binds to the lysine-specific demethylase 1 (LSD1) component of REST/CoREST complex, which has H3K4 demethylation activity. Through interacting with both PRC2 and REST/CoREST complexes, HOTAIR can coordinate H3K27 methylation and H3K4 demethylation at targeted loci within the genome (210). In contrast, another HOX- associated lncRNA, HOTTIP regulates activation of HOXA genes via recruiting a member of myeloid/lymphoid or mixed-lineage leukemia 1 (MLL1) through an interaction with WD repeat domain 5 (WDR5) (211). These lncRNAs play important epigenetic roles in gene activation or repression by coordinating recruitment and binding of chromatin modifying complexes to specific regulatory regions.

1.2.1.5 Enhancer-like activity of lncRNAs

Analysis of a new set of GENCODE annotated lncRNAs demonstrated enhancer-like mechanisms of lncRNAs and knockdown of enhancer-like lncRNAs decreased expression of neighboring protein-coding genes. A class of lncRNAs, ncRNA-activating (ncRNA-a), can activate their neighboring genes. NcRNA-a interacts with Mediator, a transcriptional co-activator complex, which induces chromatin looping between the

ncRNA-a loci and target promoters and kinases activity toward histone H3 serine 10, a histone modification that is known to be associated with transcriptional activation (212).

1.2.1.6 LncRNA-miRNA interaction

There have also been extensive studies on the interactions between lncRNAs and miRNAs. LncRNAs have been identified as competing endogenous RNAs by interacting with miRNAs. TUG1 (taurine upregulated gene1) is a lncRNA that serves as a miR26a sponge in human glioma cells (213). HOTAIR controls the expression of Rab22a by sponging mir373 in ovarian cancer (214). MiR1 was shown to bind both MALAT1 and cell division cycle 42 (cdc42) and MALAT1 induced migration and invasion of breast cancer cells, partially due to completely binding to miR1 (215). LncRNAs can also function via derived miRNA. H19 is a lncRNA that regulates intestinal epithelial barrier function via the H19-encode miR675 (216). Moreover, miRNAs can also negatively regulate lncRNA expression by direct binding to lncRNAs. It was reported that miR-21 negatively regulates growth arrest specific 5 (GAS5) in breast cancer cells in a way similar to miRNA-mediated silencing of mRNAs (217). Reciprocal regulation is also observed between lncRNA urothelial carcinoma-associated 1 (UCA1) and miR145 in bladder cancer cells. UCA1 represses miR145 expression and suppression of miR145 results in upregulation of UCA1. The binding site for miR145 within exon 2 and exon 3 of UCA1 contributed to the reciprocal negative regulation (218).

1.2.1.7 Discovery of micropeptides

Anderson et al discovered in 2015 that a small micropeptide of 46 amino acids is encoded from a lncRNA, and it impacts calcium uptake into sarcoplasmic reticulum (SR) (219). However, the discovery of micropeptides encoded by putative lncRNAs raised another question about lncRNA definition and classification: Should those RNAs encoding micropeptides be called lncRNA or it should belong to a new category?

1.2.2 LncRNAs in diseases

In the past decade, the biological role of lncRNAs in human disease, especially cancer has been extensively investigated. The first disease-related lncRNA identified is LncRNA BC200 which is specifically expressed in brain cells and may be involved in Alzheimer disease (220). The importance of lncRNAs have been observed in various diseases, such as neuropsychiatric disorder (221), coronary disease (222,223), autoimmune disease (224-227) and also cancer (228,229).

1.2.2.1 LncRNAs in cancer

Many studies have shown that lncRNAs play an important functional role in carcinogenesis, cancer metastasis and invasion (230), either as an oncogene or a tumor suppressor-like gene (231,232). Expression levels of several lncRNAs are closely associated with overall survival rates of cancer patients.

H19 is one of the first identified lncRNAs (233) and expressed in fetal tissues but not in adult tissues. However, the H19 expression has been shown to be upregulated in various cancer types, including glioma (234,235), gastric (236,237), esophageal, colorectal (238), breast (239), bladder (240), hepatocarcinoma (241), pancreatic (242), prostate (243), and nasopharyngeal cancers (244). Although it is certain that H19 plays a crucial role in the cancer development, the underlying molecular mechanism is still not fully understood. There are several reported mechanisms, for example, H19 functions via antagonism of miRNA let7 or miR874 (242,245) or H19-derived miR675 (235). LncRNAs are also associated with drug resistance in cancer. One example is that UCA1 upregulation activated Wnt signaling and in turn increased cisplatin resistance of bladder cancer cells (246).

1.2.2.2 LncRNAs as biomarkers and therapeutic targets

LncRNAs not only regulate cellular function but they also serve as diagnostic and prognostic markers (247) and expression levels of specific lncRNAs are closely associated with overall survival rates of cancer patients and are prognostic factors for several tumor types (151,248-251). Serghiou and coworkers carried out a systemic review and meta-analysis of lncRNAs as predictors of survival in cancer patients. They identified 111 studies including 127 datasets, and 96% of the studies showed a significant association of lncRNA expression with prognosis. HOTAIR is currently the most extensively studied lncRNA, and a recent meta-analysis of 19 cancer sites showed that high level of HOTAIR expression in cancer patients is correlated with poor survival

(252). Therapies targeting lncRNAs are still in the preclinical stages, however, scientists from the Cold Spring Harbor laboratory showed that by targeting lncRNA MALAT1 with an antisense oligonucleotide (ASO) the metastatic activity of aggressive primary breast tumors can be reduced (253). Although the ASO approach gave promising results, there are some limitations, such as the complicated secondary structure of lncRNAs which make them difficult to be targeted in the cells.

Another method to directly target lncRNA is through clustered regularly interspaced short palindromic repeats (CRISPR) system. The current method for gene ablations is primarily by transfection of a siRNA targeting a specific mRNA in the cytoplasm. However, most lncRNAs reside in the nucleus and this is a challenge for the RNAi method. The CRISPR system is an adaptive defense mechanism against foreign viruses in the bacteria. A short fragment of virus is inserted into the cas operon and then transcribed in conjunction with a ncRNA to guide the nuclease protein Cas9 to cleave the foreign viral DNA (254). CRISPR/CAS9 technology has been recognized as a successful tool to engineer the genome in different species (255-257). The CRISPR/CAS9 system can also target lncRNAs in human cells (258) and mouse one-cell embryos (259). For example, it was reported that CRISPR sufficiently deletes a large imprinted lncRNA RNA imprinted and accumulated in nucleus (Rian) in mice using CRISPR/CAS9 co-injection of one-cell embryos (259), suggesting that this technique is a promising approach for targeting lncRNAs.

1.2.2.3 LncRNAs in pancreatic cancer

My research has focused in the lncRNAs HOTTIP and MALAT1 and their functions in pancreatic cancer and table 5 summarizes the expression and function of total 20 lncRNAs that have been identified in pancreatic cancers, and many of these have the potential to be therapeutic targets, diagnostic or prognostic biomarkers.

Table 5. The lncRNAs studied which are expressed and functional in the pancreatic cancer.

Name	Biological function	Molecular function	References
Oncogenic			
HOTAIR	Promoted cancer cell growth and migration, HOTAIR silencing enhanced radio-sensitivity, reduced proliferation, increased apoptosis after radiation	HOTAIR silencing increased the expression of Wnt inhibitory factor-1. HOTAIR mediated gene expression is both PRC-2 dependent and independent	(260,261)
NUTF2P3-001	Promoted tumorigenesis	Depressed miR-3923/KRAS pathway	(262)
MALAT1	Correlated with tumor development and unfavorable prognosis, promoted cell growth, migration and invasion	Activated autophagy, enhanced stem cell-like phenotypes	(263-267)
LOC389641	Promoted cancer progression, increased cell invasion	Regulated E-cadherin in a TNFRSFIOA-related manner	(268)
PVT1	Increased expression associated with poor prognosis in PC patients, overexpression decreased sensitivity to gemcitabine	Unknown	(269,270)
HOTTIP	Enhanced cell proliferation, survival and migration	Regulated several HOX genes	(271,272)
HULC	Biomarker	Unknown	(273)

Table 5. Continued.

Name	Biological function	Molecular function	References
Oncogenic			
H19	Therapeutic target, promoted metastasis	Depressed let7 suppression on its target HMGA2 mediated EMT	(242,274)
ccdc26	Responsible for cell growth and apoptosis	Partly by regulating the PCNA and Bcl2 expression	(149)
LINC00673	Maintained cell homeostasis	Created a miR-1231 binding site and interferes with PTPN11 degradation	(275)
LincRNA-ROR	Promoted invasion, metastasis and tumor growth	Activated ZEB1 pathway	(276)
AFAP1-AS1	Overexpression promoted cell proliferation, migration and invasion	Unknown	(277)
Linc00675	Suppression decreased EMT process		(278)
lncRNA af339813	Knockdown of this lncRNA reduced cell proliferation and promoted apoptosis	Unknown	(279)
Tumor suppressive			
LOC285194	Low expression correlated with poor overall survival	Unknown	(280)
GAS5	Decreased in cancer, overexpression inhibited cell proliferation	Gas5 inhibition induced decrease in G0/G1 phase and an increase in S phase, negatively regulated CDK6	(281)
ENST00000480739	Downregulated expression contributed to tumor metastasis and progression	Regulated HIF-1alpha	(282)
lncRNA-ATB	Low expression indicated poor survival	Unknown	(283)
BC008363	High expression correlated with better survival		(284)
MEG3	Knockdown of MEG3 attenuated cytotoxicity induced by fen fibrate, Overexpression induced cell death	Increased p53 expression	(285)

1.2.2.3.1 Expression

The lncRNA expression in PC tissues compared to normal tissue have been investigated and in 2014, Li and coworker identified 1881 upregulated lncRNAs and 2269 downregulated lncRNAs in PC tissues compared to adjacent normal tissues from 30 PC patients with the fold change of 4.0 and p-value of less than 0.01 (284). In 2015, Wang and coworkers screened 7419 lncRNAs in the 144 PC specimens (tumor and paired adjacent normal tissues) and four chronic pancreatitis samples. Thirty three lncRNAs were differentially expressed in PC compared to normal tissues (7 upregulated and 26 downregulated, respectively) with the fold change >3 (151). Depending on the sample size and statistical analysis, there is some variability in lncRNAs expressed in PC compared to normal tissues, however, there is no doubt that there is differential expression of lncRNAs in PC.

1.2.2.3.2 Prognostic significance

LncRNAs have potential to be biomarkers for diagnosis and prognosis of cancer and most studies still detect the lncRNAs in the PC tissues either using RT-PCR or microarray. For example, HULC expression in tumor and non-tumor tissues from 304 PC patients demonstrated that higher expression of HULC is significantly correlated with increased tumor size, invasion and metastasis. Multivariate analysis suggested that HULC expression is an independent predictor of overall survival with p-value =0.032 (273). In addition, construction and analysis of dysregulated lncRNA-associated competing endogenous RNA network help to identify novel lncRNA biomarkers for

early diagnosis. Furthermore, instead of using one single lncRNA, a 7-lncRNA signature (termed lncRisk-7) was developed as a novel diagnosis of PDAC [108]. Currently, there are no reports about dysregulated lncRNAs in circulation in PC patients and further studies are need to examine whether those differentially expressed in the PC tissues are stable in circulation.

1.2.2.3.3 Function

LncRNAs regulate tumorigenesis and tumor development and exhibit oncogene or tumor suppressor-like activities. They are involved in PC cell growth, apoptosis, cell cycle progression and cell motility. Coiled-coil domain containing 26 (CCDC26) is a lncRNA overexpressed in the PC tissues and decreased expression significantly promoted growth arrest and apoptosis. Expression of CCDC26 is positively correlated with proliferating cell nuclear antigen (PCNA) and the anti-apoptotic gene Bcl2 (149). GAS5 expression is significantly decreased in the PC tissue and overexpression of GAS5 inhibited PC cell proliferation, while GAS5 inhibition induced a significant decrease in G0/G1 phase and an increase in S phase via negatively regulating CDK6 (cyclin-dependent kinase 6) expression (281). Maternally expressed gene 3 (MEG3) is a tumor suppressor lncRNA expressed in a range of human cancers. It was shown that fenofibrate inhibited PC tumor cell proliferation via upregulating MEG3 expression and subsequently increasing the p53 expression and overexpression of MEG3 in the pancreatic cancer cells increased p53 expression and induced cell death (285). In addition to the regulation of cancer cell growth and death, lncRNAs also affect tumor cell migration and invasion. LOC389641

is a novel lncRNAs which is upregulated in the PC tissues and LOC389641 knockdown inhibited PC cell migration and decreased expression of this lncRNA upregulated the E-cadherin expression possibly via affecting the LOC389641 neighboring gene, TNFRSF10A(268).

Although lncRNAs are important for pancreatic cancer growth, migration/invasion, however, the underlying mechanisms of these functions are still not well understood. LncRNAs can inhibit miRNAs expression and miR-mediated repression of target genes. H19 is overexpressed in PC tissues and knockdown of H19 in the cancer cells inhibited the epithelial-mesenchymal transition (EMT). The possible mechanism may involve antagonism of miRNA let-7 and derepressed the let-7 targeted high-mobility group A protein 2 (HMGA2) gene. It was shown that HMGA2 is involved in EMT maintenance via regulation of E-cadherin gene (286). Another important mechanism for lncRNA function is the interaction with histone modifying complex. For example, the lncRNA HOTAIR is associated with poor prognosis of different types of cancer, such as breast, liver, colon, cervical and nasopharyngeal cancer. Increased expression of HOTAIR in breast cancer cells induced the shifting of polycomb repressive complex 2 (PRC2) and alteration of histone H3 lysine 27 (H3K27) methylation (209). Our laboratory showed that HOTAIR is pro-oncogenic and a negative prognostic factor for pancreatic cancer. Knockdown of HOTAIR in different pancreatic cancer cells decreased cell proliferation, migration/invasion, and a similar effect was observed in a xenograft study. Our study of HOTAIR in pancreatic cancer cells also suggested that the mechanisms of HOTAIR

function in PC cells is different from previously reported in breast cancer cells (209). We showed that only GDF15 was co-regulated by HOTAIR and EZH2 (PRC2) in pancreatic cancer cells, suggesting that HOTAIR-mediated suppression of genes in pancreatic cancer is both PRC2-dependent and independent, and the PRC2-independent pathways of HOTAIR function are currently under investigation (261).

1.3 Microbiota-derived aryl hydrocarbon receptor ligands

1.3.1 Gut microbiota metabolites

The gut microbiota is emerging as a "new organ" in the human body. It contains 100 times more genes than the host and can be regulated by exogenous and endogenous signals, and the absence of microbiota can affect all aspects of host physiology. Over the past decade, our understanding of gut microbiota has increased exponentially due to new technologies that have helped generate a genetic and metabolic profile of the microbial community and their impacts on the host.

In the gut, there is a high density and diversity of bacteria and the data on gut microbiota is mainly acquired from fecal samples and mucosal biopsies. The abundance and composition in the small intestine is significantly different from the colon (287). The diverse bacteria in the gut perform a myriad of functions, including the transformation of substances in the gut and production of antimicrobial substances to stimulate the immune system (288). For example, short-chain fatty acids (SCFA), conjugated linoleic acid (CLA) and gamma-amino butyric acid produced by bacteria are beneficial in the

treatment of several diseases, such as cancer, obesity, and cardiovascular disease (289-292).

Gut microbiota-derived metabolites can also serve as biomarkers and therapeutic targets for different diseases. Crohn's disease (CD) is an inflammatory bowel disease (IBD) and it has been reported that specific gut microbiota metabolites can be used to differentiate between healthy and CD patients (293).

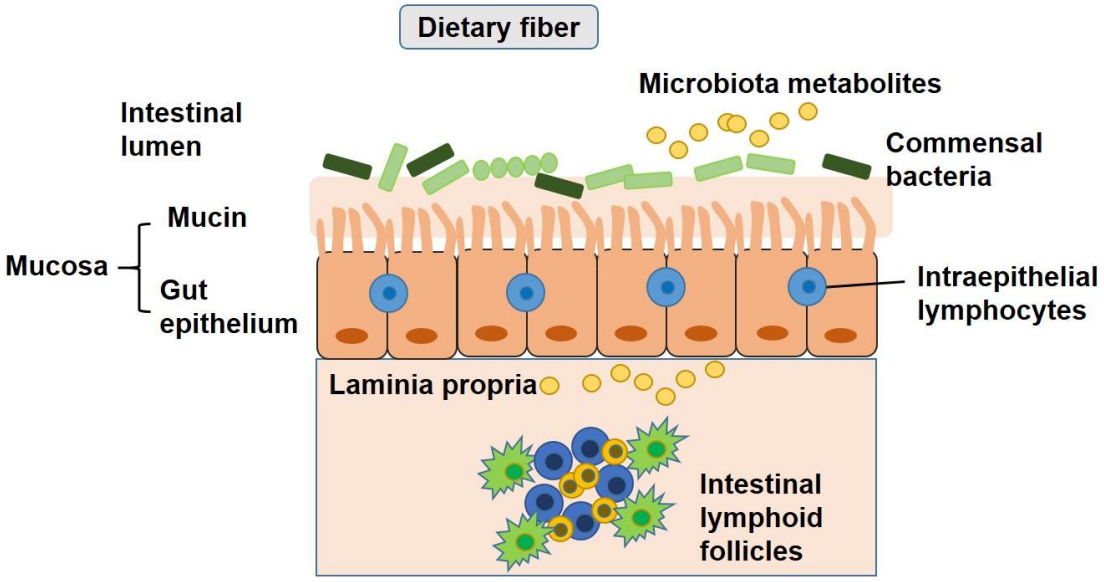


Figure 5. An illustration of the structure of the intestinal tract and the microbiota residing in the gut. The gut microbiota metabolites can affect gut epithelium, intraepithelial lymphocytes, and intestinal lymphoid follicles.

Gut microbiota provides the host with nutrients and protects against infection and the balance of pro- and anti-inflammatory pathways is important for maintaining gut

immune homeostasis. The composition of the commensal microbial community plays an important role in maintaining homeostasis (Figure 5). Butyrate, a short-chain fatty acid (SCFA), produced by commensal microbiota during carbohydrate fermentation, facilitates the extrathymic generation of Treg cells, and another SCFA, propionate potentiates the de novo generation of Treg cells, which play a key role in limiting the inflammatory responses in the intestine. Thus by regulating Treg cells, butyrate and propionate affect the balance between pro- and anti-inflammatory pathways in the gut (294) . In addition to the beneficial effects of SCFAs, protein fermentation also produces harmful metabolites, such as ammonia, phenols, and hydrogen sulfide which can result in a "leaking gut", inflammation, and cancer (295). In contrast, ingestion of dietary plant-based foods appear to inhibit gut inflammatory related diseases (296). It has been shown that microbial transformation of dietary phenolic compounds has important implications in the chemoprevention of colorectal cancer (297). Polyphenolic metabolites, such as 3-(3,4-hydroxyphenol)-propionic acid (PS) and 3,4-dihydroxyphenolacetic acid (ES) which are metabolites of chlorogenic acid/caffeic acid and quercetin, respectively, upregulate glutathione S-transferase T2 (GSTT2) and downregulate cyclooxygenase (COX2) in the human adenoma cells LT97and this may contribute to the chemopreventive potential of polyphenolic metabolites in the gut (298). Gut microbiota is also functional via different axis, such as gut-liver axis, gut-brain axis, and the gut-brain-skin axis and the microbiota-mediate defects on the gut can clearly affect distal organs/tissues (299).

Development of microbial metabolites as biomarkers may also be useful in distinguishing between health/disease states or closely related disease conditions. In addition, microbiota metabolites can help develop algorithms to predict personalized responses to dietary and pharmaceutical interventions (300,301). Community and system-level interactive optimization (CASINO) is a specialized computational platform that is developed to quantify the gut microbiota metabolites and pair this data with dietary intake characteristics and patterns. For example, CASINO was used in a clinical experiment in which 45 overweight and obese individuals were maintained on a restricted diet for 6 weeks. Using Casino, they were able to predict a decrease in carbohydrate and increase in amino acid consumption by analysis of the microbiota metabolites (299)**.** Moreover, microbial production of beneficial metabolites can also be increased, either by increasing the abundance of native species or engineering microbiota to produce higher levels of beneficial bacteria. As microbiota metabolites interact locally with host receptors to elicit diverse functions, it is possible that metabolites mimicking drugs might be novel therapies for treating different diseases (302).

1.3.2 Aryl hydrocarbon receptor

The aryl hydrocarbon receptor (AhR) is an evolutionarily old transcription factor and studies on the structure and function of the AhR were initiated in the 1970s. Early studies of the AhR mainly focused on the aryl hydrocarbon hydroxylase activity (AHH) induced by the toxic industrial byproduct 2,3,7,8-tetrachlorodibenzo-p-doxin (TCDD)

(303), and AhR was first identified as the TCDD receptor (304). Previous studies suggested that AhR is a key mediator of drug metabolism and also regulates the toxic effects induced by exposure to polychlorinated dibenzo-p-dioxin (PCDD), biphenyls (PCBs), dibenzofurans (PCDFs), polycyclic aromatic hydrocarbons (PAHs) and related compounds (305,306).

Ligand activation of the AhR induces expression of CYP1A/CYP1B and several phase II drug metabolizing enzymes and the AhR also plays an important role in the regulation of cellular process, such as apoptosis, proliferation, cell growth and differentiation (307). The AhR is required from several well characterized toxic responses induced by TCDD including chloracne, immune deficiency, thymic atrophy, liver porphyria and a wasting syndrome (308-314). In the last decade, AhR research focused on identifying physiological endogenous ligands and the roles and mechanisms of action of the AhR in normal development (315).

The AhR protein was first characterized in 1976 (303) and cloned by several groups(316-319). The AhR is a multi-domain protein, consisting of the basic helix-loop-helix (bHLH) domain, two PAS domain (PAS-A & PAS-B) and a transactivation domain (TAD) (Figure 6). The bHLH domain is required for the DNA binding and also contains a nuclear localization signal (NLS) necessary for transport of the AhR into the nucleus for DNA binding. PAS stands for Per-Arnt-Sim domains (period circadian

protein-aryl hydrocarbon receptor nuclear translocator-single minded protein). The C-terminal of PAS-B domain is required for ligand binding (320,321).

Figure 6. Schematic structure of the AhR.

In the absence of ligands activation, AhR remains in the cytoplasm and form a complex with two Hsp90 (heat-shock protein 90) molecules, an AIP protein (AhR-interacting protein, also known as ARA9), an XAP2 protein (hepatitis B Virus X-associated protein2) and the p23 co-chaperon (Hsp23). AIP interacts with AhR and Hsp90 through its tetratricopeptide repeat (TRP domains). Mutations in the TRP domains will disrupt the binding of AhR to AIP, leading to the AhR destabilization and degradation (322,323). Ligand binding to the AhR induces a conformational change in the receptor and exposes the nuclear localization signal. Within the nucleus, AhR dissociates itself from Hsp90 and forms a heterodimer with ARNT, and the resulting heterodimer regulates transcription of many genes, including phase I, phase II drug metabolites, and phase III transporter gene (Figure 7) (324). Ligand specificity and selection may be determined not only by the PAS-B domain itself but also by other parts of AhR and its interacting protein partners. It was proposed that the ligand binding pocket and access

channels leading to the pocket play equally important roles in discrimination between endogenous AhR ligand and xenobiotics such as TCDD (325). While the majority of studies on AhR report the ligand-dependent AhR activation, there have been several reports showing a ligand-independent AhR pathway (326-329). For example, Maayah and coworkers reported that Sunitinib, a tyrosine kinase inhibitor, induced CYP1A1 expression level in an AhR-dependent manner, even though Sunitinib does not directly bind to and induce transformation of cytosolic AhR (330).

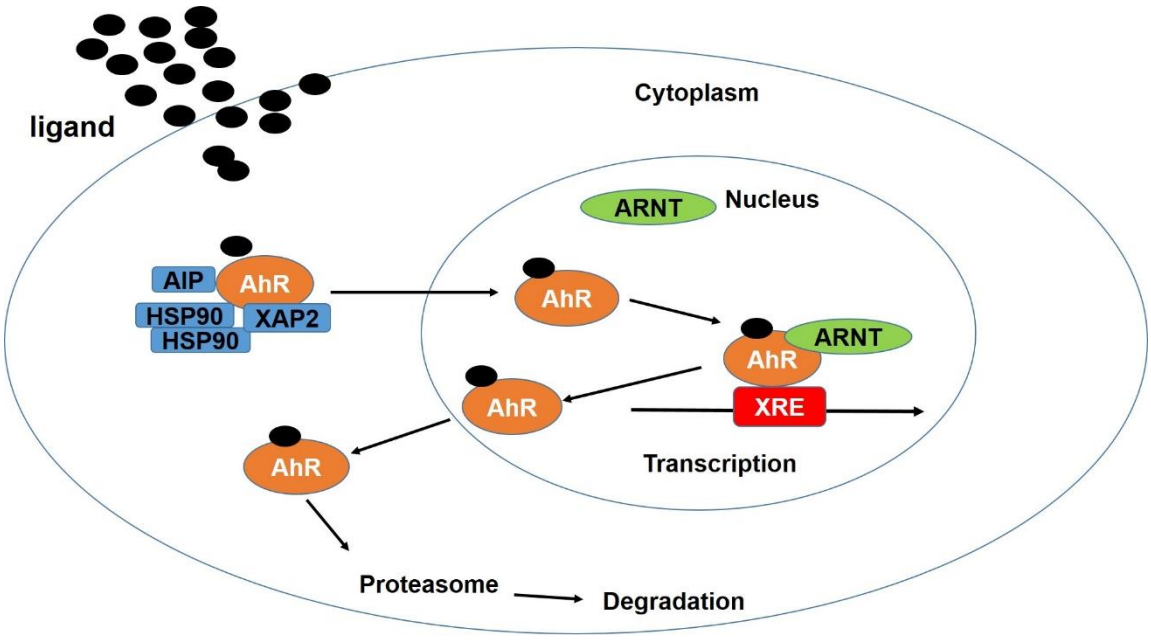


Figure 7. AhR signaling pathway.

Post-translational modification is the primary mechanism for regulating intracellular AhR levels and AhR transactivation. Upon export form the nucleus, the AhR is ubiquitinated and subsequently degraded in a proteasome-dependent manner. The AhR

is also regulated through a negative feedback loop, by upregulating a transcriptional repressor known as the aryl hydrocarbon receptor repressor (AHRR). AHRR is also a bHLH–PAS protein and has high sequence similarity to AhR. It suppresses AhR activity by competitively binding ARNT (331).

1.3.2.1 AhR ligands

At first, the most well characterized classes of AhR ligands were environmental contaminants including PAHs and HAHs. However, subsequent studies have identified natural and endogenous AhR ligands, including both biochemical and synthetic food products, phytochemicals and microbial metabolites. Figure 8 illustrates structures of selected representative AhR ligands from different sources and some possibly endogenous AhR ligands. In 1998, Chang et al observed that CYP1A1 deficient cells exhibited a higher level of baseline AhR activity than the wild-type cells and CYP1A1 specifically decreased the AhR activity in hepatoma cells (332). In addition, there were several studies showing that ectopic expression of CYP1A1, CYP1A2 or CYP1B1 decreased the AhR activity (333), indicating that some endogenous AhR ligands are CYP1A1 substrates and AhR-mediated CYP1A1 expression in turn negatively regulates AhR activity.

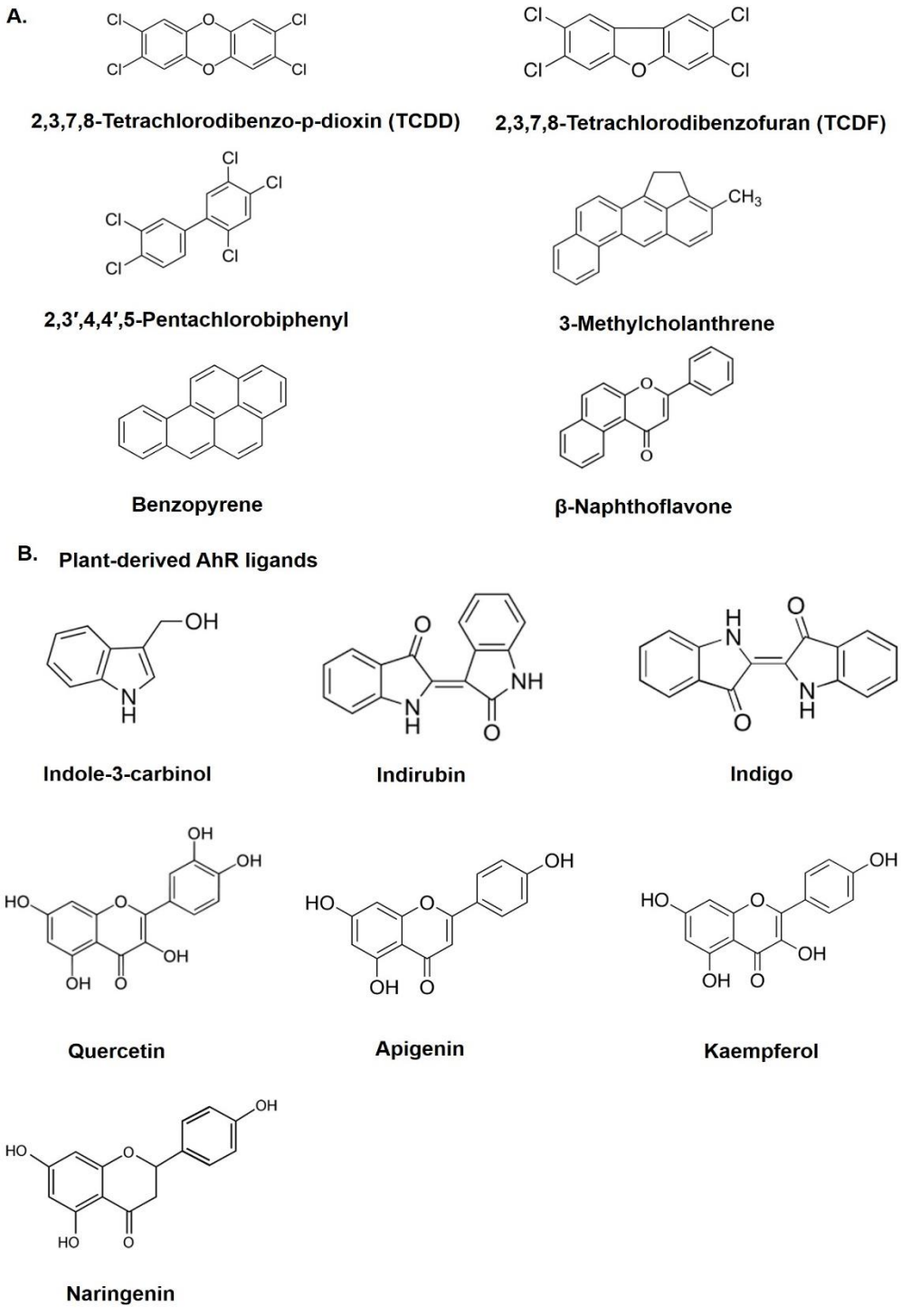


Figure 8. Selected exogenous and endogenous AhR ligands. A. classical AhR ligands; B, Plant-derived AhR ligands; C. synthetic pharmaceutical AhR ligands; D. Microbiota derived AhR ligands.

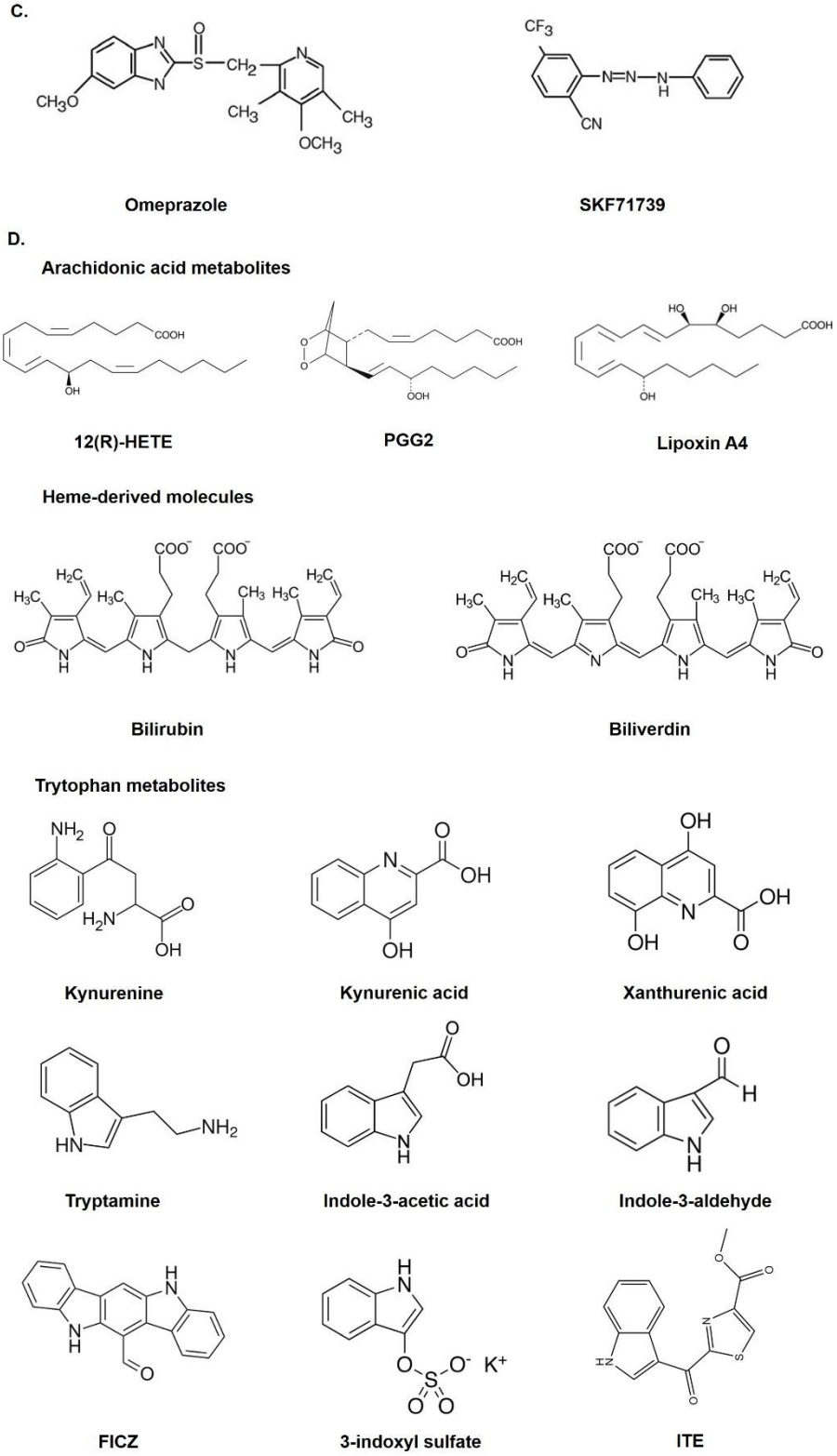


Figure 8. Continued.

1.3.2.1.1 Plant-derived AhR ligands

Plant-derived phytochemicals were among the first identified potential natural AhR ligands. Plant-derived indirubin and indigo are high potency AhR ligands (334) and indole-3-carbinol conjugates and derivatives are found in cruciferous vegetables and also exhibit high AhR binding affinity (335-341). However, there is some controversy regarding the possible role of these compounds as endogenous AhR ligands since levels of indigo and indirubin in humans are in the picomolar range which is not sufficient for activating the AhR (334,342). Nevertheless, the importance of indirubin and indigos cannot be dismissed since their local concentrations may be sufficiently high at some sites to exhibit AhR agonist activities. Another indole metabolite, 3-indoxyl sulfate is also a potent AhR ligand and plays a key role in the human disease progression. Indole can be metabolized into indoxyl sulfate (I3S) in the liver, and is secreted by the kidney via the proximal tubules. Indoxyl sulfate usually accumulates in the body of patients with kidney disease because of the failed renal function. Examination of 139 patients with chronic kidney disease (CKD) showed that high levels of serum indoxyl sulfate predicted the overall and cardiovascular mortality at p-value of 0.001 and 0.012, respectively (343). It is suggested that constitutively activation of AhR by I3S may contribute to toxicity observed in the patients with CKD (344). In addition, it has been shown that quercetin, apigenin, and kaempferol present in certain foods mediate agonist/antagonist synergistic effect of AhR activity in different cell types (345) and the dietary flavonoid naringenin induces the generation of Treg cells by AhR-mediated pathway (346).

1.3.2.1.2 Endogenous AhR ligands

1.3.2.1.2.1 Heme-derived molecules

In 1993, Kapitulnik and coworkers showed that rats defective in UDP-glucuronosyltransferase exhibited elevated CYP1A1 levels. UGT is a key enzyme for degrading heme-containing proteins to give bilirubin (347), and subsequently it was found that bilirubin induced CYP1A1 in an AhR dependent manner (348). Another heme-degradation product, biliverdin also activated the AhR (349).

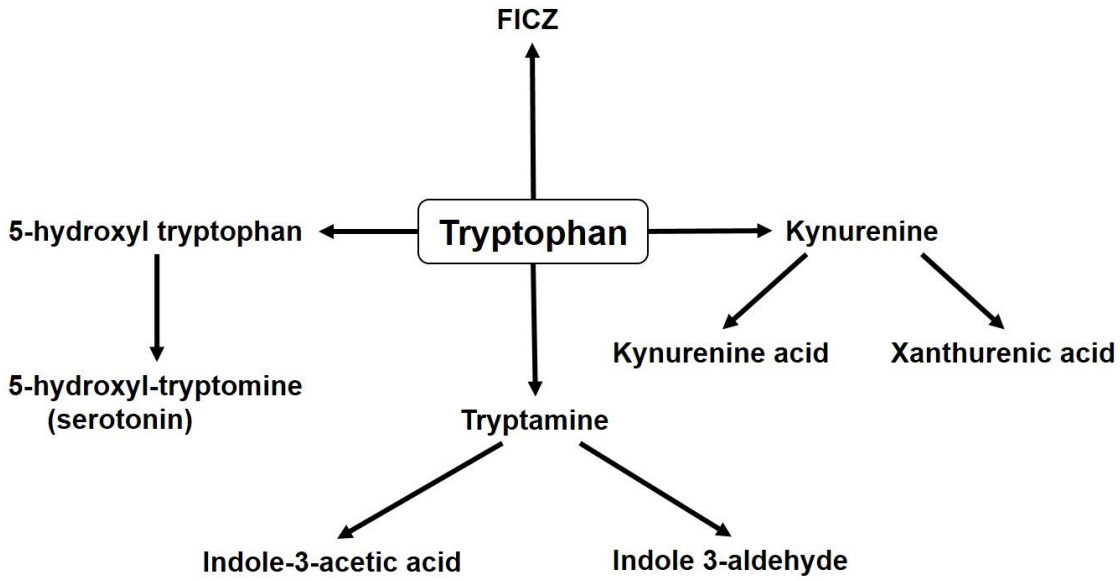


Figure 9. The schematic illustration of major pathways of tryptophan metabolism.

1.3.2.1.2.2 Arachidonic acid metabolites

Arachidonic acid derivatives have also been postulated to be AhR ligands. 12R-hydroxy-5Z,8Z, 10E, 14Z-eicosatetraenoic acid (12(R)-HETE) and prostaglandin-endoperoxide

synthase 2 (PGG2) have been identified as potent AhR agonists (350-352), and another arachidonic acid metabolite, Lipoxin 4A, which does not have any ring structures also induced CYP1A1 and CYP1A2-dpenendent activity (353).

1.3.2.1.2.3 Tryptophan metabolites

Tryptophan metabolites are considered to be among the leading candidates for designation as endogenous AhR ligands. Kynurenine pathway is the primary route of tryptophan metabolism (Figure 9). Kynurenine (Kyn) has been shown to regulate glioblastoma cell growth and migration in AhR-dependent manner. Glioblastoma cells constitutively generate Kyn and Kyn suppresses anti-tumor immune response and promotes tumor cell survival through the AhR (354). Kynurenine is further metabolized to kynurenic acid and xanthurenic acid which are also function as AhR ligands (355,356). The IDO/tryptophan-2,3-dioxygenase-independent formation of tryptamine also results in the formation of AhR ligands. The most proximal metabolite, tryptamine is a potent AhR activator (356) and also a precursor of downstream metabolites, such as indole acetic acid and indole-3- aldehyde that are also AhR ligands (356,357). The serotonin pathway may be directly or indirectly related with AhR activation since the proximal metabolite, 5-hydroxy tryptophan itself is a weak agonist (358). Upregulation of CYP1A1 enzymes could cause the depletion of tryptophan stores and reduce the production of endogenous AhR ligands for other pathways. There are also studies showing that the tryptophan photometabolite (FICZ) is also an endogenous AhR ligands (359). FICZ has a high binding affinity for the AhR and can induced AhR-mediated

genes at nanomolar concentrations (359-366). FICZ has also been shown to regulate different physiological processes. FICZ either promotes or inhibits inflammatory responses in a tissue-specific manner; FICZ induced inflammatory responses in human skin cells exposed to UV (367) while it was shown to reduce pulmonary eosinophilia and inhibit Th2 expression and differentiation via repressing GATA-3 and STAT6 expression (368). In addition, FICZ exhibits therapeutic potential and it was reported that the addition of FICZ enhanced retinoic acid (RA)-induced leukemia cell differentiation, and RA is the current standard clinical drug for treatment of acute promyelocytic leukemia (APL) (369). FICZ also affects intestinal homeostasis via regulation of intestinal epithelial cells (IEC) development. FICZ administration was shown that FICZ inhibits intestinal organoid development in vitro and reduce IECs in mice through increasing expression of Lgr5, Math1, Indian Hedgehog, and inhibiting Lgr4 (370).

In 2002, Song et al isolated an endogenous AhR ligand from porcine lung extracts and this ligand was identified as 2-(1'H-indole-3'carbonyl)-thiazole-4-carboxylic acid methyl ester (ITE) which is expressed in species from early vertebrates to humans (371). ITE was later shown to be a tryptophan metabolite however the pathway of formation is still not clear. The immunosuppressive capacity of ITE has been demonstrated in several studies, suggesting the potential application of ITE for the treatment of human immune diseases. ITE inhibited the development of type II-collagen induced arthritis in mice by suppressing the delayed hypersensitivity(372) and efficiently suppressed

experimental autoimmune encephalomyelitis in mice, by enhancing Treg activity and reducing Th17 cell function and this has subsequently been shown to be an AhR-dependent response (373). Similar, in mice with induced experimental autoimmune uveoretinitis (EAU) ITE efficiently inhibited EAU development, reduced subpopulations of Th1 and Th17 cells and secreted cytokines (IFN-gamma, and IL17). In addition, ITE may be used clinically as anti scarring agent since it was shown that ITE inhibited myofibroblast differentiation by blocking the TGFβ1 signaling pathway (374).

1.3.2.2 The function of AhR

It has been established through studies on AhR-null mice that the AhR has multiple physiological functions (375-381). In addition to the role of the AhR in normal development, this receptor is a component of various pathological and physiological disorders, including autoimmune disease (382), inflammation (383,384), cardiovascular disease (385) and cancer.

The AhR is expressed in various tumors with different prognostic significance; increased expression of AhR is associated with the development of medulloblastoma (386), while decreased expression of AhR is observed in the pituitary adenomas (387). The AhR and its ligands play a dual role in cancer development and AhR ligands exhibit both tumor-specific promotion and inhibitory functions. TCDD and 3, 3-diinodolymethane inhibited breast cancer cell growth and invasion (388), while n-butyl benzyl phthalate and dibutyl phthalate ligands increased the proliferation and invasiveness of breast cancer cells

(389). Our laboratory has reported that different AhR ligands, including 6-alkyl-1,3,8-trichlorodibenzofuran (6-MCDF), 8-alkyl-1,3,8-trichlorodibenzofuran (8-MCDF), diindolylmethane (DIM) and tetraCB inhibited ER-positive and ER-negative breast cancer cell proliferation and tumor growth. (338,390,391). These selective AhR modulators (SAhRMs) target the AhR for the treatment of hormone-dependent and hormone-independent tumors (336,337,392). The AhR is expressed widely in many histotypes of ovarian cancer tissues, and ITE, an endogenous AhR ligand suppresses ovarian cancer cell proliferation and migration (393). Our laboratory has also studied the ligand-independent function of AhR, for example, the AhR exhibits growth inhibitory effects in MCF-7 breast cancer cells, but growth-promoting effects were observed in hepatocarcinoma cells even though the AhR protects against carcinogen-induced liver cancer in vivo (394,395). Recently, we showed that AhR active pharmaceuticals such as omeprazole decreased breast and pancreatic cancer cell invasion and metastasis (396,397). Recent reviews (398) clearly show that the AhR is an important new drug target for treating cancers and development of new AhR agonists or antagonists is ongoing.

1.3.2.2.1 The AhR in immunity

The first paper suggesting the involvement of AhR in the immunity was published by Clark and coworkers in 1981 and they observed that TCDD administration markedly inhibited the generation of cytotoxic T cells in response to alloantigenic challenge (399). In recent years, the role of AhR in immune response has been extensively studied that

intrinsic AhR expression in various immune cells has been linked to allergy, autoimmunity, and mucosal immunity. Immune function disturbances have been observed in AhR $^{-/-}$ mice. Moreover, the conditional AhR knockout mice help uncover the tissue-specific function of AhR on immunity, such as gut, skin, and lung (400).

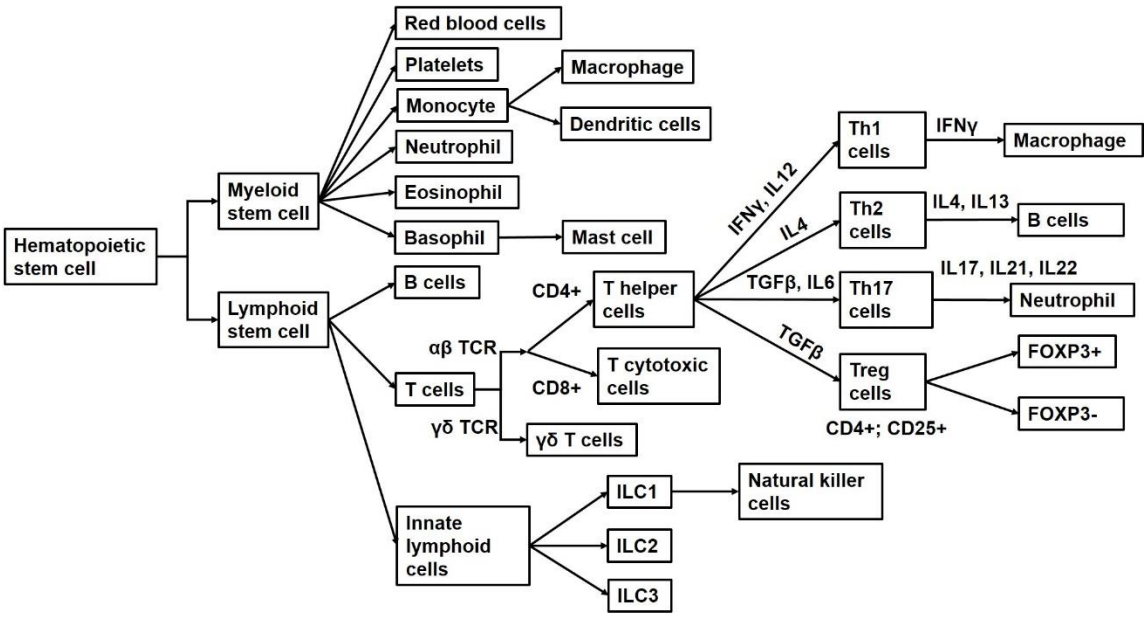


Figure 10. The schematic illustration of differentiation of hematopoietic stem cells into different immune cells and blood cells.

The hematopoietic stem cells (HTCs) can differentiate into myeloid stem cells (MSCs), which further differentiate into basophil, eosinophil, neutrophil, monocytes (further into macrophage, dendritic cells) and others. On the other hand, HTCs can also differentiate into lymphoid stem cells (LSCs), which further into B cells and T cells and innate lymphoid cells (ILCs). Naive T cells can be further differentiated into T helper (Th) cell

and T cytotoxic (Tc) cells. Th cells include Th1, Th2, Th17 and regulatory T (Treg) cells. ILCs include ILC1, ILC2, and ILC3 (Figure 10).

AhR is expressed in many hematopoietic cell types, and the presence of dioxin response element (DREs) (Cis-acting) is observed in the promoters of many genes involved in the immune cell differentiation, suggesting the potential involvement of AhR mediated effects in the immune system (401). AhR activation may regulate the hematopoietic and immune systems in a cell-specific manner (402). In addition, the AhR in other non-hematopoietic cells types such as endothelial, epithelial and stromal cells should also be considered since these cells can also be affected by immune stimulation and in turn influence immune responses (403). The AhR integrates both innate and adaptive immunities via various mechanisms. Both exogenous and endogenous AhR ligands can effectively affect host metabolism and immunity (404). Alteration in the gut microbiota through dietary or environmental contaminants intervention also disturbs intestinal homeostasis, and contributes to pathology (405). Different factors can influent the outcome of the AhR-dependent responses and these include the specific cell types where the AhR expressed, the types of AhR ligand and the presence of cofactors such as coactivators (406). Three similar factors were proposed to determine the AhR function in the lymphocytes, namely, the amount of AhR in any given cell, the abundance and potency of AhR ligands within a specific tissue or cells and the microenvironment of the AhR expressing cells (407). It was reported that activation of the AhR by different AhR ligands might cause different outcomes. For example, in the encephalomyelitis (EAE)

murine model, TCDD promoted immune suppression, while endogenous ligands such as FICZ promoted Th17 cell response and exacerbated the pathology of EAE (408). It is possible that the mode of application of a ligand rather than its nature may contribute to the outcome, with the support of results from systemic or local administration of FICZ. Systematic administration of FICZ might affect multiple tissues and cell types and lead to a strong reduction of the concurrently induced immune responses (409). In contrast, local injection of FICZ worsened the pathology of EAE, possibly by specifically targeting and promoting Th17 cells (382).

1.3.2.2.1.1 AhR in adaptive immune cells

The AhR influences adaptive immune responses by regulating both T cells and antigen-presenting cells (410). An adaptive immune response is triggered via activation, differentiation, and clonal expansion of the lymphoid lineage cells, T and B lymphocytes. Depending on the signal received, the T cells proceed to differentiate along transcription factor-specific pathways that give distinct T cell subsets. These lineage-specific transcription factors include T-bet for Th1, GATA3 for Th2 cells, retinoid-related orphan receptor gamma t (RORγt) for Th17 cells, and Foxp3 for Treg cells. Activation of AhR may directly/indirectly modulate the commitment of these T cell subsets. Several important immune cells involved in adaptive immune response are discussed below.

1.3.2.2.1.1.1 Th1 and Th2 cells

Th1 cells are a subset of T cells producing IFN-γ and inducing macrophage-driven inflammation, whereas Th2 cells secrete interleukin (IL)-4 and drive eosinophilic inflammation. The expression of AhR is not detected in mice Th1 and Th2 cells during differentiation, however, treatment with AhR agonists shifted the Th1/Th2 balance to Th1 expansion by GATA binding protein 3 (GATA3) inhibition (411). The shift of Th1/Th2 balance helped to ameliorate allergy (411) and allergic asthma (368) in mouse models.

1.3.2.2.1.1.2 Th17 and Treg cells

The AhR is pivotal in the differentiation and functions of the $CD4^+$ effector cells - Th17 and Treg which play essential roles in the autoimmunity (412,413). The AhR is highly expressed in both human and mouse Th17 cells and defective differentiation of these T cells subsets was observed in the AhR deficient mice (382,408,413). The AhR induces the production of interleukin (IL) 17 and IL22 in Th17 cells , which are play a role in intestinal mucosal immune cell regulation (42). Treatment of wild-type mice with FICZ increased production of IL17 and IL22 in Th17 cells and this enhanced encephalomyelitis (EAE) in a mouse model for this response (382).

Activation of the AhR promotes the differentiation of Treg cells. For example, the activation of AhR by TCDD induces $CD4^+FOXP3^+$ Treg cell differentiation in vitro and in vivo (408,413) and several endogenous AhR ligands such as Kyn, DIM and

naringenin promote differentiation of Treg cells (414-416). Treg cells from AhR deficient mice express lower levels of IL10 (417,418) and elevated IFN-γ which is normally inhibited by IL10 (419,420). The Treg cell expansion induced by TCDD suppressed the EAE and the suppressive effect was abolished by administration of an AhR antagonist (408).

1.3.2.2.1.1.3 B cells

The function of AhR in the B cell differentiation and its activity is not well understood. TCDD administration suppresses the production of IgM in a transformed mouse B cell line and decreases the number of IgM-secreting plasma cells and delays the differentiation of B cells by inhibition of the activator protein (AP)-1 in mice (421). Activation of AhR by ITE suppresses the differentiation of B cells into Ig-secreting plasma cells and the production of IgM, IgE and IgG1 (421).

1.3.2.2.1.2 AhR in innate immune cells

1.3.2.2.1.2.1 Neutrophils

Neutrophils are the most abundant leukocytes mediating inflammatory responses to various infections. There have been discrepancies with regard to the effect of AhR on neutrophils and this may be due to different AhR ligands and their application. TCDD administration increased the number of neutrophils accumulated in the lung upon influenza virus infection in mice and increased inflammation (422). In contrast, neutrophils numbers were not increased in the TCDD treated animals in response to

Streptococcus pneumonia infection (423), and there were decreased numbers of neutrophils in the cornea of herpesvirus simplex-infected mice treated with TCDD (424).

1.3.2.2.1.2.2 Natural killer (NK) cells

The NK cells are a diverse population of lymphocytes involved in both innate and adaptive immunity. The AhR is expressed in NK cells and the modulatory effect of AhR on NK maturation and function is still in doubt. IDO and kyn suppressed NK cells undergoing activation, but not resting cells (425).

1.3.2.2.1.2.3 Mast cell

Mast cells may also be involved in the role of AhR in chronic inflammation and autoimmunity. FICZ treatment boosted the degranulation of Mast cells and produced ROS, IL-6 and IL17 (426). In addition, AhR was expressed to present in three rat mast cell lines and Kyn treatment enhanced the production of IL6 in RBL2H3 cells (427).

1.3.2.2.1.2.4 Platelets

Platelets are found in large number in blood and involved in clotting. At the sites of vascular injury, the AhR plays a critical role in the physiologic response that platelets undergo during vascular injury (428) when platelets are recruited and their interactions with collagen, triggers formation of hemostatic plug (429). Treatment with TCDD or omeprazole results in platelet priming, as demonstrated by increased platelet aggregation and this response is inhibited by the AhR antagonist, CH-223191 (430).

1.3.2.2.1.2.5 Dendritic cells (DCs)

DCs are the group of immune cells that bridge the innate and adaptive immunities and also play an important role in autoimmune inflammation. TCDD induced differentiation of bone marrow derived immature DCs to mature DCs (431), and the AhR mediates the induction of indoleamine 2, 3-dioxygenase (IDO) in DCs and IDO converts tryptophan into Kyn and other metabolites. Kyn and IL10 production was decreased in AhR-deficient DCs challenged with lipopolysaccharides (LPS) or CpG oligodeoxynucleotides (CpG). In a coculture system of stimulated AhR-deficient DC and naive T cells, differentiation of naïve T cells into Treg cells and the immune response was inhibited (432). Similarly, ITE treatment induced tolerogenic DCs that support $FoxP3^+$ Treg differentiation (373). VAF347 is a low molecular weight compound that inhibits allergic lung inflammation and VAF347 blocks the DC-mediated generation of proinflammatory Th cells. VAF347 inhibits IL6, CD86 and Human Leukocyte Antigen - antigen D Related (HLA-DR) expression in DCs which are required for Th cell generation. AhR knockdown ablates the inhibitory effect of VAF347 and AhR deficient mice are resistant to VAF347-induced inhibition of allergic lung inflammation in vivo, suggesting that AhR is important for the effects of VAF347 (433).

1.3.2.2.1.2.6 Innate lymphoid cells (ILCs)

ILCs are required for the formation of intestinal lymphoid follicles. They are important for tissue homeostasis in both innate and adaptive immune responses. AhR is involved in

the development of intestinal lymphoid follicles and can be activated by dietary AhR-active phytochemicals (434). ILCs can be classified into group 1, 2 and 3 ILCs based on the expression of effector cytokines and the key transcriptional factors that direct their differentiation (435). ILCs are strikingly similar to Th cells, particularly Th1, Th2 and Th17 cells with respect to their effector cytokines and functions. Group 1 ILCs produce IFN-r; group 2 ILCS produce type 2 cytokines and are dependent on GATA3 and RORα and group 3 ILCs consists of lymphoid tissue inducer, natural cytotoxicity triggering receptor $(NCR)^{+}$ ILC3, and NCR^{-} ILC3 (403). All three types of ILC3 produce IL17 and/or IL22 and are dependent on RORγt, and the AhR affects the homeostasis of the ILC subsets (436-438). Without the AhR, group 3 ILCs exhibit increased apoptosis (439). And defect in the group 3 ILCs were also observed in $AhR^{-/-}$ mice, and when challenged with toxoplasma gondii, the $AhR^{-/-}$ mice lost more weight than wild type mice (437).

1.3.2.2.1.3 AhR in barrier organs

The AhR is highly expressed in barrier organs, including skin, lung, and gut. The mucosal surfaces of these barrier organs are in direct contact with environmental pollutants as well as harmful pathogens, and the AhR functions to protect against physical and immunological challenge to barrier organs (403).

1.3.2.2.1.3.1 Skin

Skin is one of primary defense organs (440). AhR plays a critical role in skin homeostasis and barrier function by inducing keratinocyte (KC) terminal differentiation and subsequently, the formation of multiple epidermal strata in the skin (441,442). AhR activation promotes terminal differentiation in the cultured human KC skin cells (443,444) and an in a three-dimensional (3D) epithelial organotypic culture system (445). Coal tar which contains AhR active PAHs has been shown to be beneficial for the treatment of atopic dermatitis. It was reported that coal tar can active AhR in an organotypic culture system and promote epidermal differentiation, and AhR knockdown abolished this effect (446). Similarly, FICZ treatment reduced the inflammation in skin lesions from psoriasis patients and AhR antagonists worsened the inflammation (447). On the other hand, constitutive aberrant expression of AhR was observed to cause inflammatory skin lesions (448). After physical damage, the skin initiates a robust inflammatory response, followed by new tissue formation and remodeling to prevent further effects on deeper body organs (440). Absence of AhR inhibits migration rates of primary endothelial cells and fibroblasts during the wound healing, indicating the AhR might play a role in this process. Moreover, the AhR promotes wound healing by increasing the production of TGFβ by fibroblasts, which in turn accelerates KC cell migration and promotes wound healing (449).

1.3.2.2.1.3.2 Lung

The AhR is abundantly expressed in the lung with expression levels similar to that observed in liver (450). Lung airways are lined with several subsets of cells, including basal stem cells, ciliated cells, brush cells and goblet cells (451). The lung epithelial cells, as well as the immune cells infiltrating the lung, control the inflammatory responses in the lung (452). Mucins are glycoproteins produced by epithelial tissues and are key components of the mucosa. Lung epithelial cells promote the mucin production upon exposure to external toxicants or pathogens, and the AhR may participate in the formation of the mucin barrier of the lung under the influence of the endogenous ligands. Mast cells also participate in the lung allergic responses, and the AhR deficient mast cells poorly respond to allergen stimulation due to dysregulated calcium signaling and mitochondrial function (453). The AhR activation by FICZ can promote or inhibit mast cell degranulation (426). Mitchell and coworkers compared the consequences of transient and sustained AhR activation and proposed that timing is critically important (454). A single administration of FICZ results in increased mast cell degranulation whereas repeated administration of this ligand leads to suppression of the process with asthma lung dysfunction.

1.3.2.2.1.3.3 Gut

The gastrointestinal tract has the largest mucosal surface of the human body and the gut content is separated from the host inner body by a mucus layer secreted by goblet cells and a single layer of intestinal epithelial cells (IECs). Interspersed among the IECs are

the intraepithelial lymphocytes. Underneath the epithelial is the lamina propria (LP) which contains a variety of immune cells, including B cells, T cells, macrophages, DCs, and innate lymphoid cells (ILCs) (451). The AhR is critical for maintaining both IEL and ILC populations in the gut (434,439,455-457) and is a critical gene for maintaining gut health.

1.3.2.2.2 AhR in gut health

The role of AhR in gut health is also closely related to the production of AhR- active gut microbiota metabolites and their role in the regulation of the intestinal mucosal immune system. The intestinal mucosal immune system comprises a well-organized and intricate network of the intestinal epithelium, innate and adaptive immune cells. Tryptophan and its metabolites are among the most important AhR-active microbiota-derived metabolites which are detected in the urine or feces. Trp metabolites have an important role in gut immune homeostasis. IL22 is a member of the family of IL-10 related cytokines and has an important role in protecting the epithelial cell barrier in the gut. Mice deficient in IL22 are highly susceptible to intestinal infection by bacteria pathogens (458). IL22 can be produced by Th17 cells, and subsequent studies showed that a subset of innate lymphoid cells also produce IL22 (459). There is a significant reduction of ILC22 cells in AhR knockout mice resulting in decreased secretion of IL22 and inadequate protection against bacterial infection (456,460). It has been shown that Trp metabolites regulate IL22 expression via AhR dependent pathway. Dietary tryptophan is metabolized

into indole-3-aldehyde, which subsequently activates the AhR pathway and enhances production of IL22 by $NKp46^+$ ILCs (404).

Inflammatory bowel disease (IBD) is relatively common and involves chronic inflammation of the digestive tract, and includes ulcerative colitis and Crohn's disease. There is increasing evidence that activation of the AhR protects against colitis in mice and humans (461-463). Dextran sulfate sodium (DSS)-induced colitis model is widely used for studying colitis because of the simplicity of the model and similarities with human colitis (464). Treatment with TCDD or FICZ protects mice from dextran sulfate sodium (DSS)-induced colitis (465,466). Several mechanisms have been proposed to explain the protective effects of activation of the AhR on colitis. AhR could activate intestinal innate lymphoid cells to release IL22 (439) or promote the functions of ILCs intestinal intraepithelial lymphocytes (IELs) (455). It was shown that AhR deficiency or lack of AhR ligands compromised the maintenance of IELs resulting in increased vulnerability to epithelial damage (455). In addition, Il-7 was recently added to the list of factors involved in attenuation of colitis by AhR activation. FICZ downregulated the epithelial cell-derived IL-7 expression in DSS-treated mice. And it has been proposed that AhR activation blocks the epithelial IL-7 expression, which subsequently affects the phenotype and function of IELs (467). Caspase recruitment domain family member 9 (CARD9) is a susceptible gene for IBD, which accelerates recovery from colitis by promoting IL-22 production. The microbiota from CARD9 knockout mice failed to metabolize tryptophan into AhR ligands, and decreased levels of AhR ligands are also

observed in microbiota from IBD patients, particularly in those with CARD9 risk alleles. It has been suggested that CARD9 regulates colitis by changing tryptophan metabolism into AhR ligands (468). FICZ, a potent endogenous AhR ligand, downregulates epithelial-derived IL7 expression and ameliorate DSS-induced colitis, suggesting that AhR active compounds might be new and promising therapies for the treatment of IBD patients (469). In patients with Crohn's disease, reduced AhR expression levels were observed in the intestinal T cells and natural killer (NK) cells (466). Fibrostrictures (FAS) are a major complication of Crohn's disease and activation of fibroblasts and excessive collagen deposition are crucial for development of FAS. FICZ inhibits activation of MAPK activation and collagen induction, suggesting that the AhR is a negative regulator of profibrotic signals in the gut (470).

Colorectal cancer (CRC) is the third cause of cancer associated death in the US (2), and AhR is involved in colorectal cancer development. In the normal intestine, the AhR is primarily localized in the stroma -containing immune cells in the lamina propria and lymphoid follicles. In the tumor tissue, AhR expression is elevated and detected in both stromal and tumor cells. A media supplemented with AhR ligands suppressed colorectal cancer cell growth. It was concluded that the AhR suppresses tumor growth via two distinct pathways: in normal tissue, the AhR is associated with tumor prevention by regulating gut immunity, while in the tumor cells, it is involved in growth suppression (471). Increased risk for CRC has been associated with chronic intestinal inflammation as observed in patients with IBD (472). AhR deficient mice develop cecal tumors at the

age of 30- 40 weeks with severe inflammation and the blockage of interleukin 1β signaling attenuates the cecal tumorigenesis in AhR deficient mice (473). The $APC^{min/+}$ mouse model has been extensively used in studies on colon cancer and those mice expressing mutations in adenomas polyposis coli (APC) gene spontaneously develop intestinal tumors (110). The $APC^{min/+}$ mice had a tumor incidence of ~50% at age of 14 weeks and reached 100% at 25 weeks while in the $Apc^{min/+}$; $AhR^{+/-}$ mice tumor progression is accelerated and ~50%incidence was observed of ages of 9-10 weeks and these mice were also more susceptible to cecal tumorigenesis. Moreover, $APC^{min/+}$ mice treated with indole-3-carbinol (I3C) exhibited few tumors, supporting the use of AhR agonists such as I3C for the treatment of patients with CRC (474). The effect of the AhR on the CRC has also been studied in the colon cancer cell lines. It was shown that FICZ inhibited prliferation of LoVo colon cancer cells by promoting G1 cell cycle arrest, associated with downregulation of cyclin D1 and upregulation of P27 expression. In addition, the effect of FICZ was abolished by cotreatment with AhR antagonist (475). Treatment with another endogenous AhR agonist, indole-3-carbinal (I3C) induced CYP1A1 mRNA expression and a dose-dependent decrease of cell viability and increased apoptosis in several colon cancer cell lines. Moreover, knockdown of the AhR significantly attenuated the chemotherapeutic effects of I3C with regard to proliferation and apoptosis (476).

In summary, the AhR has been shown to be a multi-player in gut health and endogenous AhR-ligands have potential for improving treatments of gut-associated diseases.

2. THE LONG NON-CODING RNA HOTTIP ENHANCES PANCREATIC CANCER CELL PROLIFERATION, SURVIVAL AND MIGRATION*

2.1 Introduction

Long non-coding RNAs (lncRNAs) are defined as transcripts containing >200 nucleotides and are typically transcribed by RNA polymerase II (477). Although the existence of lncRNAs has been known for several decades, it is only in the last ten years that the multiple functions of the lncRNA components of the noncoding genome have been determined. LncRNAs play important roles in maintaining cellular homeostasis during cell/tissue development and they are also critical factors in pathophysiology including cancer (197,477-481). The molecular modes of action of lncRNAs are highly variable and include their functions as molecular scaffolds for stabilizing protein-protein and protein-DNA interactions; they can also act as decoys and guides that facilitate both proximal and distal macromolecular interactions which are usually on a genome template (197,272,478).

LncRNAs have been investigated in tumors and cancer cells derived from multiple sites, and there is strong evidence that their overexpression or underexpression can influence cancer cell growth, survival and migration/invasion (209,210,482). HOX transcript

* Reprinted with permission from "The long non-coding RNA HOTTIP enhances pancreatic cancer cell proliferation, survival and migration" by Cheng Y, Jutooru I, Chadalapaka G, Corton JC, Safe S. Oncotarget. 2015; 6(13):10840-52.

antisense RNA (HOTAIR) is a 2.2 kb lncRNA in the mammalian *HOXC* locus that serves as sequence-specific scaffold for at least two histone modification complexes, namely polycomb repressive complex (PRC2) and the LSD1/CoREST/REST complex (209,210,482). In tumors and cancer cells, HOTAIR interactions with these histone modification complexes modulate expression of tumor type-dependent gene sets, and knockdown or overexpression studies show that HOTAIR is an important pro-oncogenic factor that plays a role in cancer cell proliferation, survival and migration/invasion (209,210,261,483-486). HOTAIR is also a tumor-specific negative prognostic factor for the survival of cancer patients and can be detected in serum (485).

HOXA transcript at the distal tip (HOTTIP) is another HOX-associated lncRNA transcribed from the 5' tip of the HOXA locus, and HOTTIP is associated with the PRC2 and WDR5/MLL1 chromatin modifying complexes and directly binds WDR5 (211). HOTTIP primarily coordinates expression of genes associated the HOXA locus in fibroblasts (211), and a recent paper showed a close association between HOTTIP and HOXA13 in hepatocellular carcinomas (HCCs) (487). For example, both HOTTIP and HOXA13 are upregulated in HCCs and are associated with metastasis and decreased patient survival (487); moreover, individual knockdown of HOTTIP or HOXA13 by RNA interference (RNAi) in HCC cell lines results in downregulation of HOXA13 and HOTTIP, respectively. Moreover, RNAi studies showed that knockdown of HOTTIP and HOXA13 decreased cell proliferation but did not affect apoptosis in HCC cells (487).

Previous studies in this laboratory showed that knockdown of HOTAIR in pancreatic cancer cells decreased proliferation, induced apoptosis, and inhibited invasion, and this was associated with changes in expression of genes associated with these pathways (261). We have now investigated the role of HOTTIP in pancreatic cancer cells and have observed pro-oncogenic functions similar to that reported for HOTAIR, even though both lncRNAs elicit their effects by regulating expression of different sets of genes by different pathways.

2.2 Materials and methods

2.2.1 Cell lines, reagents, and antibodies

Panc28 cells were a generous gift from Dr. Paul Chiao (University of Texas MD Anderson Cancer Center, Houston, TX), and the L3.6pL cell line was kindly provided by I. J. Fidler (University of Texas MD Anderson Cancer Center). Panc1, ASPC1, BxPC3, MiaPaCa2 cells were obtained from the American Type Culture Collection (Manassas, VA) and HPDE cells were provided by Dr. Ming Sound Tsao (Ontario Cancer Institute, Toronto, Canada). Panc1, L3.6pL, Panc28 and MiaPaCa2 cells were maintained in Dulbecco's modified Eagle medium (DMEM)-Ham's F-12 nutrient mixture (Sigma-Aldrich, St. Louis, MO) with phenol red supplemented with 0.22% sodium bicarbonate, 5% fetal bovine serum (FBS), and 10 ml/liter 100X antibiotic/antimycotic solution (Sigma-Aldrich). BxPC3 cells were maintained in RPMI-1640 medium (Sigma-Aldrich, St. Louis, MO) with phenol red supplemented with 0.15% sodium bicarbonate, 0.24

HEPES, 0.011% sodium pyruvate, 0.45% glucose, 10% FBS and 10 ml/liter 100X antibiotic/antimycotic solution (Sigma-Aldrich). Cells were grown in 150-cm^2 culture plates in an air-CO_2 (95:5) atmosphere at 37°C and passaged approximately every 3 to 5 days. Cleaved PARP (D214) antibodies were purchased from Cell Signaling Technology (Danvers, MA). β-Actin (A1978) was from Sigma-Aldrich; GDF15 (sc-377195) was from Santa Cruz Technology (Dallas, Texas) and Aurora A antibody (A300-071A) was from Bethyl Laboratories Inc. (Montgomery, TX). Chemiluminescence reagents (Immobilon Western) for Western blot imaging were purchased from Millipore (Billerica, MA). Lipofectamine 2000 was purchased from Invitrogen (Carlsbad, CA).

2.2.2 RNA interference and plasmid transfection

Small interfering RNAs (siRNAs) for HOTTIP, MLL, WDR5, AURKA, and a non-specific control (Ctrl) were purchased from Sigma-Aldrich. The siRNA complexes used in this study are listed in Appendix B-1. Panc1 and L3.6pL cells were seeded ($1x10^5$ per well) in 6-well plates in DMEM-Ham's F-12 medium supplemented with 2.5% charcoal-stripped FBS without antibiotic and left to attach for 1 day. Knockdown by RNA interference (RNAi) with siCtrl as a control was performed using Lipofectamine 2000 transfection reagent as per the manufacturer's instructions. Full length HOTTIP in pcDNA3.1+ was kindly provided by Dr. Howard Y. Chang (Stanford University, Stanford, CA) (211). Panc28 cells were seeded ($1x10^5$ per well) in 6-well plates n DMEM-Ham's F-12 medium supplemented with 2.5% charcoal-stripped FBS without antibiotic and left to attach for 1 day. PcDNA3.1-HOTTIP along with pcDNA3.1+

plasmid as a control was performed using Lipofectamine 2000 transfection reagent as per the manufacturer's instructions.

2.2.3 Real time-PCR

Total RNA was isolated using the mirVana miRNA isolation kit (Ambion, Austin, TX) according to the manufacturer's protocol. RNA was eluted with 100 μl of RNase-free water and stored at -80°C. Real-time (RT)-PCR was carried out using iTaq Universal SYBR Green One-step Kit (BioRad). The primers used are listed in Appendix B-2. The housekeeping TATA-binding protein (TBP) mRNA was use as a control for comparing relative expression of RNAs.

2.2.4 Western blot analysis

Pancreatic cancer cells were seeded in 6-well plates using 2.5% DMEM-Ham's F-12 medium, and after 24 hr, western blot analysis of whole-cell lysates was performed essentially as described previously (261).

2.2.5 Xenograft study

Female athymic nude mice, 4 to 6 weeks old, were purchased from Harlan Laboratories (Houston, TX). L3.6pL cells in culture were transfected with 100 nM siHOTTIP or siCT using Lipofectamine 2000. After 48 hr, cells were collected and $1x10^6$ cells in matrigel (1:1 ratio) were injected into either side of the flank area of female nude mice (Harlan). Tumor volumes were measured (0.5 x length x $width^2$) throughout the study and after 10

days, the mice were sacrificed and tumor weights were determined. Tumor volumes and weights were determined in mice from the siHOTTIP (5 mice) or siCT (5 mice) groups, and siHOTTIP levels were determined by real time PCR. Research involving animal experimentation was reviewed and approved by the Texas A&M University Institutional Animal Care and Use Committee.

2.2.6 Cell proliferation, death, and cycle analysis

Panc1 and L3.6pL cells were seeded in 12-well plates and permitted to attach for 24 hr, and then cells were transfected with 100 nM siRNA control or different siRNAs using Lipofectamine 2000 (Invitrogen, Grand Island, NY). Cells were trypsinized and counted at the indicated times using a Coulter Z1 cell counter (Beckman Coulter, Fullerton, CA). For cell cycle analysis, cells were stained with propidium iodide solution (50 μg/ml) and were analyzed by a FACSCalibur flow cytometer 24 hr after transfection. Apoptosis was detected using a fluorescein isothiocyanate (FITC) Annexin V staining kit (Life Technologies, Grand Island, NY) followed by fluorescence-activated cell sorter (FACS) analysis according to the manufacturer's protocol.

2.2.7 Transwell migration and scratch assays

Panc1 and L3.6pL cells were first transfected with siRNA for 24 hr, then added to the upper chamber of a transwell chamber in duplicate and allowed to migrate into the lower chamber containing Hams F12 media with 20% FBS by incubating for 24 hr at 5% CO_2 at 37°C. Cells migrating to the outer side of the upper chamber were fixed, stained and

counted, and cell migration was also determined using a scratch assay. For the scratch assay, cells were first seed in 6-well plate for 24 hr, and then a scratch through the central axis of the plate was gently made using a sterile pipette tip. Cells were transfected with siHOTTIP or siCTL, media was changed after 6 hr, and migration of the cells into the scratch was observed after 48, 72 and 96 hr. Three replicates were obtained for each time point.

2.2.8 Microarray analysis

Total RNA was extracted from Panc1 cells by using a mirVanaTM miRNA Isolation Labeling Kit (Ambion Inc.). The total RNA was quantified by using a Nanodrop ND-1000 spectrophotometer (NanoDrop Technology). The total RNA samples with adequate RNA quality index (>7) were used for microarray analysis; 700 ng of total RNA was used for labeling and hybridization on HumanHT-12 v4 expression beadchip (Illumina, Inc.) according to the manufacturer's protocols. After the beadchips were scanned with a BeadArray Reader (Illumina), the microarray data were normalized using the quantile normalization method in the Linear Models for Microarray Data (LIMMA) package in the R language (http://www.r-project.org). BRB-ArrayTools were primarily used for statistical analysis of gene expression data, and the Student’s t test was applied to identify the genes significantly different between 2 groups when compared. Differentially expressed genes were identified using >1.5 or 2 fold change cut off. Gene ontology enrichment analysis was carried out using David Functional Annotation Resources 6.7 (http://david.abcc.ncifcrf.gov/). Data for gene expression study of

pancreatic ductal adenocarcinoma were downloaded from Gene Expression Omnibus (GEO, NCBI) (http://www.ncbi.nlm.nih.gov/geoprofiles/).

2.2.9 Comparison of gene expression changes from lncRNA knockdown to a gene expression database

A rank-based nonparametric analysis strategy called the Running Fisher's algorithm and implemented within the NextBio database (http://www.nextbio.com/) environment was used to identify gene expression comparisons (biosets) which have statistically significant positive or negative correlation to the genes regulated by siHOTTIP. The Running Fisher's algorithm computes statistical significance of similarity between ranked fold-change values of two gene lists using a Fisher's exact test (488). After exporting the analysis, the list of correlated biosets were filtered to identify those that examined gene expression changes in pancreatic cancers.

2.2.10 Statistical analysis

Statistical significance of differences between the treatment groups was determined by an analysis of variance and/or Student's *t* test, and levels of probability were noted. At least 3 repeated experiments were determined for each data points and results are expressed as means ± SD.

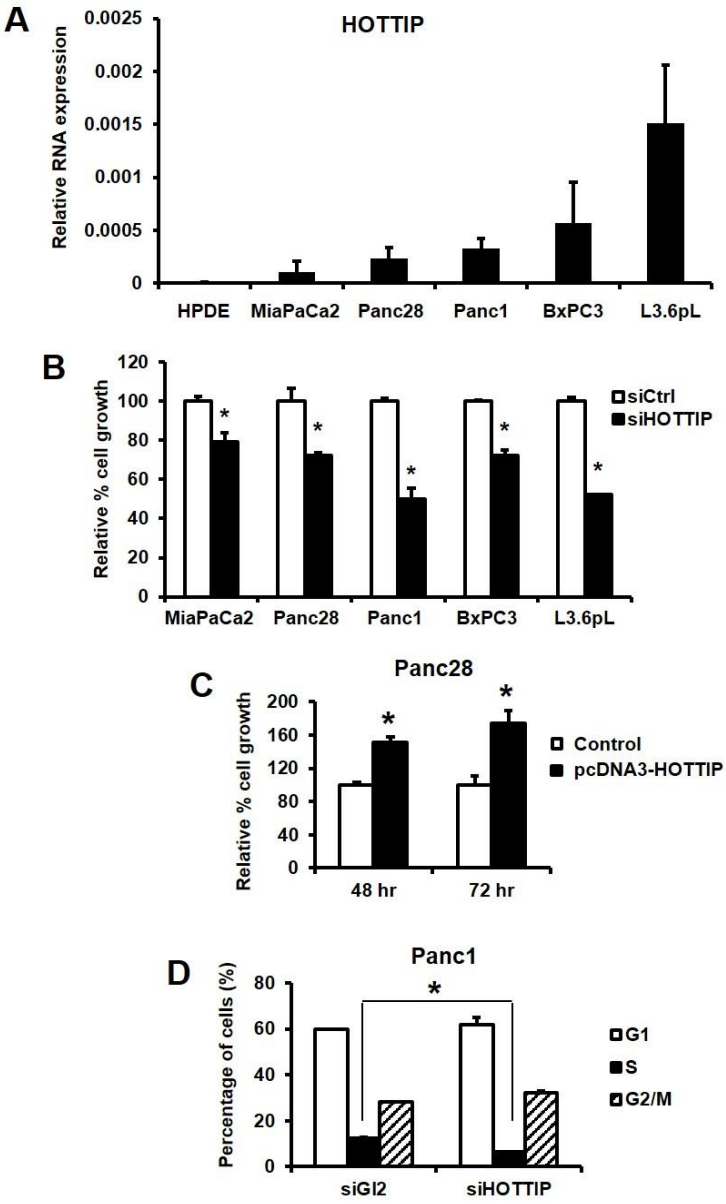


Figure 11. Effects of HOTTIP in pancreatic cell proliferation and cell cycle. (A) HOTTIP expression relative to the housekeeping gene TATA-binding protein (TBP) was determined by real time PCR as described in the Materials and Methods. HOTTIP knockdown by RNAi knockdown in Panc28, MiaPaCa2, Panc1, BxPC3 and L3.6pL cells inhibited cell growth (B), whereas HOTTIP overexpression in Panc28 cells promoted cell proliferation (C). (D) The effect of siHOTTIP (knockdown) on cell cycle progression in Panc1 cells was determined by FACS analysis as described in the Materials and Methods. Results (A-D) are expressed as means ± SD for 3 replicates. Cells transfected with a non-specific oligonucleotide (siCtrl) were used as controls and significant ($p<0.05$) changes are indicated (*).

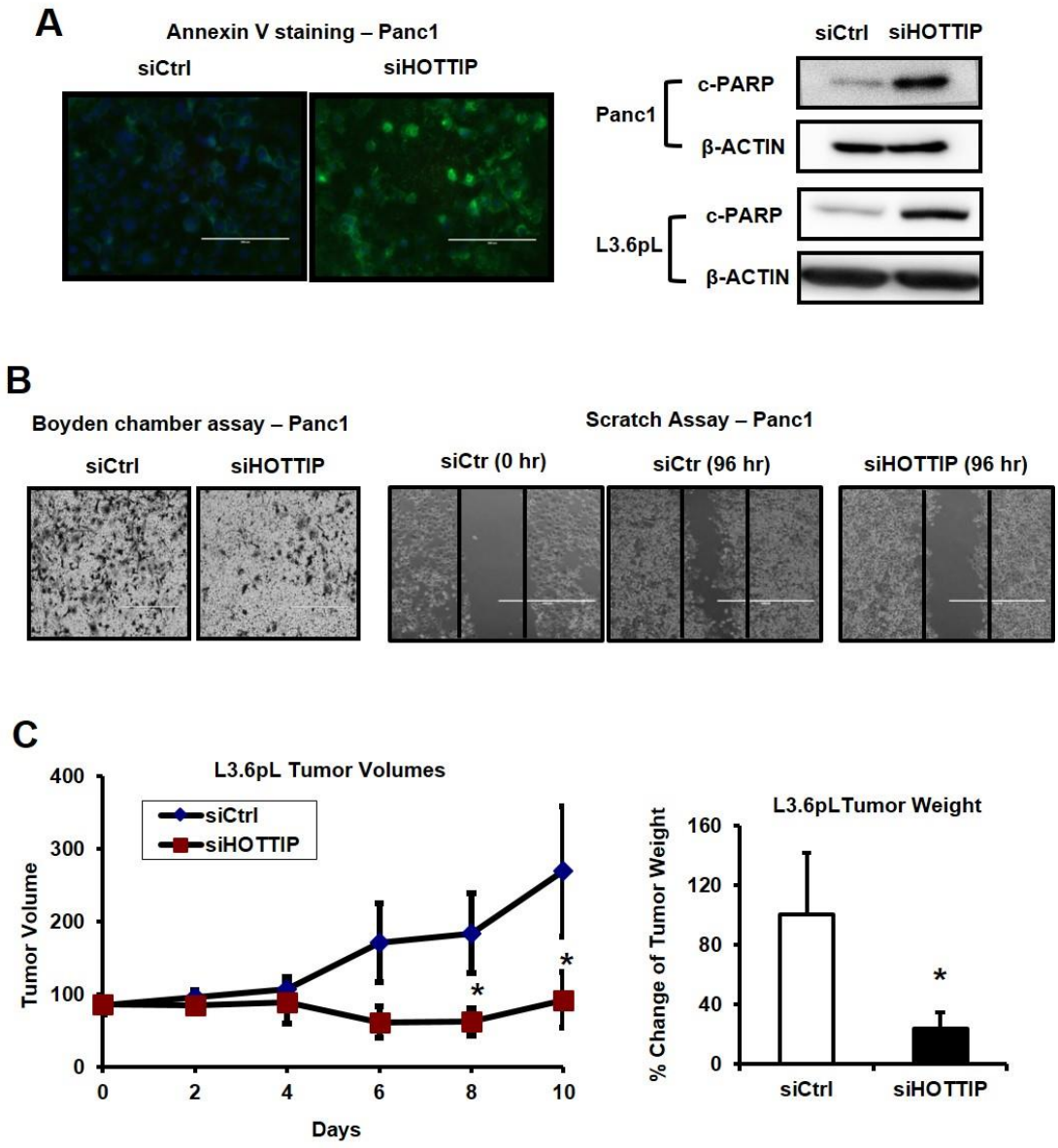


Figure 12. Effects of HOTTIP in pancreatic cell apoptosis, migration and tumor growth. (A) Panc1 cells were transfected with siHOTTIP and after 48 hr, and the increase in Annexin V staining and induction of PARP cleavage were determined by fluorescence and Western blots analysis respectively. (B) HOTTIP knockdown reduced cell migration as determined by the Boyden chamber and scratch assay as discussed in the Materials and Methods. (C) HOTTIP was silenced in L3.6pL cells which were then used in the athymic nude mice as xenografts. Tumor volumes were determined for up to 10 days, and tumor weights were measured after the animals were sacrificed at Day 10. Cells transfected with a non-specific oligonucleotide (siCtrl) were used as controls, and five mice were used in each treatment group. Significant ($p<0.05$) changes are indicated (*).

2.3 Results

2.3.1 HOTTIP: functional studies as determined by knockdown and overexpression

Figure 11A illustrates the expression of HOTTIP in five pancreatic cancer cell lines in which high expression is observed in Panc1, L3.6pL and MiaPaCa2 cells and lower (> 2-fold) expression in Panc28 and BxPC3 cells. HPDE cells are non-transformed immortalized pancreatic epithelial cells and only minimal expression of HOTTIP was observed. Knockdown of HOTTIP by RNAi significantly decreased proliferation of Panc1, L3.6pL, Panc28, BxPC3 and MiaPaCa2 cells (Figure 11B) and overexpression of HOTTIP increased proliferation of Panc28 cells (Figure 11C), and the growth promoting effects of HOTTIP were similar to those previously reported for HOTAIR in pancreatic cancer cells (261). Knockdown of HOTTIP in Panc1 cells slightly decreased the percentage of cells in S phase and increased the percentage of cells in G_2/M phase compared to Panc1 cells transfected with the scrambled siRNA (Figure 11D). Knockdown of HOTTIP by RNAi induced Annexin V staining and enhanced PARP cleavage in Panc1 cells (Figure 12A), demonstrating that HOTTIP plays a role in pancreatic cancer cell survival. Moreover, results of Boyden chamber and scratch assays (Figure 12B) show that knockdown of HOTTIP significantly decreased Panc1 cell migration and these results were similar to those previously observed in comparable experiments with HOTAIR in pancreatic cancer cells (261). Moreover, like HOTAIR, knockdown of HOTTIP in L3.6pL cells which were used in a xenograft model in athymic nude mice decreased tumor growth and tumor weights (Figure 12C) compared

to tumors in cells expressing HOTTIP (transfected with a non-specific oligonucleotide). Thus, HOTTIP plays a pro-oncogenic role in pancreatic cancer cells.

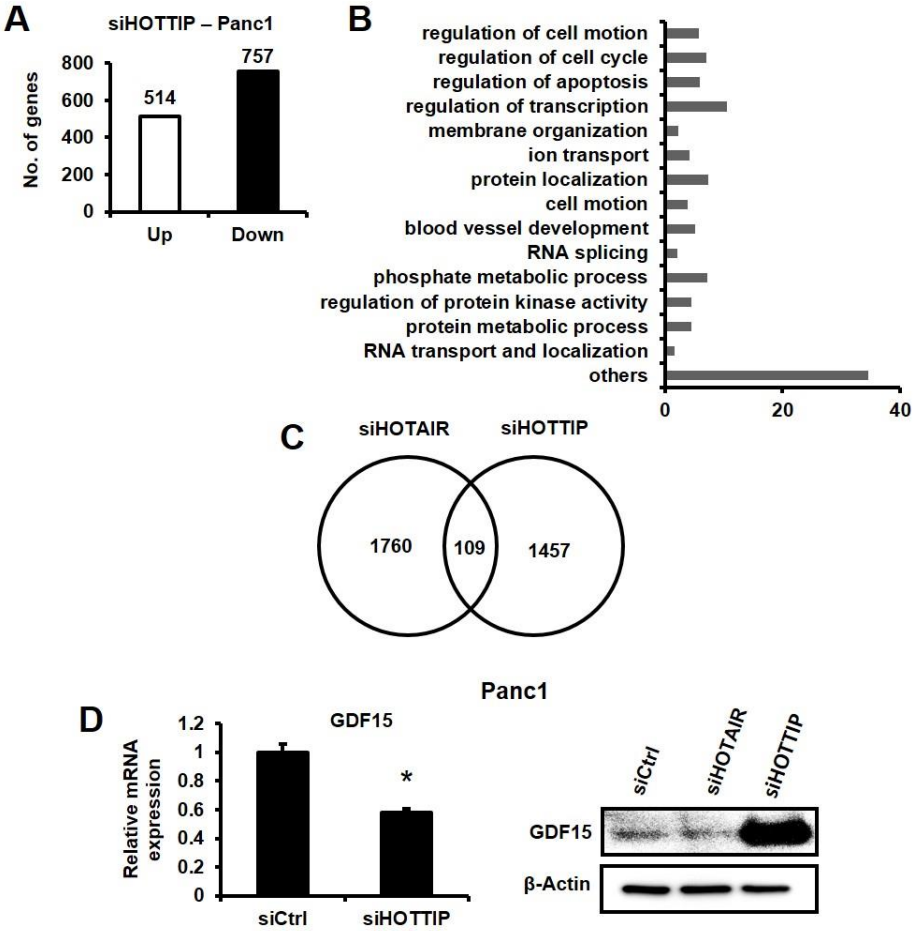


Figure 13. Gene regulation by HOTTIP and compared to HOTAIR in pancreatic cell line. (A) Panc1 cells were transfected with siHOTTIP or siCtrl and gene expression was analyzed using HumanHT-12 v4 expression beadchip (Illumina Inc.) array, and (B) the effects of siHOTTIP on different pathways were determined by gene ontology enrichment analysis. (C) Overlap of common genes was observed after treatment of Panc1 cells with siHOTTIP or siHOTAIR. (D) Panc1 cells were transfected with siHOTTIP and GDF15 mRNA, and protein levels were determined compared to the effects of siHOTAIR. Results for GDF15 mRNA are means ± SD for 3 replicated determinations, and significant ($p<0.05$) change is indicated (*). Cells were also transfected with siHOTTIP or siHOTAIR and GDF15 protein was analyzed by Western blots.

2.3.2 Regulation of gene expression by HOTTIP determined by knockdown and analysis by beadchip arrays

Regulation of gene expression by HOTTIP was investigated by comparing Panc1 cells transfected with scrambled siRNA with those that were transfected with siHOTTIP in Panc1 cells followed by analysis of gene expression using an Illumina Human V.3 HT12 beadchip array (261). Transfection of cells with siHOTTIP resulted in increased expression of 514 (HOTTIP-repressed genes) and decreased expression of 757 genes (HOTTIP-enhanced genes) (Figure 13A). Gene ontology enrichment analysis demonstrated that HOTTIP-regulated genes could be classified into several categories (Figure 13B), including those associated with the functions of HOTTIP in cell growth, survival and migration (Figures 11 and 12). Figure 13C illustrates that among the 1271 genes regulated by HOTTIP and 1006 genes regulated by HOTAIR (261), there were only 109 genes (< 5%) commonly regulated by both lncRNAs. Among the 109 commonly regulated genes, 87 genes were decreased and 22 were increased after transfecting cells with siHOTTIP or siHOTAIR; however, with a cut-off of 2-fold the number of common genes regulated by both lncRNAs was only 13 (decreased) and 2 (increased). Growth differentiation factor 15 (GDF15) is an example of the differences between HOTTIP- and HOTAIR-regulated genes; in Panc1 cells knockdown of HOTAIR induced expression of GDF15, whereas siHOTTIP decreased expression of both GDF15 mRNA and protein (Figure 13D), and induction of GDF15 protein after HOTAIR knockdown was consistent with previous studies (261).

Since HOTTIP associates with the PRC2 and WDR5/MLL1 chromatin-modifying complexes (211), we compared the overlap in genes regulated by HOTTIP or these complexes after transfection of Panc1 cells with siHOTTIP, siMLL1 and siEZH2 followed by microarray analysis of changes in gene expression (Figure 14A). The results suggest that both MLL1 and HOTTIP or EZH2 and HOTTIP coregulate expression of 547 and 209 genes, respectively, and we further investigated genes commonly regulated by MLL1 and HOTTIP. Results in Figure 14B confirm that for several genes that contribute to enhanced carcinogenesis including *AURKA*, *AHNAK*, *GDF15*, *SGK1* and *CD44* (489-494), knockdown of HOTTIP resulted in decreased expression of these genes. Since AURKA (Aurora-A kinase) plays an important multifunctional role in pancreatic cancer (491-493), we further investigated the function and regulation of AURKA in pancreatic cancer cells. Results in Figure 14B show that knockdown of AURKA decreased Panc1 cell growth, and this was accompanied by a dramatic decrease in the percentage of cells in G_0/G_1 and an increase in cells in S and G_2/M phase. Transfection of Panc1 cells with siAURKA also induced Annexin V staining and PARP cleavage (Figure 14C) and inhibited Panc1 cell migration in a scratch assay (Figure 14D). Transfection of Panc1 cells with siHOTTIP and siAURKA induced similar functional changes and this included inhibition of cell growth, induction of apoptosis, and decreased migration. However, the effects of siHOTTIP and siAURKA on cell cycle progression were different. We also observed that transfection of Panc1 cells with siHOTTIP and siMLL1 decreased expression of AURKA protein; however, transfection of cells with oligonucleotides that knockdown WDR5 (siWDR5) increased AURKA

protein levels in Panc1 cells (Figure 14D). This indicates that the effects of HOTTIP/MLL1 on enhanced *AURKA* gene expression are independent of WDR5, and the coregulation of genes by HOTTIP and other MLL1-associated chromatin-modifying complexes is currently being investigated.

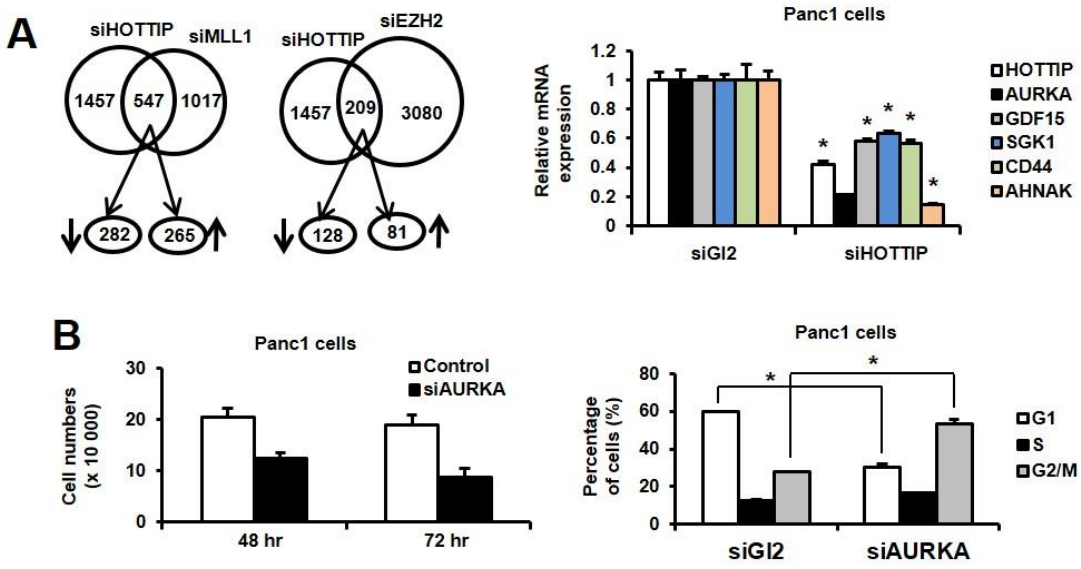


Figure 14. Interaction between HOTTIP, WDR5 and MLL1 on gene expression in Panc1 cells. (A) Cells were transfected with siHOTTIP, siMLL, siEZH2 or siCtrl, and analyzed by microarrays for changes in gene expression. Regulation of the same genes is illustrated in the Venn diagram. Panc1 cells were transfected with siHOTTIP or siCtrl and genes coregulated by HOTTIP or MLL1 were analyzed by real time PCR. (B) Panc1 cells were transfected with siAURKA or siCtrl, and effects on cell growth and cell cycle progression were determined as mentioned in the Materials and Methods. (C) Effects of Annexin V staining and PARP cleavage and (D) cell migration in a scratch assay were determined as Materials and Methods. (E) Panc1 cells were transfected with siHOTTIP, siMLL1 and siWDR5 or siCtrl and analyzed by Western blots as outlined in the Materials and Methods. Results are expressed as means ± SD for 3 replicated determinations, and significant ($p<0.05$) change is indicated (*).

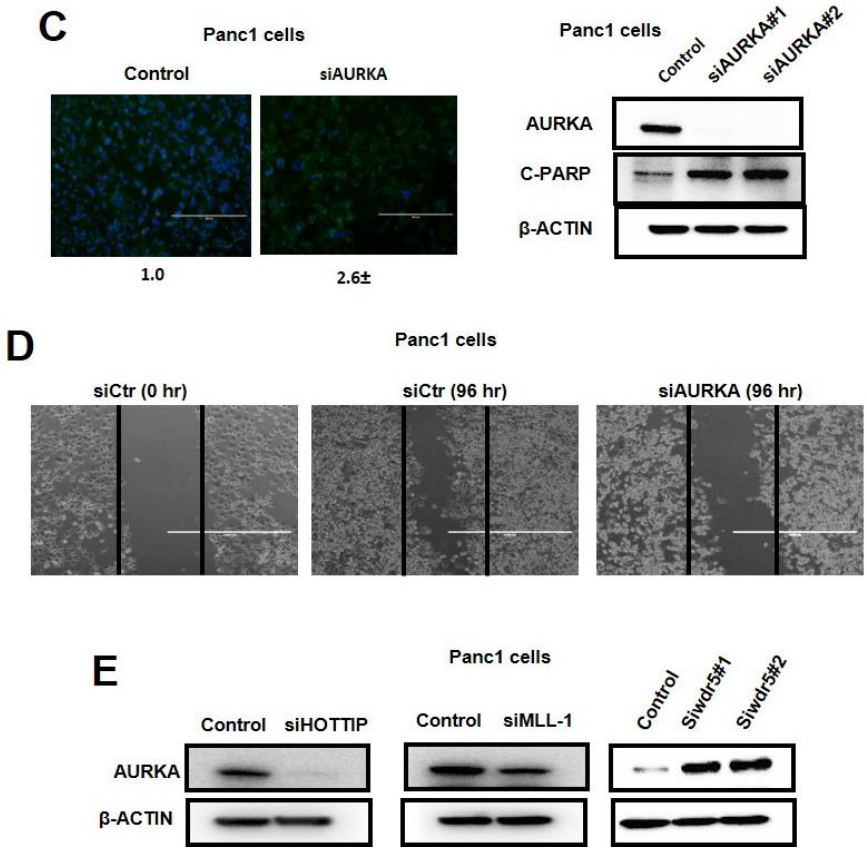


Figure 14. Continued.

We also compared the inverse expression of genes in Panc1 cells transfected with siHOTTIP with those overexpressed in pancreatic tumors compared to paired adjacent normal tissue (GSE16515) and observed 39 genes that were inversely regulated (Appendix B-3). Figure 15A summarizes the effects of siHOTTIP on expression of 4 genes upregulated in tumors (GSE16515) and decreased after knockdown of HOTTIP in Panc1 cells. Examination of gene expression comparisons from GSE16515 (pancreatic tumor vs. paired adjacent normal tissue and unpaired normal pancreatic tissue) (Figure 15B), GSE15471 (tumor vs. adjacent normal tissue) (Figure 15C), and GSE3654 (tumor

vs. normal pancreatic tissue) (Figure 15D) showed that 39, 44, 49 and 16 genes upregulated in tumors were downregulated in Panc1 cells transfected with siHOTTiP. In contrast, only a few (0-6 genes in the data sets) inversely related genes were downregulated in tumors and increased in Panc1 cells transfected with siHOTTIP and these were not included in the comparison with siHOTAIR results. Results in Figures 15B, 15C and 15D show the number of common genes regulated in the four human tumor data sets and decreased in Panc1 cells after transfection with siHOTTIP or siHOTAIR (261) and the overlap of the common siHOTTIP/siHOTAIR genes. We observed a range of 0-2 overlapping genes in the four data sets. These results further emphasize that although both HOTTIP and HOTAIR have similar functions in pancreatic cancer, they regulate very different sets of genes.

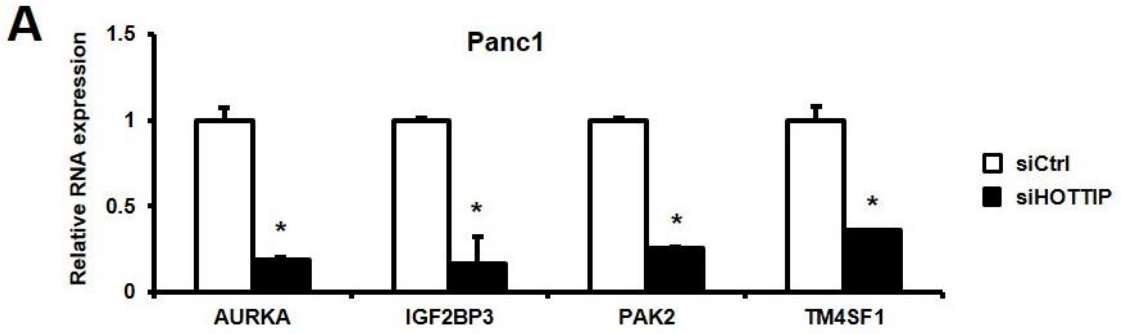


Figure 15. Inverse correlation of genes decreased by siHOTTIP and siHOTAIR and increased in pancreatic tumors. (A) Panc1 cells were transfected with siHOTTIP, and expression of 9 selected genes upregulated in pancreatic tumors (GSE16515) was determined by real time PCR as outlined in the Materials and Methods. Overlap of expression of common genes induced in human tumors in GSE1655 (B), GSE15471 (C) and GSE3654 (D) data sets and decreased in Panc1 cells transfected with siHOTTIP or siHOTAIR (261).

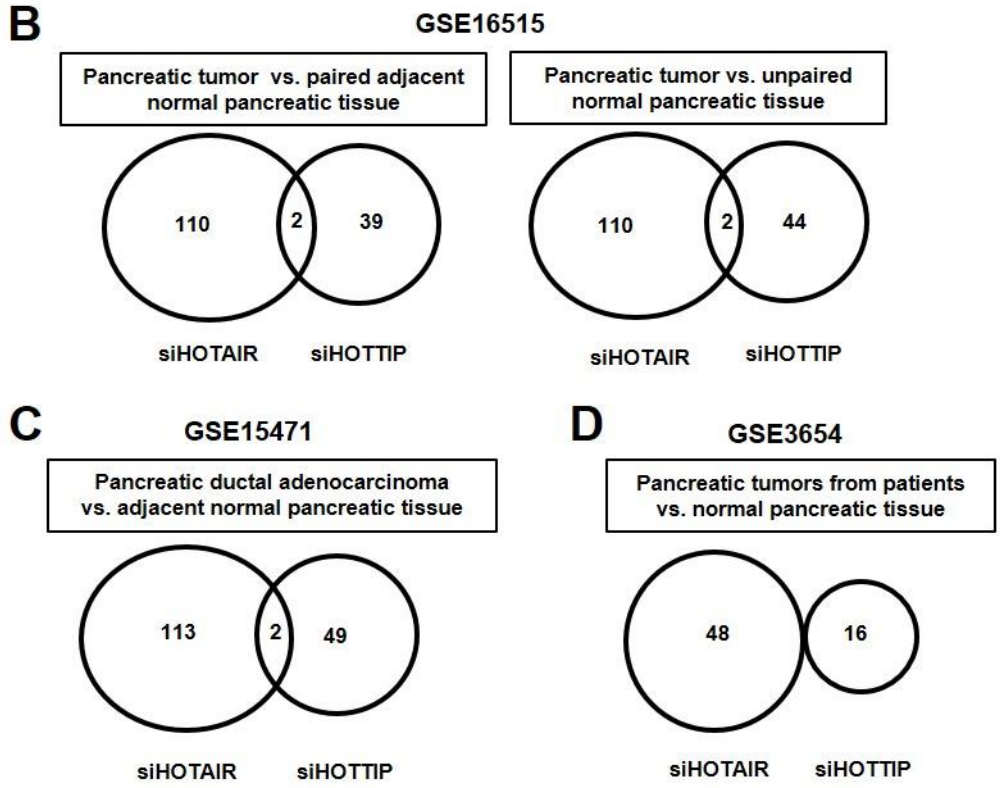


Figure 15. Continued.

2.3.3 HOTTIP regulation of specific HOX genes

Previous studies in foreskin fibroblasts show that HOTTIP knockdown decreases expression of 5' *HOXA* genes, particularly *HOXA13* (211) and in liver cancer cells and tumors, there was a parallel expression of HOTTIP and HOXA13 and siHOTTIP decreased HOXA13 mRNA levels (487). HOXA13 is more highly expressed than HOTTIP by over 2 orders of magnitude in all of the pancreatic cancer cell lines (Figure 16A); however, despite these differences in the magnitude of expression, there was a correlation between expression of HOTTIP and HOXA13 in most of these cell lines. However, in Panc1 cells transfected with siHOTTIP, there was only a slight decrease in

HOXA13 expression, whereas siHOTTIP decreased HOXA13 in SNU-499 liver cancer cells (Figure 16B) and this was consistent with previous studies in liver cancer cell lines (487). It has previously been reported that some *HOX* genes are overexpressed in pancreatic tumors and they include *HOXA10*, *HOXB7* and *HOXB2* (495-499), and transfection of Panc1 cells with siHOTTIP slightly decreased expression of *HOXB7* but significantly decreased *HOXA10* (>80%) and *HOXB2* (>60%). Transfection with siHOTTIP also decreased mRNA levels of HOXA11 (>75%), HOXA9 (>80%) and HOXA1 (>60%), whereas siHOTTIP decreased expression of HOXB7 by <25% (Figure 16C). Thus, HOTTIP regulates expression of several *HOX* genes in pancreatic cancer cells but in contrast to liver cancer cells, HOTTIP does not regulate expression of *HOXA13*. A recent study showed that HOXA10 expression in pancreatic cancer cells was associated with regulation of matrix metalloproteinase 3 (MMP-3) (496) and in Panc1 cells transfected with siHOTTIP, we also observed decreased expression of HOXA10 and MMP-3 (Figures 16C and 16D). We also observed that SMAD3, a negative prognostic factor for pancreatic cancer patients and gene that promotes epithelial-mesenchymal transition (500) was also decreased after HOTTIP knockdown (Figure 16D). It has also been reported that HOXA11 regulates MMP-2 expression (501) and transfection with siHOTTIP also decreased both HOXA11 and MMP-2 (Figures 16C and 16D), suggesting that HOTTIP regulation of HOXA10 and HOXA11 and their downstream genes contribute to the oncogenic role of HOTTIP in pancreatic cancer cell migration/invasion.

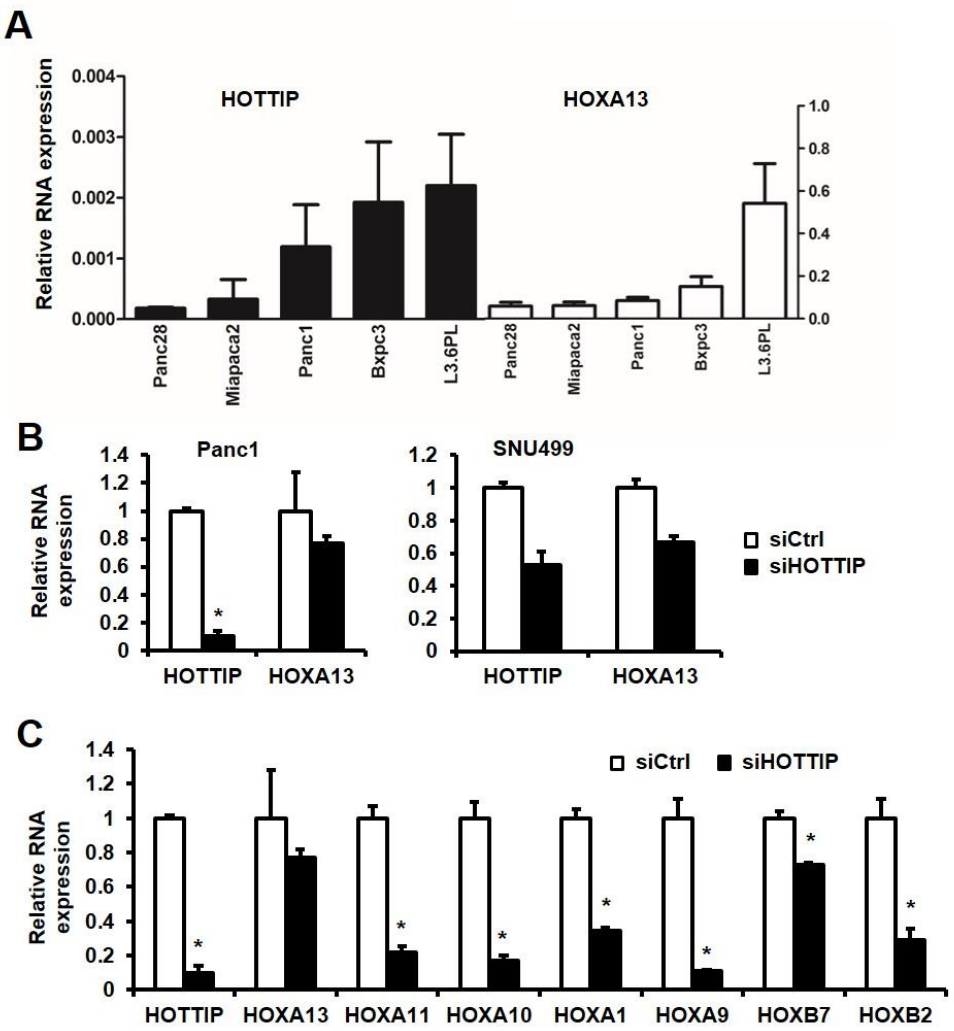


Figure 16. Coregulation of HOTTIP and *HOX* genes in pancreatic cancer cell lines. (A) The relative expression of HOTTIP and HOXA13 compared to TBP mRNA in several pancreatic cancer cell lines was determined by real time PCR. (B) Panc1 and SNU499 cells were transfected with siHOTTIP and expression of HOXA13 was determined by real time PCR as outlined in the Materials and Methods. Panc1 cells were transfected with siHOTTIP and expression of several HOX mRNAs (C) and MMP/AURKA and SMAD3 mRNA (D) were determined by real time PCR as outlined in the Materials and Methods. Results are expressed as means ± SD for 3 replicated determinations, and significant ($p<0.05$) change is indicated (*).

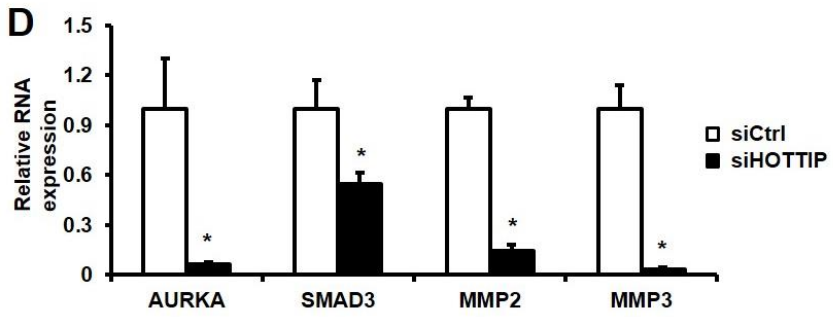


Figure 16. Continued.

2.4 Discussion

Rinn and coworkers have identified 231 non-coding RNAs associated with human *HOX* gene loci and these RNAs are spatially expressed and sequence-specific (482). HOTAIR was the first HOX-associated lncRNA that was characterized and was initially identified as a scaffold RNA associated with the chromatin-modifying PRC2 complex and the H3K27me3 histone mark which is associated with gene suppression. Subsequent studies showed that HOTAIR directly interacted with both the PRC2 and LSD1/REST/coREST repressor complexes (209,210,482), and in multiple tumor and cancer cell lines, HOTAIR-regulated gene expression enhances tumorigenesis (209,210,261,483-486). HOTAIRm2, Mistral, HOTTIP and more recently linc-HOXA1 are other HOX-associated lncRNAs that have been investigated (209,210,482,502-504) and linc-HOXA1 represses expression of HOXA1 in combination with the cofactor PURB (504).

The pro-oncogenic functions and negative prognostic significance of HOTAIR have been reported for several cancers including pancreatic cancer, whereas with the

exception of one paper on HOTTIP in liver tumors and cells (487), the expression and functions of other HOX-associated lncRNAs in cancer cell lines have not been extensively investigated. In liver cancer cell lines and tumors, HOTTIP is closely associated with expression of HOXA13 and knockdown of HOTTIP decreases expression of HOXA13 (487). In the present study, we investigated the expression and function of HOTTIP in pancreatic cancer cells and compared the results to that observed in previous studies on HOTAIR in pancreatic cancer cells (261). Results of knockdown and overexpression studies show that HOTTIP has functions comparable to that described for HOTAIR and plays a role in pancreatic cancer cell proliferation, survival and migration/invasion (Figures 11 and 12). However, a comparison of the genes regulated by HOTTIP and HOTAIR in Panc1 cells after knockdown by RNA interference showed that there was minimal gene overlap (Figure 13C), even though the pro-oncogenic functions of HOTTIP and HOTAIR are comparable. We also examined the overlap of genes overexpressed in publically available pancreatic tumor data sets and genes downregulated in Panc1 cells transfected with siHOTTIP (Figure 15). Among the human data sets (GSE16515, GSE15471 and GSE 3654), the number of overexpressed genes in tumors and genes downregulated by siHOTTIP that were in common varied among the data sets (16-49 genes in common) and similar variability (48-113 genes in common) was observed with genes downregulated by siHOTAIR in Panc1 cells (261). Some of these differences in the number of common genes may be due to the sensitivity and composition of the different arrays that were used in these studies. However, the most striking observation was the minimal overlap between the genes regulated by

HOTTIP vs. HOTAIR, further confirming the independent pro-oncogenic functions and gene regulation by these two lncRNAs in pancreatic cancer.

Since HOTTIP interacts with both the PRC2 and MLL1/WDR5 chromatin-modifying complexes (211), we also investigated by RNAi and microarrays, the overlap of genes coregulated by HOTTIP and MLL1/WDR5 (siHOTTIP/siMLL1) and by HOTTIP and PRC2 (siHOTTIP/siEZH2) (Figure 14A). Although both HOTTIP and MLL1 coregulated several genes in common, it was apparent that <40% of all HOTTIP-regulated genes were coregulated by HOTTIP and the two complexes. Moreover, among several genes that were decreased in Panc1 cells after transfection with siHOTTIP (Fig. 4B), we observed that AURKA was coregulated by HOTTIP and MLL1 but not HOTTIP and WDR5 (Figure 14E) which has been reported to bind directly to HOTTIP (211). Interestingly, knockdown of WDR5 increased AURKA protein levels, whereas siHOTTIP and siMLL1 decreased AURKA protein (Figure 14E). These results suggest regulation of gene expression by HOTTIP in pancreatic cancer cells involves interaction with complexes in addition to PRC2 and MLL1/WDR5, and this includes MLL1 complexes independent of WDR5. These interactions are currently being investigated.

Our results also showed that HOTTIP regulated expression of multiple *HOX* genes in pancreatic cancer cell cells (Figure 16C) but, in contrast to results in liver cancer cell lines (487), HOTTIP did not regulate expression of HOXA13 (Figure 16B). Previous studies indicated that HOXA10 and HOXA11 regulate expression of MMP-3 and MMP-

2, respectively (496,501), and both MMPs contribute to migration/invasion of pancreatic cancer cells (496,505-507). Figures 16C and 16D show that siHOTTIP decreased both HOXA10 and HOXA11 and this was paralleled by decreased expression of MMP-3 and MMP-2, demonstrating that HOTTIP regulates specific *HOX* genes that play a role in the migration/invasion of pancreatic cancer cells. Thus, HOTTIP functions in pancreatic cancer cells are due, in part, to regulation of some *HOX* genes but not *HOXA13* as previously observed in liver cancer cells (487). In addition, the HOX gene targets in HOTTIP in pancreatic cancer cells are different from those regulated by HOTTIP in primary human fibroblasts (482). Results of our study demonstrate a novel pro-oncogenic role for HOTTIP in pancreatic cancer cells, and we are currently investigating HOTTIP expression in tumors and both the *HOX*-dependent and -independent pro-oncogenic functions of HOTTIP in pancreatic and other cancer cell lines. Current studies are focused on the regulation of HOTTIP expression and the discovery of agents that target this lncRNA.

3. ROLE OF METASTASIS-ASSOCIATED LUNG ADENOCARCINOMA TRANSCRIPT 1 (MALAT1) IN PANCREATIC CANCER*

3.1 Introduction

Results of high throughput sequencing technologies show that less than 2% of the human genome encodes for proteins, whereas up to 75% of the genome transcribes non-coding RNAs (ncRNAs) which are highly variable in length, function and regulation (508). MicroRNAs (miRNAs) are 21-23 bp in length and there is now extensive evidence showing that miRNAs play important roles in maintaining cellular homeostasis and in diseases such as cancer through their sequence-specific regulation (primarily repression) of genes (509,510). There is also evidence that long ncRNAs (lncRNAs) greater than 200 bp play an equally important role in normal cell functions and disease, and estimations from the Encyclopedia of DNA Element Project Consortium indicates that the human genome contains up to 16,000 genes encoding 28,000 lncRNA transcripts (508,511).

Metastasis-Associated-in-Lung-Adenocarcinoma-Transcript-1 (MALAT1) is a lncRNA that is overexpressed in multiple cancer cell lines and tumors, and initial studies show

* Reprint from "Role of metastasis-associated lung adenocarcinoma transcript-1 (MALAT-1) in pancreatic cancer" by Cheng Y, Imanirad P, Jutooru I, Hedrick E, Jin UH, *et al*. Carcinogenesis (In review) 2016

that MALAT1 expression in early stage non-small cell lung cancer patients predicted patient survival and metastasis (512). The expression and prognostic value of MALAT1 has now been extensively investigated in multiple tumor types. The results demonstrate that high tumor expression of MALAT1 is a negative prognostic factor for lung, liver, pancreatic, melanoma, cervical, colorectal, gastric, multiple myeloma, clear cell, renal cell, glioma and adrenocortical cancer patients (262-264,266,512-521). The negative prognosis associated with overexpression of MALAT1 in tumors correlates with the functions of MALAT1 in cancer cells. For example, results of knockdown studies in lung cancer cells indicate that loss of MALAT1 decreased migration/wound healing and injection of MALAT1-deficient A549 lung cancer cells in mice resulted in significantly decreased formation of lung nodules. In mice bearing A549 cells as xenografts, injection of MALAT1 antisense oligonucleotides significantly decreased tumor growth (521).

The mechanisms associated with the pro-oncogenic functions of MALAT1 involve multiple pathways and genes. Initial studies reported that MALAT1 stably localizes to nuclear speckles and interacts with pre-mRNA splicing factors (SR proteins) to modulate gene expression in some cells (e.g. HeLa cells) (201,522-524). However, in MALAT1 knockout mice, the formation of nuclear speckles and pre-mRNA splicing were unaffected (525-527), suggesting that pro-oncogenic function of MALAT1 may be independent of nuclear speckles. MALAT1 exhibits multiple pro-oncogenic functions in cancer cells and plays a role in cell proliferation, survival, epithelial to mesenchymal transition (EMT), migration and metastasis through diverse mechanisms which include

acting as a decoy sponge and direct or indirect interactions with DNA, RNA and protein (528-536). Recent studies show that MALAT1 is a negative prognostic factor for pancreatic cancer (263,264) and our preliminary results also demonstrate that the pro-oncogenic functions of MALAT1 in pancreatic cancer are similar to those observed in other tumors. Results of microarray and knockdown studies also suggest a role for MALAT1 as a scaffold for chromatin modifying complexes. Although loss of MALAT1 decreased pancreatic tumor growth and metastasis in a mouse orthotopic model (262,267), the loss of MALAT1 in highly aggressive p53L/L; LSL-KrasG12DL/+; p48Cre+/- and the p53L/+ heterozygotes has minimal effects on mouse survival and tumor pathology.

3.2 Materials and methods

3.2.1 Cell lines, reagents, and antibodies

Panc28 cells were a generous gift from Dr. Paul Chiao (University of Texas MD Anderson Cancer Center, Houston, TX), and the L3.6pL cell line was kindly provided by I. J. Fidler (University of Texas MD Anderson Cancer Center). Panc1, ASPC1, BxPC3, MiaPaCa2 cells were obtained from the American Type Culture Collection (Manassas, VA) and HPDE cells were provided by Dr. Ming Sound Tsao (Ontario Cancer Institute, Toronto, Canada). Panc1, L3.6pL, Panc28 and MiaPaCa2 cells were maintained in Dulbecco's modified Eagle medium (DMEM) (GenDEPOT, Barker, TX, USA) supplemented with 10% fetal bovine serum (FBS). BxPC3 and ASPC1 cells were maintained in RPMI-1640 medium (GenDEPOT, Barker, TX, USA) supplemented with

10% FBS. Cells were grown in 150-cm2 culture plates in an air-CO2 (95:5) atmosphere at 37°C and passaged approximately every 3 to 5 days. Cleaved PARP (D214), cleaved caspase-7 (Asp198) (D6H1), cleaved caspase-9 (Asp330) (D2D4), and GAPDH antibodies were purchased from Cell Signaling Technology (Danvers, MA). Sp3 and Sp4 antibodies were purchased from Santa Cruz Biotech (Santa Cruz, CA, USA) and a Sp1 antibody was purchased from Abcam (Cambridge, MA, USA). β-Actin (A1978) antibody was obtained from Sigma-Aldrich. 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT); glutathione was purchased from ThermoFisher Scientific (Waltham, MA, USA). Chemiluminescence reagents (Immobilon Western) for Western blot imaging were purchased from Millipore (Billerica, MA). Lipofecatmine 2000 was purchased from Invitrogen (Carlsbad, CA).

Table 6. List of siRNAs.

Name	Sequence
SiGL2	Sigma CGUACGCGGAAUACUUCGA
SiMALAT1#1	Sigma SASI_Hs02_00377093
SiMALAT1#2	Ambion cat#4390771
SiAPAF1	Sigma SASI_Hs02_00331274
SiEZH2	Sigma SASI_Hs01_00147882
SiLSD1	Sigma SASI_Hs01_00213078
SiMLL-1	Sigma SASI_Hs01_00090459
SiNDRG1	Sigma SASI_Hs01_00034470
SiSP1	Sigma SASI_Hs02_00333289
SiSP3	Sigma SASI_Hs01_00211941
SiSP4	Sigma SASI_Hs01_00114420

3.2.2 RNA interference

Pancreatic cancer cells were seeded ($1x10^5$ per well) in 6-well plates in DMEM medium supplemented with 2.5% FBS and left to attach for 1 day. Knockdown by RNA interference (RNAi) with siCtrl as a control was performed using lipofectamine 2000 transfection reagent as per the manufacturer's instructions. The siRNAs were listed in Table 6.

3.2.3 Real time-PCR

Total RNA was isolated using Zymo Quick RNA MiniPrep Kit (Zymo Research, Irvine, CA) according to the manufacturer's protocol. RNA was eluted with 35 μl of RNase-free water and stored at -80°C. Real-time (RT)-PCR was carried out using iTaq Universal SYBR Green One-step Kit (Bio-Rad, Hercules, CA). The following primers were used (Tables 7 and 8).

Table 7. Human primers used for real Time-PCR.

Gene Name	Forward Primer	Reverse Primer
NRAS	GCACCATAGGTACATCATCCG	GCTTCCTCTGTGTATTTGCCA
PCNA	AAGAGAGTGGAGTGGCTTTTG	TGTCGATAAAGAGGAGGAAGC
SPRY2	GAAGTGTGGTCACTCCAGCA	TTGCACATCGCAGAAAGAAG
SMAD3	CGGCAGTAGATGACATGAGG	TCAACACCAAGTGCATCACC
APAF1	CCTCTCATTTGCTGATGTCG	TCACTGCAGATTTTCACCAGA

Table 8. Mouse primers used for real Time-PCR.

Gene Name	Forward Primer	Reverse Primer
PCNA	GGAGACAGTGGAGTGGCTTT	TGGATAAAGAAGAGGAGGCG
MALAT1	TGAAAAAGGAAATGAGGAGAAAAG	CTTCACAAAACCTCCCTTTACAAT
GAPDH	TTGATGGCAACAATCTCCAC	CGTCCCGTAGACAAAATGGT

3.2.4 Western blot analysis

Seventy-two hours after siRNA transfection, cells were collected using high-salt buffer (50 mM HEPES, 0.5 mol/l NaCl, 1.5 mM $MgCl_2$, 1 mM EGTA, 10% glycerol, and 1% Triton-X-100) and 10 μl/ml Protease Inhibitor Cocktail (Sigma-Aldrich). Protein lysates were incubated for 5 min at 95°C before electrophoresis and then separated on 10% SDS-polyacrylamide gel electrophoresis 120 V for 2 to 3 hr. Proteins were transferred onto polyvinylidene difluoride membranes by wet electroblotting in a buffer containing 25 mM Tris, 192 mM glycine, and 20% methanol for 1.5 hours at 900 mA. Membranes were then blocked for 30 min with specific antibodies. Detection of specific proteins was performed using Chemiluminescence and then exposed to Kodak image station 4000 mm Pro (Carestream Health, Rochester, NY).

3.2.5 Cell proliferation assays

Cell counting: Pancreatic cancer cells were seeded in 12-well plates, and 72 hrs after siRNA transfection, cells were trypsinized and counted using a Coulter Z1 cell counter (Beckman Coulter, Fullerton, CA). MTT assay: Pancreatic cells were seed into a 96-well

plate and 72 hours after siRNA transfection, medium was removed, and MTT solution diluted in PBS was added to cell cultures. After 2 hours incubation, the medium was aspirated and washed with PBS. Dimethyl sulfoxide (DMSO) was added and incubate at 37° for 10 mins and absorbance was measured at 570 nM.

3.2.6 Apoptosis and cell cycle analysis assays

For cell cycle analysis, 48 hours after transfection, cells were stained with propidium iodide solution (50 μg/ml) and were analyzed by fluorescence-activated cell sorter (FACS). Apoptosis was detected using Alexa Fluor 488 Annexin V/Dead Cell apoptosis kit followed by FACS analysis according to the manufacturer's protocol.

3.2.7 Migration and invasion assays

Transwell migration/invasion: Pancreatic cancer cells were first transfected with siRNAs for 48 hours, then added to the upper chamber of a transwell chamber (with or without matrigel) in duplicate and allowed to migrate/invade into the lower chamber containing DMEM media with 20% FBS by incubating for 24 hr. Cells migrating/invading to the outer side of the upper chamber were fixed, stained and counted. Scratch assay: Cells were first seed in 6-well plates for 24 hours, and then a scratch through the central axis of the plate was gently made using a sterile pipette tip. Cells were transfected with siRNAs, and media was changed after 6 hr. Migration of the cells into the scratch was observed after 24, 48, and 72 hr. Ibidi assay: Pancreatic cancer cells were first transfected with siRNAs for 24 hours, then seeded in the silicone cell culture inserts

(Ibidi, Martinsried, Germany) which are attached to culture plates. Inserts were removed with tweezers and cells were rinsed with PBS and new medium was added into the cell after 24 hrs. The movement of cells to the middle gap was observed.

3.2.8 Microarray analysis

Total RNA was isolated using Zymo Quick RNA MiniPrep Kit (Zymo Research, Irvine, CA) according to the manufacturer's protocol. RNA was eluted with 35 μl of RNase-free water and stored at -80°C. The total RNA was quantified by using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technology). The total RNA samples with adequate RNA quality index (>7) were used for microarray analysis; 700 ng of total RNA was used for labeling and hybridization on Human HT-12 v4 expression beadchip (Illumina, Inc.) according to the manufacturer's protocols. After the beadchips were scanned with a BeadArray Reader (Illumina), the microarray data were normalized using the quantile normalization method in the Linear Models for Microarray Data (LIMMA) package in the R language (http://www.r-project.org). BRB-Array Tools were primarily used for statistical analysis of gene expression data, and the Student's *t* test was applied to identify the genes significantly different between 2 groups when compared. Function and pathway analysis of differentially regulated genes was determined using Ingenuity Pathway Analysis (IPA) database (Invitrogen, Carlsbad, CA).

3.2.9 RNA sequencing analysis

Total RNA was isolated using Zymo Quick RNA MiniPrep Kit (Zymo Research, Irvine,

CA) according to the manufacturer's protocol. RNA was eluted with 35 μl of RNase-free water and stored at -80°C. The total RNA was quantified by using a Nanodrop ND-1000 spectrophotometer (NanoDrop Technology). RNA sequencing was carried out in the genomics and bioinformatics service at Texas A&M University and the sequencing results were analyzed using sequencing-pipeline developed by the lab of Dr. Robert Chapkin at Texas A&M University. Functional and pathway analysis of differentially regulated genes was determined using Ingenuity Pathway Analysis (IPA) database (Invitrogen, Carlsbad, CA).

3.2.10 Animal pathology

A total of 17 mice were euthanized with carbon dioxide (CO_2). Of these, one was genotype p53L/L LSL-KrasG12D L/+ P48Cre +/-; two were genotype Malat1 -/- p53L/L LSL-KrasG12D L/+ P48Cre +/-; nine were genotype p53L/+ LSL-KrasG12D L/+ P48Cre+/-; three were genotype Malat1+/- p53L/+ LSL-KrasG12D L/+ P48Cre+/-; and two were genotype Malat1-/- p53L/+ LSL-KrasG12D L/+ P48Cre+/-. The following tissues were harvested from all these mice and fixed in 10% neutral buffered formalin: heart, lungs, diaphragm, pancreas, liver, kidney, spleen, stomach, small intestine, large intestine, testis/ovaries and brain. Fixed tissues were processed for routine histology and paraffin embedded. Histological sections were stained with hematoxylin and eosin, and examined by a board certified pathologist (ARH).

3.2.11 Production of Malat1 KO mice

Malat1 knockout mice were generated using a gene-trapping technique (537). Mice (strain C57BL/6) were cloned from an ES cell line (IST14461G11; Texas A&M Institute for Genomic Medicine, TIGM). The ES cell clone contained a retroviral insertion in the Malat1 gene identified from the TIGM gene trap database, and was microinjected into C57BL/6 host blastocysts to generate germline chimeras using standard procedures (538). The retroviral OmniBank Vector 74 contained a splice acceptor sequence (SA) followed by a 5' selectable marker neomycin resistance genes, for identification of successful gene trap events followed by a polyadenylation signal (pA). Insertion of the retroviral vector into the Malat1 gene led to the splicing of the endogenous upstream exons into this cassette to produce a fusion transcript and terminate expression of the RNA downstream. Chimeric males were bred to C57BL/6 females for germline transmission of the mutant Malat1 allele. Ablation of Malat1 expression in homozygous mice was confirmed by RT-PCR.

3.2.12 Statistical analysis

Statistical significance of differences between the treatment groups was determined by an analysis of variance and/or Student's *t* test, and levels of probability were noted. At least 3 repeated experiments were determined for each data points and results are expressed as means ± SD.

3.3 Results

3.3.1 Expression and pro-oncogenic functions of MALAT1 in pancreatic cancer cells

Previous studies showed that MALAT1 is overexpressed in pancreatic cancer cell lines and tumors compared to non-transformed pancreatic cells/tissue (266,267) and results of knockdown studies by RNA interference confirmed the pro-oncogenic activity of MALAT1. We also observed high expression of MALAT1 in pancreatic cancer cells (Appendix A-1); knockdown of MALAT1 (siMALAT1) in Panc1 and MiaPaCa2 cells decreased cell proliferation and induced G2/M arrest (Figures 17A and 17B) and induced apoptosis as indicated by increased Annexin V staining and expression of cleaved PARP (Figure 17C). Moreover, knockdown of MALAT1 also decreased cell migration and invasion in scratch and Boyden chamber assays (Figures 17D and 17E), and decreased migration was also observed in ibidi and Boyden chamber assays (Appendix A-1). These results confirm that MALAT1 plays a role in pancreatic cancer cell proliferation, survival and migration/invasion.

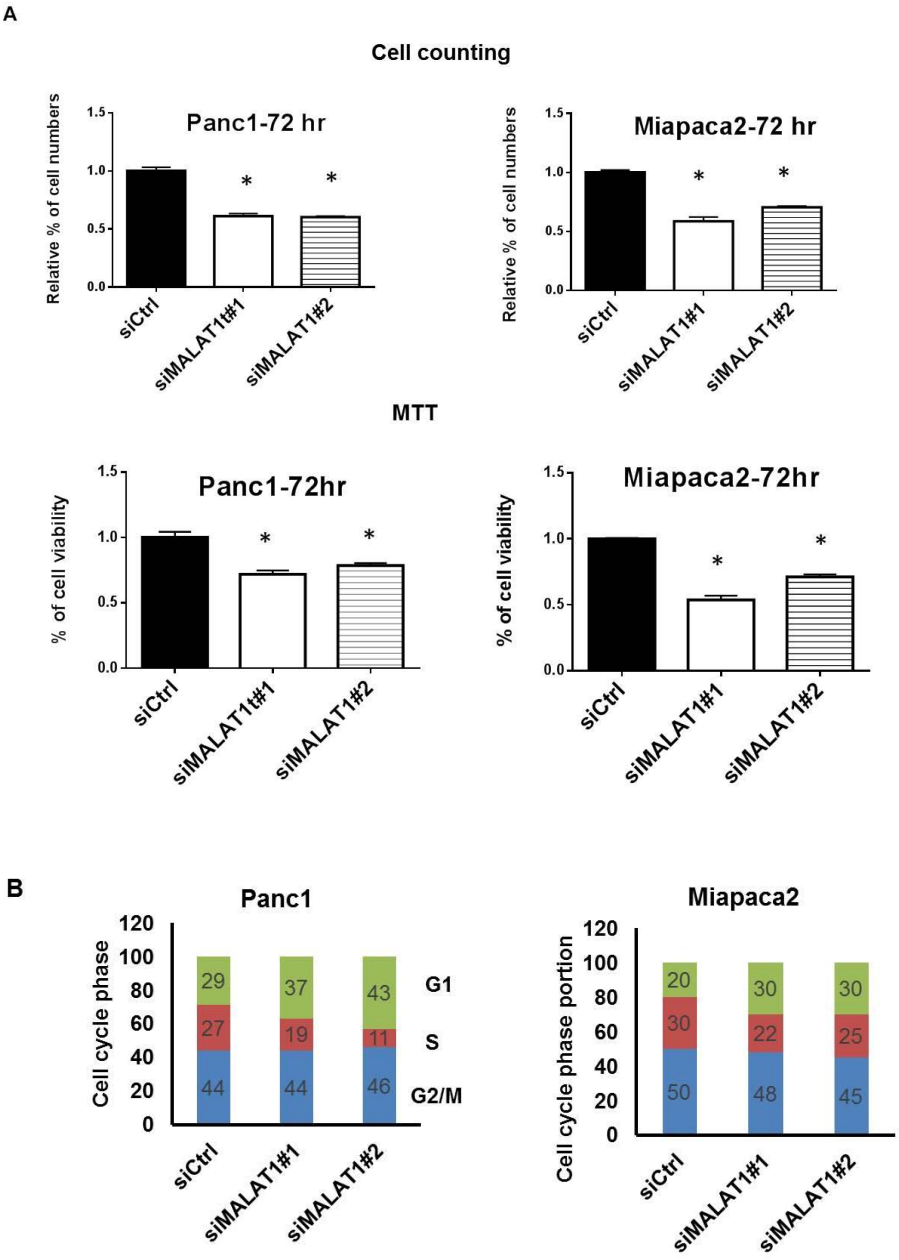


Figure 17. Effects of MALAT1 in pancreatic cell proliferation, cell cycle, apoptosis, migration and invasion. (A) MALAT1 knockdown by RNAi in Panc1, MiaPaCa2 inhibited cell growth. (B) The effect of siMALAT1 (knockdown) on cell cycle progression in Panc1 and MiaPaCa2 cells was determined by FACS analysis. (C) The apoptotic cells were quantified using FACS analysis and induction of PARP cleavage was determined by western blot analysis. MALAT1 knockdown reduced cell migration (D) and cell invasion (E) as determined by scratch assay and Boyden chamber assay, respectively. Significant ($p<0.05$) changes are indicated (*).

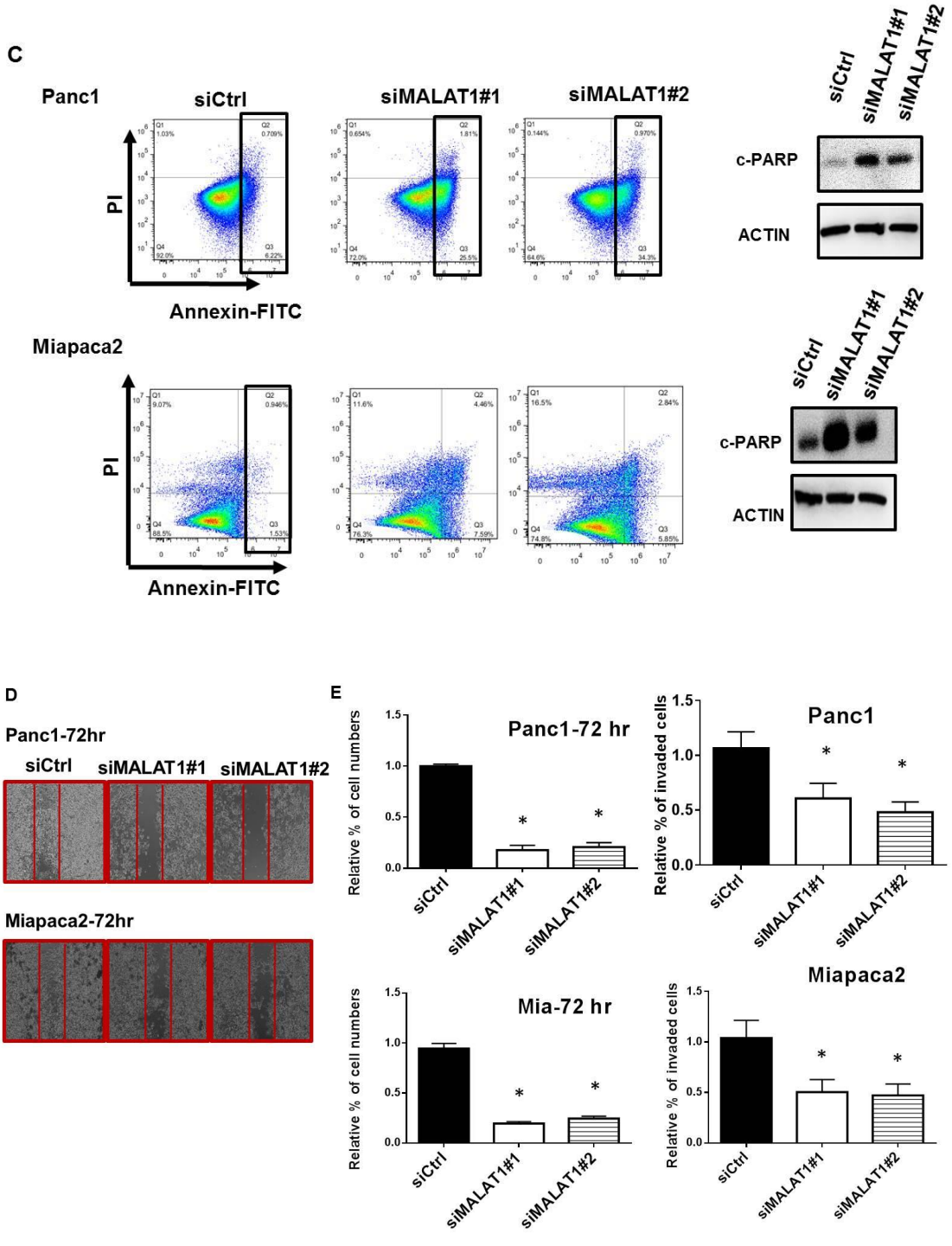
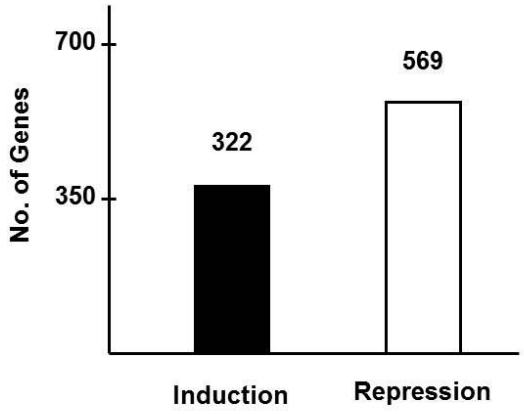


Figure 17. Continued.

3.3.2 Analysis of MALAT1 regulated gene expression in Panc1 cells

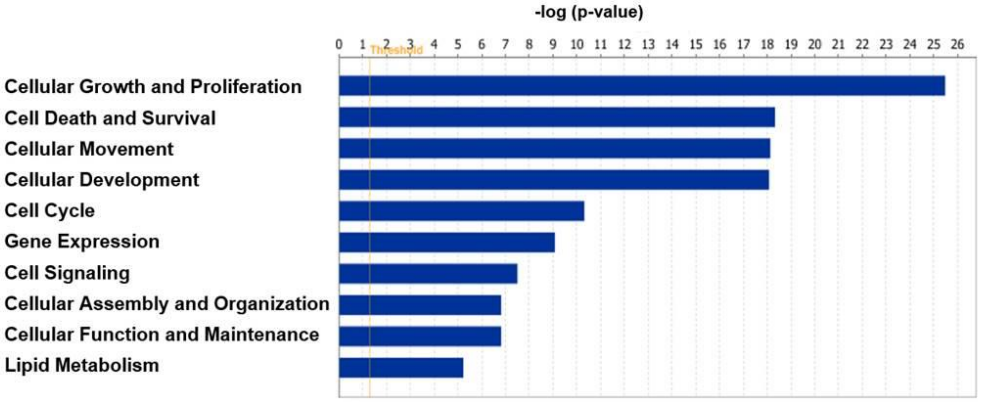
Knockdown of MALAT1 by RNAi in Panc1 cells resulted in the induction of 352 and repression of 611 genes (Figure 18A). Further analysis of these genes by Ingenuity Pathway Analysis demonstrated that MALAT1-regulated genes are involved in multiple functions (Figure 18B). Not surprising, many of the MALAT1-regulated genes were associated with cell growth and proliferation, cell death and survival, and cell movement (motility, migration and invasion) and these correlated with functional responses observed after MALAT1 knockdown (Figure 17). Figure 18C summarizes analysis of the overall changes in gene expression after MALAT1 knockdown using causal Ingenuity Pathway Analysis which is a quantitative approach that integrates both changes in gene expression and pathways to predict biologic function. The low *p* values and activation scores (> 2.0 or < -2.0) obtained from causal IPA strongly predicted that loss of MALAT1 was associated with decreased cell movement and proliferation and increased cell death, and this analysis correlated with results of functional studies (Figure 17). Previous studies in this laboratory showed that like MALAT1, the lncRNAs HOTAIR and HOTTIP also regulate cell proliferation, survival and movement in Panc1 pancreatic cancer cells (261,272), and Figure 19A shows the overlap of total changes in gene expression observed for MALAT1, HOTTAIR and HOTTIP. Based on the total number of genes regulated by MALAT1 (963), HOTAIR (1628) and HOTTIP (1125), the common genes coregulated by MALAT1/HOTTIP and MALAT1/HOTAIR were only 8.5 and 16.1%, respectively, suggesting that the common pro-oncogenic responses regulated by the three lncRNAs are primarily due to different sets of genes.

A



B

Differentially expressed genes are involved in cell proliferation, death and movement



C

Diseases or Functions Annotation	p-Value	Predicted Activation State	Activation z-score
cell movement of tumor cell lines	7.35E-19	Decreased	-2.636
proliferation of tumor cell lines	8.29E-19	Decreased	-2.632
cell death of tumor cells	7.41E-06	Increased	2.393

Figure 18. Gene regulation by MALAT1. (A) Panc1 cells were transfected with siMALAT1 or siCtrl and gene expression was analyzed using Human HT-12 v4 expression beadchip (Illumina, Inc.) array. (B) The effects of siMALAT1 on different function categories and the predicted activation state of cell proliferation, death and movement after MALAT1 knockdown were determined by IPA.

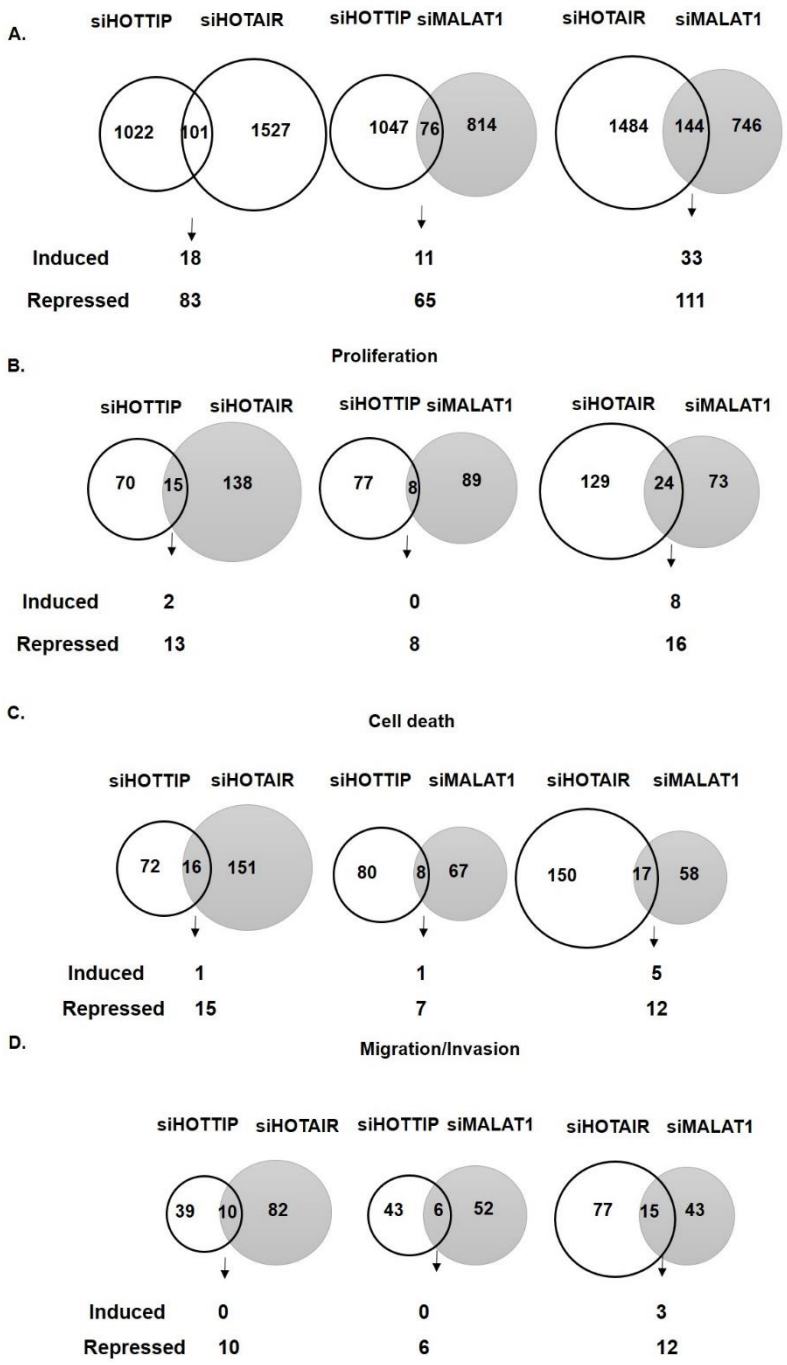


Figure 19. Comparison of regulated genes by HOTAIR, HOTTIP and MALAT1 in Panc1 cells. Panc1 cells were transfected with siHOTTIP, siHOTAIR or siMALAT1, and changes in gene expression were determined using Human HT-12 V4 expression bead chip arrays. The overlap of total genes (A), proliferation inhibition (B), cell death (C) and inhibition of migration/invasion (D) genes coregulated by HOTTIP/HOTAIR, HOTTIP/MALAT1 and HOTAIR/MALAT1 was determined by IPA.

Venn diagram analysis of the overlap in gene expression associated with cell proliferation, cell death and migration/invasion by MALAT1, HOTAIR and HOTTIP (Figures 19B-D) shows that there was minimal gene overlap (<10%), demonstrating that the lncRNA-regulation of these responses was primarily due to different sets of genes.

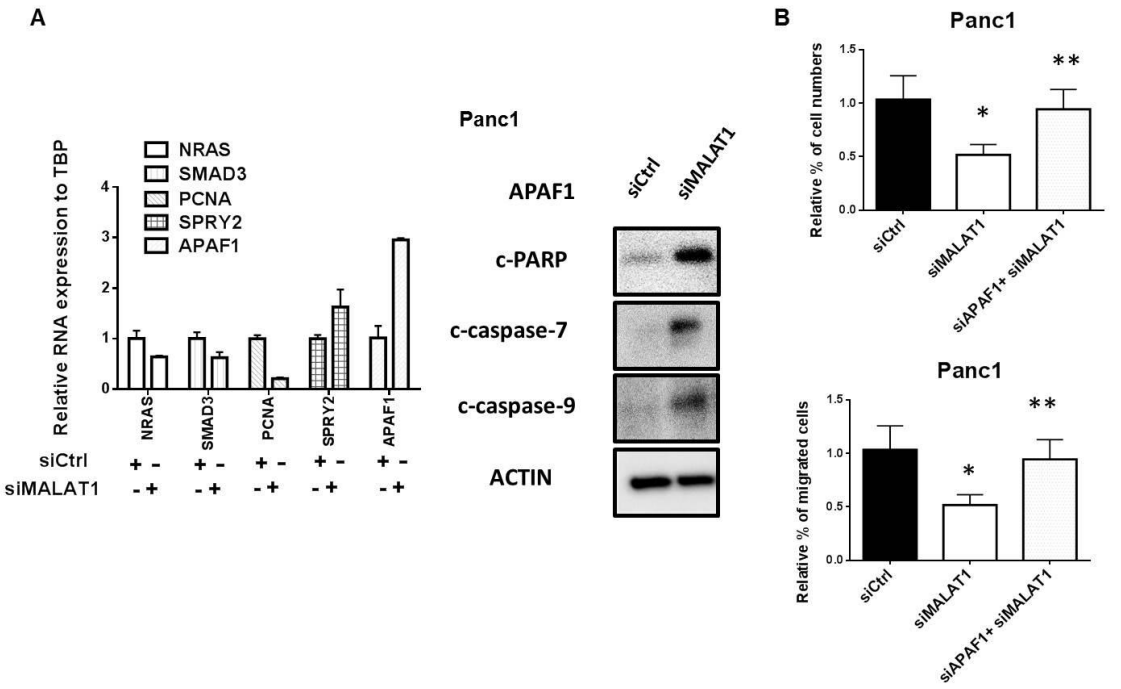


Figure 20. Interaction between MALAT1 and EZH2 on gene expression in Panc1 cells. (A) Cells were transfected with siMALAT1 or siCtrl. Genes regulated by MALAT1 were analyzed by real time PCR and the expression of apoptosis associated proteins was analyzed by western blot. (B) Panc1 cells were transfected with siMALAT1, siMALAT1+siAPAF1 or siCtrl, and effects on cell growth and migration were determined. (C) Panc1 cells were transfected with siMALAT1, siEZH2, siLSD1, siMLL1 or siCtrl, analyzed by microarrays for changes in gene expression. (D) Cells were transfected with siMALAT1, siEZH2, or siCtrl, and the mRNA and protein expression of NDRG1 were determined by real time PCR and western blot. (E) Panc1 cells were transfected with siMALAT1, siMALAT1+siNDRG1 or siCtrl and effects on cell growth and migration were determined. Significant ($p<0.05$) changes are indicated (*) or (**).

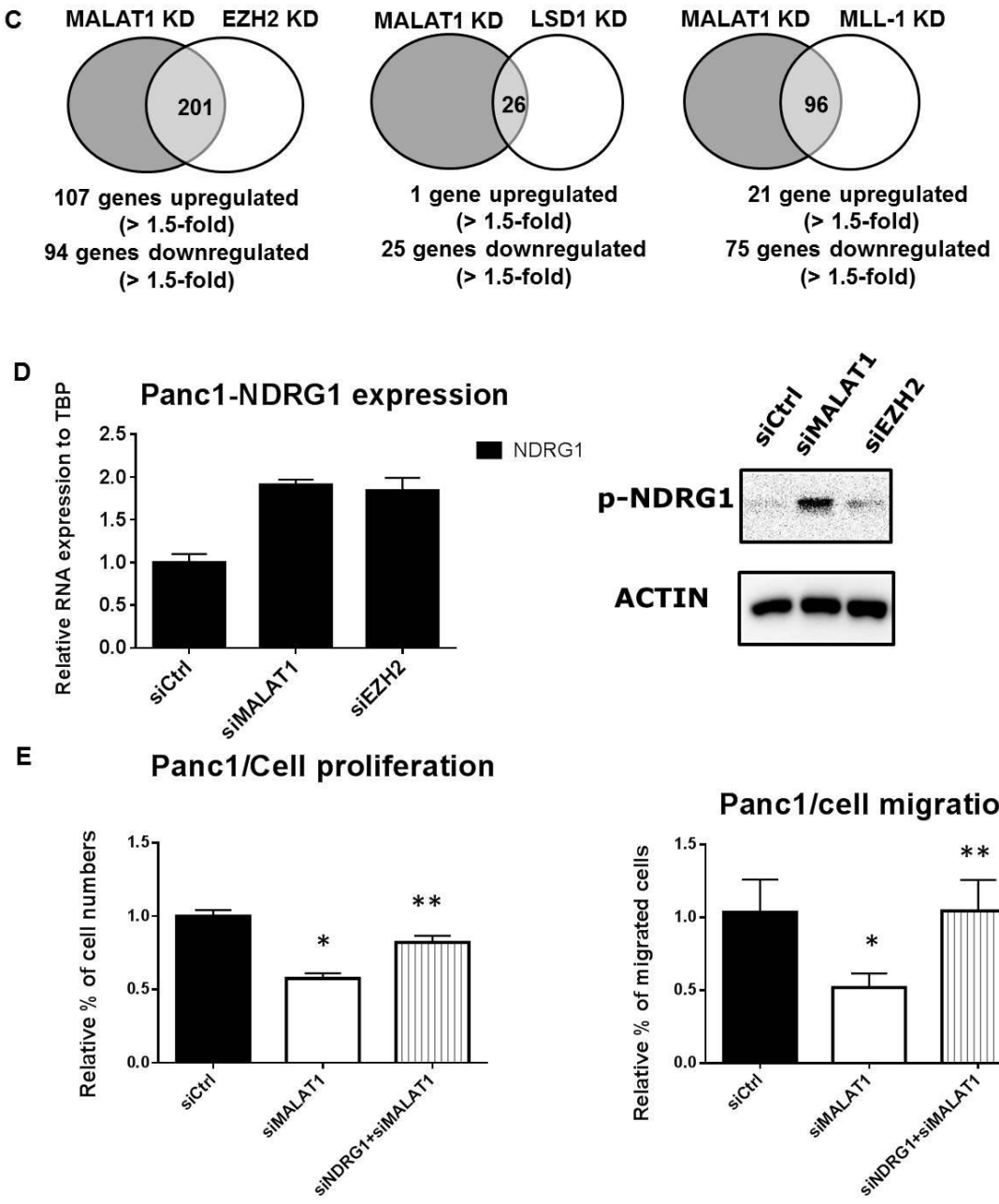


Figure 20. Continued.

Figure 20A illustrates induction and repression of representative genes observed in the arrays after MALAT1 knockdown in Panc1 cells as determined by real time PCR. One

of these genes, apoptotic protease activating factor 1 (APAF1), is a key protein component of the apoptosome, and knockdown of MALAT1 induced expression of APAF1 protein and activation (cleavage) of caspases 7 and 9 and PARP in Panc1 cells (Figure 20A). This suggests that the pro-apoptotic activity observed after MALAT1 knockdown was due, in part, to induction of APAF1. Figure 20B shows that in Panc1 cells transfected with siMALAT1, there was a decrease in cell migration and number which could be partially reversed by knocking down APAF-1 (induced by siMALAT1), suggesting that APAF-1 may also play a role in Panc1 cell growth and migration.

We also used RNAi coupled with analysis of array data to investigate the overlap of MALAT1 with EZH2-, LSD1- and MLL-1-regulated genes to identify genes coregulated by PRC2, REST/coREST and MLL-1 chromatin modifying complexes, respectively. There was overlap of genes coregulated by MALAT1 and EZH2, LSD1 and MLL-1 with the highest gene overlap observed with EZH2 (a component of the PRC2 complex) in which 107 common genes were upregulated after Panc1 cells were transfected with siMALAT1 and siEZH2 (Figure 20C). One of these coregulated genes was N-myc downregulated gene-1 (NDRG1) which exhibits tumor suppressor-like activity in pancreatic cancer (539-541). Knockdown of MALAT1 or EZH2 by RNAi in Panc1 cells induced expression of NDRG1 mRNA as determined by real time PCR (Figure 20D), thus confirming the array results and these same treatments also increased NDRG1 protein. The functional role of MALAT1-dependent suppression of NDRG1 was further investigated in Panc1 cells transfected with siMALAT1 which decreased cell

proliferation and migration (Figure 20E) and this was partially reversed by cotransfection with siNDRG1. Thus, NDRG1 suppression by MALAT1 also plays a role in the pro-oncogenic functions of this lncRNA associated with cell proliferation and migration.

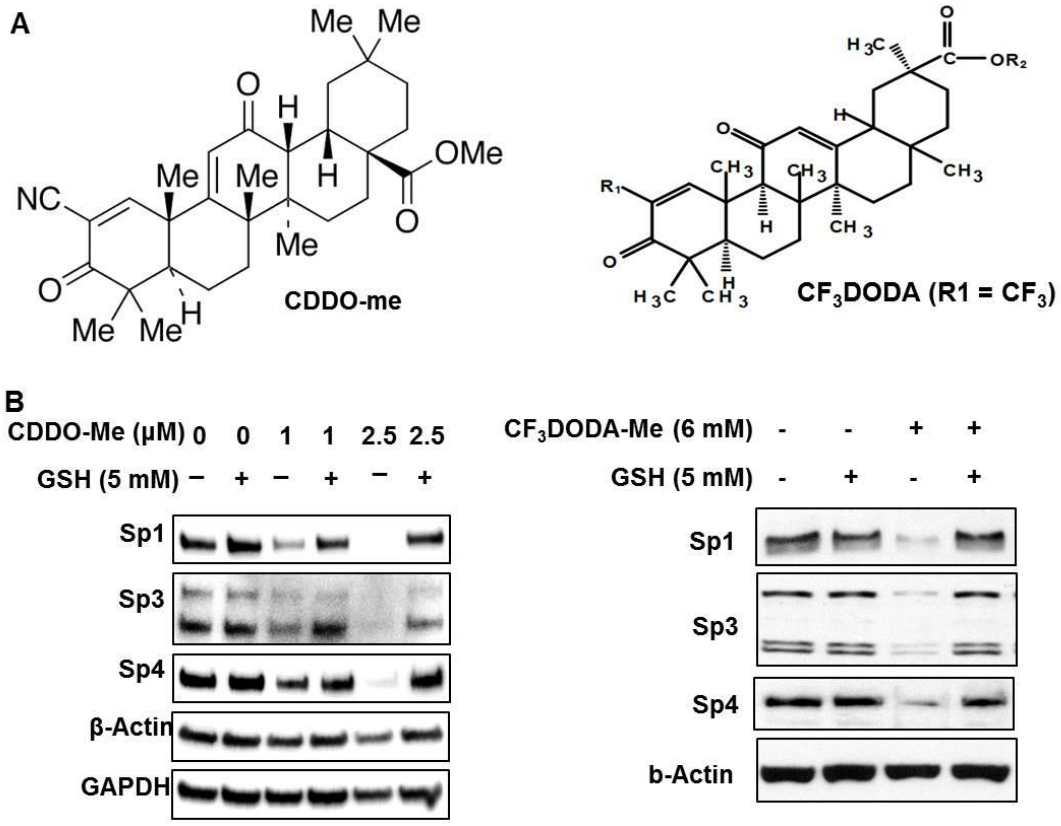


Figure 21. Regulation of MALAT1 by Specific proteins and ROS inducers. (A) The structure of CDDO-Me and CF_3DODA-Me. Panc1 cells were treated with different concentrations of CDDO-Me or in combination of GSH, CF_3DODA-Me or in combination with GSH, and the changes of different proteins (B) and MALAT1 expression (C) were determined by western blot and real time PCR, respectively. (D) Panc1 cells were transfected with siSp1/3/4 or siCtrl, and the MALAT1 RNA expression was determined by real time PCR. Significant ($p<0.05$) changes are indicated (*) or (**).

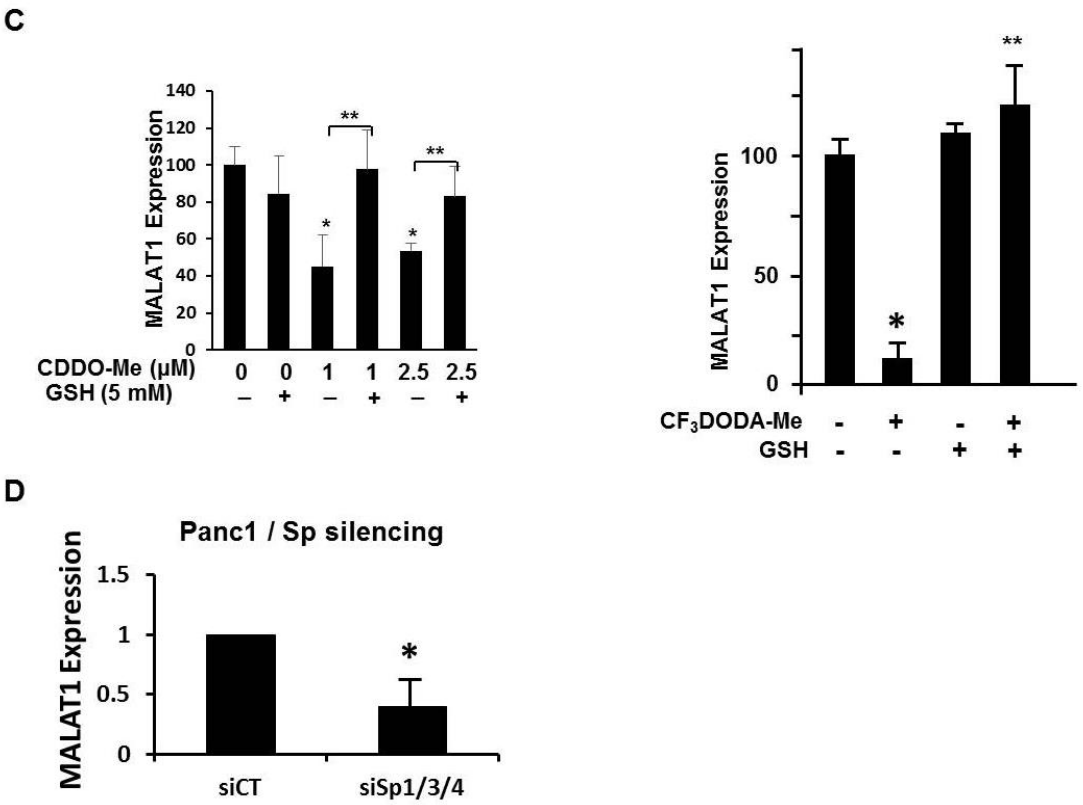


Figure 21. Continued.

The transcription factors Sp1, Sp3 and Sp4 are overexpressed in pancreatic cancer cells and like MALAT1, knockdown of Sp1, Sp3 and Sp4 or Sp1/Sp3/Sp4 (combined) results in decreased cell growth, induces apoptosis, and decreases migration (542-546). A recent report showed that MALAT1 is an Sp1-regulated gene (518) and we therefore determined if ROS-inducing anticancer agents that downregulate Sp1, Sp3 and Sp4 in pancreatic and other cancer cell lines also decrease MALAT1 expression. Treatment of Panc1 cells with two triterpenoid ROS inducers, methyl-2-cyano-3,12-dioxooleana-1,9-dien-28-oate (CDDO-Me, bardoxolone-methyl) and methyl-2-trifluoromethyl-3,11-

dioxo-18-olean-1,12-dien-30-oate (CF3DODA-Me) (Figure 21A), decreased expression of Sp1, Sp3 and Sp4, and these effects were reversed in Panc1 cells cotreated with the antioxidant glutathione (GSH) (Figure 21B). These results are consistent with previous reports on ROS inducers that downregulate Sp proteins in pancreatic and other cancer cell lines. Treatment of PANC1 cells with the two triterpenoids alone also decreased MALAT1 expression and cotreatment with GSH also attenuated this response (Figure 21C). Moreover, silencing Sp1, Sp3 and Sp4 combined (siSp1/3/4) also decreased MALAT1 expression (Figure 21D), confirming that MALAT1 is a Sp-regulated gene.

MALAT1-/- mice have previously been reported (525-527) and were also generated in our laboratory using a gene trapping technique (537), and results in Figure 22A show that real time PCR analysis did not detect MALAT1 expression in multiple tissues of MALAT1-/- mice. Using founder mice provided by the DePinho laboratory (83,125), we generated p53L/L; LSLKrasG12DL/+; p48Cre+/- (p53L/L/KrasG12D) and p53L/+; LSL-KrasG12DL/+; p48Cre+/- (p53L/+/KrasG12D) mice which are p53 homo- and heterozygous, respectively. We observed high expression of Sp1, Sp3 and Sp4 in pancreatic tumors from these mice (Figure 22B). These transgenic mice rapidly develop tumors and typically present with adverse symptoms prior to tumor-induced lethality. Results in Figures 22C and 22D show that p53 heterozygous mice live longer than the corresponding p53-/- mice; moreover in these two mouse models, the loss of one or two MALAT1 alleles does not significantly affect the time to death, although the results suggest that the loss of MALAT1 in the p53-/+ mice results in an increased (not

significant) lifespan. RNAseq was used to determine differences in pancreatic cancer gene expression in Ras overexpressing/p53+/- mice ± MALAT1 expression and we observed that >1000 genes were differentially expressed. A comparison of the Panc1 gene analysis (Figure 18) vs. the in vivo indicates that 50 genes were commonly altered after loss of MALAT1 (Appendix B-4) and this included NDRG1. Most experimental mice presented with a high grade ductular adenocarcinoma (DAC) (Figure 22E), independent of their phenotype which was composed of infiltrative and unencapsulated neoplasms arranged in variably sized ducts separated by abundant fibrovascular stroma. Neoplastic cells were cuboidal, with moderate amounts of eosinophilic cytoplasm and round nuclei. Anisocytosis and aniskaryosis were marked with high mitotic activity. There were multifocal areas of necrosis within the neoplasm and in some cases the neoplasm infiltrated adjacent lymph nodes, the muscularis of the stomach and small intestine and rarely the diaphragm. Implantation and metastasis to the liver and lungs were seen in 4 and one mouse in the p53L/+ LSL-KrasG12D L/+ P48Cre +/- genotype, respectively. One mouse had hepatic and another had renal implantation, and two mice presented pulmonary metastasis in the Malat1+/- p53L/+ LSL-KrasG12D L/+ P48Cre +/- genotype group. In 4 mice in 3 different groups, the DAC had transformed into a carcinosarcoma, characterized by malignancy also involving the epithelial as well as the mesenchymal tissue. In two mice within the p53L/+ LSL-KrasG12D L/+ P48Cre +/- genotype an early and well differentiated DAC was observed; however, one of these mice also exhibited a primitive neoplasm within the lateral aspect of the head, and compressing the brain. Besides the neoplasm, in 6 mice there were multifocal areas of

hepatic necrosis, with 2 mice having grossly apparent icterus and hemoglobin casts within the kidneys, indicating hemolysis. This histological finding was not seen in any of Malat1 +/- (or -/-) p53L/+ LSL-KrasG12D L/+ P48Cre +/- genotype groups. Thus, the loss of MALAT1 has minimal effect on pancreatic tumors in transgenic mice where there is activation of Kras and a loss of p53 or activation of Ras in the pancreas, and this contrasts a recent report showing that loss of MALAT-1 in BxPC3 pancreatic cancer cells resulted in decreased tumor growth and metastasis in an orthotopic mouse model (262).

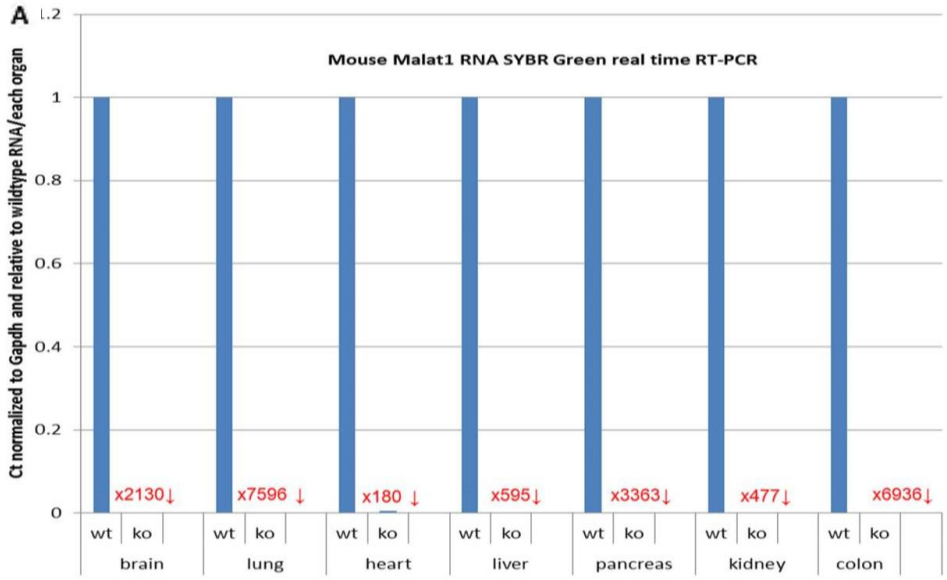


Figure 22. MALAT1 function in the transgenic mouse model of pancreatic cancer. (A) Fold-change in Malat1 gene expression in knockout mice as compared to the wild type. (B) Sp1, Sp3, Sp4 and c-Myc expression in homozygous floxed p53/KrasGD12 mice. (C) Survival of homozygous floxed p53L/L: KrasGD12: p48Cre+/- mice expression MALAT1 (+) or with loss of MALAT1 (-/+). (D) Survival of heterozygous floxed p53L/+: KrasGD12: p48Cre+/- mice expression Malat1 (+) or with loss of Malat1 (-/+). (E) Histology analysis of tumor samples from different strains of mice. Significant ($p<0.05$) changes are indicated (*) or (**).

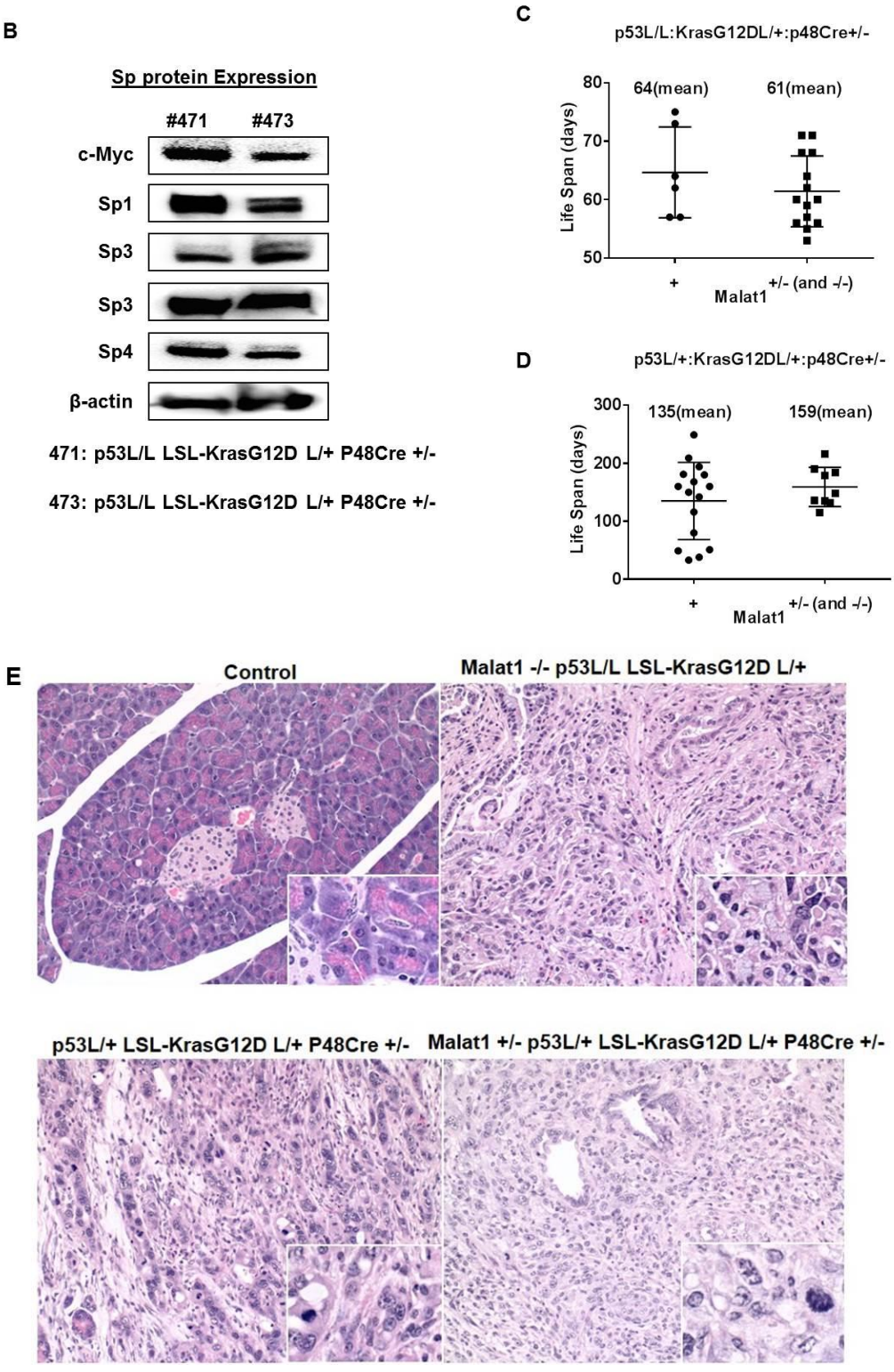


Figure 22. Continued.

3.4 Discussion

MALAT1 expression in many tumors is a negative prognostic factor for patients, and results obtained in this study and previous reports confirm that MALAT1 is a pro-oncogenic factor for pancreatic cancer (Figure 17). We have previously reported that the lncRNAs HOTAIR and HOTTIP regulate expression of genes associated with pancreatic cancer cell proliferation, survival and migration (261,272). A comparison of genes regulated by MALAT1 and HOTTIP and HOTAIR in Panc1 cells indicates <6% overlap in coregulated genes, and similar results were observed after comparing specific gene sets associated with proliferation, survival and migration (Figures 19B-19D). APAF1 is an example of a gene regulated (suppressed) by MALAT1 and not HOTAIR or HOTTIP, and APAF1 induction after MALAT1 knockdown plays an important role in activating apoptosis and other pathways in Panc1 cells (Figure 20B). Interestingly, although MALAT1, HOTAIR and HOTTIP exhibit similar functional pro-oncogenic activities in pancreatic cancer cells (Figure 17), the targeted (RNAi) loss of any one of these lncRNAs cannot be rescued by the other two lncRNAs and this supports the array data showing their regulation of different sets of genes, and we are currently investigating the functions of key genes differentially regulated by MALAT1, HOTTIP and HOTAIR.

Several genetic mouse models of pancreatic cancer have been developed (547) and in this study, we used mice with p53 mutations and overexpression of activated Kras in the pancreas since they rapidly develop pancreatic tumors that resemble human PDAC (83).

However, p53 heterozygotes survive longer than homozygous mice and the loss of MALAT1 only slightly extends the survival of the latter (not statistically significant) but not the former mice (Figures 22C and 22D). Thus, the loss of MALAT1 in the transgenic mice driven by Kras expression and p53 deletion (-/- or +/-) had minimal impact on this aggressive tumor model, whereas the loss of MALAT1 in an orthotopic and xenograft mouse models of pancreatic cancer decreased tumor growth and invasion (262,523). Since Ras activation is observed in most pancreatic tumors, it is unlikely that drugs targeting MALAT1 alone would be effective. Liby and coworkers (548) reported that treatment of a similar transgenic mouse model [LSL-KrasG12D/+; LSL-Trp53R127H/+; Pdx-1-Cre (KPC)] with CDDO-Me (15 mg/kg body weight) increased survival from 20.5 ± 0.9 (control) to 24.2 ± 2.7 weeks. Moreover, they also reported that CDDO-Me induced ROS in pancreatic cancer cells (548). This was consistent with studies in this laboratory showing that CDDO-Me induced ROS and ROS-dependent downregulation of Sp1, Sp3 and Sp4 and pro-oncogenic Sp-regulated genes (544) and has been observed with other ROS-inducing anticancer agents in pancreatic cancer cells (543,545). Results illustrated in Figure 21 show that the ROS-inducing triterpenoids CDDO-Me and CF3DODA-Me induce ROS-dependent downregulation of Sp1, Sp3 and Sp4 in Panc1 cells (Figure 21B) and also MALAT1 (Figure 21C), which is a Sp-regulated gene (Figure 21D). High levels of Sp1, Sp3 and Sp4 protein and MALAT1 are expressed in transgenic mice (Figure 22B); however, the effectiveness of CDDO-Me in extending the life of these mice (548) is due not only to decreased MALAT1 expression but also several pro-oncogenic Sp-regulated genes including bcl-2, survivin, epidermal growth

factor receptor, other receptor tyrosine kinases and p65 (NFκB) (542-546). Current studies are focused on using genetic mouse models for pancreatic cancer in which both MALAT1 and Sp transcription factors can be deleted (tissue-specific) or specifically targeted to confirm this important role of these genes in tumor development and growth and thereby demonstrate the utility of ROS-inducing anticancer agents such as CF3DODA-Me and CDDO-Me for treating this disease.

4. ARYL HYDROCARBON RECEPTOR ACTIVITY OF TRYPTOPHAN METABOLITES IN YOUND ADULT MOUSE COLONOCYTES[*]

4.1 Introduction

The aryl hydrocarbon receptor (AhR) is a ligand-activated transcription factor that was initially identified as the intracellular protein that bound the environmental toxicant 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD), related halogenated aromatics, and polynuclear aromatic hydrocarbons (303,549,550). The role of the AhR in mediating the biochemical and toxic responses induced by TCDD and related compounds has been confirmed in AhR knockout ($AhR^{-/-}$) mice which are resistant to the effects of TCDD (381,400,550,551). Ligand-dependent activation of the AhR results in formation of a nuclear complex with the AhR nuclear translocator (Arnt) protein which binds cis-xenobiotic response elements (XREs) in the *Cyp1a1* and other target gene promoters (303,552,553). However, several non-classical pathways have been discovered and these include AhR interactions with other nuclear partners, binding to non-consensus cis-promoter elements and also responses that are associated with the extranuclear AhR (554-563).

---

[*] Reprinted with permission of the American Society for Pharmacology and Experimental Therapeutics from "Aryl hydrocarbon receptor activity of tryptophan metabolites in young adult mouse colonocytes" by Cheng Y, Jin UH, Allred CD, Jayaraman A, Chapkin RS, *et al*. Drug Metab Dispos 2015;43(10):1536-43. All rights reserved.

Since the initial discovery that the AhR binds toxic polychlorinated and polynuclear aromatic hydrocarbons, it has subsequently been shown that the AhR also binds structurally and functionally diverse ligands including health-promoting phytochemicals such indole-3-carbinol, flavonoids and extracts from fruits and vegetables, and a growing list of pharmaceuticals including omeprazole and other benzimidazoles (339,371,564-568). In addition, structurally diverse "endogenous" biochemicals have been identified as AhR ligands, and there is evidence that the tryptophan photoproduct 6-formylindolo[3,2-b]carbazole (FICZ) and kynurenine may function as endogenous ligands for the AhR (354,361,364,371,568). The development of $AhR^{-/-}$ and tissue specific AhR knockout mice has been instrumental in showing that this receptor plays an essential role in various tissues and is a critical regulator of inflammation, autoimmune and immune responses and is a potential drug target for treating multiple diseases including cancer (398,569-572). For example, there is extensive evidence that the AhR and its agonists including AhR-active cruciferous vegetables play a protective in mouse models of intestinal cancer and inflammation (434,455,456,466,474,573-576).

Several studies have reported that the gut microbiota produces metabolites including AhR-active compounds that could potentially modulate AhR-mediated intestinal resiliency and responses to inflammatory stimuli (404,577-580). Research in our laboratories has previously investigated the AhR activities of the tryptophan metabolites indole, indole-3-acetate, tryptamine and 3-indoxyl sulfate, using CYP1A1 induction as a prototypical AhR-dependent response in human CaCo2 colon cancer cells (581). In this

report, we determined the AhR activity of tryptophan metabolites in a non-transformed young adult mouse coloncyte (YAMC) cell line (582) and there were significant differences between YAMC vs CaCo2 cells with respect to the gene-specific AhR agonist and antagonist activities of tryptophan metabolites.

4.2 Materials and methods

4.2.1 Cell Lines, antibodies, and reagents

The YAMC cell line was initially generated from the Immorto mouse (583) and has been previously used in our studies (584-586). Cells were maintained in RPMI 1640 medium with 5% fetal bovine serum, 5 units/ml mouse interferon-γ (IF005) (EMD Millipore, Massachusetts), 1% ITS "-" minus (Insulin, Transferrin, Selenium) (41-400-045) (Life Technologies, Grand Island, NY) at 33°C (permissive conditions). In preparation for experiments, cells were transferred to 37°C (nonpermissive conditions). AhR antibody (BML-SA210) was purchased by Enzo (Enzo Life Sciences Inc., Farmingdale, NY). The MTT assay for metabolic activity of the tryptophan metabolites was determined using the MTT (3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyltetrazolium bromide) assay essentially as described (581) (Appendix A-2). β-ACTIN (A1978) was purchased from Sigma-Aldrich (St. Louis, MO), and Cyp1a1 antibody was kindly provided by Dr. Paul Thomas (Rutgers University). Indole (>99%), indole-3-acetate (98%), indole-3-aldehyde (97%), and tryptamine (99%) were purchased from Sigma-Aldrich and TCDD (99%) was synthesized in our laboratory. CH223191 (Cat.No.3858) was purchased from Tocris

Bioscience (Bristol, United Kingdom). GelRed Nucleic Acid Stain (RGB-4103) was purchased from Phenix (Phenix Research Products, Candler, NC).

4.2.2 Chromatin immunoprecipitation assay

The chromatin immunoprecipitation (ChIP) assay was performed using the ChIP-IT Express Magnetic Chromatin Immunoprecipitation kit (Active Motif, Carlsbad, CA) according to the manufacturer's protocol. YAMC cells (1.2×10^7 cells) were treated with TCDD and/or compounds for 2 or 24 hr. The cells were then fixed with 1% formaldehyde, and the cross-linking reaction was stopped by addition of 0.125 M glycine. After washing with phosphate-buffered saline (PBS), cells were scraped and pelleted. The collected cells were hypotonically lysed, and nuclei were collected and then sonicated to the desired chromatin length (200-1500 bp). The sonicated chromatin was immunoprecipitated with normal rabbit IgG or AhR antibodies and protein A-conjugated magnetic beads at 4°C for overnight. After the magnetic beads were extensively washed, protein-DNA crosslinks were reversed and eluted. DNA was prepared by proteinase K digestion followed by polymerase chain reaction (PCR) amplification. The *Cyp1a1* primers were 5'-AGG CTC TTC TCA CGC AAC TC-3' (sense) and 5'-CGG GTG CAG AGC TAT CTA AGT-3' (antisense); we then amplified a 207-bp region of mouse *Cyp1a1* promoter, which contained the AhR-binding sequences. The PCR products were analyzed on a 2% agarose gel in the presence of GelRed Nucleic Acid Stain.

4.2.3 Quantitative real-time PCR

Total RNA was isolated using Zymo Quick RNA MiniPrep Kit (Zymo Research, Irvine, CA) according to the manufacturer's protocol. RNA was eluted with 35 μl of RNase-free water and stored at -80°C. Real-time (RT)-PCR was carried out using iTaq Universal SYBR Green One-step Kit (Bio-Rad, Hercules, CA). The following primers were used (Table 9).

Table 9. Primers used in quantitative real-time PCR.

Name	Forward Primer	Reverse Primer
TBP	GAACAATCCAGACTAGCAGCA	GGGAACTTCACATCACAGCTC
Cyp1a1	CTGAAGTGGTTCTGAGCGG	TCCACTCCATCTTCCGACTT
Cyp1b1	GGATATCAGCCACGACGAAT	ATTATCTGGGCAAAGCAACG
TiParp	GCCAGACTGTGTAGTACAGCC	GGGTTCCAGTTCCCAATCTTTT
Ahrr	ACATACGCCGGTAGGAAGAGA	GGTCCAGCTCTGTATTGAGGC

4.2.4 Western blot analysis

Cells (1×10^5) were plated in six-well plates in RPMI media containing 2.5% FBS for 16 hr and then treated with different concentrations of the compounds for 24 hr. Cells were collected using high-salt buffer (50 mM HEPES, 0.5 mol/l NaCl, 1.5 mM $MgCl_2$, 1 mM EGTA, 10% glycerol, and 1% Triton-X-100) and 10 μl/ml Protease Inhibitor Cocktail (Sigma-Aldrich). Protein lysates were incubated for 5 min at 95°C before electrophoresis and then separated on 10% SDS-polyacrylamide gel electrophoresis 120 V for 2 to 3 hr. Proteins were transferred onto polyvinylidene difluoride membranes by wet electroblotting in a buffer containing 25 mM Tris, 192 mM glycine, and 20% methanol

for 1.5 hr at 180 mA. Membranes were then blocked for 30 min with specific antibodies. Detection of specific proteins was performed using Chemiluminescence and then exposed to Kodak image station 4000 mm Pro (Carestream Health, Rochester, NY)

4.2.5 Statistical analysis

Statistical significance of differences between the treatment groups was determined by an analysis of variance and/or Student's t test, and levels of probability were noted. At least 3 repeated experiments were determined for each data point, and results are expressed as means ± SD.

4.3 Results

YAMC cells were treated with different concentrations of tryptamine (10-100 μM), indole (50-500 μM), indole-3-acetate (50-500 μM) and indole-3-aldehyde (50-500 μM) and induction of *Cyp1a1* mRNA was determined (Figure 23A). Tryptamine and indole-3-acetate significantly induced *Cyp1a1* mRNA levels (>10 fold) at concentrations of 50 and 500 μM respectively whereas indole and indole-3-aldehyde were inactive. In contrast 0.01-10 nM TCDD, the most potent AhR agonist induced a >600-fold increase in *Cyp1a1* mRNA levels with maximal induction by 10 nM TCDD (Figure 23B) as previously observed in CaCo2 cells (581). Induction of *Cyp1a1* mRNA by the tryptophan metabolites (Figure 23C) and TCDD was inhibited after cotreatment with the AhR antagonist CH223191 (CH) (Figure 23D). In the inhibition experiment we observed some induction of *Cyp1a1* by indole and indole-3-aldehyde alone (Figure 23C) and over

several experiments low level induction responses by these compounds were variable (0-4 fold) but <1% of that observed for TCDD.

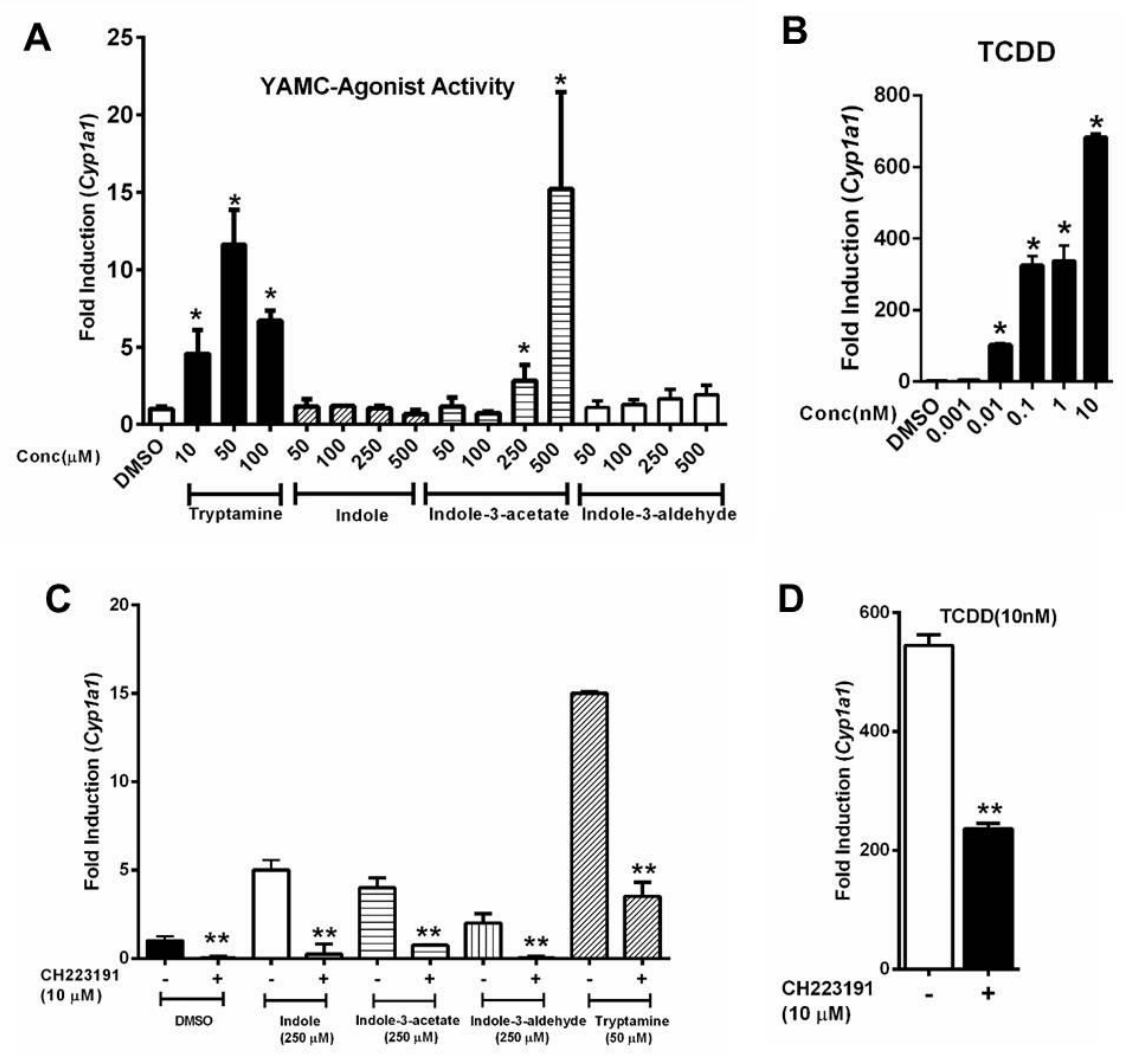


Figure 23. Tryptophan metabolites and TCDD as inducers of *Cyp1a1* in YAMC cells. YAMC cells were treated for 24 hr with tryptophan metabolites (A), TCDD (B), tryptophan metabolites plus CH (C), and TCDD plus CH (D). Expression of *Cyp1a1* mRNA was determined by real time PCR. Results are expressed as means ± SE for three replicate determinations and significant ($p<0.05$) induction (*) (A and B) or inhibition by CH (**) (C and D) is indicated.

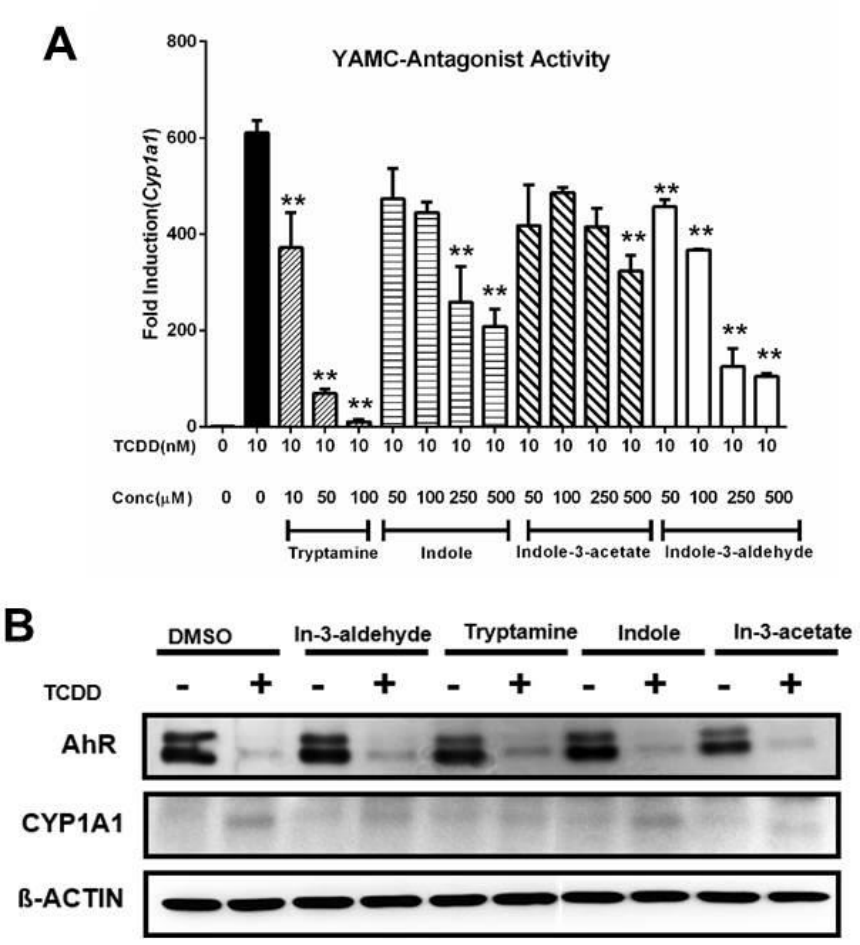


Figure 24. Tryptophan metabolites as AhR antagonists. YAMC cells were treated with tryptophan metabolites, TCDD and their combination, and effects on *Cyp1a1* mRNA (A) and CYP1A1/AhR proteins (B) were determined by real time PCR and Western blot analysis. (C) YAMC cells were treated with 10 nM TCDD, 50 μM tryptamine, and their combination for 2 and 24 hr and real time PCR was used to determine interactions of the AhR with the *Cyp1a1* promoter (containing XRE) in a ChIP assay. (D) YAMC cells were treated with DMSO or 10 nM for 24 hr and also cotreated with 50 μM tryptamine after 18, 20, 22 and 23 hr, and *Cyp1a1* mRNA was determined by real time PCR. Results (A and D) are expressed as means ± SE (3 replicates) and significant ($p<0.05$), inhibition is indicated (**).

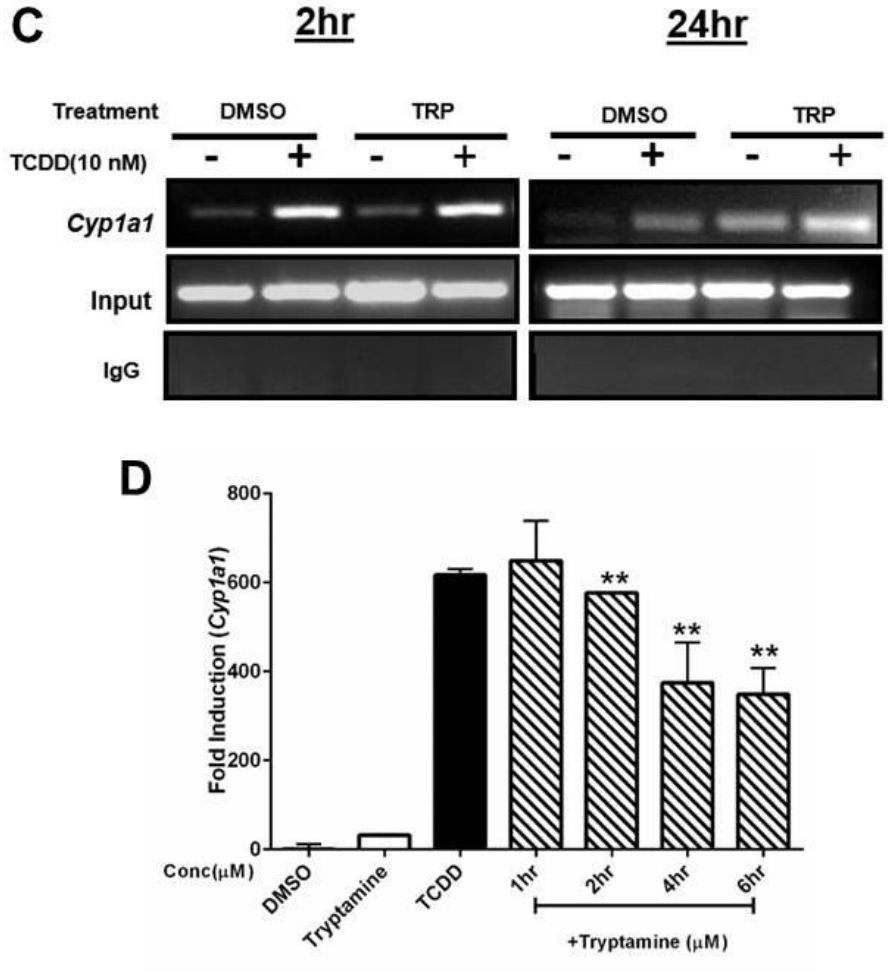


Figure 24. Continued.

Previous studies in CaCo2 cells showed that indole was an AhR antagonist (581) and we further investigated the inhibitory effect of the tryptophan metabolites on induction of *Cyp1a1* by TCDD (Figure 24A). All four compounds exhibited AhR antagonist activity, and both tryptamine and indole-3-aldehyde decreased induction of *Cyp1a1* mRNA by TCDD by >75% which was more effective than observed for CH (Figure 23D). Western blot analysis (Figure 24B) showed that TCDD but not the tryptophan metabolites decreased AhR protein expression and in combination experiments AhR levels

resembled that observed for TCDD alone. TCDD induced CYP1A1 protein in YAMC cells, whereas minimal induction was observed for the tryptophan metabolites and in combination experiments indole-3-acetate appeared to be the most effective inhibitor of TCDD-induced CYP1A1 protein. Unfortunately, CYP1A1 protein levels in YAMC cells were low and results of TCDD+tryptophan metabolites studies were difficult to interpret. We also examined the effects of TCDD, tryptamine and their combination on recruitment of the AhR to the *Cyp1a1* XRE in a ChIP assay. After treatment for 2 hr, TCDD alone or in combination with tryptamine induced AhR interactions with the *Cyp1a1* promoter, whereas minimal effects were observed in YAMC cells treated with tryptamine alone (Figure 24C). These results contrasted to those observed in CaCo2 cells where indole, the most effective AhR antagonist, blocked TCDD-induced AhR interactions with the *Cyp1a1* promoter (581). Analysis of these interactions were also investigated after treatment for 24 hr; significant AhR recruitment to the *Cyp1a1* promoter was observed after treatment with tryptamine alone and a comparison of the results of 2 and 24 hr treatments suggested that the tryptamine-induced AhR recruitment was a relatively slow process. In contrast, after treatment with TCDD for 24 hr, the AhR binding to the *Cyp1a1* promoter was decreased and this was consistent with the observed TCDD-induced degradation of the AhR protein (Figure 24B). Despite the inhibition of TCDD-induced *Cyp1a1* mRNA levels by tryptamine treatment for 24 hr (Figure 24A), AhR binding to the *Cyp1a1* promoter in the combined treatment group was essentially additive (Figure 24C). Therefore, it is possible that the inhibition of TCDD-induced *Cyp1a1* by tryptamine is post-transcriptional and AhR-independent. YAMC cells were

treated with TCDD alone for 24 hr and cotreated with tryptamine after 18, 20, 22 and 23 hr after addition of TCDD. The results showed that there was a time-dependent decrease in induced *Cyp1a1* mRNA (Figure 24D), suggesting that some of the inhibitory effects of tryptamine were post-transcriptional and may be due to destabilization of *Cyp1a1* mRNA.

The tryptophan metabolites exhibited structure-dependent AhR agonist/antagonist activities with respect to induction of *Cyp1a1* in YAMC cells and this pattern of activity was investigated with other Ah-responsive genes (582,587-589). Results in Figure 25A show that TCDD but not indole-3-aldehyde induced *Cyp1b1* expression in YAMC cells and in combination studies indole-3-aldehyde partially inhibited TCDD-induced *Cyp1b1* expression. Indole was a partial agonist for induction of *Cyp1b1* but did not inhibit TCDD-induced *Cyp1b1* mRNA levels (Figure 25B). Indole-3-acetate (Figure 25C) and tryptamine (Figure 25D) induced *Cyp1b1* mRNA levels similar to that of TCDD and did not inhibit induction by TCDD, indicating that both compounds were full AhR agonists for induction of *Cyp1b1*.

A similar approach was used in examine the AhR agonist/antagonist activities of the tryptophan metabolites with respect to regulation of *Ahrr* and *TiParp* gene expression. Indole-3-aldehyde minimally induced *Ahrr* (<2-fold) at the highest concentration (500 μM); TCDD induced *Ahrr* (<7- fold) and in combination indole-3-aldehyde was a weak AhR antagonist (Figure 26A). Indole (Figure 26B), indole-3-acetate (Figure 26C), and

tryptamine (Figure 26D) were partial agonists and induced *Ahrr*, and only tryptamine exhibited partial AhR antagonist activity. Indole-3-aldehyde, indole and indole-3-acetate (Figures 27A-27C) did not induce *TiParp* or inhibit induction of *TiParp* by TCDD, whereas tryptamine (Figure 27D) exhibited partial agonist/antagonist activity. Thus, the effects of the tryptophan metabolites as AhR agonist and antagonists were highly gene specific in YAMC cells and these differences are summarized in Table 10.

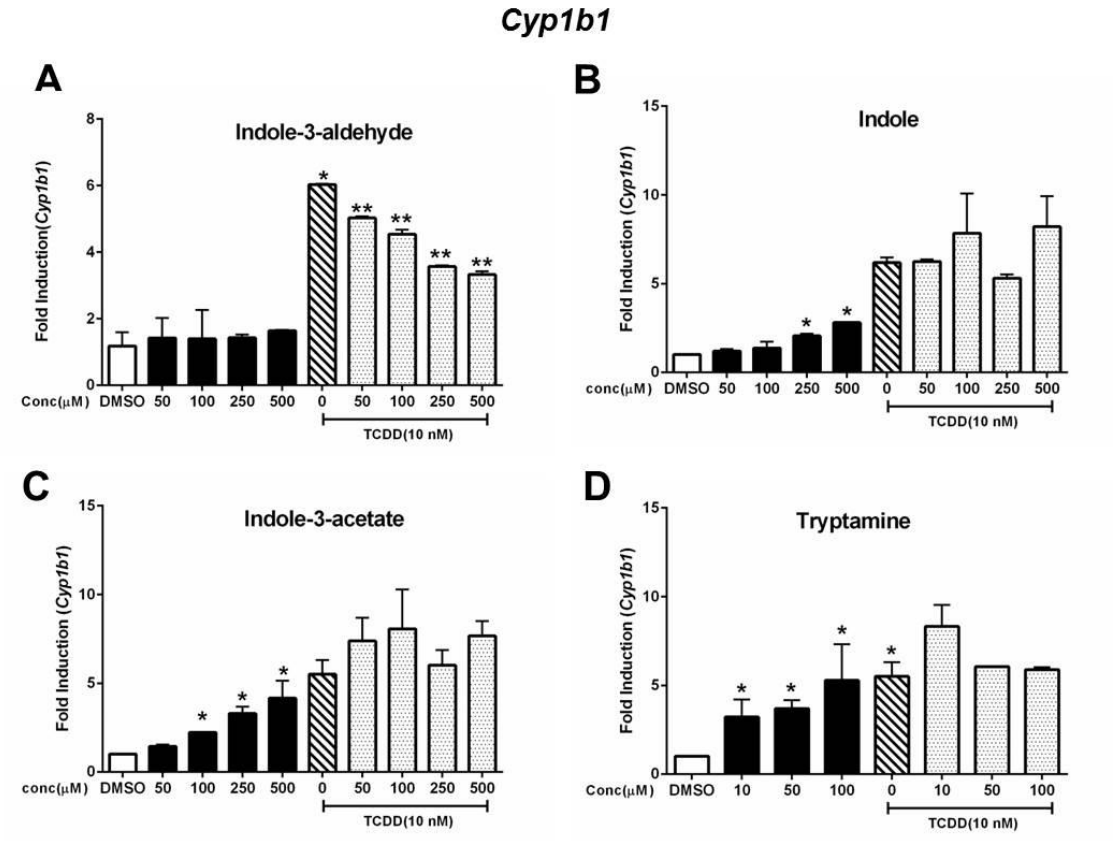


Figure 25. Induction of *Cyp1b1*. YAMC cells were treated with TCDD (alone), indole-3-aldehyde (A), indole (B), indole-3-acetate (C), and tryptamine (D) alone and in combination with TCDD for 24 hr, and *Cyp1b1* was determined by real time PCR. Results are expressed as means ± SE (3 replicates) and significant ($p<0.05$) induction (*) or inhibition (**) of TCDD-induced *Cyp1b1* is indicated.

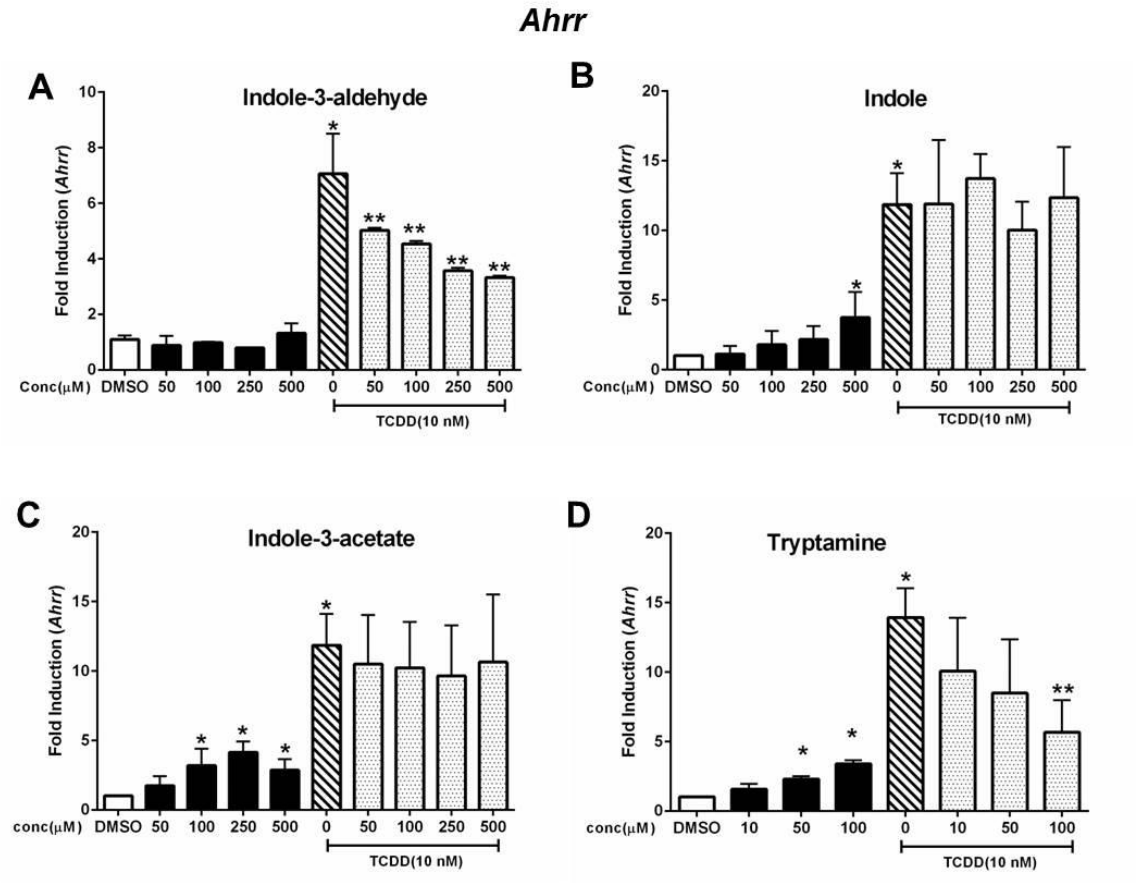


Figure 26. Induction of *Ahrr*. YAMC cells were treated with TCDD (alone), indole-3-aldehyde (A), indole (B), indole-3-acetate (C), and tryptamine (D) alone and in combination with TCDD for 24 hr, and *Cyp1b1* was determined by real time PCR. Results are expressed as means ± SE (3 replicates) and significant ($p<0.05$) induction (*) or inhibition (**) of TCDD-induced *Ahrr* is indicated.

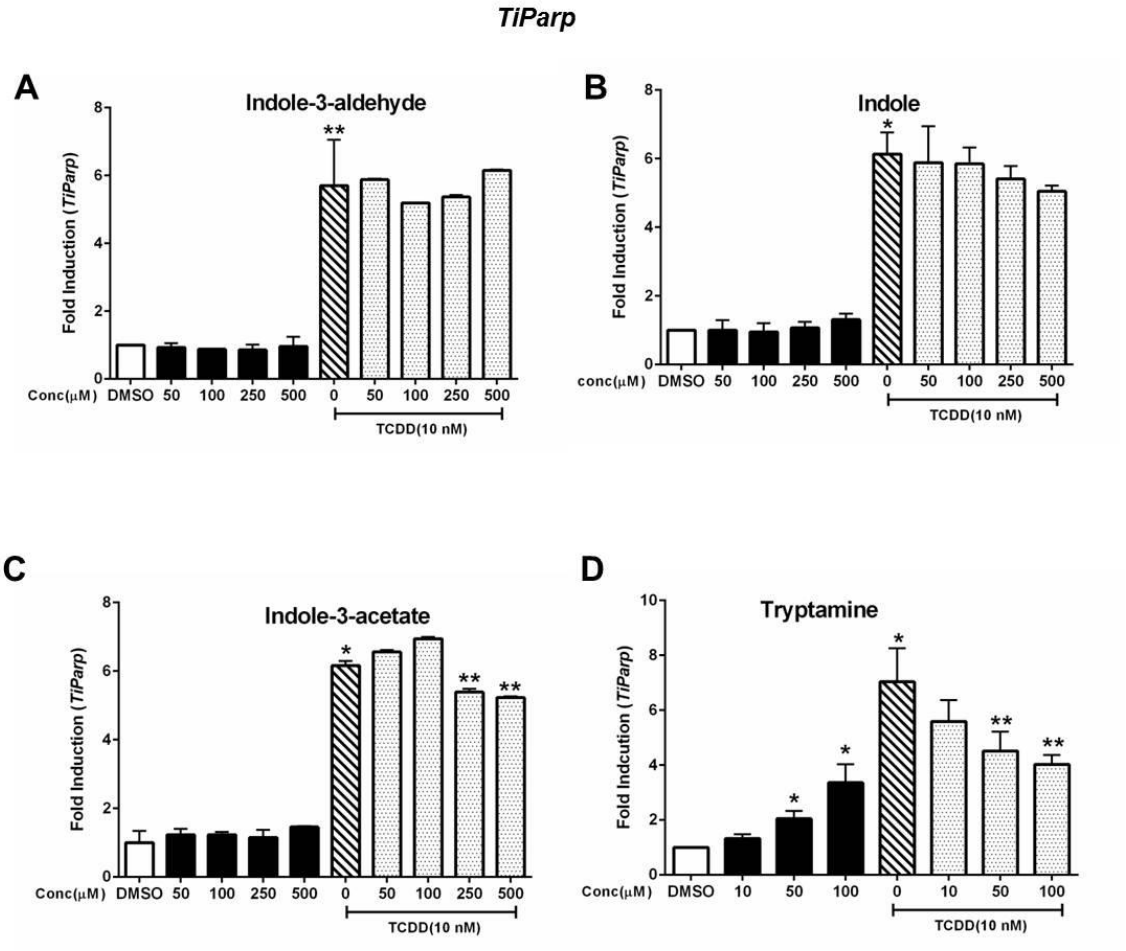


Figure 27. Induction of *TiParp*. YAMC cells were treated with TCDD (alone), indole-3-aldehyde (A), indole (B), indole-3-acetate (C), and tryptamine (D) alone and in combination with TCDD for 24 hr, and *TiParp* was determined by real time PCR. Results are expressed as means ± SE (3 replicates) and significant ($p<0.05$) induction (*) or inhibition (**) of TCDD-induced *TiParp* is indicated.

Table 10. Effects of the tryptophan metabolites as AhR agonists and antagonists are highly gene specific in YAMC cells.

	Cyp1a1		*Cyp1b1*		*Ahrr*		*TiParp*	
	Ag	Ant	Ag	Ant	Ag	Ant	Ag	Ant
TCDD	+	-	+	-	+	-	+	-
Indole-3-aldehyde	-*	+	-	+	-	+	-	-
Indole	-*	+	+	-	+	-	-	-
Indole-3-acetate	+	+	+	-	+	-	-	-
Tryptamine	+	+	+	-	+	-	+	+

***** Weak agonist activity and somewhat variable.

4.4 Discussion

The AhR is expressed in the GI tract, and studies in animal models demonstrate that this receptor and its ligands play an important role in gut health and response to stressors and disease (434,455,456,466,474,573-576). Loss of the AhR results in formation of colon tumors at the cecum and this is accompanied by increased expression of β-catenin in the small intestine, whereas wild-type $AhR^{+/+}$ mice do not develop tumors or overexpress β-catenin (474). $Apc^{min/+}$ mice which express a mutation in the Apc tumor suppressor gene were crossed with $AhR^{+/-}$ (heterozygote) mice and the resulting $Apc^{min/+}/AhR^{+/-}$ mice animals were more susceptible to cecal tumorigenesis (474). However, AhR-active botanical compound such as indole-3-carbinol and diindolylmethane (from cruciferous vegetables) significantly suppressed intestinal tumorigenesis in $Apc^{min/+}$ and $Apc^{min/+}$

$AhR^{+/-}$ mice (474). Gut ILC22 cells and postnatal lymphoid tissue-inducer-like (LTi) subsets express the AhR which is essential for many of their functions. For example, the loss of the AhR results in decreased expression of ILC22 and decreased protection against bacterial infections (456). The AhR agonist TCDD has been shown to induce Notch1 which is differentially required for the development of various ILC22 and LTi cell sub-types. The important functions of the AhR in maintaining intestinal function and health and protection against bacterial infections has been described in several reports showing that the AhR and its ligands also protect against intestinal damage/inflammation in experimental models of colitis and Crohn's disease (466,573-575). The severity of the effects of 2,4-trinitrobenzene sulfonic acid-induced colitis (resembles Crohn's disease) in mice was significantly decreased by treatment with the AhR agonists FICZ (456) and TCDD (576) and this was accompanied by suppression of several markers of inflammation. The severity of dextran sodium sulfate-induced colitis in mice was also decreased by the AhR agonists β-naphthoflavone (575), TCDD (576) and FICZ (456), and in the latter study, the AhR antagonist 2-methyl-2H-pyrazole-3-carboxylic acid enhanced the severity of the colitis (456).

Previously, we investigated the tryptophan metabolites in CaCo2 human colon cancer cells and demonstrated their ligand-dependent AhR agonist and antagonist activities based primarily on modulation of *CYP1A1* gene expression (581). We also observed similar responses in non-transformed YAMC cells where the most active AhR agonists for induction of *Cyp1a1* were tryptamine and indole-3-acetate; however, the fold

induction by both compounds was <3% of that observed for TCDD. In contrast, both tryptamine and indole-3-aldehyde were potent inhibitors of TCDD-induced *Cyp1a1* in YAMC cells, whereas tryptamine was primarily a full AhR agonist in CaCo2 cells using *CYP1A1* mRNA as an endpoint, demonstrating the importance of cell context.

We also investigated the AhR activity of the four tryptophan metabolites using three additional Ah-responsive genes, namely, *Ahrr*, *Cyp1b1* and *TiParp*, and the results indicated that the AhR agonist and antagonist activities were both compound- and gene-specific (Table 10). For example, tryptamine was a weak AhR agonist and partial antagonist for *Cyp1a1* mRNA expression; however, examination of the ligand-dependent recruitment of the AhR complex to the *Cyp1a1* promoter (Fig. 2C) did not readily explain a mechanism for the activity of tryptamine as an AhR agonist (Fig. 2A). In a separate experiment, we observed that treatment of YAMC cells with TCDD alone for 24 hr maximally induced *Cyp1a1* mRNA which could then be significantly decreased by addition of tryptamine 18, 20 or 22 hr after treatment with TCDD, suggesting that some of the inhibitory effects of tryptamine on induced *Cyp1a1* mRNA may be post-transcriptional. In contrast, tryptamine and TCDD induced similar levels of *Cyp1b1* mRNA and tryptamine did not affect TCDD-induced *Cyp1b1*, indicating that tryptamine was a full AhR agonist for this response. The cell context- and gene-specific AhR agonist and antagonist activities of the tryptophan metabolites are not unique and have been observed for other AhR ligands including 6-methyl-1,3,8-trichlorobenzofuran, flavonoids and pharmaceuticals (567,590-594). We also observed that the AhR

antagonist CH inhibited *Cyp1a1* induction by TCDD and the tryptophan metabolites, indicating that CH inhibited induction of *Cyp1a1* by a diverse spectrum of AhR ligands as previously reported (595). Ongoing studies show that CH also antagonized induction of *Cyp1b1* by TCDD and the tryptophan metabolites; however, CH did not antagonize induction of Ahrr or TiParp by TCDD and tryptophan metabolites (data not shown) and this is currently being investigated.

In summary, our results show that for a limited set of Ah-responsive genes the AhR agonist and antagonist activities of the tryptophan metabolites are gene-specific in non-transformed YAMC cells and different from that previously observed in CaCo2 cancer cells (581). Differences in the AhR agonist and antagonist activities of the tryptophan metabolites are due not only to the transformed vs. non-transformed phenotype of CaCo2 and YAMC cells but also to their different human vs. mouse origins. These results suggest that indole-3-aldehyde, indole, indole-3-acetate and tryptamine are selective AhR modulators (398,596) based on the results observed in this study. Although indole-3-aldehyde exhibited minimal AhR agonist activity, a recent report indicated that AhR-dependent induction of interleukin-22 by indole-3-aldehyde plays a key role in microbiota-mediated protection from fungal infection and colitis (404); however, indole-3-aldehyde did not induce interleukin 22 in YAMC cells (data not shown). Current studies are evaluating the contributions of AhR-active tryptophan metabolites in YAMC and other mouse- and human-derived cell lines to identify *in vitro* models that mimic *in vivo* effects of these compounds and identify relevant endpoints

such as interleukin 22 induction that will predict the effects of AhR-active microbiota metabolites on gut health.

5. MICROBIAL-DERIVED 1,4-DIHYDROXY-2-NAPHTHOIC ACID AND RELATED COMPOUNDS AS ARYL HYDROCARBON RECEPTOR AGONISTS/ANTAGONISTS: STRUCTURE-ACTIVITY RELATIONSHIPS AND RECEPTOR MODELING [*]

5.1 Introduction

Bifidobacteria are prominent in the gastrointestinal tract, and these bacteria and their metabolites have been associated with promotion of good health and are used as probiotic agents (597,598). For example, bifidobacteria alone or in combination with other bacterial species (e.g. *Lactobacillus*) result in decreased inflammation associated with Crohn's disease and ulcerative colitis (599-602). It was reported that cell-free filtrate from the bifidobacteria *Propionibacterium freudenreichii* stimulates bifodobacterial growth and this is attributed primarily to two microbial metabolites, namely 2-amino-3-carboxy-1,4-naphthoquinone (minor) and 1,4-dihydroxy-2-naphthoic acid (1,4-DHNA) (major) (603,604). 1,4-DHNA is an intermediate in the biosynthesis of menaquinone (vitamin K2) (605), and 1,4-DHNA has also been identified in lactic acid-producing bacterial *Lactobacillus casei* LP1 and in Korea traditional rice wine (606,607). Subsequent studies showed that 1,4-DHNA inhibits dextran sodium sulfate-induced colitis in mice and also decreases induced inflammation and colitis in

[*] Reprinted with permission from "Microbial-derived 1,4-dihydroxy-2-naphthoic acid and related compounds as aryl hydrocarbon receptor agonists/antagonists: structure-activity relationships and receptor modeling" by Cheng Y, Jin UH, Davidson LA, Chapkin RS, Jayaraman A, *et al*. Toxicol Sci (Accepted).

interleukin 10-deficient mice by suppressing macrophage-derived pro-inflammatory cytokines (608,609). These results suggest that 1,4-DHNA contributes to the health-promoting effects of bifidobacteria. It was also reported that 1,4-DHNA inhibits growth of *Helicobacter pylori* and induces apoptosis in human keratinocytes, indicating a potential application for treating psoriasis (610-612).

Several studies show that the aryl hydrocarbon receptor (AhR) and its receptor agonists also play a role as inhibitors of colitis and development of colorectal cancer in rodent models (466,574-576). *In vitro* studies in human Caco2 colon cancer cells showed that 1,4-DHNA induces *CYP1A1* gene expression, a marker of Ah responsiveness, and similar results were observed in the small intestine of wild-type but not AhR knockout mice (578). 1,4-DHNA also inhibits DSS-induced colitis and this response is attenuated after cotreatment with 1,4-DHNA plus CH-223191, an AhR antagonist (578). Thus, the health-promoting effects of 1,4-DHNA in the gut are due, in part, to its activity as an AhR agonist.

2,3,7,8-Tetrachlorodibenzo-*p*-dioxin (TCDD) is a prototypical and highly potent AhR agonist and environmental toxicant, and several halogenated aromatic industrial compounds and by-products and polynuclear aromatic hydrocarbons also act through the AhR (613). In addition, other classes of AhR ligands include endogenous biochemicals such as indolo-2,3[b]-carbazole, kynurenine and microbiota-derived tryptophan metabolites, health promoting phytochemicals, and pharmaceuticals (562,563,581,614).

The structure-dependent effects of TCDD and related halogenated aromatics have been extensively investigated; however, less is known about other structural classes of AhR ligands. Therefore, in this study we further elucidated the structure-activity relationships (SARs) of 1,4-DHNA and related naphthalene analogs as AhR ligands and demonstrates the important roles of both the hydroxyl- and carboxyl substituents and their location in the naphthalene ring. Moreover, since the structures of TCDD and 1,4-DHNA are different, we have also used computational modeling approaches to investigate differences in their interactions with the AhR.

5.2 Materials and methods

5.2.1 Cell lines, antibodies, and reagents

The young adult mouse colonic (YAMC) cell line was initially generated from the Immorto mouse (583) and has been previously used in our studies (584,585). Cells were maintained in RPMI 1640 medium with 5% fetal bovine serum, 5 units/ml mouse interferon-γ (IF005) (EMD Millipore, Massachusetts), 1% ITS "-" minus (insulin, transferrin, selenium) (41-400-045) (Life Technologies, Grand Island, NY) at 33°C (permissive conditions). In preparation for experiments, cells were transferred to 37°C (nonpermissive conditions). Caco2 human colon cancer cell line was obtained from the American Type Culture Collection (ATCC, Manassas, VA). Caco2 cells were maintained in Dulbecco's modified Eagle's medium (DMEM) nutrient mixture supplemented with 20% fetal bovine serum (FBS), 10 ml/L 100X MEM non-essential amino acid solution (Gibco), and 10 ml/L 100X antibiotic/antimycotic solution (Sigma-

Aldrich). Caco2 cells were maintained at 37°C in the presence of 5% CO_2, and the solvent (dimethyl sulfoxide, DMSO) used in the experiments was ≤ 0.2%. Mouse AhR antibody (BML-SA210) was purchased by Enzo (Enzo Life Sciences Inc., Farmingdale, NY). β-Actin (A1978) was purchased from Sigma-Aldrich (St. Louis, MO), and mouse CYP1A1 antibody was kindly provided by the late Dr. Paul Thomas (Rutgers University) and Dr. B. Moorthy (Baylor College of Medicine, Houston). Human CYP1A1, AHR, and GAPDH antibodies were purchased from Santa Cruz Biotechnology (Santa Cruz, CA). 1,4-Dihydroxy-2-naphthoic acid (1,4-DHNA), 3,5-dihydroxy-2-naphthoic acid (3,5-DHNA), 3,7-dihydroxy-2-naphthoic acid (3,7-DHNA), 1,4-dimethoxy-2-naphthoic acid (1,4-DMNA), 1-naphthoic acid (1-NA), 2-naphthoic acid (2-NA), 1-naphthol (1-NOH), and 2-naphthol (2-NOH) used in this study were purchased from Sigma-Aldrich (St. Louis, MO). 1-Hydroxy-2-naphthoic acid (1-HNA) was purchased from Alfa Aesar (Ward Hill, MA, USA) and 4-hydroxy-2-naphthoic acid (4-HNA) was purchased from Chem Scene (www.chemscene.com).

5.2.2 Chromatin immunoprecipitation assay

The chromatin immunoprecipitation (ChIP) assay was performed using the ChIP-IT Express Magnetic Chromatin Immunoprecipitation kit (Active Motif, Carlsbad, CA) according to the manufacturer's protocol. YAMC cells ($1.2x10^7$ cells) were treated with TCDD and/or compounds for 2 or 24 hr. Caco2 cells ($5x10^6$ cells) were treated with TCDD and/or compounds for 2 hr. The cells were then fixed with 1% formaldehyde, and the fixation was stopped by 0.125 M glycine. After washing with phosphate-buffered

saline (PBS), cells were scraped and pelleted. The cell pellets were hypotonically lysed to release nuclei and then sonicated to the desired chromatin length (200-1500 bp). The sonicated chromatin was immunoprecipitated with normal rabbit IgG or AhR antibodies and protein A-conjugated magnetic beads at 4°C for overnight. After the magnetic beads were extensively washed, protein-DNA crosslinks were reversed and eluted. DNA was prepared by proteinase K digestion followed by polymerase chain reaction (PCR) amplification. The mouse *Cyp1a1* primers were 5'-CAG GAG AGC TGG CCC TTT A-3' (sense) and 5'-TAA GCC TGC TC ATC CTG TG-3' (antisense), and subsequently amplified by targeting a 215-bp region of mouse *Cyp1a1* promoter, which contained the AhR-binding sequences. The human *CYP1A1* primers were 5'-TCA GGG CTG GGG TCG CAG CGC TTC T-3' (sense) and 5'-GCT ACA GCC TAC CAG GAC TCG GCA G-3' (antisense) which amplified a 112-bp region of the human *Cyp1A1* promoter which containing the AhR binding sequences. PCR products were resolved on a 2% agarose gel in the presence of ETBR.

5.2.3 Quantitative real-time PCR

Total RNA was isolated using Zymo Quick RNA MiniPrep Kit (Zymo Research, Irvine, CA) according to the manufacturer's protocol. RNA was eluted with RNase-free water and stored at -80°C. Real-time (RT)-PCR was carried out using iTaq Universal SYBR Green One-step Kit (Bio-Rad, Hercules, CA). The primers were listed in Tables 11 and 12.

Table 11. Mouse primers list.

Name	Forward Primer	Reverse Primer
TBP	GAACAATCCAGACTAGCAGCA	GGGAACTTCACATCACAGCTC
Cyp1a1	ATCCAAGGCAGAATACGGTG	TCCACTCCATCTTCCGACTT
Cyp1b1	GGATATCAGCCACGACGAAT	ATTATCTGGGCAAAGCAACG

Table 12. Human primers list.

Name	Forward Primer	Reverse Primer
TBP	GATCAGAACAACAGCCTGCC	TTCTGAATAGGCTGTGGGGT
CYP1A1	GACCACAACCACCAAGAAC	AGCGAAGAATAGGGATGAAG
CYP1B1	CACTGACATCTTCGGCG	ACCTGATCCAATTCTGCCTG

5.2.4 Western blot analysis

Cells were treated with different concentrations of the compounds for 18 hr and then collected using high-salt buffer (50 mM HEPES, 0.5 mol/l NaCl, 1.5 mM $MgCl_2$, 1 mM EGTA, 10% glycerol, and 1% Triton-X-100) and 10 μl/ml Protease Inhibitor Cocktail (Sigma-Aldrich, St. Louis, MO). Protein lysates were incubated for 5 min at 95°C before electrophoresis and then separated on 10% SDS-polyacrylamide gel electrophoresis 120 V for 2 to 3 hr. Proteins were transferred onto polyvinylidene difluoride membranes by wet electroblotting in a buffer containing 25 mM Tris, 192 mM glycine, and 20% methanol for 1.5 hr at 180 mA. Membranes were then blocked for 30 min with specific antibodies. Detection of specific proteins was performed using Chemiluminescence and then exposed to Kodak image station 4000 mm Pro (Carestream Health, Rochester, NY).

5.2.5 Gel retardation

Gel retardation experiments were performed using guinea pig hepatic cytosol according to a previously published standard protocol (615).

5.2.6 Generation of AhR-deficient YAMC cells

Two AhR CRISPR guide RNAs, in a Cas9 vector which also expresses GFP, were purchased from GenScript (Piscataway, NJ). Sequences of the guide RNAs were CGGTCTCTGTGTCGCTTAGA and GAACACAGAGTTAGACCGCC. YAMC cells were cotransfected with both plasmids and 48 hr later, cells were FACS sorted to collect the 5% highest GFP expressing cells into individual wells of a 96-well plate. Clonal cells were grown into larger cultures and tested for knock-out of AhR protein.

5.2.7 Statistical analysis

Statistical significance of differences between the treatment groups was determined by an analysis of variance and/or Student's t test, and levels of probability were noted. At least 3 repeated experiments were determined for each data point, and results are expressed as means ± SD.

5.2.8 Computational homology modeling of AhR

Residues 241 through 400 (sequence HGQNKKGKDG-ALLPPQLALF-AIATPLQPPS-ILEIRTKNFI-FRTKHKLDFT-PIGCDAKGQL-ILGYTEVELC-TRGSGYQFIH-AADMLHCAES-HIRMIKTGES-GMTVFRLFAK-HSRWRWVQSN-ARLIYRNGRP-

DYIIATQRPL-TDEEGREHLQ-KRSTSLPFMF) of the mouse AhR were investigated through homology modeling. The homology model of the AhR was derived using I-TASSER (616). All binding site residues characterized by mutagenesis studies and known to be critical or influence TCDD binding (617-619), which were also investigated in a previous study (620), are included in our model. The homology model was built using the crystal structure of the hypoxia-inducible factor-2α:AhR nuclear translocator complex (PDB ID: PZP4 (621), chain B) as an initial template. The N- and C- terminal ends of the modeled protein were acetylated and amidated to avoid any artifacts which could occur due to the artificial placement of positively and negatively charged groups at the backbone termini of the truncated ends of the modeled systems under investigation.

5.2.9 Generation of docking poses

TCDD and 1,4-DHNA were independently positioned into the binding sites of AhR using AutoDock Vina (622). The structures for both TCDD and 1,4-DHNA were obtained from the ZINC database (623). The search space used was 20 Å x 24 Å x 20 Å so as to include in the binding pocket AhR residues involved in TCDD binding according to mutagenesis studies (617-619). During the initial AutoDock Vina (622) docking, the side-chains of residues identified as binding pocket for TCDD binding to AhR according to mutagenesis studies (617-619) were treated as flexible. The produced complex conformations of TCDD and 1,4-DHNA in complex with AhR with the lowest binding free energy according to AutoDock Vina (622) were used as initial structures for docking simulation runs performed in CHARMM (624). Six separate docking simulation

protocols were introduced independently to investigate the binding of TCDD and 1,4-DHNA to AhR. In summary, during the docking simulations each ligand independently was constrained using harmonic or quartic potential energy functions to the docked binding site through the MMFP module of CHARMM [42]; a quartic potential energy function was used so as to avoid bias toward the initial positioning of the molecules performed by Autodock Vina (622). In each 20 independent runs comprising 200 of short 2 ps simulations were performed. In each step, prior to the short MD simulation run, the ligands were independently rotated about a randomly generated axis and posterior to the short MD simulation run, the complex conformation was minimized and was saved for evaluation. This procedure resulted in the generation of 4,000 binding conformations for each ligand in complex with AhR per protocol. Additional information on the protocols used in the docking simulation runs are provided in the Supplemental Materials and Methods (Appendix C). As an initial screening, from each docking simulation protocol, out of the 4,000 complex structures produced, we extracted the 3 complex structures with the lowest interaction energy for further analysis. This translated to 18 docking conformations of TCDD in complex with AhR and 18 docking conformations of 1,4-DHNA in complex with AhR were extracted in total.

5.2.10 Molecular dynamics simulations of selected TCDD:AhR and 1,4-DHNA:AhR complexes

In order to refine the ligand:receptor structures, optimize intermolecular interactions, determine the structural stability of the selected binding modes, and assess the most

energetically favored binding modes of the TCDD:AhR and 1,4-DHNA:AhR complexes, we performed 36 independent MD simulation runs of which the initial structures corresponded to the 18 selected TCDD:AhR docking conformations, and the 18 selected 1,4-DHNA: AhR docking conformations. All MD simulations of the 36 systems under investigation were performed in explicit solvent using CHARMM (624) and CHARMM36 topology and parameters (625) with periodic boundary conditions. Additional information on the MD simulations is provided in the Supplemental Materials and Methods (Appendix C).

5.2.11 MM GBSA association free energy calculations

To identify the most energetically favorable conformation of TCDD and 1,4-DHNA in complex with AhR, we calculated the association free energy of the 18 complexes per ligand over the 10 ns production runs using the Molecular Mechanics Generalized Born Surface Area (MM GBSA approximation (626-628), by extracting snapshots from the simulations every 20 ps. Additional information on the MM GBSA calculations is provided in the Supplemental Materials and Methods (Appendix C).

5.2.12 Selection and analysis of the binding modes with lowest association free energy

The simulations of the TCDD:AhR and 1,4-DHNA:AhR binding modes with the most favorable MM GBSA association free energies were selected as the ones representing the most likely naturally occurring binding conformations of the two ligands in complex

with AhR; the selection was performed similarly to the studies as previously described (629-632). The selected simulations of both the TCDD:AhR and 1,4-DHNA:AhR binding modes were extended for an additional 20 ns, for a total production run of 30 ns each. To determine the stability of the ligands in the AhR binding pockets, the entire complexes were structurally aligned by the backbone atoms of the pocket residues and the RMSD of the ligand heavy atoms was calculated with respect to the average conformation over the entire 10 ns production run duration using VMD (633,634). To determine the key interactions occurring in the lowest association free energy binding modes, the average per AhR residue interaction free energies between the AhR protein and each ligand of the structures with the lowest MM GBSA binding free energies were calculated for the entire 30 ns production runs (626,635,636). Additional information on the per AhR residue interaction free energy calculations is provided in the Supplemental Materials and Methods (Appendix C).

5.3 Results

5.3.1 Induction of *Cyp1a1* and *Cyp1b1* in mouse YAMC cells by TCDD and naphthalene compounds

Both TCDD and 1,4-DHNA induce *Cyp1a1* and inhibit DSS-induced colitis; however, 1,4-DHNA is approximately three orders of magnitude lower in potency (575,576,578). In contrast to the well-known structure-activity relationships for dioxin-like compounds, the contributions of the hydroxyl and carboxylic acid substituents and their positions on the naphthalene ring has not previously been reported. In this study, we used YAMC

cells as a model for investigating the Ah-responsiveness of naphthalene derivatives on normal colon, and also human Caco2 colon cancer cells which are frequently used as an *in vitro* model for colonic responses and 1,4-DHNA induced AhR-dependent *Cyp1a1* gene expression in this cell line has previously been reported (578). The substituted naphthalene derivatives used in this study included 1,4-DHNA, 3,5- and 3,7-DHNA, 1,4-dimethoxy-2-naphthoic acid (DMNA), 1-hydroxy-2-naphthoic acid (1-HNA), 4-HNA, 2-naphthoic acid (2-NA), 1-NA, 2-naphthol (2-NOH), and 1-NOH (Appendix A-3). The effects of these compounds on YAMC and Caco2 cell viability are summarized in Appendix A-4. In YAMC cells, 1,4-DHNA was the most cytotoxic of the substituted naphthalenes, and only 500 and 1000 μM 1- and 2-NOH were more cytotoxic than DHNA. *Cyp1a1* induction is widely used as a marker of Ah-responsiveness, and 5-50 μM DHNA induced a concentration-dependent increase in *Cyp1a1* mRNA levels (Figure 28A). TCDD (10 nM) induced approximately a 600-fold increase in *Cyp1a1* mRNA levels, and the maximal induction by 1,4-DHNA (50 μM) was approximately 450-fold. In contrast, 3,5- and 3,7-DHNA exhibited minimal activity as *Cyp1a1* inducers (15- to 40-fold lower induction than TCDD) (Figures 28B and 28C), and even lower inducibility was observed for 1,4-DMNA (Figure 28D). Thus, maximal activity was observed for the 1,4-dihydroxy substitution and methylation of these hydroxyl groups resulted in loss of activity. Both 1-HNA and 4-HNA contain a single hydroxyl substituent and these compounds induced *Cyp1a1* mRNA (1-HNA > 4-HNA) (Figures 28E and 28F), whereas 1- and 2-NA (containing no hydroxyl groups) exhibited low to non-detectable induction (Figures 28G and 28H). 1- and 2-NOH maximally induced a 43- and 50-fold

enhancement of *Cyp1a1* mRNA compared to the approximately 600-fold induction response observed for TCDD (Figures 28I and 28J), indicating that loss of the 2-carboxyl substituent resulted in decreased activity. 1,4-Dihydroxynaphthalene is unstable in solution and is oxidized to the 1,4-quinone; however, induction of *Cyp1a1* by this compound was also observed (Appendix A-5). Western blot analysis showed that TCDD induced CYP1A1 protein in YAMC cells with only minimal changes in protein levels by 1,4-DHNA, 1- and 4-HNA; however, TCDD, 1,4-DHNA, 1- and 4-HNA, 1- and 2- NOH decreased expression of the AhR protein (Figure 28K). TCDD, 1,4-DHNA, 1- and 4-HNA, 1- and 2-NOH treatment downregulated AhR expression and minimal effects were observed for the other analogs. Thus, the induction response (CYP1A1) for 1,4-DHNA and related compounds was maximal for 1,4-DHNA, and the loss of one or both hydroxyl or carboxyl groups decreased potency.

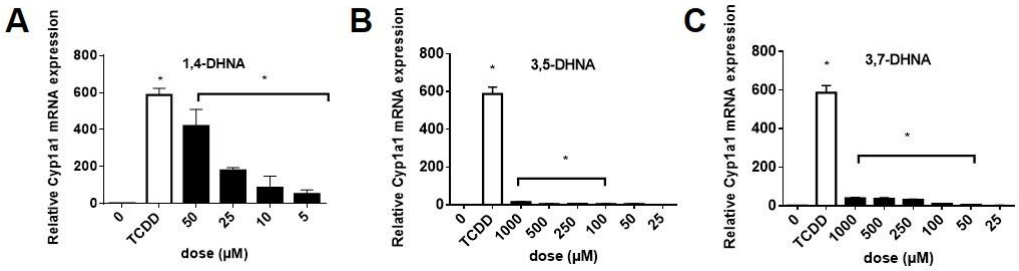


Figure 28. Induction of *Cyp1a1* in mouse YAMC cells. YAMC cells were treated with different concentrations of 1,4-DHNA (A), 3,5-DHNA (B), 3,7-DHNA (C), 1,4-DMNA (D), 1-HNA (E), 4-HNA (F), 1-NA (G), 2-NA (H), 1-NOH (I) and 2-NOH (J) for 18 hr, and *Cyp1a1* mRNA levels were determined (in triplicate) by real time PCR as outlined in the Materials and Methods. (K) Western blot analysis. YAMC cells were treated with a single concentration of the naphthalene compounds for 24 hr, and whole cell lysates were then analyzed by western blots. TCDD (10 nM) was used a positive control. Significant ($p<0.05$) induction is indicated (*).

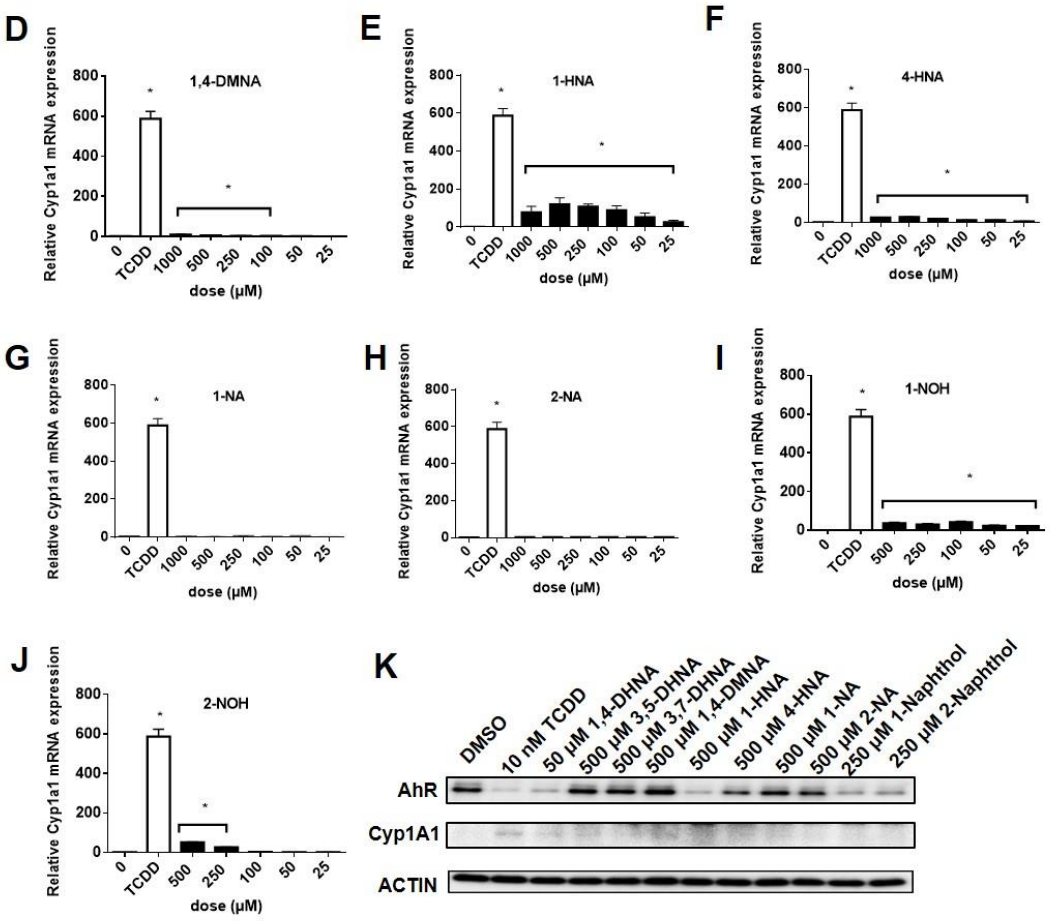


Figure 28. Continued.

Basal levels of *Cyp1b1* mRNA in YAMC cells were higher than observed for *Cyp1a1*, and 10 nM TCDD induced a 10-fold increase in *Cyp1b1* mRNA in this cell line and 50 μM DHNA induced a similar fold induction response (Figure 29A). 3,5- and 3,7-DHNA, 1,4-DMNA, 1- and 4-HNA induced *Cyp1b1* mRNA (Figures 29B-29F), and with the exception of 3,7-DHNA, the maximal induction response for the naphthalene derivatives was similar to that observed for TCDD. Both naphthoic acids (1-NA and 2-NA) exhibited minimal induction of *Cyp1b1* mRNA (Figures 29G and 29H), whereas 1- and

2-NOH induced *Cyp1b1* levels > 50% of that observed for TCDD (Figures 29I and 29J). Thus, the fold induction of *Cyp1b1* by TCDD was much lower than observed for *Cyp1a1* in YAMC cells and although the SARs for the naphthalene compounds were similar for both responses, their fold induction responses compared to TCDD were significantly higher for *Cyp1b1* compared to *Cyp1a1*. The most striking difference in compound-induced gene expression was observed for 1,4-DMNA which did not induce *Cyp1a1* but induced levels of *Cyp1b1* mRNA >80% of the level observed for 10 nM TCDD (Figure 29D).

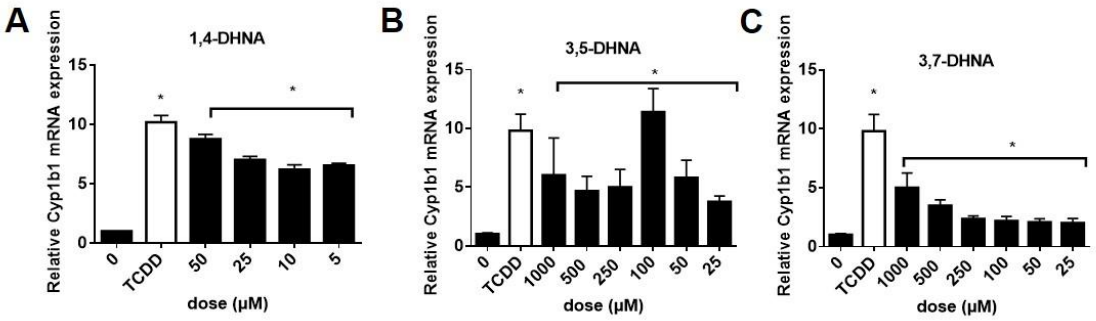


Figure 29. Induction of *Cyp1b1* in mouse YAMC cells. YAMC cells were treated with different concentrations of 1,4-DHNA (A), 3,5-DHNA (B), 3,7-DHNA (C), 1,4-DMNA (D), 1-HNA (E), 4-HNA (F), 1-NA (G), 2-NA (H), 1-NOH (I) and 2-NOH (J) for 18 hr and *Cyp1b1* mRNA levels were determined (in triplicate) by real time PCR as outlined in the Materials and Methods. Significant ($p<0.05$) induction is indicated (*).

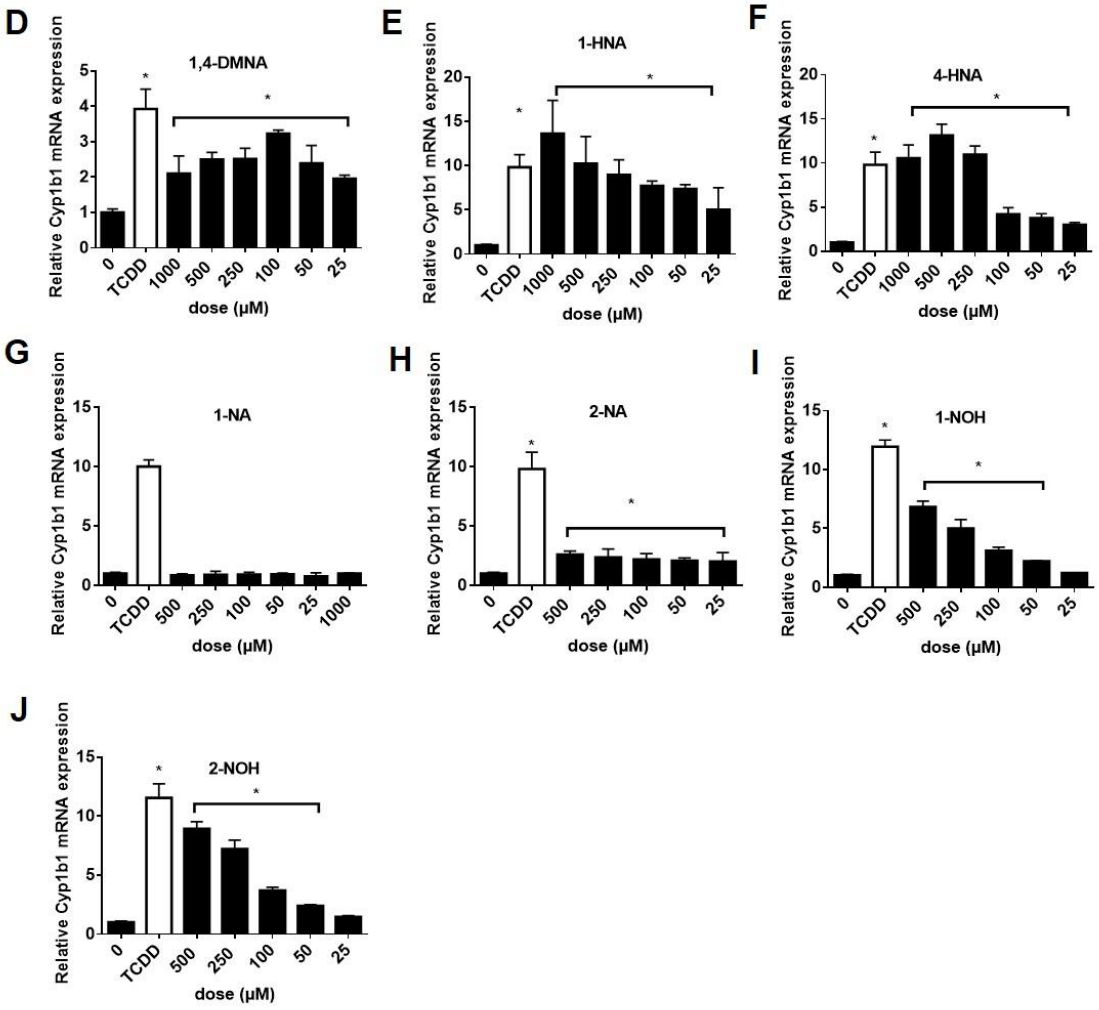


Figure 29. Continued.

5.3.2 Induction of *CYP1A1* and *CYP1B1* in human Caco2 cells by TCDD and naphthalene compounds

The SARs for 1,4-DHNA and structurally-related analogs were also carried out in human Caco2 cells, a colon cancer cell line used extensively as a model for investigating colonic effects of various drugs and dietary factors (627,637). 1,4-DHNA (Figure 30A) but not 3,5-DHNA, 3,7-DHNA or 1,4-DMNA (Figures 30B-30D) induced *CYP1A1*

mRNA levels >50% of that observed for TCDD (140-fold induction). In contrast, both 1-HNA and 4-HNA maximally induced *CYP1A1* mRNA (Figures 30E and 30F); 1- and 2-NA (Figures 30G and 30H) were relatively inactive, and both 1- and 2-NOH (Figures 30I and 30J) induced <15% of the maximal response observed for TCDD. 1,4-Dihydroxynaphthalene also induced *CYP1A1* in Caco2 cells (Appendix A-6). Western blot analysis showed that TCDD and 1,4-DHNA induced CYP1A1 protein, and both the 1- and 4-HNA compounds also induced these responses and all 4 compounds decreased AhR protein levels. TCDD, 1,4-DHNA, 1- and 4-HNA induced AhR downregulation and these results were different than their effects on the AhR in YAMC cells (Figure 28).

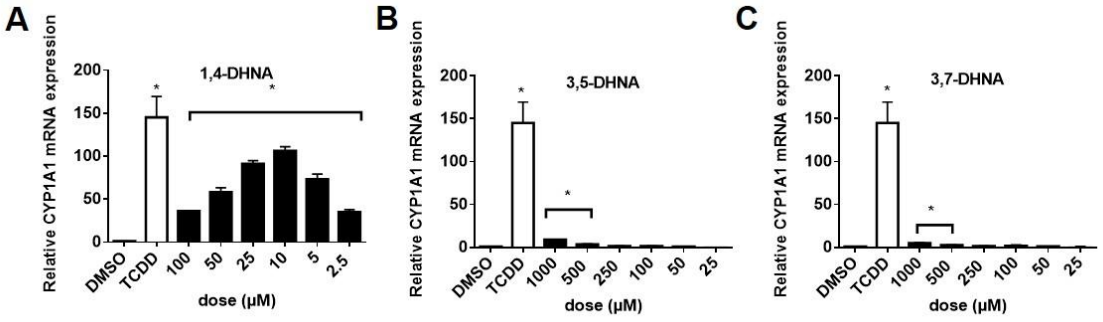


Figure 30. Induction of *CYP1A1* in human Caco2 cells. Caco2 cells were treated with different concentrations of 1,4-DHNA (A), 3,5-DHNA (B), 3,7-DHNA (C), 1,4-DMNA (D), 1-HNA (E), 4-HNA (F), 1-NA (G), 2-NA (H), 1-NOH (I) and 2-NOH (J) for 18 hr, and *CYP1A1* mRNA levels were determined (in triplicate) by real time PCR as outlined in the Materials and Methods. (K) Western blot analysis. Caco2 cells were treated with the various compounds for 24 hr, and whole cell lysates were then analyzed by western blots. TCDD (10 nM) was used a positive control. Significant ($p<0.05$) induction is indicated (*).

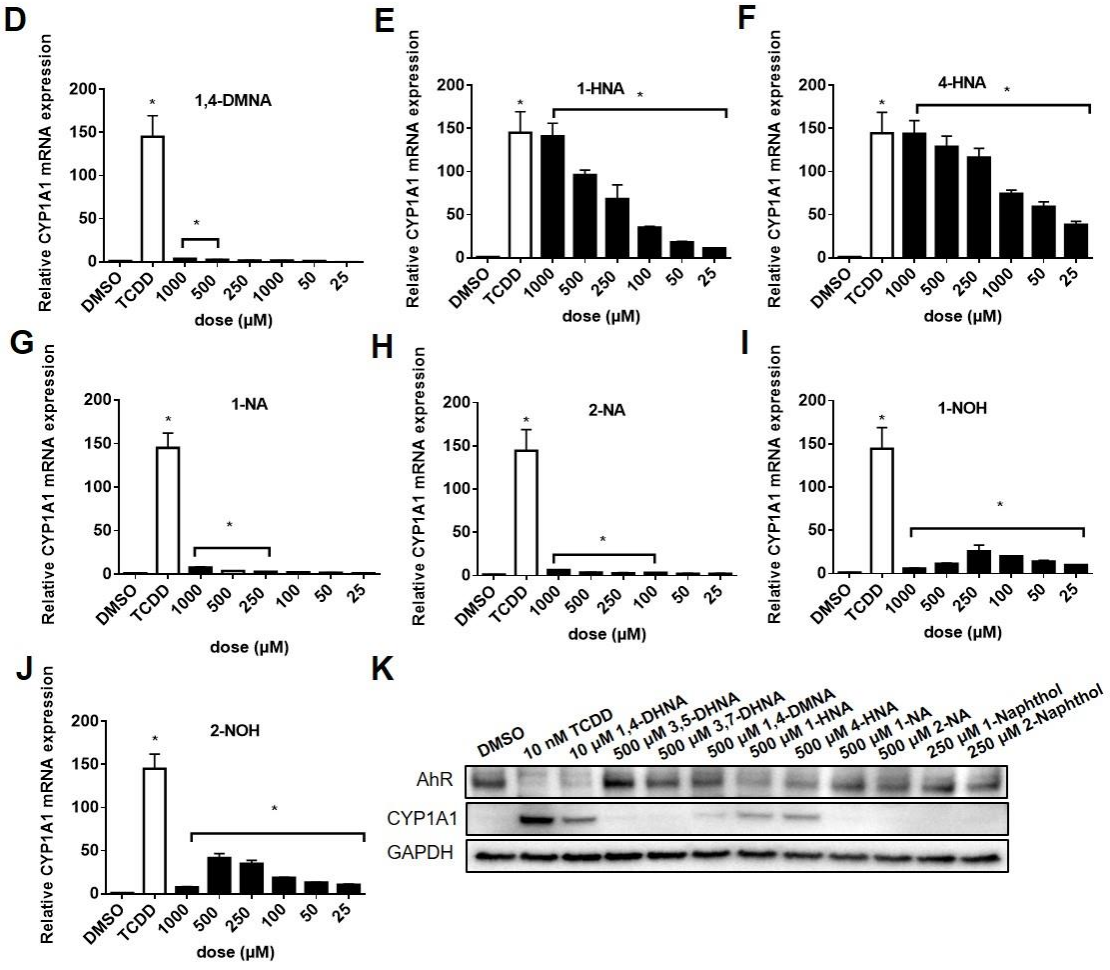


Figure 30. Continued.

In Caco2 cells, TCDD induced a 31-fold induction of *CYP1B1* compared to controls (Figure 31A), and only minimal to non-detectable induction was observed for 3,5-DHNA, 3,7-DHNA or 1,4-DMNA (Figures 31B-31D), whereas maximal induction responses were observed for 1- and 2- HNA (Figures 31E and 31F). Minimal induction was observed for 1- and 2-NA (Figures 31G and 31H), whereas both 1-NOH and 2-NOH induced *CYP1B1* (50% of TCDD-induced response) (Figures 31I and 31J) (2-NOH > 1-

NOH). Appendix A-7 compares the induction of *CYP1A1/CYP1B1* in YAMC and Caco2 cells after treatment for 6 or 18 hr. Induction responses for *CYP1A1* (YAMC and Caco2) and *CYP1B1* (Caco2) were significantly higher at the latter time point and this is consistent with previous studies on induction of *CYP1A1* mRNA in this cell line (638). In contrast, comparable compound-dependent induction of *Cyp1b1* mRNA levels was observed YAMC cells after treatment for 6 and 18 hr. The SARs for induction of *CYP1A1* by 1,4-DHNA and related compounds (compared to TCDD) were similar in Caco2 and YAMC cells; however, both 1- and 4-HNA induction responses were lower in YAMC vs. Caco2 cells, suggestive of species differences in the AhR and/or cell-specific differences in metabolism of these compounds. In contrast, SARs for induction of *Cyp1b1/CYP1B1* were highly variable and cell context-dependent for the naphthalene compounds.

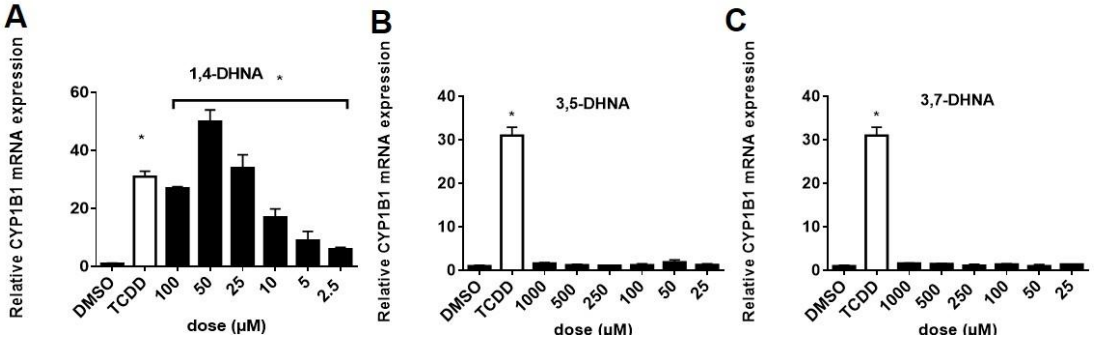


Figure 31. Induction of *CYP1B1* in human Caco2 cells. Caco2 cells were treated with different concentrations of 1,4-DHNA (A), 3,5-DHNA (B), 3,7-DHNA (C), 1,4-DMNA (D), 1-HNA (E), 4-HNA (F), 1-NA (G), 2-NA (H), 1-NOH (I) and 2-NOH (J) for 18 hr, and *CYP1B1* mRNA levels were determined (in triplicate) by real time PCR as outlined in the Materials and Methods. Significant ($p<0.05$) induction is indicated (*).

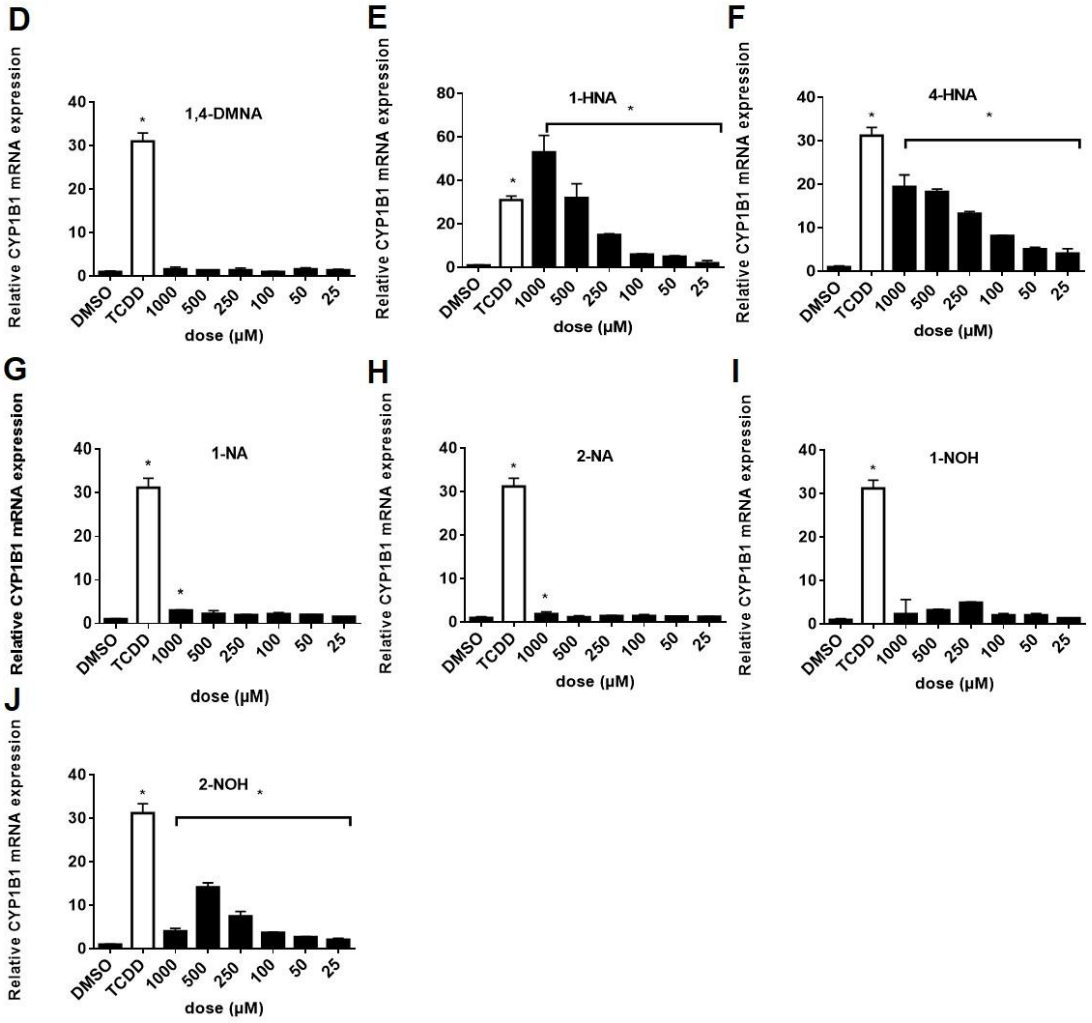


Figure 31. Continued.

5.3.3 Effects of naphthalene compounds as AhR antagonists and in transformation of guinea pig cytosol and AhR DRE localization in ChIP assays

In order to confirm the application of YAMC cells for investigating the Ah-responsiveness of compounds (563), we used the CRISPR/Cas9 technology to generate AhR knockout cells and expression of AhR in one of these cell lines used in this study is illustrated in Figure 32A. Treatment of the knockout cells with TCDD, 1,4-DHNA and

related compounds did not induce *Cyp1a1* (Figure 32B) or *Cyp1b1* (Figure 32C), confirming the AhR-dependence of the induction response in wild-type YAMC cells (Figures 28 and 29). We also investigated the potential AhR antagonist activities of 1,4-DHNA and related compounds in mouse YAMC and human Caco2 cells by determining their inhibition of TCDD-induced *Cyp1a1* gene expression (Figure 32D). In YAMC cells, all compounds, with the exception of 1,4- and 3,7-DHNA, inhibited TCDD-induced *Cyp1a1* mRNA expression at one or more of the higher concentrations. However, AhR antagonist activity for these substituted naphthalenes was observed at concentrations that induced some level of cytotoxicity (Appendix A-4). The only compound that did not act as an AhR antagonist (3,7-DHNA) was not cytotoxic. In contrast, with the exception of 1- and 2-NOH, this series of substituted naphthalenes exhibited minimal cytotoxicity in Caco2 cells (Appendix A-5), suggesting that this cell line may be more suitable for determining AhR antagonist activity. The results (Figure 32E) show that 1,4-DHNA (20-100 μM) significantly inhibited TCDD-induced *CYP1A1* gene expression in Caco2 cells, providing an explanation for the reduced reduction response at higher 1,4-DHNA concentrations (Figure 30A). Non-cytotoxic concentrations of both 1- and 2-NOH (e.g. 250 μM) also exhibited AhR antagonist activity with 1-NOH being highly active (>80% inhibition).

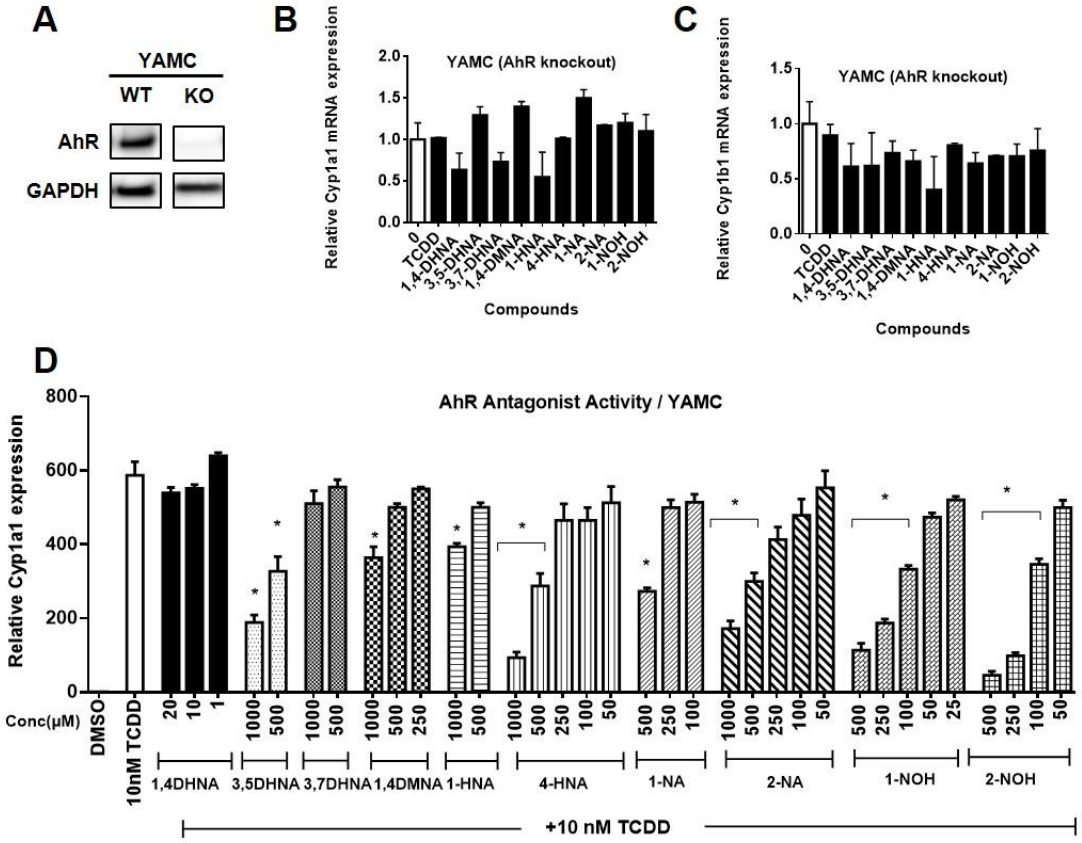


Figure 32. 1,4-DHNA and related compounds do not activate AhR-deficient YAMC cells and their partial AhR antagonist activity. (A) AhR knockout YAMC cells. The AhR was knocked out (ko) in YAMC cells using CRISPR/Cas9, and expression of the AhR in wild type (wt) and ko cells was determined by western blots as outlined in the Materials and Methods. ko-YAMC cells were treated with 10 nM TCDD, 10 μM 1,4-DHNA, 500 μM 3,5-DHNA, 3,7-DHNA, 1,4-DMNA, and 1-HNA, 4-HNA, 1-NA, 2-NA, and 250 μM 1-NOH and 2-NOH for 18 hr and analyzed for expression of *Cyp1a1* (B) and *Cyp1b1* (C) mRNA (in triplicate) by real time PCR. YAMC (D) and Caco2 (E) cells were treated with 10 nM TCDD alone and in combination with the hydroxyl/naphthoic acids for 18 hr, and *Cyp1a1*/*CYP1A1* mRNA levels were determined (in triplicate) by real time PCR. Significant ($p<0.05$) antagonist activity is indicated (*). In the YAMC knockout cells, 3,5-DHNA, 1,4-DMNA and 1-NA slightly induced (<50%) *Cyp1a1* mRNA levels.

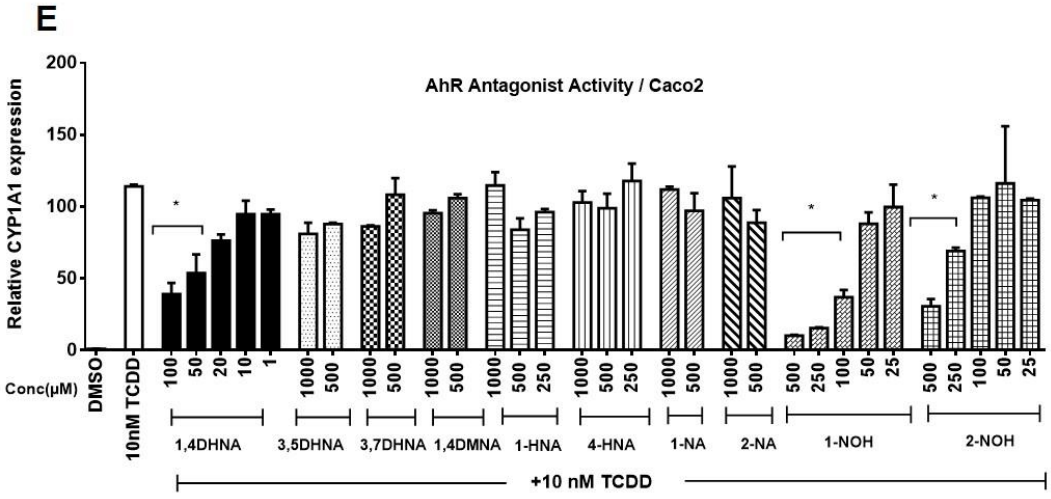


Figure 32. Continued.

Figures 33A and 33B summarize the effects of 1,4-DHNA and related compounds on transformation of hepatic cytosol (guinea pig) to its DNA binding form in a gel mobility shift assay. 1,4-DHNA (10 and 100 μM) and 100 μM 3,5-DHNA and 1- and 4-HNA, 1- and 2-NOH alone significantly stimulated AhR transformation/DNA binding of guinea pig hepatic cytosol, whereas AhR transformation/DNA binding was not observed for 3,7-DHNA, 1,4-DMNA, 1- and 2-NA. In combination studies with TCDD, only 1- and 2-NOH inhibited TCDD-induced transformation and this correlated with the AhR antagonist activity observed in the transaction assays (Figures 32D and 32E). As a positive control, we show that the AhR antagonist CH229131 did not stimulate AhR transformation/DNA binding of guinea pig cytosol but inhibited TCDD-induced transformation/DNA binding. These AhR transformation/DNA binding assays confirmed that 1,4-DHNA, 1- and 4-HNA which induced *CYP1A1* and *CYP1B1* also induced AhR transformation. The results obtained for 3,5-DHNA were somewhat

surprising based on the lack of *CYP1A1* induction; however, this compound induced *Cyp1b1* in YAMC cells (Figure 29). The effects of 1- and 2-NOH (100 μM) alone and in combination with TCDD show that both compounds alone induced transformation/DNA binding but also inhibited TCDD-induced transformation/DNA binding to an extent similar to that observed for the well-characterized AhR antagonist CH229131. These data further confirm the partial AhR agonist/antagonist activities observed for 1- and 2-NA for induction of *Cyp1a1* in colon cells.

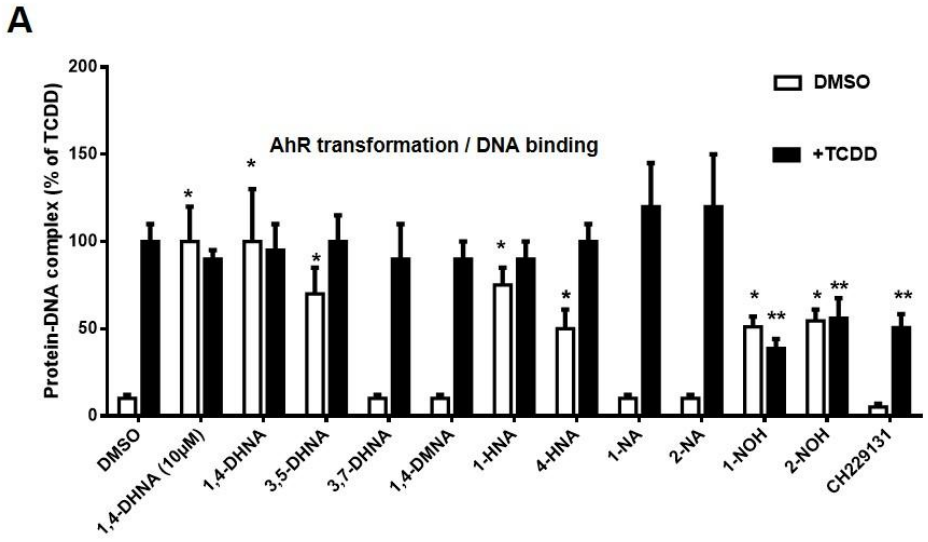


Figure 33. Effects of 1,4-DHNA and related compounds on transformation and DNA binding of guinea pig cytosol and ChIP analysis of the *Cyp1a1* promoter. (A) Transformation and DNA binding of guinea pig hepatic cytosolis AhR. DMSO (solvent control), 10 μM CH229131, 10 and 100 μM 1,4-DHNA, and 100 μM concentrations of the remaining compounds alone and in combination with TCDD were incubated with guinea pig cytosol and analyzed by gel mobility shift assays as outlined in the Materials and Methods. (B) A representative gel showing the ligand-induced transformed AhR-DRE complex is illustrated in this panel. (C) ChIP assay. YAMC and Caco2 cells were treated with various compounds and analysis of recruitment of pol II and the AhR to the *Cyp1a1*/*CYP1A1* promoters was determined in a ChIP assay as outlined in the Materials and Methods.

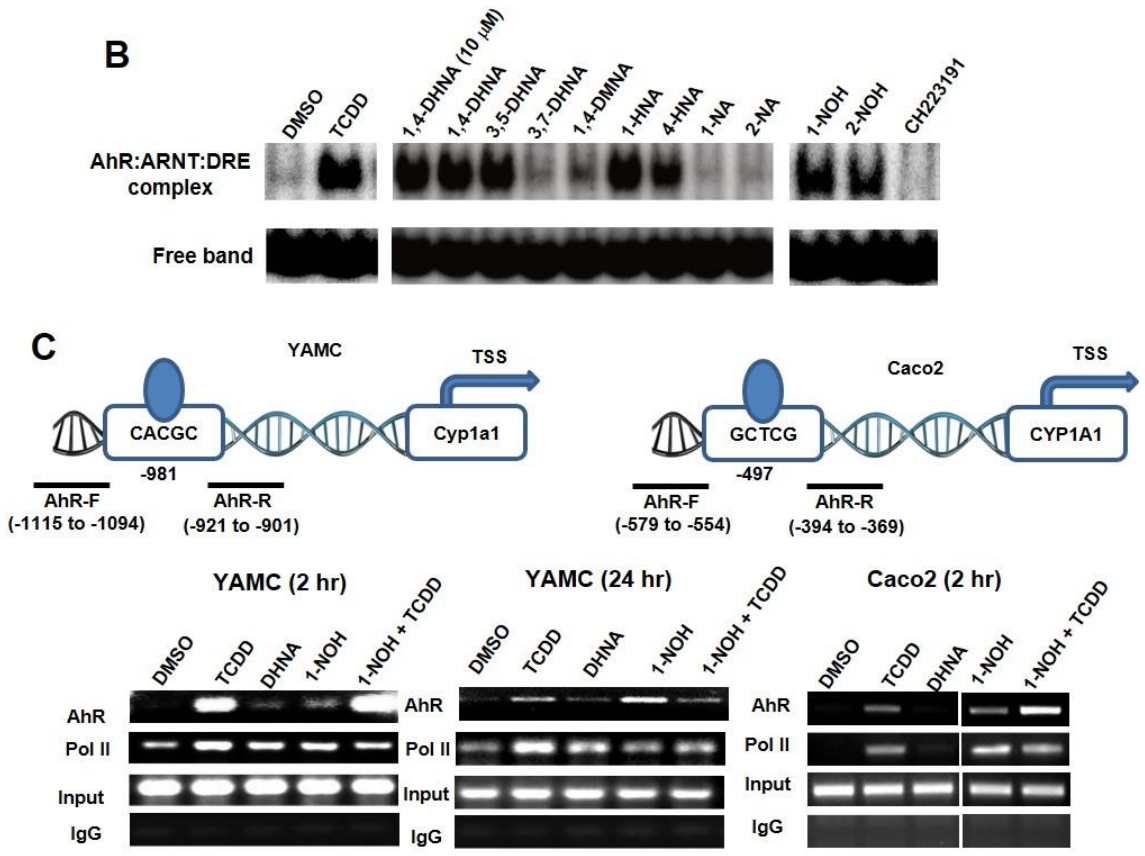


Figure 33. Continued.

We also examined the effects of TCDD, 1,4-DHNA and 1-NOH alone and in combination with TCDD in a ChIP assay. Treatment with TCDD for 2 hr resulted in the recruitment of the AhR and pol II to the DRE region of the *CYP1A1* promoter in Caco2 and YAMC cells (Figure 33C). Similar effects were observed for 1-NA and 1,4-DHNA in YAMC and for 1-NOH (but not 1,4-DHNA) in Caco2 cells. Surprisingly, results observed in cells treated with 1-NOH plus TCDD for 2 hr appeared to be additive with respect to AhR interactions with the *Cyp1a1* promoter. These unexpected ligand-induced effects after treatment for 2 hr were further investigated in YAMC cells after treatment

for 24 hr. TCDD and to a lesser extent 1-NOH recruited the AhR and pol II to the *Cyp1a1* promoter, and in the combination treatment (1-NOH plus TCDD), the 1-NOH compound decreased the effects of TCDD on AhR and pol II recruitment. 1-NOH alone which induced minimal expression of *Cyp1a1* mRNA (Figure 28I) recruited relatively high levels of the AhR to the *Cyp1a1* promoter. Currently, we are further investigating both the time- and compound-dependent recruitment of the AhR, pol II and other nuclear cofactors, including coactivators and corepressors, to the DRE region of the *Cyp1a1* and other Ah-responsive genes.

5.3.4 Modeling of TCDD and 1,4-DHNA interactions with the AhR in order to identifying the most energetically favored binding conformations

We introduced the MM GBSA approximation and identified the TCDD:AhR and 1,4-DHNA:AhR binding modes which acquire the lowest MM GBSA association free energy across all 18 simulated TCDD:AhR and 1,4-DHNA:AhR binding modes, respectively. The average association free energies of the simulated TCDD:AhR and 1,4-DHNA:AhR binding modes are tabulated in Appendix B5 and B6, respectively. The simulations encompassing the most energetically favored binding conformations of TCDD:AhR and 1,4-DHNA:AhR according to MM GBSA were both derived from the docking protocols using quartic potential energy functions and are the most likely to correspond to the naturally occurring binding conformations.

5.3.4.1 Structural stability of binding modes

The stability of lowest association free energy binding modes of TCDD and 1,4-DHNA in complex with AhR was confirmed through RMSD calculations over the 10 ns simulation production runs. The RMSD of the heavy atoms of TCDD in complex with AhR is 1.0 ± 0.4 Å with respect to the average structure of TCDD in complex with AhR, and the RMSD of heavy atoms of 1,4-DHNA in complex with AHR is 0.6 ± 0.3 Å with respect to the average structure of 1,4 in complex with AhR.

5.3.4.2 Interactions between TCDD and AhR

The average per-residue interaction free energy between AhR residues and TCDD are decomposed into polar and non-polar contributions, and selected interactions are presented in Figure 34. Residues contributing the most interaction free energies are presented in Figure 35A. In the TCDD:AhR binding mode, the medial oxygen atom of TCDD forms a hydrogen bond with the NE group of Gln377, indicated with a black dotted line in Figure 35A. As predicted by previous studies (617-619), the TCDD binding pocket of AhR is highly hydrophobic, and the binding of TCDD in AhR is primarily stabilized by non-polar interactions (Figure 35A). Residues Phe289, Cys327, Met342, Ile319, Phe345, and Leu347 form hydrophobic walls around the left side of the ligand, in the perspective of Figure 35A. The aromatic rings residues Phe289, Phe345, and, less frequently, Phe318 form π-π interactions with the aromatic rings of TCDD. Van der Waals interactions are also formed between TCDD and the side-chain atoms of residues Thr283, His285, Pro291, Cys294, Leu302, Leu309, Ile319, Cys327, Ser330,

Met342, Leu347, Ser359, Ala361, and Ala375 as well as the backbone atoms of residues Gly315 and Tyr316 due to their close proximity to the bound TCDD molecule.

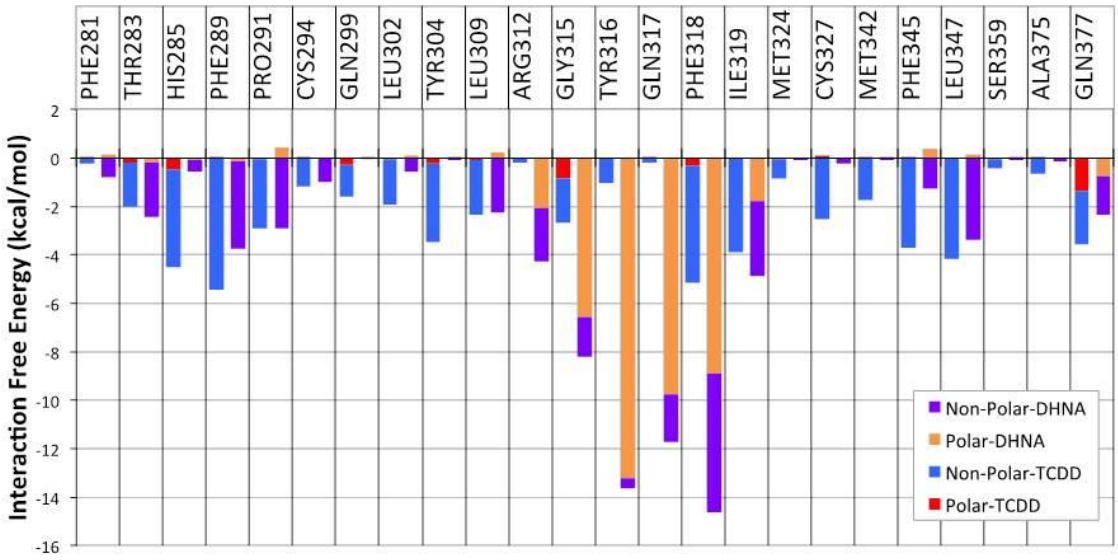


Figure 34. Average interaction free energies (kcal/mol) decomposed into polar (red or orange) and non-polar (blue or purple) contributions for AhR interacting residues in complex with TCDD (first bar per residue) and in complex with DHNA (second bar per residue). The sum of polar and non-polar contributions corresponds to the total average interaction free energy per AHR residue. Only residues with at less than -0.5 kcal/mol average interaction free energy are presented. Results were calculated using the ensemble of snapshots extracted from simulation trajectories of the most energetically favored binding conformations. The average and standard deviation values for the polar and non-polar components of the interaction free energies were calculated over four "measurements", where the first, second, third, and fourth measurement corresponds to the individual average interaction free energy components of the first, second, third, and fourth 2.5 ns segment of the 10 ns MD simulation production run.

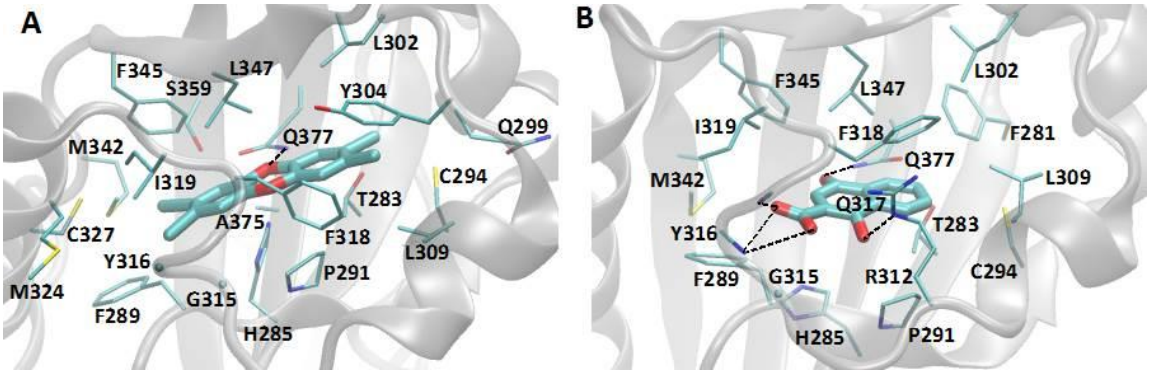


Figure 35. Molecular graphics images of TCDD (panel A) and 1,4-DHNA (panel B) in complex with AhR, which correspond to snapshots extracted from the most energetically favored binding conformations. The ligand molecules are shown in licorice representation in both panels. Interacting AHR protein residues are shown in thin licorice representation, and the entire AHR protein is shown in transparent, gray new cartoon representation. Hydrogen bonds are indicated using black dotted lines. Q377 forms hydrogen bonds with both TCDD and 1,4-DHNA.

5.3.4.3 Interactions between 1,4-DHNA and AhR

As with the TCDD:AhR complex, the average per residue interaction free energy between AhR residues and 1,4-DHNA are decomposed into polar and non-polar contributions, and selected interactions are presented in Figure 34. Residues with the largest associated interaction free energy contributions are presented in Figure 35B. The hydroxyl group of 1,4-DHNA farthest from its carboxylic acid group forms a hydrogen bond with the NE group of Gln377; the hydroxyl group of 1,4-DHNA closest to its carboxylic acid group forms a hydrogen bond with the NE group of Arg312, and the oxygen atoms of the carboxylic acid group of 1,4-DHNA forms hydrogen bonds with the backbone amino groups of Tyr316 and Gln317 as well as a low interacting salt-bridge with the NH group of Arg312. These hydrogen bonds are indicated using black dotted lines in Figure 35B. Hydrophobic residues Phe281, Phe289, Pro291, Cys294, Leu302,

Leu309, Phe318, Ile319, Phe345, and Leu347 predominantly form the 1,4-DHNA binding pocket of AhR (Figure 35B). The aromatic rings of residues Phe318 and, occasionally, Phe289 participate in π-π interactions with the aromatic rings of 1,4-DHNA. Due to their close proximity, strong van der Waals interactions are formed between 1,4-DHNA and the side-chains of residues Phe281, Thr283, Pro291, Cys294, Leu302, Leu309, Ile319, Phe345, and Leu347 as well as the backbone atoms of Gly315 and Tyr316.

5.4 Discussion

TCDD and structurally-related halogenated aromatic compounds have been characterized as widespread and persistent environmental contaminants, and risk assessment of dioxin-like compounds have been developed and are based on well-known structure-activity relationships (562,613). In contrast, SARs for other structural classes of AhR ligands including pharmaceuticals, microbiota-derived compounds such as 1,4-DHNA, endogenous AhR ligands, and food-derived compounds have not been determined and despite some insights on their intake and levels of exposure, it is difficult to predict their potencies and also their interactions with "dioxin-like" compounds.

Among PAHs, naphthalene is not an AhR ligand (639) and therefore the 1- and 4-hydroxyl and carboxylic acid groups are responsible for the activity of 1,4-DHNA which induces near maximal expression (compared to 10 nM TCDD) of *Cyp1a1* and *Cyp1b1* in YAMC and Caco2 cells. The loss of both hydroxyl groups to give 2-NA (or 1-NA) or

replacement with methoxyl substituent (1,4-DMNA) resulted in significant loss of *Cyp1a1*/*Cyp1b1* inducibility in both cell lines. Both 3,5- and 3,7-DHNA exhibited minimal activity as Cyp1a1 inducers in both cell lines; however, in YAMC but not Caco2 cells, these compounds induced *Cyp1b1* expression, indicating that the requirement of the dihydroxy groups was both gene- and cell context-specific to the cell context of YAMC and Caco2 cell lines. Differences in structure-activity relationships for induction of *CYP1B1* may also be species-dependent and due not only to differences between the human and mouse AhR but also in expression of cofactors. Hydroxy naphthoic acids are bacterial metabolites of PAHs, and previous studies on their developmental toxicity in medaka fish embryos show that 1-HNA was not only the most toxic compound but also induced *Cyp1a1* expression in medaka and mouse Hepa1c1c7 liver cancer cell lines (628). We also observed that both 1- and 4-HNA induced *Cyp1a1/Cyp1b1* in YAMC and Caco2 cells and with the exception of relative low induction of Cyp1a1 in YAMC cells, their maximal induction responses were comparable to 1,4-DHNA and TCDD but 5-20 fold less potent than the former compounds. We further confirmed the role of the AhR in mediating compound-induced Cyp1a1 and Cyp1b1 using AhR knockout YAMC cells obtained using the CRISPR/Cas9 technology (Figure 32). It was interesting to note that compounds such as 3,5-DHNA, 3,7-DHNA and 1,4-DMNA that exhibited minimal to non-detectable induction of *Cyp1a1* were effective inducers of *Cyp1b1* (>50% of the response induced by 10 nM TCDD). The differential induction of two AhR-responsive genes in the same cell line is

typical of selective AhR modulators and due, in part, to gene/promoter histone and chromatin differences and is currently being investigated.

1,4-Dihydroxynaphthalene is readily oxidized to the quinone; however, our results show that although this dihydroxy/quinone mixture is cytotoxic, we also observed induction of *CYP1A1* in Caco2 and YAMC cells (Appendix A-6). Thus, the loss of the carboxyl group from 1,4-DHNA to give 1,4-dihydroxynaphthalene decreases but does not abrogate AhR activity of this compound. 1- and 2-NOH are biomarkers of human exposure to PAH and smoking (640-644), and both compounds exhibited weak induction of *Cyp1a1*/*CYP1A1* (YAMC and Caco2) and *CYP1B1* (Caco2) mRNA; however, *Cyp1b1* was induced to >50% of the maximal TCDD-induced response in YAMC cells. Both 1- and 2-NOH inhibited TCDD-induced *Cyp1a1* gene expression in YAMC and Caco2 cells (Figures 32D and 32E), and their AhR antagonist activity was also observed in the gel mobility shift assay (Figure 33A), suggesting that 1- and 2-NOH represent a new class of partial AhR antagonists. We also observed some inconsistencies in the effects of 1-NOH in ChIP assays in YAMC and Caco2 cells (Figure 33C) with respect to recruitment of the AhR to the DRE region of the *Cyp1a1* promoter, and this is currently being investigated.

The most energetically favored binding mode of TCDD in complex with AhR according to our computational studies is in strong accordance with previous experimental and computational studies examining TCDD binding to AhR. Recent experiments suggest

that residues Phe318, Ile319, and Ala375 are key to ligand selectivity (645). Other mutagenesis studies revealed that substitutions in residues Thr283, His285, Phe289, Pro291, Leu302, Leu309, Cys327, Phe345, and Leu347 result in a reduction in TCDD binding (562,617,618). Finally, both computational methods and further mutagenesis studies performed on Gln377 support the contention that the polar side chain atoms of Gln377 form hydrogen bonds with the medial oxygen of TCDD (619). The excellent agreement of our work in comparison to previous studies support the validity of the computational protocol introduced here, and suggest that the in silico identified binding modes most likely represent the naturally occurring binding modes of TCDD and 1,4-DHNA with residues in the AhR ligand binding pocket.

The presence of more polar groups which include the negatively charged carboxylic group in 1,4-DHNA compared to TCDD contributes in general to stronger polar interactions between 1,4-DHNA and AhR compared to TCDD and AhR. While in the TCDD:AhR complex, only one hydrogen bond is formed with the NE group of Gln377 and the medial oxygen of TCDD, 1,4-DHNA forms a hydrogen bond with the NE group of Gln377, the NE group of Arg312, and the backbone amino groups of Tyr316 and Gln317. Additionally, the negatively charged group of 1,4-DHNA forms a low interacting salt-bridge with the NH group of Arg312. The binding of both TCDD and 1,4-DHNA in AhR is also stabilized by non-polar interactions. Both the aromatic rings of TCDD and 1,4-DHNA form π-π interactions with the aromatic rings of Phe289 and Phe318. Both TCDD and 1,4-DHNA also form strong van der Waals interactions with

the side-chains of residues Phe281, Thr283, Pro291, Cys294, Leu302, Leu309, Ile319, Phe345, and Leu347 are critical for TCDD binding (618,646). These similarities are largely due to TCDD and 1,4-DHNA sharing the same binding site and both ligands containing aromatic rings. Owing to the smaller size of DHNA, and thus a lesser amount of possible van der Waals interaction and π-π interaction sites, interactions between 1,4-DHNA and the side-chains of residues His285, Cys327, Ser330, Met342, Ser359, Ala361, and Ala375 are weaker compared to the TCDD:AhR complex. Thus, the modeling studies show that both 1,4-DHNA and TCDD interact within the same binding pocket of the AhR but 1,4-DHNA binds with lower affinity. This observation is consistent with similar efficacies of 1,4-DHNA and TCDD for induction of Cyp1a1 but differences in their potencies. To further understand the differences between both 1,4-DHNA and TCDD in comparison to a minimally active compound, 3,7-DHNA, we performed a preliminary study of the latter in complex with AhR using a similar strategy to the one used in the study of TCDD and 1,4-DHNA, with the only difference that the simulation entailing the binding mode with the lowest association free energy was not extended for an additional 20 ns. In comparison with both 1,4-DHNA and TCDD, 3,7-DHNA forms weaker interactions with Thr283, Pro291, Cys294, Leu302, Leu309, and Phe318. Alanine mutagenesis experiments suggest that Thr283, Leu302, Leu309 and Phe318 are critical to TCDD binding (618,646), and Phe318 has been shown to be an "agonist/antagonist switch"(645).

In summary, this study confirms that 1,4-DHNA is a relatively potent AhR agonist in both YAMC and Caco2 cells; however, the potential impact of endogenous 1,4-DHNA alone or in combination with other AhR agonists/antagonists on gut health is unknown and in the future, we hope to more accurately quantitate these compounds and determine their combined effects. Structure activity studies show that both the hydroxyl and carboxyl groups and their positions on the naphthalene ring are important for AhR activity. Both TCDD and 1,4-DHNA are ligands for the AhR and predominantly form strong interactions with the same AhR residues. More polar interactions occur in the AhR:1,4-DHNA complex in comparison to the AhR:TCDD complex due, in part, to differences in overall charge and substituent interactions with various amino acid side-chains. Our results demonstrate that for some of the hydroxyl naphthoic acid analogs, there are differences in their activation of *Cyp1a1* vs. *Cyp1b1* in the same cell line and also differences in the mouse (YAMC) vs. human (Caco2) colon-derived cells. However, 1,4-DHNA induced >70% of the response observed for 10 nM TCDD for *Cyp1a1* and *Cyp1b1* in both the mouse and human cell lines. This suggests possible efficacy for 1,4-DHNA in humans as an AhR agonist that is protective in the gut; however, this will require more definitive proof from dietary studies and manipulation of gut microorganisms.

6. SUMMARY

Cancer is a severe lethal disease and leading cause of death and numerous studies have been carried out to determine the origin and progression of cancer, in the hope of finding prevention strategies or cures. Although my dissertation consists of two different topics, they all convergent into one final goal, namely, promoting human health and curing cancer.

In the first part, we examined the role of lncRNAs in pancreatic cancer development and proved that lncRNAs, previously considered as the dark matter of the genome, actually play an important role in cancer development. Due to the complicated structure of lncRNAs and differences across species, the study of lncRNAs is challenging. Pancreatic cancer is the fourth leading cause of cancer associated death in the United States with the lowest five-year survival rate of 5% and effective early detection and treatment are essential to treat this lethal disease. Our laboratory was the first to show that HOTTIP, a recently identified novel lncRNA (211), is a pro-oncogenic lncRNA in pancreatic cancer. Our studies showed that decreased HOTTIP expression by RNAi inhibited pancreatic cancer cell growth/migration and induced apoptosis. Chang et al showed that HOTTIP binds to WDR5 and recruited MLL1 and the complex regulated H3K4 trimethylation and the HOXA gene cluster expression (211). We proposed that HOTTIP functions in pancreatic cancer cells via regulating an array of HOXA genes. Several HOXA genes play a role in the pancreatic cancer development; for example, HOXA10

promotes pancreatic cancer cell migration via regulation of TGFβ pathway (647). In the human fibroblast cell, HOTTIP regulates HOXA genes in a distance-dependent manner, and HOXA13 is the most significantly regulated while HOXA1 is minimally affected (211). However, in the Panc1 cells, HOXA13 mRNA levels were not significantly decreased after knocking down of HOTTIP. Another group showed that HOTTIP promotes pancreatic cancer progression, in part via HOTTIP-HOXA13 axis (271), as proposed in other cancer studies (648). They used two different pancreatic cancer cell lines, and this might be the reason for differences between their studies and ours. There is no doubt that HOTTIP plays a role in the progression of different cancers, however, the mechanisms of the role of cell/tissue-context still need to be further investigated.

Another lncRNA we studied is MALAT1. MALAT1 also plays an important role in the pancreatic cancer progression and knockdown of MALAT1 by RNAi decreased cell proliferation, migration/invasion, and induced apoptosis. Several mechanisms have previously been proposed for the function of MALAT1, namely, regulation of alternative splicing (201), recruitment of SF2/ASF (649), interaction with miRNA (519,650), interactions with histone modification complexes (519). In our studies, microarray data indicated the MALAT1 regulates downstream genes via EZH2-dependent and independent pathways. We showed that MALAT1 or EZH2 knockdown increased the expression of NDRG1 and knockdown of NDRG1 partially compensates the effects of MALAT1 knockdown. MALAT1 also regulated other essential genes via EZH2-independent pathways. Apoptotic peptidase activating factor 1 (APAF1) is an important

component of the apoptosome. Decreased expression of MALAT1 induced APAF1 expression and downstream caspase activation and eventually apoptosis. MALAT1 xenograft study showed that decreased MALAT1 expression decreased the tumor size and volume in the xenograft model (262), however, our transgenic mouse study indicated that MALAT1 is dispensable in the KRAS-p53 pancreatic cancer mouse model. The results from our animal studies did not rule out the importance of MALAT1 in the pancreatic cancer development. Since we only harvested the tumor samples immediately before the mice died of cancer, there were no significant differences between the tumor samples in MALAT1 knockout and wide-type MALAT1 in the transgenic mouse model. Harvesting tumor sample at an early stage or use of other transgenic mouse models which developed PanIN instead of full development of PDAC might help clarify the MALAT1 functions in the development of pancreatic cancer, especially at early stages of the diseases.

In the second part of the dissertation, we extensively studied the AhR activity of different microbiota metabolites. The gut microbiome used to be “a forgotten organ” of the human body but it has recently has drawn a lot of attention. The microbiota metabolites are the mediator between human system and gut microbiome. The microbiome and its metabolites are important for human health, however, the underlying molecular mechanisms have only been partially determined. The AhR is an evolutional transcription factor and the initial studies focused on AhR-mediated induction of drug metabolizing enzymes and the metabolism of xenobiotics, particularly the environmental

contaminants, such as PAHs AhR research has been significantly expanded with the discovery of endogenous functions of the AhR. In our projects, we investigated the indole, indole-3-acetate, indole-3-aldehyde, tryptamine of tryptophan metabolites as well as DHNA and related compounds. Tryptophan metabolites are considered to be the largest group of metabolites produced by gut microbiota. Our lab previously tested those typical tryptophan metabolites in human Caco2 cells (651), and late on we tested those metabolites in a normal mouse colonocytes cell lines. The results reflected that tryptophan metabolites are selected AhR modulators, and the AhR induction is cell type and gene specific. Recently, we extensively examined the AhR activity of dihydroxy-2-naphthoic acid (DHNA) and its analogs. DHNA was shown to attenuate colonic inflammation via balancing intestinal bacterial flora and suppressing lymphocyte infiltration (608). Another study in 2013 also demonstrated the anti-inflammatory role of DHNA in colitis and the proposed mechanism is the DHNA effect on intestinal macrophages, leading to the decreased proinflammatory cytokine production. Subsequently, the researcher identified DHNA as an AhR activator. They showed that DHNA activated AhR pathway in Caco2 cell line and also in the mouse intestine (609). Our extensive studies on the AhR activities of different microbiota metabolites facilitate the understanding the molecular mechanisms of those metabolites. In future, we will further dig into the inflammation-related gene regulation by those metabolites via AhR pathway. The proposed in the vivo study would help expand our understanding of the role of AhR ligands in the inflammation process.

In summary, we explored the function of two lncRNAs and their possible mechanisms in pancreatic cancer development. The application of lncRNAs research into a clinical trial is still on its early and premature stage, our study of the molecular mechanism of those lncRNAs would help promote the usage of lncRNAs in future cancer treatment. In addition, we also checked the AhR activity of microbiota metabolites. AhR, as another promising drug targets, has been studied in various cancer and the other diseases. Our hope is to find a connection between the important microbiota metabolites and the AhR pathway. Our study has shown the AhR activity of those microbiota metabolites and our future studies will be focused on the crosstalk between those AhR activity induced by microbiota metabolites and immune response in vivo.

REFERENCES

1. Forman D, Ferlay J. The global and regional burden of cancer. In: Stewart B, Wild C, editors. World cancer report 2014. Lyon (France): International Agency for Research on Cancer; 2014. p 16-53.
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J Clin 2016;66(1):7-30.
3. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell 2000;100(1):57-70.
4. Groner B, Hynes NE. Mutations in human breast cancer cells: dominantly-acting oncogenes and tumor suppressor genes suggest strategies for targeted interference. Int J Cancer Suppl 1990;5:40-6.
5. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell 1990;61(5):759-67.
6. Sledge GW, Jr., Miller KD. Exploiting the hallmarks of cancer: the future conquest of breast cancer. Eur J Cancer 2003;39(12):1668-75.
7. Howe JR, Conlon KC. The molecular genetics of pancreatic cancer. Surg Oncol 1997;6(1):1-18.
8. Sidransky D. Molecular genetics of head and neck cancer. Curr Opin Oncol 1995;7(3):229-33.
9. Greenman C, Stephens P, Smith R, Dalgliesh GL, Hunter C, *et al.* Patterns of somatic mutation in human cancer genomes. Nature 2007;446(7132):153-8.

10. Hanahan D, Weinberg Robert A. Hallmarks of cancer: the next generation. Cell 2011;144(5):646-74.

11. Bhowmick NA, Neilson EG, Moses HL. Stromal fibroblasts in cancer initiation and progression. Nature 2004;432(7015):332-7.

12. Cheng N, Chytil A, Shyr Y, Joly A, Moses HL. Transforming growth factor-beta signaling-deficient fibroblasts enhance hepatocyte growth factor signaling in mammary carcinoma cells to promote scattering and invasion. Mol Cancer Res 2008;6(10):1521-33.

13. Davies MA, Samuels Y. Analysis of the genome to personalize therapy for melanoma. Oncogene 2010;29(41):5545-55.

14. Jiang BH, Liu LZ. PI3K/PTEN signaling in angiogenesis and tumorigenesis. Adv Cancer Res 2009;102:19-65.

15. Kastan MB, Onyekwere O, Sidransky D, Vogelstein B, Craig RW. Participation of p53 protein in the cellular response to DNA damage. Cancer Res 1991;51(23 Pt 1):6304-11.

16. Kuerbitz SJ, Plunkett BS, Walsh WV, Kastan MB. Wild-type p53 is a cell cycle checkpoint determinant following irradiation. Proc Natl Acad Sci U S A 1992;89(16):7491-5.

17. Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, *et al.* Mutational landscape and significance across 12 major cancer types. Nature 2013;502(7471):333-9.

18. Hallstrom TC, Nevins JR. Balancing the decision of cell proliferation and cell fate. Cell Cycle 2009;8(4):532-5.

19. Chau BN, Wang JY. Coordinated regulation of life and death by RB. Nat Rev Cancer 2003;3(2):130-8.

20. Nevins JR. The Rb/E2F pathway and cancer. Hum Mol Genet 2001;10(7):699-703.

21. Junttila MR, Evan GI. p53--a Jack of all trades but master of none. Nat Rev Cancer 2009;9(11):821-9.

22. Adams JM, Cory S. The Bcl-2 apoptotic switch in cancer development and therapy. Oncogene 2007;26(9):1324-37.

23. Lipponen P. Apoptosis in breast cancer: relationship with other pathological parameters. Endocr Relat Cancer 1999;6(1):13-6.

24. Lipponen P, Aaltomaa S, Kosma VM, Syrjanen K. Apoptosis in breast cancer as related to histopathological characteristics and prognosis. Eur J Cancer 1994;30a(14):2068-73.

25. Wang RA, Li ZS, Yan QG, Bian XW, Ding YQ, *et al.* Resistance to apoptosis should not be taken as a hallmark of cancer. Chin J Cancer 2014;33(2):47-50.

26. Mizushima N. Autophagy: process and function. Genes Dev 2007;21(22):2861-73.

27. Levine B, Kroemer G. Autophagy in the pathogenesis of disease. Cell 2008;132(1):27-42.

28. Apel A, Zentgraf H, Buchler MW, Herr I. Autophagy-a double-edged sword in oncology. Int J Cancer 2009;125(5):991-5.

29. White E, DiPaola RS. The double-edged sword of autophagy modulation in cancer. Clin Cancer Res 2009;15(17):5308-16.

30. Yang ZJ, Chee CE, Huang S, Sinicrope FA. The role of autophagy in cancer: therapeutic implications. Mol Cancer Ther 2011;10(9):1533-41.

31. Acosta JC, Gil J. Senescence: a new weapon for cancer therapy. Trends Cell Biol;22(4):211-9.

32. Coppé J-P, Desprez P-Y, Krtolica A, Campisi J. The senescence-associated secretory phenotype: the dark side of tumor suppression. Annu Rev Pathol 2010;5:99-118.

33. Mc Gee MM. Targeting the mitotic catastrophe signaling pathway in cancer. Mediators Inflamm 2015;2015:146282.

34. Kuniyasu H, Chihara Y, Takahashi T. Co-expression of receptor for advanced glycation end products and the ligand amphoterin associates closely with metastasis of colorectal cancer. Oncol Rep 2003;10(2):445-8.

35. Taguchi A, Blood DC, del Toro G, Canet A, Lee DC, *et al.* Blockade of RAGE-amphoterin signalling suppresses tumour growth and metastases. Nature 2000;405(6784):354-60.

36. Kuniyasu H, Chihara Y, Kondo H. Differential effects between amphoterin and advanced glycation end products on colon cancer cells. Int J Cancer 2003;104(6):722-7.

37. Luo Y, Chihara Y, Fujimoto K, Sasahira T, Kuwada M, *et al.* High mobility group box 1 released from necrotic cells enhances regrowth and metastasis of cancer cells that have survived chemotherapy. Eur J Cancer 2013;49(3):741-51.

38. Low KC, Tergaonkar V. Telomerase: central regulator of all of the hallmarks of cancer. Trends Biochem Sci 2013;38(9):426-34.

39. Hanahan D, Folkman J. Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis. Cell 1996;86(3):353-64.

40. Bergers G, Benjamin LE. Tumorigenesis and the angiogenic switch. Nat Rev Cancer 2003;3(6):401-10.

41. Ren B, Yee KO, Lawler J, Khosravi-Far R. Regulation of tumor angiogenesis by thrombospondin-1. Biochim Biophys Acta 2006;1765(2):178-88.

42. Wang Y, Zhou BP. Epithelial-mesenchymal transition---a hallmark of breast cancer metastasis. Cancer Hallm 2013;1(1):38-49.

43. Du B, Shim JS. Targeting epithelial-mesenchymal transition (EMT) to overcome drug resistance in cancer. Molecules 2016;21(7):pii: E965.

44. Fischer KR, Durrans A, Lee S, Sheng J, Li F, *et al.* Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. Nature 2015;527(7579):472-6.

45. Zheng X, Carstens JL, Kim J, Scheible M, Kaye J, *et al.* Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. Nature 2015;527(7579):525-30.

46. Tsai JH, Donaher JL, Murphy DA, Chau S, Yang J. Spatiotemporal regulation of epithelial-mesenchymal transition is essential for squamous cell carcinoma metastasis. Cancer Cell 2012;22(6):725-36.

47. Ocana OH, Corcoles R, Fabra A, Moreno-Bueno G, Acloque H, *et al.* Metastatic colonization requires the repression of the epithelial-mesenchymal transition inducer Prrx1. Cancer Cell 2012;22(6):709-24.

48. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. Nat Rev Mol Cell Biol 2014;15(3):178-96.

49. Mi Z, Bhattacharya SD, Kim VM, Guo H, Talbot LJ, *et al.* Osteopontin promotes CCL5-mesenchymal stromal cell-mediated breast cancer metastasis. Carcinogenesis 2011;32(4):477-87.

50. Borczuk AC, Papanikolaou N, Toonkel RL, Sole M, Gorenstein LA, *et al.* Lung adenocarcinoma invasion in TGFbetaRII-deficient cells is mediated by CCL5/RANTES. Oncogene 2008;27(4):557-64.

51. Hsu PP, Sabatini DM. Cancer cell metabolism: Warburg and beyond. Cell 2008;134(5):703-7.

52. DeBerardinis RJ, Lum JJ, Hatzivassiliou G, Thompson CB. The biology of cancer: metabolic reprogramming fuels cell growth and proliferation. Cell Metab 2008;7(1):11-20.

53. Jones RG, Thompson CB. Tumor suppressors and cell metabolism: a recipe for cancer growth. Genes Dev 2009;23(5):537-48.

54. Pavlova Natalya N, Thompson Craig B. The emerging hallmarks of cancer metabolism. Cell Metab 2016;23(1):27-47.

55. Kim R, Emi M, Tanabe K. Cancer immunoediting from immune surveillance to immune escape. Immunology 2007;121(1):1-14.

56. Smyth MJ, Dunn GP, Schreiber RD. Cancer immunosurveillance and immunoediting: the roles of immunity in suppressing tumor development and shaping tumor immunogenicity. Adv Immunol 2006;90:1-50.

57. Pages F, Galon J, Dieu-Nosjean MC, Tartour E, Sautes-Fridman C, *et al.* Immune infiltration in human tumors: a prognostic factor that should not be ignored. Oncogene 2010;29(8):1093-102.

58. Macheret M, Halazonetis TD. DNA replication stress as a hallmark of cancer. Annu Rev Pathol 2015;10:425-48.

59. Grivennikov SI, Greten FR, Karin M. Immunity, inflammation, and cancer. Cell 2010;140(6):883-99.

60. Tas SW, Maracle CX, Balogh E, Szekanecz Z. Targeting of proangiogenic signalling pathways in chronic inflammation. Nat Rev Rheumatol 2016;12(2):111-22.

61. Shimizu T, Marusawa H, Endo Y, Chiba T. Inflammation-mediated genomic instability: roles of activation-induced cytidine deaminase in carcinogenesis. Cancer Sci 2012;103(7):1201-6.

62. Yeh AC, Ramaswamy S. Mechanisms of cancer cell dormancy--another hallmark of cancer? Cancer Res 2015;75(23):5014-22.

63. Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet 2008;40(12):1413-5.

64. Aoki K, Taketo MM. Adenomatous polyposis coli (APC): a multi-functional tumor suppressor gene. J Cell Sci 2007;120(Pt 19):3327-35.

65. De Rosa M, Morelli G, Cesaro E, Duraturo F, Turano M, *et al.* Alternative splicing and nonsense-mediated mRNA decay in the regulation of a new adenomatous polyposis coli transcript. Gene 2007;395(1-2):8-14.

66. Floor SL, Dumont JE, Maenhaut C, Raspe E. Hallmarks of cancer: of all cancer cells, all the time? Trends Mol Med 2012;18(9):509-15.

67. Baker SG. A cancer theory kerfuffle can lead to new lines of research. J Natl Cancer Inst 2015;107(2):pii: dju405.

68. Soto AM, Sonnenschein C. The tissue organization field theory of cancer: a testable replacement for the somatic mutation theory. Bioessays 2011;33(5):332-40.

69. Sonnenschein C, Soto AM, Rangarajan A, Kulkarni P. Competing views on cancer. J Biosci 2014;39(2):281-302.

70. Rosenfeld S. Are the somatic mutation and tissue organization field theories of carcinogenesis incompatible? Cancer Inform 2013;12:221-9.

71. Bedessem B, Ruphy S. SMT or TOFT? How the two main theories of carcinogenesis are made (artificially) incompatible. Acta Biotheor 2015;63(3):257-67.

72. Bizzarri M, Cucina A. SMT and TOFT: why and how they are opposite and incompatible paradigms. Acta Biotheor 2016;64(3):221-39.

73. Bedessem B, Ruphy S. SMT and TOFT integrable after all: a reply to Bizzarri and Cucina. Acta Biotheoretica 2016:1-5.

74. Hruban RH. Pancreatic cancer. In: Stewart B, Wild C, editors. World cancer report 2014. Lyon (France): International Agency for Reaserch on Cancer; 2014. p 413-21.

75. Longnecker D. Anatomy and histology of the pancreas. <http://www.pancreapedia.org/?q=node/8098>. Accessed 2016 October 23.

76. Hruban RH, Adsay NV, Albores-Saavedra J, Compton C, Garrett ES, *et al.* Pancreatic intraepithelial neoplasia: a new nomenclature and classification system for pancreatic duct lesions. Am J Surg Pathol 2001;25(5):579-86.

77. Hruban RH, Takaori K, Klimstra DS, Adsay NV, Albores-Saavedra J, *et al.* An illustrated consensus on the classification of pancreatic intraepithelial neoplasia and intraductal papillary mucinous neoplasms. Am J Surg Pathol 2004;28(8):977-87.

78. Hezel AF, Kimmelman AC, Stanger BZ, Bardeesy N, Depinho RA. Genetics and biology of pancreatic ductal adenocarcinoma. Genes Dev 2006;20(10):1218-49.

79. Smit VT, Boot AJ, Smits AM, Fleuren GJ, Cornelisse CJ, *et al.* KRAS codon 12 mutations occur very frequently in pancreatic adenocarcinomas. Nucleic Acids Res 1988;16(16):7773-82.

80. Almoguera C, Shibata D, Forrester K, Martin J, Arnheim N, *et al.* Most human carcinomas of the exocrine pancreas contain mutant c-K-ras genes. Cell 1988;53(4):549-54.

81. Lohr M, Kloppel G, Maisonneuve P, Lowenfels AB, Luttges J. Frequency of K-ras mutations in pancreatic intraductal neoplasias associated with pancreatic ductal adenocarcinoma and chronic pancreatitis: a meta-analysis. Neoplasia 2005;7(1):17-23.

82. Moskaluk CA, Hruban RH, Kern SE. p16 and K-ras gene mutations in the intraductal precursors of human pancreatic adenocarcinoma. Cancer Res 1997;57(11):2140-3.

83. Ying H, Kimmelman AC, Lyssiotis CA, Hua S, Chu GC, *et al.* Oncogenic Kras maintains pancreatic tumors through regulation of anabolic glucose metabolism. Cell 2012;149(3):656-70.

84. Collins MA, Bednar F, Zhang Y, Brisset JC, Galban S, *et al.* Oncogenic Kras is required for both the initiation and maintenance of pancreatic cancer in mice. J Clin Invest 2012;122(2):639-53.

85. Schutte M, Hruban RH, Geradts J, Maynard R, Hilgers W, *et al.* Abrogation of the Rb/p16 tumor-suppressive pathway in virtually all pancreatic carcinomas. Cancer Res 1997;57(15):3126-30.

86. Caldas C, Hahn SA, da Costa LT, Redston MS, Schutte M, *et al.* Frequent somatic mutations and homozygous deletions of the p16 (MTS1) gene in pancreatic adenocarcinoma. Nat Genet 1994;8(1):27-32.

87. Lukas J, Parry D, Aagaard L, Mann DJ, Bartkova J, *et al.* Retinoblastoma-protein-dependent cell-cycle inhibition by the tumour suppressor p16. Nature 1995;375(6531):503-6.

88. Maitra A, Adsay NV, Argani P, Iacobuzio-Donahue C, De Marzo A, *et al.* Multicomponent analysis of the pancreatic adenocarcinoma progression model using a pancreatic intraepithelial neoplasia tissue microarray. Mod Pathol 2003;16(9):902-12.

89. Hahn SA, Schutte M, Hoque AT, Moskaluk CA, da Costa LT, *et al.* DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. Science 1996;271(5247):350-3.

90. Ottenhof NA, Morsink FH, Ten Kate F, van Noorden CJ, Offerhaus GJ. Multivariate analysis of immunohistochemical evaluation of protein expression in pancreatic ductal adenocarcinoma reveals prognostic significance for persistent Smad4 expression only. Cell Oncol (Dordr) 2012;35(2):119-26.

91. Oshima M, Okano K, Muraki S, Haba R, Maeba T, *et al.* Immunohistochemically detected expression of 3 major genes (CDKN2A/p16, TP53, and SMAD4/DPC4) strongly predicts survival in patients with resectable pancreatic cancer. Ann Surg 2013;258(2):336-46.

92. Izeradjene K, Combs C, Best M, Gopinathan A, Wagner A, *et al.* Kras(G12D) and Smad4/Dpc4 haploinsufficiency cooperate to induce mucinous cystic neoplasms and invasive adenocarcinoma of the pancreas. Cancer Cell 2007;11(3):229-43.

93. American Cancer Society. Pancreatic cancer stages. <http://www.cancer.org/cancer/pancreaticcancer/detailedguide/pancreatic-cancer-staging>. Accessed 2016 October 23.

94. Deer EL, Gonzalez-Hernandez J, Coursen JD, Shea JE, Ngatia J, *et al.* Phenotype and genotype of pancreatic cancer cell lines. Pancreas 2010;39(4):425-35.

95. Arumugam T, Simeone DM, Van Golen K, Logsdon CD. S100P promotes pancreatic cancer growth, survival, and invasion. Clin Cancer Res 2005;11(15):5356-64.

96. Marchesi F, Monti P, Leone BE, Zerbi A, Vecchi A, *et al.* Increased survival, proliferation, and migration in metastatic human pancreatic tumor cells expressing functional CXCR4. Cancer Res 2004;64(22):8420-7.

97. Arumugam T, Ramachandran V, Fournier KF, Wang H, Marquis L, *et al.* Epithelial to mesenchymal transition contributes to drug resistance in pancreatic cancer. Cancer Res 2009;69(14):5820-8.

98. Pan C, Kumar C, Bohl S, Klingmueller U, Mann M. Comparative proteomic phenotyping of cell lines and primary cells to assess preservation of cell type-specific functions. Mol Cell Proteomics 2009;8(3):443-50.

99. Alge CS, Hauck SM, Priglinger SG, Kampik A, Ueffing M. Differential protein profiling of primary versus immortalized human RPE cells identifies expression patterns associated with cytoskeletal remodeling and cell survival. J Proteome Res 2006;5(4):862-78.

100. Boj SF, Hwang CI, Baker LA, Chio, II, Engle DD, *et al.* Organoid models of human and mouse ductal pancreatic cancer. Cell 2015;160(1-2):324-38.

101. Qiu W, Su GH. Challenges and advances in mouse modeling for human pancreatic tumorigenesis and metastasis. Cancer Metastasis Rev 2013;32(1):83-107.

102. Giovanella BC, Stehlin JS, Wall ME, Wani MC, Nicholas AW, *et al.* DNA topoisomerase I--targeted chemotherapy of human colon cancer in xenografts. Science 1989;246(4933):1046-8.

103. Pantazis P, Kozielski AJ, Mendoza JT, Early JA, Hinz HR, *et al.* Camptothecin derivatives induce regression of human ovarian carcinomas grown in nude mice and distinguish between non-tumorigenic and tumorigenic cells in vitro. Int J Cancer 1993;53(5):863-71.

104. Pantazis P, Hinz HR, Mendoza JT, Kozielski AJ, Williams LJ, Jr., *et al.* Complete inhibition of growth followed by death of human malignant melanoma cells in vitro and regression of human melanoma xenografts in immunodeficient mice induced by camptothecins. Cancer Res 1992;52(14):3980-7.

105. Takimoto CH. Why drugs fail: of mice and men revisited. Clin Cancer Res 2001;7(2):229-30.

106. Feig C, Gopinathan A, Neesse A, Chan DS, Cook N, *et al.* The pancreas cancer microenvironment. Clin Cancer Res 2012;18(16):4266-76.

107. Cheon DJ, Orsulic S. Mouse models of cancer. Annu Rev Pathol 2011;6:95-119.

108. Stewart TA, Pattengale PK, Leder P. Spontaneous mammary adenocarcinomas in transgenic mice that carry and express MTV/myc fusion genes. Cell 1984;38(3):627-37.

109. Brinster RL, Chen HY, Messing A, van Dyke T, Levine AJ, *et al.* Transgenic mice harboring SV40 T-antigen genes develop characteristic brain tumors. Cell 1984;37(2):367-79.

110. Moser AR, Luongo C, Gould KA, McNeley MK, Shoemaker AR, *et al.* ApcMin: a mouse model for intestinal and mammary tumorigenesis. Eur J Cancer 1995;31a(7-8):1061-4.

111. Greenberg NM, DeMayo FJ, Sheppard PC, Barrios R, Lebovitz R, *et al.* The rat probasin gene promoter directs hormonally and developmentally regulated expression of a heterologous gene specifically to the prostate in transgenic mice. Mol Endocrinol 1994;8(2):230-9.

112. Greenberg NM, DeMayo F, Finegold MJ, Medina D, Tilley WD, *et al.* Prostate cancer in a transgenic mouse. Proc Natl Acad Sci U S A 1995;92(8):3439-43.

113. Xu X, Wagner KU, Larson D, Weaver Z, Li C, *et al.* Conditional mutation of Brca1 in mammary epithelial cells results in blunted ductal morphogenesis and tumour formation. Nat Genet 1999;22(1):37-43.

114. Jackson EL, Willis N, Mercer K, Bronson RT, Crowley D, *et al.* Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. Genes Dev 2001;15(24):3243-8.

115. Jaffee EM, Hruban RH, Canto M, Kern SE. Focus on pancreas cancer. Cancer Cell 2002;2(1):25-8.

116. Hingorani SR, Petricoin EF, Maitra A, Rajapakse V, King C, *et al.* Preinvasive and invasive ductal pancreatic cancer and its early detection in the mouse. Cancer Cell 2003;4(6):437-50.

117. Offield MF, Jetton TL, Labosky PA, Ray M, Stein RW, *et al.* PDX-1 is required for pancreatic outgrowth and differentiation of the rostral duodenum. Development 1996;122(3):983-95.

118. Kawaguchi Y, Cooper B, Gannon M, Ray M, MacDonald RJ, *et al.* The role of the transcriptional regulator Ptf1a in converting intestinal to pancreatic progenitors. Nat Genet 2002;32(1):128-34.

119. Kenzelmann Broz D, Attardi LD. In vivo analysis of p53 tumor suppressor function using genetically engineered mouse models. Carcinogenesis 2010;31(8):1311-8.

120. Lang GA, Iwakuma T, Suh YA, Liu G, Rao VA, *et al.* Gain of function of a p53 hot spot mutation in a mouse model of Li-Fraumeni syndrome. Cell 2004;119(6):861-72.

121. Olive KP, Tuveson DA, Ruhe ZC, Yin B, Willis NA, *et al.* Mutant p53 gain of function in two mouse models of Li-Fraumeni syndrome. Cell 2004;119(6):847-60.

122. Hingorani SR, Wang L, Multani AS, Combs C, Deramaudt TB, *et al.* Trp53R172H and KrasG12D cooperate to promote chromosomal instability and

widely metastatic pancreatic ductal adenocarcinoma in mice. Cancer Cell 2005;7(5):469-83.

123. Majumder K, Arora N, Modi S, Chugh R, Nomura A, *et al.* A novel immunocompetent mouse model of pancreatic cancer with robust stroma: a valuable tool for preclinical evaluation of new therapies. J Gastrointest Surg 2016;20(1):53-65; discussion

124. De Latouliere L, Manni I, Iacobini C, Pugliese G, Grazi GL, *et al.* A bioluminescent mouse model of proliferation to highlight early stages of pancreatic cancer: a suitable tool for preclinical studies. Ann Anat 2016;207:2-8.

125. Aguirre AJ, Bardeesy N, Sinha M, Lopez L, Tuveson DA, *et al.* Activated Kras and Ink4a/Arf deficiency cooperate to produce metastatic pancreatic ductal adenocarcinoma. Genes Dev 2003;17(24):3112-26.

126. Bardeesy N, Aguirre AJ, Chu GC, Cheng KH, Lopez LV, *et al.* Both p16(Ink4a) and the p19(Arf)-p53 pathway constrain progression of pancreatic adenocarcinoma in the mouse. Proc Natl Acad Sci U S A 2006;103(15):5947-52.

127. Bardeesy N, Cheng KH, Berger JH, Chu GC, Pahler J, *et al.* Smad4 is dispensable for normal pancreas development yet critical in progression and tumor biology of pancreas cancer. Genes Dev 2006;20(22):3130-46.

128. Kojima K, Vickers SM, Adsay NV, Jhala NC, Kim HG, *et al.* Inactivation of Smad4 accelerates Kras(G12D)-mediated pancreatic neoplasia. Cancer Res 2007;67(17):8121-30.

129. Ijichi H, Chytil A, Gorska AE, Aakre ME, Fujitani Y, *et al.* Aggressive pancreatic ductal adenocarcinoma in mice caused by pancreas-specific blockade of transforming growth factor-beta signaling in cooperation with active Kras expression. Genes Dev 2006;20(22):3147-60.

130. Guerra C, Schuhmacher AJ, Canamero M, Grippo PJ, Verdaguer L, *et al.* Chronic pancreatitis is essential for induction of pancreatic ductal adenocarcinoma by K-Ras oncogenes in adult mice. Cancer Cell 2007;11(3):291-302.

131. Siveke JT, Einwachter H, Sipos B, Lubeseder-Martellato C, Kloppel G, *et al.* Concomitant pancreatic activation of Kras(G12D) and Tgfa results in cystic papillary neoplasms reminiscent of human IPMN. Cancer Cell 2007;12(3):266-79.

132. Skoulidis F, Cassidy LD, Pisupati V, Jonasson JG, Bjarnason H, *et al.* Germline Brca2 heterozygosity promotes Kras(G12D) -driven carcinogenesis in a murine model of familial pancreatic cancer. Cancer Cell 2010;18(5):499-509.

133. Rowley M, Ohashi A, Mondal G, Mills L, Yang L, *et al.* Inactivation of Brca2 promotes Trp53-associated but inhibits KrasG12D-dependent pancreatic cancer development in mice. Gastroenterology 2011;140(4):1303-13 e1-3.

134. Morton JP, Jamieson NB, Karim SA, Athineos D, Ridgway RA, *et al.* LKB1 haploinsufficiency cooperates with Kras to promote pancreatic cancer through suppression of p21-dependent growth arrest. Gastroenterology 2010;139(2):586-97, 97 e1-6.

135. Mazur PK, Einwachter H, Lee M, Sipos B, Nakhai H, *et al.* Notch2 is required for progression of pancreatic intraepithelial neoplasia and development of pancreatic ductal adenocarcinoma. Proc Natl Acad Sci U S A 2010;107(30):13438-43.

136. Guerra C, Collado M, Navas C, Schuhmacher AJ, Hernandez-Porras I, *et al.* Pancreatitis-induced inflammation contributes to pancreatic cancer by inhibiting oncogene-induced senescence. Cancer Cell 2011;19(6):728-39.

137. Perez-Mancera PA, Rust AG, van der Weyden L, Kristiansen G, Li A, *et al.* The deubiquitinase USP9X suppresses pancreatic ductal adenocarcinoma. Nature 2012;486(7402):266-70.

138. Shakya R, Reid LJ, Reczek CR, Cole F, Egli D, *et al.* BRCA1 tumor suppression depends on BRCT phosphoprotein binding, but not its E3 ligase activity. Science 2011;334(6055):525-8.

139. Carriere C, Gore AJ, Norris AM, Gunn JR, Young AL, *et al.* Deletion of Rb accelerates pancreatic carcinogenesis by oncogenic Kras and impairs senescence in premalignant lesions. Gastroenterology 2011;141(3):1091-101.

140. Tinder TL, Subramani DB, Basu GD, Bradley JM, Schettini J, *et al.* MUC1 enhances tumor progression and contributes toward immunosuppression in a mouse model of spontaneous pancreatic adenocarcinoma. J Immunol 2008;181(5):3116-25.

141. Marengo E, Robotti E. Biomarkers for pancreatic cancer: recent achievements in proteomics and genomics through classical and multivariate statistical methods. World J Gastroenterol 2014;20(37):13325-42.

142. Chang M-C, Wong J-M, Chang Y-T. Screening and early detection of pancreatic cancer in high risk population. World J Gastroenterol 2014;20(9):2358-64.

143. Distler M, Pilarsky E, Kersting S, Grutzmann R. Preoperative CEA and CA 19-9 are prognostic markers for survival after curative resection for ductal adenocarcinoma of the pancreas - a retrospective tumor marker prognostic study. Int J Surg 2013;11(10):1067-72.

144. Piagnerelli R, Marrelli D, Roviello G, Ferrara F, Di Mare G, *et al.* Clinical value and impact on prognosis of peri-operative CA 19-9 serum levels in stage I and II adenocarcinoma of the pancreas. Tumour Biol 2016;37(2):1959-66.

145. Poruk KE, Gay DZ, Brown K, Mulvihill JD, Boucher KM, *et al.* The clinical utility of CA 19-9 in pancreatic adenocarcinoma: diagnostic and prognostic updates. Curr Mol Med 2013;13(3):340-51.

146. Wingren C, Sandstrom A, Segersvard R, Carlsson A, Andersson R, *et al.* Identification of serum biomarker signatures associated with pancreatic cancer. Cancer Res 2012;72(10):2481-90.

147. Shaw VE, Lane B, Jenkinson C, Cox T, Greenhalf W, *et al.* Serum cytokine biomarker panels for discriminating pancreatic cancer from benign pancreatic disease. Mol Cancer 2014;13:114.

148. Yu X, Koenig MR, Zhu Y. Plasma miRNA, an emerging biomarker for pancreatic cancer. Ann Transl Med 2015;3(19):297.

149. Peng W, Jiang A. Long noncoding RNA CCDC26 as a potential predictor biomarker contributes to tumorigenesis in pancreatic cancer. Biomed Pharmacother 2016;83:712-7.

150. Xie Z, Yin X, Gong B, Nie W, Wu B, *et al.* Salivary microRNAs show potential as a noninvasive biomarker for detecting resectable pancreatic cancer. Cancer Prev Res (Phila) 2015;8(2):165-73.

151. Wang Y, Li Z, Zheng S, Zhou Y, Zhao L, *et al.* Expression profile of long non-coding RNAs in pancreatic cancer and their clinical significance as biomarkers. Oncotarget 2015;6(34):35684-98.

152. Pantel K, Brakenhoff RH. Dissecting the metastatic cascade. Nat Rev Cancer 2004;4(6):448-56.

153. Polyak K, Weinberg RA. Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. Nat Rev Cancer 2009;9(4):265-73.

154. Mataki Y, Takao S, Maemura K, Mori S, Shinchi H, *et al.* Carcinoembryonic antigen messenger RNA expression using nested reverse transcription-PCR in the peripheral blood during follow-up period of patients who underwent curative surgery for biliary-pancreatic cancer: longitudinal analyses. Clin Cancer Res 2004;10(11):3807-14.

155. Kurihara T, Itoi T, Sofuni A, Itokawa F, Tsuchiya T, *et al.* Detection of circulating tumor cells in patients with pancreatic cancer: a preliminary result. J Hepatobiliary Pancreat Surg 2008;15(2):189-95.

156. Laird PW. The power and the promise of DNA methylation markers. Nat Rev Cancer 2003;3(4):253-66.

157. Sidransky D. Emerging molecular markers of cancer. Nat Rev Cancer 2002;2(3):210-9.

158. Kisiel JB, Raimondo M, Taylor WR, Yab TC, Mahoney DW, *et al.* New DNA methylation markers for pancreatic cancer: discovery, tissue validation, and pilot testing in pancreatic juice. Clin Cancer Res 2015;21(19):4473-81.

159. Kahlert C, Melo SA, Protopopov A, Tang J, Seth S, *et al.* Identification of double-stranded genomic DNA spanning all chromosomes with mutated KRAS and p53 DNA in the serum exosomes of patients with pancreatic cancer. J Biol Chem 2014;289(7):3869-75.

160. Su G, Meyer K, Nandini CD, Qiao D, Salamat S, *et al.* Glypican-1 is frequently overexpressed in human gliomas and enhances FGF-2 signaling in glioma cells. Am J Pathol 2006;168(6):2014-26.

161. Matsuda K, Maruyama H, Guo F, Kleeff J, Itakura J, *et al.* Glypican-1 is overexpressed in human breast cancer and modulates the mitogenic effects of multiple heparin-binding growth factors in breast cancer cells. Cancer Res 2001;61(14):5562-9.

162. Melo SA, Luecke LB, Kahlert C, Fernandez AF, Gammon ST, *et al.* Glypican-1 identifies cancer exosomes and detects early pancreatic cancer. Nature 2015;523(7559):177-82.

163. Honda K, Yamada T, Endo R, Ino Y, Gotoh M, *et al.* Actinin-4, a novel actin-bundling protein associated with cell motility and cancer invasion. J Cell Biol 1998;140(6):1383-93.

164. Kikuchi S, Honda K, Tsuda H, Hiraoka N, Imoto I, *et al.* Expression and gene amplification of actinin-4 in invasive ductal carcinoma of the pancreas. Clin Cancer Res 2008;14(17):5348-56.

165. Watanabe T, Ueno H, Watabe Y, Hiraoka N, Morizane C, *et al.* ACTN4 copy number increase as a predictive biomarker for chemoradiotherapy of locally advanced pancreatic cancer. Br J Cancer 2015;112(4):704-13.

166. Pant S, Martin LK, Geyer S, Wei L, Van Loon K, *et al.* Baseline serum albumin is a predictive biomarker for patients with advanced pancreatic cancer treated with bevacizumab: a pooled analysis of 7 prospective trials of gemcitabine-based therapy with or without bevacizumab. Cancer 2014;120(12):1780-6.

167. Thomas G, Chardes T, Gaborit N, Mollevi C, Leconet W, *et al.* HER3 as biomarker and therapeutic target in pancreatic cancer: new insights in pertuzumab therapy in preclinical models. Oncotarget 2014;5(16):7138-48.

168. Ritchie SA, Akita H, Takemasa I, Eguchi H, Pastural E, *et al.* Metabolic system alterations in pancreatic cancer patient serum: potential for early detection. BMC Cancer 2013;13:416.

169. Ritchie SA, Chitou B, Zheng Q, Jayasinghe D, Jin W, *et al.* Pancreatic cancer serum biomarker PC-594: diagnostic performance and comparison to CA19-9. World J Gastroenterol 2015;21(21):6604-12.

170. Kuhlmann KF, van Till JW, Boermeester MA, de Reuver PR, Tzvetanova ID, *et al.* Evaluation of matrix metalloproteinase 7 in plasma and pancreatic juice as a biomarker for pancreatic cancer. Cancer Epidemiol Biomarkers Prev 2007;16(5):886-91.

171. Chang DK, Johns AL, Merrett ND, Gill AJ, Colvin EK, *et al.* Margin clearance and outcome in resected pancreatic cancer. J Clin Oncol 2009;27(17):2855-62.

172. Strasberg SM, Drebin JA, Linehan D. Radical antegrade modular pancreatosplenectomy. Surgery 2003;133(5):521-7.

173. Zuckerman DS, Ryan DP. Adjuvant therapy for pancreatic cancer: a review. Cancer 2008;112(2):243-9.

174. Oettle H, Neuhaus P, Hochhaus A, Hartmann JT, Gellert K, *et al.* Adjuvant chemotherapy with gemcitabine and long-term outcomes among patients with resected pancreatic cancer: the CONKO-001 randomized trial. JAMA 2013;310(14):1473-81.

175. Takahashi Y, Kitadai Y, Bucana CD, Cleary KR, Ellis LM. Expression of vascular endothelial growth factor and its receptor, KDR, correlates with vascularity, metastasis, and proliferation of human colon cancer. Cancer Res 1995;55(18):3964-8.

176. Ferrara N. Vascular endothelial growth factor: basic science and clinical progress. Endocr Rev 2004;25(4):581-611.

177. Van Cutsem E, Vervenne WL, Bennouna J, Humblet Y, Gill S, *et al.* Phase III trial of bevacizumab in combination with gemcitabine and erlotinib in patients with metastatic pancreatic cancer. J Clin Oncol 2009;27(13):2231-7.

178. Ciliberto D, Staropoli N, Chiellino S, Botta C, Tassone P, *et al.* Systematic review and meta-analysis on targeted therapy in advanced pancreatic cancer. Pancreatology 2016;16(2):249-58.

179. Crick F. Central dogma of molecular biology. Nature 1970;227(5258):561-3.

180. Stein LD. Human genome: end of the beginning. Nature 2004;431(7011):915-6.

181. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, *et al.* Initial sequencing and analysis of the human genome. Nature 2001;409(6822):860-921.

182. Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR, *et al.* Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 2007;447(7146):799-816.

183. Mattick JS. Non-coding RNAs: the architects of eukaryotic complexity. EMBO Rep 2001;2(11):986-91.

184. Antharam VC, McEwen DC, Garrett TJ, Dossey AT, Li EC, *et al.* An integrated metabolomic and microbiome analysis identified specific gut Microbiota associated with fecal cholesterol and coprostanol in clostridium difficile infection. PLoS One 2016;11(2):e0148824.

185. Da Sacco L, Baldassarre A, Masotti A. Bioinformatics tools and novel challenges in long non-coding RNAs (lncRNAs) functional analysis. Int J Mol Sci 2012;13(1):97-114.

186. Al-Shehri SS, Knox CL, Liley HG, Cowley DM, Wright JR, *et al.* Breastmilk-saliva interactions boost innate immunity by regulating the oral microbiome in early infancy. PLoS One 2015;10(9):e0135047.

187. Alolga RN, Fan Y, Chen Z, Liu LW, Zhao YJ, *et al.* Significant pharmacokinetic differences of berberine are attributable to variations in gut microbiota between Africans and Chinese. Sci Rep 2016;6:27671.

188. Al-Lahham SH, Peppelenbosch MP, Roelofsen H, Vonk RJ, Venema K. Biological effects of propionic acid in humans; metabolism, potential applications and underlying mechanisms. Biochim Biophys Acta 2010;1801(11):1175-83.

189. Lipovich L, Johnson R, Lin CY. MacroRNA underdogs in a microRNA world: evolutionary, regulatory, and biomedical significance of mammalian long non-protein-coding RNA. Biochim Biophys Acta 2010;1799(9):597-615.

190. Kapranov P, Cheng J, Dike S, Nix DA, Duttagupta R, *et al.* RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science 2007;316(5830):1484-8.

191. Ma L, Bajic VB, Zhang Z. On the classification of long non-coding RNAs. RNA Biol 2013;10(6):925-33.

192. Heo JB, Sung S. Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. Science 2011;331(6013):76-9.

193. Aoki-Yoshida A, Aoki R, Moriya N, Goto T, Kubota Y, *et al.* Omics studies of the murine intestinal ecosystem exposed to subchronic and mild social defeat stress. J Proteome Res 2016;15(9):3126-38.

194. Sanbonmatsu KY. Towards structural classification of long non-coding RNAs. Biochim Biophys Acta 2016;1859(1):41-5.

195. Bernalier-Donadille A. Fermentative metabolism by the human gut microbiota. Gastroenterol Clin Biol 2010;34 Suppl 1:S16-22.

196. Bertini I, Calabro A, De Carli V, Luchinat C, Nepi S, *et al.* The metabonomic signature of celiac disease. J Proteome Res 2009;8(1):170-7.

197. Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. Mol Cell 2011;43(6):904-14.

198. Huarte M, Guttman M, Feldser D, Garber M, Koziol MJ, *et al.* A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. Cell 2010;142(3):409-19.

199. Hung T, Wang Y, Lin MF, Koegel AK, Kotake Y, *et al.* Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. Nat Genet 2011;43(7):621-9.

200. Moore MJ. From birth to death: the complex lives of eukaryotic mRNAs. Science 2005;309(5740):1514-8.

201. Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, *et al.* The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Mol Cell 2010;39(6):925-38.

202. Geisler S, Coller J. RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. Nat Rev Mol Cell Biol 2013;14(11):699-712.

203. Hu G, Lou Z, Gupta M. The long non-coding RNA GAS5 cooperates with the eukaryotic translation initiation factor 4E to regulate c-Myc translation. PLoS One 2014;9(9):e107016.

204. Yuan S, Tao Q, Wang J, Yang F, Liu L, *et al.* Antisense long non-coding RNA PCNA-AS1 promotes tumor growth by regulating proliferating cell nuclear antigen in hepatocellular carcinoma. Cancer Lett 2014;349(1):87-94.

205. Joh RI, Palmieri CM, Hill IT, Motamedi M. Regulation of histone methylation by noncoding RNAs. Biochim Biophys Acta 2014;1839(12):1385-94.

206. Van Werven FJ, Neuert G, Hendrick N, Lardenois A, Buratowski S, *et al.* Transcription of two long noncoding RNAs mediates mating-type control of gametogenesis in budding yeast. Cell 2012;150(6):1170-81.

207. Brown CJ, Hendrich BD, Rupert JL, Lafreniere RG, Xing Y, *et al.* The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. Cell 1992;71(3):527-42.

208. Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science 2008;322(5902):750-6.

209. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, *et al.* Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature 2010;464(7291):1071-6.

210. Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, *et al.* Long noncoding RNA as modular scaffold of histone modification complexes. Science 2010;329(5992):689-93.

211. Wang KC, Yang YW, Liu B, Sanyal A, Corces-Zimmerman R, *et al.* A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature 2011;472(7341):120-4.

212. Lai F, Orom UA, Cesaroni M, Beringer M, Taatjes DJ, *et al.* Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. Nature 2013;494(7438):497-501.

213. Li J, An G, Zhang M, Ma Q. Long non-coding RNA TUG1 acts as a miR-26a sponge in human glioma cells. Biochem Biophys Res Commun 2016;477(4):743-8.

214. Zhang Z, Cheng J, Wu Y, Qiu J, Sun Y, *et al.* LncRNA HOTAIR controls the expression of Rab22a by sponging miR-373 in ovarian cancer. Mol Med Rep 2016;14(3):2465-72.

215. Chou J, Wang B, Zheng T, Li X, Zheng L, *et al.* MALAT1 induced migration and invasion of human breast cancer cells by competitively binding miR-1 with cdc42. Biochem Biophys Res Commun 2016;472(1):262-9.

216. Zou T, Jaladanki SK, Liu L, Xiao L, Chung HK, *et al.* H19 long noncoding RNA regulates intestinal epithelial barrier function via microRNA 675 by interacting with RNA-binding protein HuR. Mol Cell Biol 2016;36(9):1332-41.

217. Zhang Z, Zhu Z, Watabe K, Zhang X, Bai C, *et al.* Negative regulation of lncRNA GAS5 by miR-21. Cell Death Differ 2013;20(11):1558-68.

218. Xue M, Pang H, Li X, Li H, Pan J, *et al.* Long non-coding RNA urothelial cancer-associated 1 promotes bladder cancer cell migration and invasion by way of the hsa-miR-145-ZEB1/2-FSCN1 pathway. Cancer Sci 2016;107(1):18-27.

219. Anderson DM, Anderson KM, Chang CL, Makarewich CA, Nelson BR, *et al.* A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. Cell 2015;160(4):595-606.

220. Lukiw WJ, Handley P, Wong L, Crapper McLachlan DR. BC200 RNA in normal human neocortex, non-Alzheimer dementia (NAD), and senile dementia of the Alzheimer type (AD). Neurochem Res 1992;17(6):591-7.

221. Lin M, Pedrosa E, Shah A, Hrabovsky A, Maqbool S, *et al.* RNA-Seq of human neurons derived from iPS cells reveals candidate long non-coding RNAs involved in neurogenesis and neuropsychiatric disorders. PLoS One 2011;6(9):e23356.

222. Yang Y, Cai Y, Wu G, Chen X, Liu Y, *et al.* Plasma long non-coding RNA, CoroMarker, a novel biomarker for diagnosis of coronary artery disease. Clin Sci (Lond) 2015;129(8):675-85.

223. Cai Y, Yang Y, Chen X, He D, Zhang X, *et al.* Circulating "LncPPARdelta" from monocytes as a novel biomarker for coronary artery diseases. Medicine (Baltimore) 2016;95(6):e2360.

224. Shirasawa S, Harada H, Furugaki K, Akamizu T, Ishikawa N, *et al.* SNPs in the promoter of a B cell-specific antisense transcript, SAS-ZFAT, determine susceptibility to autoimmune thyroid disease. Hum Mol Genet 2004;13(19):2221-31.

225. Stuhlmuller B, Kunisch E, Franz J, Martinez-Gamboa L, Hernandez MM, *et al.* Detection of oncofetal h19 RNA in rheumatoid arthritis synovial tissue. Am J Pathol 2003;163(3):901-11.

226. Sonkoly E, Bata-Csorgo Z, Pivarcsi A, Polyanka H, Kenderessy-Szabo A, *et al.* Identification and characterization of a novel, psoriasis susceptibility-related noncoding RNA gene, PRINS. J Biol Chem 2005;280(25):24159-67.

227. Tsoi LC, Iyer MK, Stuart PE, Swindell WR, Gudjonsson JE, *et al.* Analysis of long non-coding RNAs highlights tissue-specific expression patterns and epigenetic profiles in normal and psoriatic skin. Genome Biol 2015;16:24.

228. Sahu A, Singhal U, Chinnaiyan AM. Long noncoding RNAs in cancer: from function to translation. Trends Cancer;1(2):93-109.

229. Schmitt Adam M, Chang Howard Y. Long noncoding RNAs in cancer pathways. Cancer Cell;29(4):452-63.

230. Xu Q, Deng F, Qin Y, Zhao Z, Wu Z, *et al.* Long non-coding RNA regulation of epithelial-mesenchymal transition in cancer metastasis. Cell Death Dis 2016;7(6):e2254.

231. Gibb EA, Brown CJ, Lam WL. The functional role of long non-coding RNA in human carcinomas. Mol Cancer 2011;10:38.

232. Prensner JR, Chinnaiyan AM. The emergence of lncRNAs in cancer biology. Cancer Discov 2011;1(5):391-407.

233. Brannan CI, Dees EC, Ingram RS, Tilghman SM. The product of the H19 gene may function as an RNA. Mol Cell Biol 1990;10(1):28-36.

234. Li C, Lei B, Huang S, Zheng M, Liu Z, *et al.* H19 derived microRNA-675 regulates cell proliferation and migration through CDK6 in glioma. Am J Transl Res 2015;7(10):1747-64.

235. Shi Y, Wang Y, Luan W, Wang P, Tao T, *et al.* Long non-coding RNA H19 promotes glioma cell invasion by deriving miR-675. PLoS One 2014;9(1):e86295.

236. Zhuang M, Gao W, Xu J, Wang P, Shu Y. The long non-coding RNA H19-derived miR-675 modulates human gastric cancer cell proliferation by targeting tumor suppressor RUNX1. Biochem Biophys Res Commun 2014;448(3):315-22.

237. Zhou X, Ye F, Yin C, Zhuang Y, Yue G, *et al.* The interaction between miR-141 and lncRNA-H19 in regulating cell proliferation and migration in gastric cancer. Cell Physiol Biochem 2015;36(4):1440-52.

238. Han D, Gao X, Wang M, Qiao Y, Xu Y, *et al.* Long noncoding RNA H19 indicates a poor prognosis of colorectal cancer and promotes tumor growth by recruiting and binding to eIF4A3. Oncotarget 2016;7(16):22159-73.

239. Sun H, Wang G, Peng Y, Zeng Y, Zhu QN, *et al.* H19 lncRNA mediates 17beta-estradiol-induced cell proliferation in MCF-7 breast cancer cells. Oncol Rep 2015;33(6):3045-52.

240. Byun HM, Wong HL, Birnstein EA, Wolff EM, Liang G, *et al.* Examination of IGF2 and H19 loss of imprinting in bladder cancer. Cancer Res 2007;67(22):10753-8.

241. Zhang L, Yang F, Yuan JH, Yuan SX, Zhou WP, *et al.* Epigenetic activation of the miR-200 family contributes to H19-mediated metastasis suppression in hepatocellular carcinoma. Carcinogenesis 2013;34(3):577-86.

242. Ma C, Nong K, Zhu H, Wang W, Huang X, *et al.* H19 promotes pancreatic cancer metastasis by derepressing let-7's suppression on its target HMGA2-mediated EMT. Tumour Biol 2014;35(9):9163-9.

243. Zhu M, Chen Q, Liu X, Sun Q, Zhao X, *et al.* LncRNA H19/miR-675 axis represses prostate cancer metastasis by targeting TGFBI. FEBS J 2014;281(16):3766-75.

244. Li X, Lin Y, Yang X, Wu X, He X. Long noncoding RNA H19 regulates EZH2 expression by interacting with miR-630 and promotes cell invasion in nasopharyngeal carcinoma. Biochem Biophys Res Commun 2016;473(4):913-9.

245. Su Z, Zhi X, Zhang Q, Yang L, Xu H, *et al.* LncRNA H19 functions as a competing endogenous RNA to regulate AQP3 expression by sponging miR-874 in the intestinal barrier. FEBS Lett 2016;590(9):1354-64.

246. Fan Y, Shen B, Tan M, Mu X, Qin Y, *et al.* Long non-coding RNA UCA1 increases chemoresistance of bladder cancer cells by regulating Wnt signaling. FEBS J 2014;281(7):1750-8.

247. Zhang R, Xia LQ, Lu WW, Zhang J, Zhu JS. LncRNAs and cancer. Oncol Lett 2016;12(2):1233-9.

248. Tu Z, He D, Deng X, Xiong M, Huang X, *et al.* An eight-long non-coding RNA signature as a candidate prognostic biomarker for lung cancer. Oncol Rep 2016;36(1):215-22.

249. Zhang A, Zhang J, Kaipainen A, Lucas JM, Yang H. Long non-coding RNA: a newly deciphered "code" in prostate cancer. Cancer Lett 2016;375(2):323-30.

250. Qi P, Zhou XY, Du X. Circulating long non-coding RNAs in cancer: current status and future perspectives. Mol Cancer 2016;15(1):39.

251. Serghiou S, Kyriakopoulou A, Ioannidis JP. Long noncoding RNAs as novel predictors of survival in human cancer: a systematic review and meta-analysis. Mol Cancer 2016;15(1):50.

252. Serghiou S, Kyriakopoulou A, Ioannidis JPA. Long noncoding RNAs as novel predictors of survival in human cancer: a systematic review and meta-analysis. Mol Cancer 2016;15(1):1-14.

253. Arun G, Diermeier S, Akerman M, Chang KC, Wilkinson JE, *et al.* Differentiation of mammary tumors and reduction in metastasis upon Malat1 lncRNA loss. Genes Dev 2016;30(1):34-51.

254. Wiedenheft B, Sternberg SH, Doudna JA. RNA-guided genetic silencing systems in bacteria and archaea. Nature 2012;482(7385):331-8.

255. Yang H, Wang H, Shivalila CS, Cheng AW, Shi L, *et al.* One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. Cell 2013;154(6):1370-9.

256. Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, *et al.* One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell 2013;153(4):910-8.

257. Auer TO, Duroure K, De Cian A, Concordet JP, Del Bene F. Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. Genome Res 2014;24(1):142-53.

258. Ho TT, Zhou N, Huang J, Koirala P, Xu M, *et al.* Targeting non-coding RNAs with the CRISPR/Cas9 system in human cell lines. Nucleic Acids Res 2015;43(3):e17.

259. Han J, Zhang J, Chen L, Shen B, Zhou J, *et al.* Efficient in vivo deletion of a large imprinted lncRNA by CRISPR/Cas9. RNA Biol 2014;11(7):829-35.

260. Jiang Y, Li Z, Zheng S, Chen H, Zhao X, *et al.* The long non-coding RNA HOTAIR affects the radiosensitivity of pancreatic ductal adenocarcinoma by

regulating the expression of Wnt inhibitory factor 1. Tumour Biol 2016;37(3):3957-67.

261. Kim K, Jutooru I, Chadalapaka G, Johnson G, Frank J, *et al.* HOTAIR is a negative prognostic factor and exhibits pro-oncogenic activity in pancreatic cancer. Oncogene 2013;32(13):1616-25.

262. Li X, Deng SJ, Zhu S, Jin Y, Cui SP, *et al.* Hypoxia-induced lncRNA-NUTF2P3-001 contributes to tumorigenesis of pancreatic cancer by derepressing the miR-3923/KRAS pathway. Oncotarget 2016;7(5):6000-14.

263. Pang EJ, Yang R, Fu XB, Liu YF. Overexpression of long non-coding RNA MALAT1 is correlated with clinical progression and unfavorable prognosis in pancreatic cancer. Tumour Biol 2015;36(4):2403-7.

264. Liu JH, Chen G, Dang YW, Li CJ, Luo DZ. Expression and prognostic significance of lncRNA MALAT1 in pancreatic cancer tissues. Asian Pac J Cancer Prev 2014;15(7):2971-7.

265. Li L, Chen H, Gao Y, Wang YW, Zhang GQ, *et al.* Long noncoding RNA MALAT1 promotes aggressive pancreatic cancer proliferation and metastasis via the stimulation of autophagy. Mol Cancer Ther 2016;15(9):2232-43.

266. Jiao F, Hu H, Han T, Yuan C, Wang L, *et al.* Long noncoding RNA MALAT-1 enhances stem cell-like phenotypes in pancreatic cancer cells. Int J Mol Sci 2015;16(4):6677-93.

267. Jiao F, Hu H, Yuan C, Wang L, Jiang W, *et al.* Elevated expression level of long noncoding RNA MALAT-1 facilitates cell growth, migration and invasion in pancreatic cancer. Oncol Rep 2014;32(6):2485-92.

268. Zheng S, Chen H, Wang Y, Gao W, Fu Z, *et al.* Long non-coding RNA LOC389641 promotes progression of pancreatic ductal adenocarcinoma and increases cell invasion by regulating E-cadherin in a TNFRSF10A-related manner. Cancer Lett 2016;371(2):354-65.

269. You L, Chang D, Du HZ, Zhao YP. Genome-wide screen identifies PVT1 as a regulator of Gemcitabine sensitivity in human pancreatic cancer cells. Biochem Biophys Res Commun 2011;407(1):1-6.

270. Huang C, Yu W, Wang Q, Cui H, Wang Y, *et al.* Increased expression of the lncRNA PVT1 is associated with poor prognosis in pancreatic cancer patients. Minerva Med 2015;106(3):143-9.

271. Li Z, Zhao X, Zhou Y, Liu Y, Zhou Q, *et al.* The long non-coding RNA HOTTIP promotes progression and gemcitabine resistance by regulating HOXA13 in pancreatic cancer. J Transl Med 2015;13:84.

272. Cheng Y, Jutooru I, Chadalapaka G, Corton JC, Safe S. The long non-coding RNA HOTTIP enhances pancreatic cancer cell proliferation, survival and migration. Oncotarget 2015;6(13):10840-52.

273. Peng W, Gao W, Feng J. Long noncoding RNA HULC is a novel biomarker of poor prognosis in patients with pancreatic cancer. Med Oncol 2014;31(12):346.

274. Hanna N, Ohana P, Konikoff FM, Leichtmann G, Hubert A, *et al.* Phase 1/2a, dose-escalation, safety, pharmacokinetic and preliminary efficacy study of intratumoral administration of BC-819 in patients with unresectable pancreatic cancer. Cancer Gene Ther 2012;19(6):374-81.

275. Zheng J, Huang X, Tan W, Yu D, Du Z, *et al.* Pancreatic cancer risk variant in LINC00673 creates a miR-1231 binding site and interferes with PTPN11 degradation. Nat Genet 2016;48(7):747-57.

276. Zhan HX, Wang Y, Li C, Xu JW, Zhou B, *et al.* LincRNA-ROR promotes invasion, metastasis and tumor growth in pancreatic cancer through activating ZEB1 pathway. Cancer Lett 2016;374(2):261-71.

277. Ye Y, Chen J, Zhou Y, Fu Z, Zhou Q, *et al.* High expression of AFAP1-AS1 is associated with poor survival and short-term recurrence in pancreatic ductal adenocarcinoma. J Transl Med 2015;13(1):1-11.

278. Li DD, Fu ZQ, Lin Q, Zhou Y, Zhou QB, *et al.* Linc00675 is a novel marker of short survival and recurrence in patients with pancreatic ductal adenocarcinoma. World J Gastroenterol 2015;21(31):9348-57.

279. Hu P, Shangguan J, Zhang L. Downregulation of NUF2 inhibits tumor growth and induces apoptosis by regulating lncRNA AF339813. Int J Clin Exp Pathol 2015;8(3):2638-48.

280. Ding YC, Yu W, Ma C, Wang Q, Huang CS, *et al.* Expression of long non-coding RNA LOC285194 and its prognostic significance in human pancreatic ductal adenocarcinoma. Int J Clin Exp Pathol 2014;7(11):8065-70.

281. Lu X, Fang Y, Wang Z, Xie J, Zhan Q, *et al.* Downregulation of gas5 increases pancreatic cancer cell proliferation by regulating CDK6. Cell Tissue Res 2013;354(3):891-6.

282. Sun YW, Chen YF, Li J, Huo YM, Liu DJ, *et al.* A novel long non-coding RNA ENST00000480739 suppresses tumour cell invasion by regulating OS-9 and HIF-1alpha in pancreatic ductal adenocarcinoma. Br J Cancer 2014;111(11):2131-41.

283. Qu S, Yang X, Song W, Sun W, Li X, *et al.* Downregulation of lncRNA-ATB correlates with clinical progression and unfavorable prognosis in pancreatic cancer. Tumour Biol 2016;37(3):3933-8.

284. Li J, Liu D, Hua R, Zhang J, Liu W, *et al.* Long non-coding RNAs expressed in pancreatic ductal adenocarcinoma and lncRNA BC008363 an independent prognostic factor in PDAC. Pancreatology 2014;14(5):385-90.

285. Hu D, Su C, Jiang M, Shen Y, Shi A, *et al.* Fenofibrate inhibited pancreatic cancer cells proliferation via activation of p53 mediated by upregulation of LncRNA MEG3. Biochem Biophys Res Commun 2016;471(2):290-5.

286. Watanabe S, Ueda Y, Akaboshi S, Hino Y, Sekita Y, *et al.* HMGA2 maintains oncogenic RAS-induced epithelial-mesenchymal transition in human pancreatic cancer cells. Am J Pathol 2009;174(3):854-68.

287. Zoetendal EG, Raes J, Van den Bogert B, Arumugam M, Booijink C, *et al.* The human small intestinal microbiota is driven by rapid uptake and conversion of simple carbohydrates. ISME J 2012;6(7):1415-26.

288. Guarner F, Malagelada JR. Gut flora in health and disease. Lancet 2003;361(9356):512-9.

289. Bhattacharya A, Banu J, Rahman M, Causey J, Fernandes G. Biological effects of conjugated linoleic acids in health and disease. J Nutr Biochem 2006;17(12):789-810.

290. Topping DL, Clifton PM. Short-chain fatty acids and human colonic function: roles of resistant starch and nonstarch polysaccharides. Physiol Rev 2001;81(3):1031-64.

291. Hinnebusch BF, Meng S, Wu JT, Archer SY, Hodin RA. The effects of short-chain fatty acids on human colon cancer cell phenotype are associated with histone hyperacetylation. J Nutr 2002;132(5):1012-7.

292. Joseph J, Niggemann B, Zaenker KS, Entschladen F. The neurotransmitter gamma-aminobutyric acid is an inhibitory regulator for the migration of SW 480 colon carcinoma cells. Cancer Res 2002;62(22):6467-9.

293. Jansson J, Willing B, Lucio M, Fekete A, Dicksved J, *et al.* Metabolomics reveals metabolic biomarkers of Crohn's disease. PLoS One 2009;4(7):e6386.

294. Arpaia N, Campbell C, Fan X, Dikiy S, van der Veeken J, *et al.* Metabolites produced by commensal bacteria promote peripheral regulatory T-cell generation. Nature 2013;504(7480):451-5.

295. Windey K, De Preter V, Verbeke K. Relevance of protein fermentation to gut health. Mol Nutr Food Res 2012;56(1):184-96.

296. Ierardi E, Sorrentino C, Principi M, Giorgio F, Losurdo G, *et al.* Intestinal microbial metabolism of phosphatidylcholine: a novel insight in the cardiovascular risk scenario. Hepatobiliary Surg Nutr 2015;4(4):289-92.

297. Russell WR, Scobbie L, Chesson A, Richardson AJ, Stewart CS, *et al.* Anti-inflammatory implications of the microbial transformation of dietary phenolic compounds. Nutr Cancer 2008;60(5):636-42.

298. Miene C, Weise A, Glei M. Impact of polyphenol metabolites produced by colonic microbiota on expression of COX-2 and GSTT2 in human colon cells (LT97). Nutr Cancer 2011;63(4):653-62.

299. Clarke G, Stilling RM, Kennedy PJ, Stanton C, Cryan JF, *et al.* Minireview: Gut microbiota: the neglected endocrine organ. Mol Endocrinol 2014;28(8):1221-38.

300. Shoaie S, Ghaffari P, Kovatcheva-Datchary P, Mardinoglu A, Sen P, *et al.* Quantifying diet-induced metabolic changes of the human gut microbiome. Cell Metab 2015;22(2):320-31.

301. Zeevi D, Korem T, Zmora N, Israeli D, Rothschild D, *et al.* Personalized nutrition by prediction of glycemic responses. Cell 2015;163(5):1079-94.

302. Saha S, Rajpal DK, Brown JR. Human microbial metabolites as a source of new drugs. Drug Discov Today 2016;21(4):692-8.

303. Poland A, Glover E, Kende AS. Stereospecific, high affinity binding of 2,3,7,8-tetrachlorodibenzo-p-dioxin by hepatic cytosol. Evidence that the binding species is receptor for induction of aryl hydrocarbon hydroxylase. J Biol Chem 1976;251(16):4936-46.

304. Carlstedt-Duke J, Elfström G, Snochowski M, Högberg B, Gustafsson J-å. Detection of the 2,3,7,8-tetpachlorodibenzo-p-dioxin (TCDD) receptor in rat liver by isoelectric focusing in polyacrylamide gels. Toxicol Lett 1978;2(6):365-73.

305. Denison MS, Pandini A, Nagy SR, Baldwin EP, Bonati L. Ligand binding and activation of the Ah receptor. Chem Biol Interact 2002;141(1-2):3-24.

306. Denison MS, Nagy SR. Activation of the aryl hydrocarbon receptor by structurally diverse exogenous and endogenous chemicals. Annu Rev Pharmacol Toxicol 2003;43:309-34.

307. Hao N, Whitelaw ML. The emerging roles of AhR in physiology and immunity. Biochem Pharmacol 2013;86(5):561-70.

308. Tuomisto JT, Pohjanvirta R, Unkila M, Tuomisto J. TCDD-induced anorexia and wasting syndrome in rats: effects of diet-induced obesity and nutrition. Pharmacol Biochem Behav 1999;62(4):735-42.

309. Goldstein JA, Linko P, Bergman H. Induction of porphyria in the rat by chronic versus acute exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin. Biochem Pharmacol 1982;31(8):1607-13.

310. Staples JE, Murante FG, Fiore NC, Gasiewicz TA, Silverstone AE. Thymic alterations induced by 2,3,7,8-tetrachlorodibenzo-p-dioxin are strictly dependent on aryl hydrocarbon receptor activation in hemopoietic cells. J Immunol 1998;160(8):3844-54.

311. Fernandez-Salguero PM, Hilbert DM, Rudikoff S, Ward JM, Gonzalez FJ. Aryl-hydrocarbon receptor-deficient mice are resistant to 2,3,7,8-tetrachlorodibenzo-p-dioxin-induced toxicity. Toxicol Appl Pharmacol 1996;140(1):173-9.

312. Vorderstrasse BA, Steppan LB, Silverstone AE, Kerkvliet NI. Aryl hydrocarbon receptor-deficient mice generate normal immune responses to model antigens and are resistant to TCDD-induced immune suppression. Toxicol Appl Pharmacol 2001;171(3):157-64.

313. Gawkrodger DJ, Harris G, Bojar RA. Chloracne in seven organic chemists exposed to novel polycyclic halogenated chemical compounds (triazoloquinoxalines). Br J Dermatol 2009;161(4):939-43.

314. Scerri L, Zaki I, Millard LG. Severe halogen acne due to a trifluoromethylpyrazole derivative and its resistance to isotretinoin. Br J Dermatol 1995;132(1):144-8.

315. Mulero-Navarro S, Fernandez-Salguero PM. New trends in aryl hydrocarbon receptor biology. Front Cell Dev Biol 2016;4:45.

316. Poland A, Glover E, Bradfield CA. Characterization of polyclonal antibodies to the Ah receptor prepared by immunization with a synthetic peptide hapten. Mol Pharmacol 1991;39(1):20-6.

317. Bradfield CA, Glover E, Poland A. Purification and N-terminal amino acid sequence of the Ah receptor from the C57BL/6J mouse. Mol Pharmacol 1991;39(1):13-9.

318. Perdew GH, Poland A. Purification of the Ah receptor from C57BL/6J mouse liver. J Biol Chem 1988;263(20):9848-52.

319. Poland A, Glover E, Taylor BA. The murine Ah locus: a new allele and mapping to chromosome 12. Mol Pharmacol 1987;32(4):471-8.

320. Kelley LA, Gardner SP, Sutcliffe MJ. An automated approach for clustering an ensemble of NMR-derived protein structures into conformationally related subfamilies. Protein Eng 1996;9(11):1063-5.

321. Fukunaga BN, Probst MR, Reisz-Porszasz S, Hankinson O. Identification of functional domains of the aryl hydrocarbon receptor. J Biol Chem 1995;270(49):29270-8.

322. Lees MJ, Peet DJ, Whitelaw ML. Defining the role for XAP2 in stabilization of the dioxin receptor. J Biol Chem 2003;278(38):35878-88.

323. Morgan RM, Hernandez-Ramirez LC, Trivellin G, Zhou L, Roe SM, *et al.* Structure of the TPR domain of AIP: lack of client protein interaction with the C-terminal alpha-7 helix of the TPR domain of AIP is sufficient for pituitary adenoma predisposition. PLoS One 2012;7(12):e53339.

324. Xu C, Li CY, Kong AN. Induction of phase I, II and III drug metabolism/transport by xenobiotics. Arch Pharm Res 2005;28(3):249-68.

325. Szollosi D, Erdei A, Gyimesi G, Magyar C, Hegedus T. Access path to the ligand binding pocket may play a role in xenobiotics selection by AhR. PLoS One 2016;11(1):e0146066.

326. Fujisawa Y, Li W, Wu D, Wong P, Vogel C, *et al.* Ligand-independent activation of the arylhydrocarbon receptor by ETK (Bmx) tyrosine kinase helps MCF10AT1 breast cancer cells to survive in an apoptosis-inducing environment. Biol Chem 2011;392(10):897-908.

327. Chuang CY, Chang H, Lin P, Sun SJ, Chen PH, *et al.* Up-regulation of osteopontin expression by aryl hydrocarbon receptor via both ligand-dependent and ligand-independent pathways in lung cancer. Gene 2012;492(1):262-9.

328. Chang X, Fan Y, Karyala S, Schwemberger S, Tomlinson CR, *et al.* Ligand-independent regulation of transforming growth factor beta1 expression and cell cycle progression by the aryl hydrocarbon receptor. Mol Cell Biol 2007;27(17):6127-39.

329. Lee CC, Yang WH, Li CH, Cheng YW, Tsai CH, *et al.* Ligand independent aryl hydrocarbon receptor inhibits lung cancer cell invasion by degradation of Smad4. Cancer Lett 2016;376(2):211-7.

330. Maayah ZH, El Gendy MA, El-Kadi AO, Korashy HM. Sunitinib, a tyrosine kinase inhibitor, induces cytochrome P450 1A1 gene in human breast cancer MCF7 cells through ligand-independent aryl hydrocarbon receptor activation. Arch Toxicol 2013;87(5):847-56.

331. Haarmann-Stemmann T, Abel J. The arylhydrocarbon receptor repressor (AhRR): structure, expression, and function. Biol Chem 2006;387(9):1195-9.

332. Levine-Fridman A, Chen L, Elferink CJ. Cytochrome P4501A1 promotes G1 phase cell cycle progression by controlling aryl hydrocarbon receptor activity. Mol Pharmacol 2004;65(2):461-9.

333. Chiaro CR, Patel RD, Marcus CB, Perdew GH. Evidence for an aryl hydrocarbon receptor-mediated cytochrome p450 autoregulatory pathway. Mol Pharmacol 2007;72(5):1369-79.

334. Adachi J, Mori Y, Matsui S, Takigami H, Fujino J, *et al.* Indirubin and indigo are potent aryl hydrocarbon receptor ligands present in human urine. J Biol Chem 2001;276(34):31475-8.

335. Ociepa-Zawal M, Rubis B, Lacinski M, Trzeciak WH. The effect of indole-3-carbinol on the expression of CYP1A1, CYP1B1 and AhR genes and proliferation of MCF-7 cells. Acta Biochim Pol 2007;54(1):113-7.

336. Safe S, Qin C, McDougal A. Development of selective aryl hydrocarbon receptor modulators for treatment of breast cancer. Expert Opin Investig Drugs 1999;8(9):1385-96.

337. Safe S, McDougal A. Mechanism of action and development of selective aryl hydrocarbon receptor modulators for treatment of hormone-dependent cancers (Review). Int J Oncol 2002;20(6):1123-8.

338. Chen I, McDougal A, Wang F, Safe S. Aryl hydrocarbon receptor-mediated antiestrogenic and antitumorigenic activity of diindolylmethane. Carcinogenesis 1998;19(9):1631-9.

339. Bjeldanes LF, Kim JY, Grose KR, Bartholomew JC, Bradfield CA. Aromatic hydrocarbon responsiveness-receptor agonists generated from indole-3-carbinol in vitro and in vivo: comparisons with 2,3,7,8-tetrachlorodibenzo-p-dioxin. Proc Natl Acad Sci U S A 1991;88(21):9543-7.

340. Chen I, Safe S, Bjeldanes L. Indole-3-carbinol and diindolylmethane as aryl hydrocarbon (Ah) receptor agonists and antagonists in T47D human breast cancer cells. Biochem Pharmacol 1996;51(8):1069-76.

341. Degner SC, Papoutsis AJ, Selmin O, Romagnolo DF. Targeting of aryl hydrocarbon receptor-mediated activation of cyclooxygenase-2 expression by the indole-3-carbinol metabolite 3,3'-diindolylmethane in breast cancer cells. J Nutr 2009;139(1):26-32.

342. Peter Guengerich F, Martin MV, McCormick WA, Nguyen LP, Glover E, *et al.* Aryl hydrocarbon receptor response to indigoids in vitro and in vivo. Arch Biochem Biophys 2004;423(2):309-16.

343. Barreto FC, Barreto DV, Liabeuf S, Meert N, Glorieux G, *et al.* Serum indoxyl sulfate is associated with vascular disease and mortality in chronic kidney disease patients. Clin J Am Soc Nephrol 2009;4(10):1551-8.

344. Schroeder JC, Dinatale BC, Murray IA, Flaveny CA, Liu Q, *et al.* The uremic toxin 3-indoxyl sulfate is a potent endogenous agonist for the human aryl hydrocarbon receptor. Biochemistry 2010;49(2):393-400.

345. Van der Heiden E, Bechoux N, Muller M, Sergent T, Schneider YJ, *et al.* Food flavonoid aryl hydrocarbon receptor-mediated agonistic/antagonistic/synergic

activities in human and rat reporter gene assays. Anal Chim Acta 2009;637(1-2):337-45.

346. Wang HK, Yeh CH, Iwamoto T, Satsu H, Shimizu M, *et al.* Dietary flavonoid naringenin induces regulatory T cells via an aryl hydrocarbon receptor mediated pathway. J Agric Food Chem 2012;60(9):2171-8.

347. Kapitulnik J, Gonzalez FJ. Marked endogenous activation of the CYP1A1 and CYP1A2 genes in the congenitally jaundiced Gunn rat. Mol Pharmacol 1993;43(5):722-5.

348. Sinal CJ, Bend JR. Aryl hydrocarbon receptor-dependent induction of cyp1a1 by bilirubin in mouse hepatoma hepa 1c1c7 cells. Mol Pharmacol 1997;52(4):590-9.

349. Phelan D, Winter GM, Rogers WJ, Lam JC, Denison MS. Activation of the Ah receptor signal transduction pathway by bilirubin and biliverdin. Arch Biochem Biophys 1998;357(1):155-63.

350. Chiaro CR, Morales JL, Prabhu KS, Perdew GH. Leukotriene A4 metabolites are endogenous ligands for the Ah receptor. Biochemistry 2008;47(32):8445-55.

351. Seidel SD, Winters GM, Rogers WJ, Ziccardi MH, Li V, *et al.* Activation of the Ah receptor signaling pathway by prostaglandins. J Biochem Mol Toxicol 2001;15(4):187-96.

352. Chiaro CR, Patel RD, Perdew GH. 12(R)-Hydroxy-5(Z),8(Z),10(E),14(Z)-eicosatetraenoic acid [12(R)-HETE], an arachidonic acid derivative, is an activator of the aryl hydrocarbon receptor. Mol Pharmacol 2008;74(6):1649-56.

353. Schaldach CM, Riby J, Bjeldanes LF. Lipoxin A4: a new class of ligand for the Ah receptor. Biochemistry 1999;38(23):7594-600.

354. Opitz CA, Litzenburger UM, Sahm F, Ott M, Tritschler I, *et al.* An endogenous tumour-promoting ligand of the human aryl hydrocarbon receptor. Nature 2011;478(7368):197-203.

355. DiNatale BC, Murray IA, Schroeder JC, Flaveny CA, Lahoti TS, *et al.* Kynurenic acid is a potent endogenous aryl hydrocarbon receptor ligand that synergistically induces interleukin-6 in the presence of inflammatory signaling. Toxicol Sci 2010;115(1):89-97.

356. Heath-Pagliuso S, Rogers WJ, Tullis K, Seidel SD, Cenijn PH, *et al.* Activation of the Ah receptor by tryptophan and tryptophan metabolites. Biochemistry 1998;37(33):11508-15.

357. Vikstrom Bergander L, Cai W, Klocke B, Seifert M, Pongratz I. Tryptamine serves as a proligand of the AhR transcriptional pathway whose activation is dependent of monoamine oxidases. Mol Endocrinol 2012;26(9):1542-51.

358. Bittinger MA, Nguyen LP, Bradfield CA. Aspartate aminotransferase generates proagonists of the aryl hydrocarbon receptor. Mol Pharmacol 2003;64(3):550-6.

359. Rannug A, Rannug U, Rosenkranz HS, Winqvist L, Westerholm R, *et al.* Certain photooxidized derivatives of tryptophan bind with very high affinity to the Ah receptor and are likely to be endogenous signal substances. J Biol Chem 1987;262(32):15422-7.

360. Wincent E, Bengtsson J, Mohammadi Bardbori A, Alsberg T, Luecke S, *et al.* Inhibition of cytochrome P4501-dependent clearance of the endogenous agonist FICZ as a mechanism for activation of the aryl hydrocarbon receptor. Proc Natl Acad Sci U S A 2012;109(12):4479-84.

361. Wincent E, Amini N, Luecke S, Glatt H, Bergman J, *et al.* The suggested physiologic aryl hydrocarbon receptor activator and cytochrome P4501 substrate 6-formylindolo[3,2-b]carbazole is present in humans. J Biol Chem 2009;284(5):2690-6.

362. Mukai M, Tischkau SA. Effects of tryptophan photoproducts in the circadian timing system: searching for a physiological role for aryl hydrocarbon receptor. Toxicol Sci 2007;95(1):172-81.

363. Goryo K, Suzuki A, Del Carpio CA, Siizaki K, Kuriyama E, *et al.* Identification of amino acid residues in the Ah receptor involved in ligand binding. Biochem Biophys Res Commun 2007;354(2):396-402.

364. Oberg M, Bergander L, Hakansson H, Rannug U, Rannug A. Identification of the tryptophan photoproduct 6-formylindolo[3,2-b]carbazole, in cell culture medium, as a factor that controls the background aryl hydrocarbon receptor activity. Toxicol Sci 2005;85(2):935-43.

365. Wei YD, Rannug U, Rannug A. UV-induced CYP1A1 gene expression in human cells is mediated by tryptophan. Chem Biol Interact 1999;118(2):127-40.

366. Wei YD, Bergander L, Rannug U, Rannug A. Regulation of CYP1A1 transcription via the metabolism of the tryptophan-derived 6-formylindolo[3,2-b]carbazole. Arch Biochem Biophys 2000;383(1):99-107.

367. Kostyuk V, Potapovich A, Stancato A, De Luca C, Lulli D, *et al.* Photo-oxidation products of skin surface squalene mediate metabolic and inflammatory responses to solar UV in human keratinocytes. PLoS One 2012;7(8):e44472.

368. Jeong KT, Hwang SJ, Oh GS, Park JH. FICZ, a tryptophan photoproduct, suppresses pulmonary eosinophilia and Th2-type cytokine production in a mouse model of ovalbumin-induced allergic asthma. Int Immunopharmacol 2012;13(4):377-85.

369. Bunaciu RP, Yen A. 6-Formylindolo (3,2-b)carbazole (FICZ) enhances retinoic acid (RA)-induced differentiation of HL-60 myeloblastic leukemia cells. Mol Cancer 2013;12:39.

370. Park JH, Choi AJ, Kim SJ, Cheong SW, Jeong SY. AhR activation by 6-formylindolo[3,2-b]carbazole and 2,3,7,8-tetrachlorodibenzo-p-dioxin inhibit the development of mouse intestinal epithelial cells. Environ Toxicol Pharmacol 2016;43:44-53.

371. Song J, Clagett-Dame M, Peterson RE, Hahn ME, Westler WM, *et al.* A ligand for the aryl hydrocarbon receptor isolated from lung. Proc Natl Acad Sci U S A 2002;99(23):14694-9.

372. Ma T, Cao YL, Xu BB, Zhou XM. Effect of (3,5,6-trimethylpyrazin-2-yl)methyl 2-[4-(2-methylpropyl)phenyl]propanoate (ITE), a newly developed anti-

inflammatory drug, on type II collagen-induced arthritis in mice. Biol Pharm Bull 2004;27(6):810-2.

373. Quintana FJ, Murugaiyan G, Farez MF, Mitsdoerffer M, Tukpah AM, *et al.* An endogenous aryl hydrocarbon receptor ligand acts on dendritic cells and T cells to suppress experimental autoimmune encephalomyelitis. Proc Natl Acad Sci U S A 2010;107(48):20768-73.

374. Lehmann GM, Xi X, Kulkarni AA, Olsen KC, Pollock SJ, *et al.* The aryl hydrocarbon receptor ligand ITE inhibits TGFbeta1-induced human myofibroblast differentiation. Am J Pathol 2011;178(4):1556-67.

375. Gonzalez FJ, Fernandez-Salguero P. The aryl hydrocarbon receptor: studies using the AHR-null mice. Drug Metab Dispos 1998;26(12):1194-8.

376. Thackaberry EA, Bedrick EJ, Goens MB, Danielson L, Lund AK, *et al.* Insulin regulation in AhR-null mice: embryonic cardiac enlargement, neonatal macrosomia, and altered insulin regulation and response in pregnant and aging AhR-null females. Toxicol Sci 2003;76(2):407-17.

377. Harstad EB, Guite CA, Thomae TL, Bradfield CA. Liver deformation in Ahr-null mice: evidence for aberrant hepatic perfusion in early development. Mol Pharmacol 2006;69(5):1534-41.

378. Esser C. The immune phenotype of AhR null mouse mutants: not a simple mirror of xenobiotic receptor over-activation. Biochem Pharmacol 2009;77(4):597-607.

379. Boutros PC, Bielefeld KA, Pohjanvirta R, Harper PA. Dioxin-dependent and dioxin-independent gene batteries: comparison of liver and kidney in AHR-null mice. Toxicol Sci 2009;112(1):245-56.

380. Aleksunes LM, Klaassen CD. Coordinated regulation of hepatic phase I and II drug-metabolizing genes and transporters using AhR-, CAR-, PXR-, PPARalpha-, and Nrf2-null mice. Drug Metab Dispos 2012;40(7):1366-79.

381. Schmidt JV, Su GH, Reddy JK, Simon MC, Bradfield CA. Characterization of a murine Ahr null allele: involvement of the Ah receptor in hepatic growth and development. Proc Natl Acad Sci U S A 1996;93(13):6731-6.

382. Veldhoen M, Hirota K, Westendorf AM, Buer J, Dumoutier L, *et al.* The aryl hydrocarbon receptor links TH17-cell-mediated autoimmunity to environmental toxins. Nature 2008;453(7191):106-9.

383. Totlandsdal AI, Ovrevik J, Cochran RE, Herseth JI, Bolling AK, *et al.* The occurrence of polycyclic aromatic hydrocarbons and their derivatives and the proinflammatory potential of fractionated extracts of diesel exhaust and wood smoke particles. J Environ Sci Health A Tox Hazard Subst Environ Eng 2014;49(4):383-96.

384. Podechard N, Lecureur V, Le Ferrec E, Guenon I, Sparfel L, *et al.* Interleukin-8 induction by the environmental contaminant benzo(a)pyrene is aryl hydrocarbon receptor-dependent and leads to lung inflammation. Toxicol Lett 2008;177(2):130-7.

385. Savouret JF, Berdeaux A, Casper RF. The aryl hydrocarbon receptor and its xenobiotic ligands: a fundamental trigger for cardiovascular diseases. Nutr Metab Cardiovasc Dis 2003;13(2):104-13.

386. Dever DP, Opanashuk LA. The aryl hydrocarbon receptor contributes to the proliferation of human medulloblastoma cells. Mol Pharmacol 2012;81(5):669-78.

387. Jaffrain-Rea ML, Angelini M, Gargano D, Tichomirowa MA, Daly AF, *et al.* Expression of aryl hydrocarbon receptor (AHR) and AHR-interacting protein in pituitary adenomas: pathological and clinical implications. Endocr Relat Cancer 2009;16(3):1029-43.

388. Hall JM, Barhoover MA, Kazmin D, McDonnell DP, Greenlee WF, *et al.* Activation of the aryl-hydrocarbon receptor inhibits invasive and metastatic features of human breast cancer cells and promotes breast cancer cell differentiation. Mol Endocrinol 2010;24(2):359-69.

389. Hsieh TH, Tsai CF, Hsu CY, Kuo PL, Lee JN, *et al.* Phthalates induce proliferation and invasiveness of estrogen receptor-negative breast cancer through the AhR/HDAC6/c-Myc signaling pathway. FASEB J 2012;26(2):778-87.

390. McDougal A, Wilson C, Safe S. Inhibition of 7,12-dimethylbenz[a]anthracene-induced rat mammary tumor growth by aryl hydrocarbon receptor agonists. Cancer Lett 1997;120(1):53-63.

391. Ramamoorthy K, Gupta MS, Sun G, McDougal A, Safe SH. 3,3'4,4'-Tetrachlorobiphenyl exhibits antiestrogenic and antitumorigenic activity in the rodent uterus and mammary cells and in human breast cancer cells. Carcinogenesis 1999;20(1):115-23.

392. Zhang S, Lei P, Liu X, Li X, Walker K, *et al.* The aryl hydrocarbon receptor as a target for estrogen receptor-negative breast cancer chemotherapy. Endocr Relat Cancer 2009;16(3):835-44.

393. Wang K, Li Y, Jiang YZ, Dai CF, Patankar MS, *et al.* An endogenous aryl hydrocarbon receptor ligand inhibits proliferation and migration of human ovarian cancer cells. Cancer Lett 2013;340(1):63-71.

394. Richardson HL, Stier AR, Borsos-Nachtnebel E. Liver tumor inhibition and adrenal histologic responses in rats to which 3′-methyl-4-dimethylaminoazobenzene and 20-methylcholanthrene were simultaneously administered. Cancer Res 1952;12(5):356-61.

395. Abdelrahim M, Smith R, 3rd, Safe S. Aryl hydrocarbon receptor gene silencing with small inhibitory RNA differentially modulates Ah-responsiveness in MCF-7 and HepG2 cancer cells. Mol Pharmacol 2003;63(6):1373-81.

396. Jin UH, Lee SO, Pfent C, Safe S. The aryl hydrocarbon receptor ligand omeprazole inhibits breast cancer cell invasion and metastasis. BMC Cancer 2014;14:498.

397. Jin UH, Kim SB, Safe S. Omeprazole inhibits pancreatic cancer cell invasion through a nongenomic aryl hydrocarbon receptor pathway. Chem Res Toxicol 2015;28(5):907-18.

398. Safe S, Lee SO, Jin UH. Role of the aryl hydrocarbon receptor in carcinogenesis and potential as a drug target. Toxicol Sci 2013;135(1):1-16.

399. Clark DA, Gauldie J, Szewczuk MR, Sweeney G. Enhanced suppressor cell activity as a mechanism of immunosuppression by 2,3,7,8-tetrachlorodibenzo-p-dioxin. Proc Soc Exp Biol Med 1981;168(2):290-9.

400. Fernandez-Salguero P, Pineau T, Hilbert DM, McPhail T, Lee SS, *et al.* Immune system impairment and hepatic fibrosis in mice lacking the dioxin-binding Ah receptor. Science 1995;268(5211):722-6.

401. Sun YV, Boverhof DR, Burgoon LD, Fielden MR, Zacharewski TR. Comparative analysis of dioxin response elements in human, mouse and rat genomic sequences. Nucleic Acids Res 2004;32(15):4512-23.

402. Smith BW, Rozelle SS, Leung A, Ubellacker J, Parks A, *et al.* The aryl hydrocarbon receptor directs hematopoietic progenitor cell expansion and differentiation. Blood 2013;122(3):376-85.

403. Stockinger B, Di Meglio P, Gialitakis M, Duarte JH. The aryl hydrocarbon receptor: multitasking in the immune system. Annu Rev Immunol 2014;32:403-32.

404. Zelante T, Iannitti RG, Cunha C, De Luca A, Giovannini G, *et al.* Tryptophan catabolites from microbiota engage aryl hydrocarbon receptor and balance mucosal reactivity via interleukin-22. Immunity 2013;39(2):372-85.

405. Zhang L, Nichols RG, Correll J, Murray IA, Tanaka N, *et al.* Persistent organic pollutants modify gut microbiota-host metabolic homeostasis in mice through aryl hydrocarbon receptor activation. Environ Health Perspect 2015;123(7):679-88.

406. Gargaro M, Pirro M, Romani R, Zelante T, Fallarino F. Aryl hydrocarbon receptor-dependent pathways in immune regulation. Am J Transplant 2016;16(8):2270-6.

407. Zhou L. AHR function in lymphocytes: emerging concepts. Trends Immunol 2016;37(1):17-31.

408. Quintana FJ, Basso AS, Iglesias AH, Korn T, Farez MF, *et al.* Control of T(reg) and T(H)17 cell differentiation by the aryl hydrocarbon receptor. Nature 2008;453(7191):65-71.

409. Duarte JH, Di Meglio P, Hirota K, Ahlfors H, Stockinger B. Differential influences of the aryl hydrocarbon receptor on Th17 mediated responses in vitro and in vivo. PLoS One 2013;8(11):e79819.

410. Sakaguchi S, Miyara M, Costantino CM, Hafler DA. FOXP3+ regulatory T cells in the human immune system. Nat Rev Immunol 2010;10(7):490-500.

411. Negishi T, Kato Y, Ooneda O, Mimura J, Takada T, *et al.* Effects of aryl hydrocarbon receptor signaling on the modulation of TH1/TH2 balance. J Immunol 2005;175(11):7348-56.

412. Veldhoen M, Hirota K, Westendorf AM, Buer J, Dumoutier L, *et al.* The aryl hydrocarbon receptor links TH17-cell-mediated autoimmunity to environmental toxins. Nature 2008;453(7191):106-9.

413. Kimura A, Naka T, Nohara K, Fujii-Kuriyama Y, Kishimoto T. Aryl hydrocarbon receptor regulates Stat1 activation and participates in the development of Th17 cells. Proc Natl Acad Sci U S A 2008;105(28):9721-6.

414. Mezrich JD, Fechner JH, Zhang X, Johnson BP, Burlingham WJ, *et al.* An interaction between kynurenine and the aryl hydrocarbon receptor can generate regulatory T cells. J Immunol 2010;185(6):3190-8.

415. Hauben E, Gregori S, Draghici E, Migliavacca B, Olivieri S, *et al.* Activation of the aryl hydrocarbon receptor promotes allograft-specific tolerance through direct and dendritic cell–mediated effects on regulatory T cells. Blood 2008;112(4):1214-22.

416. Rouse M, Singh NP, Nagarkatti PS, Nagarkatti M. Indoles mitigate the development of experimental autoimmune encephalomyelitis by induction of reciprocal differentiation of regulatory T cells and Th17 cells. Br J Pharmacol 2013;169(6):1305-21.

417. Apetoh L, Quintana FJ, Pot C, Joller N, Xiao S, *et al.* The aryl hydrocarbon receptor interacts with c-Maf to promote the differentiation of type 1 regulatory T cells induced by IL-27. Nat Immunol 2010;11(9):854-61.

418. Gandhi R, Kumar D, Burns EJ, Nadeau M, Dake B, *et al.* Activation of the aryl hydrocarbon receptor induces human type 1 regulatory T cell-like and Foxp3+ regulatory T cells. Nat Immunol 2010;11(9):846-53.

419. Head JL, Lawrence BP. The aryl hydrocarbon receptor is a modulator of anti-viral immunity. Biochem Pharmacol 2009;77(4):642-53.

420. Elizondo G, Rodriguez-Sosa M, Estrada-Muniz E, Gonzalez FJ, Vega L. Deletion of the aryl hydrocarbon receptor enhances the inflammatory response to Leishmania major infection. Int J Biol Sci 2011;7(9):1220-9.

421. Yoshida T, Katsuya K, Oka T, Koizumi S, Wakita D, *et al.* Effects of AhR ligands on the production of immunoglobulins in purified mouse B cells. Biomed Res 2012;33(2):67-74.

422. Teske S, Bohn AA, Hogaboam JP, Lawrence BP. Aryl hydrocarbon receptor targets pathways extrinsic to bone marrow cells to enhance neutrophil recruitment during influenza virus infection. Toxicol Sci 2008;102(1):89-99.

423. Wang T, Wyrick KL, Pecka MR, Wills TB, Vorderstrasse BA. Mechanistic exploration of AhR-mediated host protection against Streptococcus pneumoniae infection. Int Immunopharmacol 2012;13(4):490-8.

424. Veiga-Parga T, Suryawanshi A, Rouse BT. Controlling viral immuno-inflammatory lesions by modulating aryl hydrocarbon receptor signaling. PLoS Pathog 2011;7(12):e1002427.

425. Frumento G, Rotondo R, Tonetti M, Damonte G, Benatti U, *et al.* Tryptophan-derived catabolites are responsible for inhibition of T and natural killer cell proliferation induced by indoleamine 2,3-dioxygenase. J Exp Med 2002;196(4):459-68.

426. Sibilano R, Frossi B, Calvaruso M, Danelli L, Betto E, *et al.* The aryl hydrocarbon receptor modulates acute and late mast cell responses. J Immunol 2012;189(1):120-7.

427. Maaetoft-Udsen K, Shimoda LM, Frokiaer H, Turner H. Aryl hydrocarbon receptor ligand effects in RBL2H3 cells. J Immunotoxicol 2012;9(3):327-37.

428. Lindsey S, Jiang J, Woulfe D, Papoutsakis ET. Platelets from mice lacking the aryl hydrocarbon receptor exhibit defective collagen-dependent signaling. J Thromb Haemost 2014;12(3):383-94.

429. Nieswandt B, Watson SP. Platelet-collagen interaction: is GPVI the central receptor? Blood 2003;102(2):449-61.

430. Pombo M, Lame MW, Walker NJ, Huynh DH, Tablin F. TCDD and omeprazole prime platelets through the aryl hydrocarbon receptor (AhR) non-genomic pathway. Toxicol Lett 2015;235(1):28-36.

431. Vogel CF, Wu D, Goth SR, Baek J, Lollies A, *et al.* Aryl hydrocarbon receptor signaling regulates NF-kappaB RelB activation during dendritic-cell differentiation. Immunol Cell Biol 2013;91(9):568-75.

432. Nguyen NT, Kimura A, Nakahama T, Chinen I, Masuda K, *et al.* Aryl hydrocarbon receptor negatively regulates dendritic cell immunogenicity via a kynurenine-dependent mechanism. Proc Natl Acad Sci U S A 2010;107(46):19961-6.

433. Lawrence BP, Denison MS, Novak H, Vorderstrasse BA, Harrer N, *et al.* Activation of the aryl hydrocarbon receptor is essential for mediating the anti-inflammatory effects of a novel low-molecular-weight compound. Blood 2008;112(4):1158-65.

434. Kiss EA, Vonarbourg C, Kopfmann S, Hobeika E, Finke D, *et al.* Natural aryl hydrocarbon receptor ligands control organogenesis of intestinal lymphoid follicles. Science 2011;334(6062):1561-5.

435. Spits H, Artis D, Colonna M, Diefenbach A, Di Santo JP, *et al.* Innate lymphoid cells--a proposal for uniform nomenclature. Nat Rev Immunol 2013;13(2):145-9.

436. Chea S, Perchet T, Petit M, Verrier T, Guy-Grand D, *et al.* Notch signaling in group 3 innate lymphoid cells modulates their plasticity. Sci Signal 2016;9(426):ra45.

437. Wagage S, Harms Pritchard G, Dawson L, Buza EL, Sonnenberg GF, *et al.* The group 3 innate lymphoid cell defect in aryl hydrocarbon receptor deficient mice

is associated with T cell hyperactivation during intestinal infection. PLoS One 2015;10(5):e0128335.

438. Qiu J, Zhou L. Aryl hydrocarbon receptor promotes RORgammat(+) group 3 ILCs and controls intestinal immunity and inflammation. Semin Immunopathol 2013;35(6):657-70.

439. Qiu J, Heller JJ, Guo X, Chen ZM, Fish K, *et al.* The aryl hydrocarbon receptor regulates gut immunity through modulation of innate lymphoid cells. Immunity 2012;36(1):92-104.

440. Di Meglio P, Perera GK, Nestle FO. The multitasking organ: recent insights into skin immune function. Immunity 2011;35(6):857-69.

441. Furue M, Takahara M, Nakahara T, Uchi H. Role of AhR/ARNT system in skin homeostasis. Arch Dermatol Res 2014;306(9):769-79.

442. Wanner R, Panteleyev A, Henz BM, Rosenbach T. Retinoic acid affects the expression rate of the differentiation-related genes aryl hydrocarbon receptor, ARNT and keratin 4 in proliferative keratinocytes only. Biochim Biophys Acta 1996;1317(2):105-11.

443. Sutter CH, Yin H, Li Y, Mammen JS, Bodreddigari S, *et al.* EGF receptor signaling blocks aryl hydrocarbon receptor-mediated transcription and cell differentiation in human epidermal keratinocytes. Proc Natl Acad Sci U S A 2009;106(11):4266-71.

444. Osborne R, Greenlee WF. 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) enhances terminal differentiation of cultured human epidermal cells. Toxicol Appl Pharmacol 1985;77(3):434-43.

445. Loertscher JA, Sattler CA, Allen-Hoffmann BL. 2,3,7,8-Tetrachlorodibenzo-p-dioxin alters the differentiation pattern of human keratinocytes in organotypic culture. Toxicol Appl Pharmacol 2001;175(2):121-9.

446. Van den Bogaard EH, Bergboer JG, Vonk-Bergers M, Van Vlijmen-Willems IM, Hato SV, *et al.* Coal tar induces AHR-dependent skin barrier repair in atopic dermatitis. J Clin Invest 2013;123(2):917-27.

447. Di Meglio P, Duarte JH, Ahlfors H, Owens ND, Li Y, *et al.* Activation of the aryl hydrocarbon receptor dampens the severity of inflammatory skin conditions. Immunity 2014;40(6):989-1001.

448. Tauchi M, Hida A, Negishi T, Katsuoka F, Noda S, *et al.* Constitutive expression of aryl hydrocarbon receptor in keratinocytes causes inflammatory skin lesions. Mol Cell Biol 2005;25(21):9360-8.

449. Rey-Barroso J, Colo GP, Alvarez-Barrientos A, Redondo-Munoz J, Carvajal-Gonzalez JM, *et al.* The dioxin receptor controls beta1 integrin activation in fibroblasts through a Cbp-Csk-Src pathway. Cell Signal 2013;25(4):848-59.

450. Hayashi S, Watanabe J, Nakachi K, Eguchi H, Gotoh O, *et al.* Interindividual difference in expression of human Ah receptor and related P450 genes. Carcinogenesis 1994;15(5):801-6.

451. Esser C, Rannug A. The aryl hydrocarbon receptor in barrier organ physiology, immunology, and toxicology. Pharmacol Rev 2015;67(2):259-79.

452. Chantry D, Burgess LE. Chemokines in allergy. Curr Drug Targets Inflamm Allergy 2002;1(1):109-16.

453. Zhou Y, Tung HY, Tsai YM, Hsu SC, Chang HW, *et al.* Aryl hydrocarbon receptor controls murine mast cell homeostasis. Blood 2013;121(16):3195-204.

454. Mitchell KA, Elferink CJ. Timing is everything: consequences of transient and sustained AhR activity. Biochem Pharmacol 2009;77(6):947-56.

455. Li Y, Innocentin S, Withers DR, Roberts NA, Gallagher AR, *et al.* Exogenous stimuli maintain intraepithelial lymphocytes via aryl hydrocarbon receptor activation. Cell 2011;147(3):629-40.

456. Lee JS, Cella M, McDonald KG, Garlanda C, Kennedy GD, *et al.* AHR drives the development of gut ILC22 cells and postnatal lymphoid tissues via pathways dependent on and independent of Notch. Nat Immunol 2012;13(2):144-51.

457. Nakahama T, Hanieh H, Nguyen NT, Chinen I, Ripley B, *et al.* Aryl hydrocarbon receptor-mediated induction of the microRNA-132/212 cluster promotes interleukin-17-producing T-helper cell differentiation. Proc Natl Acad Sci U S A 2013;110(29):11964-9.

458. Zheng Y, Valdez PA, Danilenko DM, Hu Y, Sa SM, *et al.* Interleukin-22 mediates early host defense against attaching and effacing bacterial pathogens. Nat Med 2008;14(3):282-9.

459. Sonnenberg GF, Fouser LA, Artis D. Border patrol: regulation of immunity, inflammation and tissue homeostasis at barrier surfaces by IL-22. Nat Immunol 2011;12(5):383-90.

460. Spits H, Di Santo JP. The expanding family of innate lymphoid cells: regulators and effectors of immunity and tissue remodeling. Nat Immunol 2011;12(1):21-7.

461. Esser C, Rannug A, Stockinger B. The aryl hydrocarbon receptor in immunity. Trends Immunol 2009;30(9):447-54.

462. Monteleone I, MacDonald TT, Pallone F, Monteleone G. The aryl hydrocarbon receptor in inflammatory bowel disease: linking the environment to disease pathogenesis. Curr Opin Gastroenterol 2012;28(4):310-3.

463. Qiu J, Guo X, Chen ZM, He L, Sonnenberg GF, *et al.* Group 3 innate lymphoid cells inhibit T-cell-mediated intestinal inflammation through aryl hydrocarbon receptor signaling and regulation of microflora. Immunity 2013;39(2):386-99.

464. Chassaing B, Aitken JD, Malleshappa M, Vijay-Kumar M. Dextran sulfate sodium (DSS)-induced colitis in mice. Curr Protoc Immunol 2014;104:Unit 15.25.

465. Takamura T, Harama D, Matsuoka S, Shimokawa N, Nakamura Y, *et al.* Activation of the aryl hydrocarbon receptor pathway may ameliorate dextran sodium sulfate-induced colitis in mice. Immunol Cell Biol 2010;88(6):685-9.

466. Monteleone I, Rizzo A, Sarra M, Sica G, Sileri P, *et al.* Aryl hydrocarbon receptor-induced signals up-regulate IL-22 production and inhibit inflammation in the gastrointestinal tract. Gastroenterology 2011;141(1):237-48, 48 e1.

467. Nakao A. IL-7: AhR we ready for a new cytokine to fight colitis? Dig Dis Sci 2015;60(7):1876-7.

468. Lamas B, Richard ML, Leducq V, Pham HP, Michel ML, *et al.* CARD9 impacts colitis by altering gut microbiota metabolism of tryptophan into aryl hydrocarbon receptor ligands. Nat Med 2016;22(6):598-605.

469. Ji T, Xu C, Sun L, Yu M, Peng K, *et al.* Aryl hydrocarbon receptor activation down-regulates IL-7 and reduces inflammation in a mouse model of DSS-induced colitis. Dig Dis Sci 2015;60(7):1958-66.

470. Monteleone I, Zorzi F, Marafini I, Di Fusco D, Dinallo V, *et al.* Aryl hydrocarbon receptor-driven signals inhibit collagen synthesis in the gut. Eur J Immunol 2016;46(4):1047-57.

471. Ikuta T, Kurosumi M, Yatsuoka T, Nishimura Y. Tissue distribution of aryl hydrocarbon receptor in the intestine: Implication of putative roles in tumor suppression. Exp Cell Res 2016;343(2):126-34.

472. Ullman TA, Itzkowitz SH. Intestinal inflammation and cancer. Gastroenterology 2011;140(6):1807-16.

473. Ikuta T, Kobayashi Y, Kitazawa M, Shiizaki K, Itano N, *et al.* ASC-associated inflammation promotes cecal tumorigenesis in aryl hydrocarbon receptor-deficient mice. Carcinogenesis 2013;34(7):1620-7.

474. Kawajiri K, Kobayashi Y, Ohtake F, Ikuta T, Matsushima Y, *et al.* Aryl hydrocarbon receptor suppresses intestinal carcinogenesis in ApcMin/+ mice with natural ligands. Proc Natl Acad Sci U S A 2009;106(32):13481-6.

475. Yin J, Sheng B, Han B, Pu A, Yang K, *et al.* The AhR is involved in the regulation of LoVo cell proliferation through cell cycle-associated proteins. Cell Biol Int 2016;40(5):560-8.

476. Megna BW, Carney PR, Nukaya M, Geiger P, Kennedy GD. Indole-3-carbinol induces tumor cell death: function follows form. J Surg Res 2016;204(1):47-54.

477. Flynn RA, Chang HY. Long noncoding RNAs in cell-fate programming and reprogramming. Cell Stem Cell 2014;14(6):752-61.

478. Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. Annu Rev Biochem 2012;81:145-66.

479. Huarte M, Rinn JL. Large non-coding RNAs: missing links in cancer? Hum Mol Genet 2010;19(R2):R152-61.

480. Lee JT. Epigenetic regulation by long noncoding RNAs. Science 2012;338(6113):1435-9.

481. Batista PJ, Chang HY. Long noncoding RNAs: cellular address codes in development and disease. Cell 2013;152(6):1298-307.

482. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, *et al.* Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell 2007;129(7):1311-23.

483. Yang Z, Zhou L, Wu LM, Lai MC, Xie HY, *et al.* Overexpression of long non-coding RNA HOTAIR predicts tumor recurrence in hepatocellular carcinoma patients following liver transplantation. Ann Surg Oncol 2011;18(5):1243-50.

484. Kogo R, Shimamura T, Mimori K, Kawahara K, Imoto S, *et al.* Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. Cancer Res 2011;71(20):6320-6.

485. Svoboda M, Slyskova J, Schneiderova M, Makovicky P, Bielik L, *et al.* HOTAIR long non-coding RNA is a negative prognostic factor not only in primary tumors, but also in the blood of colorectal cancer patients. Carcinogenesis 2014;35(7):1510-5.

486. Li X, Wu Z, Mei Q, Li X, Guo M, *et al.* Long non-coding RNA HOTAIR, a driver of malignancy, predicts negative prognosis and exhibits oncogenic activity in oesophageal squamous cell carcinoma. Br J Cancer 2013;109(8):2266-78.

487. Quagliata L, Matter MS, Piscuoglio S, Arabi L, Ruiz C, *et al.* Long noncoding RNA HOTTIP/HOXA13 expression is associated with disease progression and predicts outcome in hepatocellular carcinoma patients. Hepatology 2014;59(3):911-23.

488. Kupershmidt I, Su QJ, Grewal A, Sundaresh S, Halperin I, *et al.* Ontology-based meta-analysis of global collections of high-throughput public data. PLoS One 2010;5(9): pii: e13066.

489. Zoller M. CD44: can a cancer-initiating cell profit from an abundantly expressed molecule? Nat Rev Cancer 2011;11(4):254-67.

490. Orian-Rousseau V. CD44, a therapeutic target for metastasising tumours. Eur J Cancer 2010;46(7):1271-7.

491. Hata T, Furukawa T, Sunamura M, Egawa S, Motoi F, *et al.* RNA interference targeting aurora kinase a suppresses tumor growth and enhances the taxane chemosensitivity in human pancreatic cancer cells. Cancer Res 2005;65(7):2899-905.

492. Furukawa T, Kanai N, Shiwaku HO, Soga N, Uehara A, *et al.* AURKA is one of the downstream targets of MAPK1/ERK2 in pancreatic cancer. Oncogene 2006;25(35):4831-9.

493. Zhu J, Abbruzzese JL, Izzo J, Hittelman WN, Li D. AURKA amplification, chromosome instability, and centrosome abnormality in human pancreatic carcinoma cells. Cancer Genet Cytogenet 2005;159(1):10-7.

494. Sommer EM, Dry H, Cross D, Guichard S, Davies BR, *et al.* Elevated SGK1 predicts resistance of breast cancer cells to Akt inhibitors. Biochem J 2013;452(3):499-508.

495. Gray S, Pandha HS, Michael A, Middleton G, Morgan R. HOX genes in pancreatic development and cancer. JOP 2011;12(3):216-9.

496. Cui Z, Ren S, Lu J, Wang F, Xu W, *et al.* The prostate cancer-up-regulated long noncoding RNA PlncRNA-1 modulates apoptosis and proliferation through reciprocal regulation of androgen receptor. Urol Oncol 2012;31(7):1117-23.

497. Chile T, Fortes MA, Correa-Giannella ML, Brentani HP, Maria DA, *et al.* HOXB7 mRNA is overexpressed in pancreatic ductal adenocarcinomas and its knockdown induces cell cycle arrest and apoptosis. BMC Cancer 2013;13:451.

498. Nguyen Kovochich A, Arensman M, Lay AR, Rao NP, Donahue T, *et al.* HOXB7 promotes invasion and predicts survival in pancreatic adenocarcinoma. Cancer 2013;119(3):529-39.

499. Segara D, Biankin AV, Kench JG, Langusch CC, Dawson AC, *et al.* Expression of HOXB2, a retinoic acid signaling target in pancreatic cancer and pancreatic intraepithelial neoplasia. Clin Cancer Res 2005;11(9):3587-96.

500. Yamazaki K, Masugi Y, Effendi K, Tsujikawa H, Hiraoka N, *et al.* Upregulated SMAD3 promotes epithelial-mesenchymal transition and predicts poor prognosis in pancreatic ductal adenocarcinoma. Lab Invest 2014;94(6):683-91.

501. Yilmaz N, Ozaksit G, Terzi YK, Yilmaz S, Budak B, *et al.* HOXA11 and MMP2 gene expression in uterosacral ligaments of women with pelvic organ prolapse. J Turk Ger Gynecol Assoc 2014;15(2):104-8.

502. Zhang X, Lian Z, Padden C, Gerstein MB, Rozowsky J, *et al.* A myelopoiesis-associated regulatory intergenic noncoding RNA transcript within the human HOXA cluster. Blood 2009;113(11):2526-34.

503. Bertani S, Sauer S, Bolotin E, Sauer F. The noncoding RNA Mistral activates Hoxa6 and Hoxa7 expression and stem cell differentiation by recruiting MLL1 to chromatin. Mol Cell 2011;43(6):1040-6.

504. Maamar H, Cabili MN, Rinn J, Raj A. linc-HOXA1 is a noncoding RNA that represses Hoxa1 transcription in cis. Genes Dev 2013;27(11):1260-71.

505. Zhi YH, Song MM, Wang PL, Zhang T, Yin ZY. Suppression of matrix metalloproteinase-2 via RNA interference inhibits pancreatic carcinoma cell invasiveness and adhesion. World J Gastroenterol 2009;15(9):1072-8.

506. Vihinen P, Kahari VM. Matrix metalloproteinases in cancer: prognostic markers and therapeutic targets. Int J Cancer 2002;99(2):157-66.

507. Roy R, Yang J, Moses MA. Matrix metalloproteinases as novel biomarkers and potential therapeutic targets in human cancer. J Clin Oncol 2009;27(31):5287-97.

508. Huarte M. The emerging role of lncRNAs in cancer. Nat Med 2015;21(11):1253-61.

509. Adams BD, Anastasiadou E, Esteller M, He L, Slack FJ. The inescapable influence of noncoding RNAs in cancer. Cancer Res 2015;75(24):5206-10.

510. Mendell JT, Olson EN. MicroRNAs in stress signaling and human disease. Cell 2012;148(6):1172-87.

511. Cheng Y, Imanirad P, Jutooru I, Hedrick E, Jin UH, *et al.* Role of metastasis-associated lung adenocarcinoma transcript-1 (MALAT-1) in pancreatic cancer. Carcinogenesis (In review) 2016.

512. Ji P, Diederichs S, Wang W, Boing S, Metzger R, *et al.* MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. Oncogene 2003;22(39):8031-41.

513. Schmidt LH, Spieker T, Koschmieder S, Schaffers S, Humberg J, *et al.* The long noncoding MALAT-1 RNA indicates a poor prognosis in non-small cell lung

cancer and induces migration and tumor growth. J Thorac Oncol 2011;6(12):1984-92.

514. Lin R, Maeda S, Liu C, Karin M, Edgington TS. A large noncoding RNA is a marker for murine hepatocellular carcinomas and a spectrum of human carcinomas. Oncogene 2007;26(6):851-8.

515. Zheng HT, Shi DB, Wang YW, Li XX, Xu Y, *et al.* High expression of lncRNA MALAT1 suggests a biomarker of poor prognosis in colorectal cancer. Int J Clin Exp Pathol 2014;7(6):3174-81.

516. Okugawa Y, Toiyama Y, Hur K, Toden S, Saigusa S, *et al.* Metastasis-associated long non-coding RNA drives gastric cancer development and promotes peritoneal metastasis. Carcinogenesis 2014;35(12):2731-9.

517. Ma KX, Wang HJ, Li XR, Li T, Su G, *et al.* Long noncoding RNA MALAT1 associates with the malignant status and poor prognosis in glioma. Tumour Biol 2015;36(5):3355-9.

518. Li S, Wang Q, Qiang Q, Shan H, Shi M, *et al.* Sp1-mediated transcriptional regulation of MALAT1 plays a critical role in tumor. J Cancer Res Clin Oncol 2015;141(11):1909-20.

519. Hirata H, Hinoda Y, Shahryari V, Deng G, Nakajima K, *et al.* Long noncoding RNA MALAT1 promotes aggressive renal cell carcinoma through Ezh2 and interacts with miR-205. Cancer Res 2015;75(7):1322-31.

520. Tano K, Mizuno R, Okada T, Rakwal R, Shibato J, *et al.* MALAT-1 enhances cell motility of lung adenocarcinoma cells by influencing the expression of motility-related genes. FEBS Lett 2010;584(22):4575-80.

521. Gutschner T, Hammerle M, Eissmann M, Hsu J, Kim Y, *et al.* The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. Cancer Res 2013;73(3):1180-9.

522. Zong X, Tripathi V, Prasanth KV. RNA splicing control: yet another gene regulatory role for long nuclear noncoding RNAs. RNA Biol 2011;8(6):968-77.

523. Lin R, Roychowdhury-Saha M, Black C, Watt AT, Marcusson EG, *et al.* Control of RNA processing by a large non-coding RNA over-expressed in carcinomas. FEBS Lett 2011;585(4):671-6.

524. Miyagawa R, Tano K, Mizuno R, Nakamura Y, Ijiri K, *et al.* Identification of cis- and trans-acting factors involved in the localization of MALAT-1 noncoding RNA to nuclear speckles. RNA 2012;18(4):738-51.

525. Eissmann M, Gutschner T, Hammerle M, Gunther S, Caudron-Herger M, *et al.* Loss of the abundant nuclear non-coding RNA MALAT1 is compatible with life and development. RNA Biol 2012;9(8):1076-87.

526. Zhang B, Arun G, Mao YS, Lazar Z, Hung G, *et al.* The lncRNA Malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult. Cell Rep 2012;2(1):111-23.

527. Nakagawa S, Ip JY, Shioi G, Tripathi V, Zong X, *et al.* Malat1 is not an essential component of nuclear speckles in mice. RNA 2012;18(8):1487-99.

528. Bernard D, Prasanth KV, Tripathi V, Colasse S, Nakamura T, *et al.* A long nuclear-retained non-coding RNA regulates synaptogenesis by modulating gene expression. EMBO J 2010;29(18):3082-93.

529. Yang L, Lin C, Liu W, Zhang J, Ohgi KA, *et al.* ncRNA- and Pc2 methylation-dependent gene relocation between nuclear structures mediates gene activation programs. Cell 2011;147(4):773-88.

530. Fan Y, Shen B, Tan M, Mu X, Qin Y, *et al.* TGF-beta-induced upregulation of malat1 promotes bladder cancer metastasis by associating with suz12. Clin Cancer Res 2014;20(6):1531-41.

531. Zhao Y, Yang Y, Trovik J, Sun K, Zhou L, *et al.* A novel wnt regulatory axis in endometrioid endometrial cancer. Cancer Res 2014;74(18):5103-17.

532. Shen L, Chen L, Wang Y, Jiang X, Xia H, *et al.* Long noncoding RNA MALAT1 promotes brain metastasis by inducing epithelial-mesenchymal transition in lung cancer. J Neurooncol 2015;121(1):101-8.

533. Ji Q, Liu X, Fu X, Zhang L, Sui H, *et al.* Resveratrol inhibits invasion and metastasis of colorectal cancer cells via MALAT1 mediated Wnt/beta-catenin signal pathway. PLoS One 2013;8(11):e78700.

534. Li B, Chen P, Qu J, Shi L, Zhuang W, *et al.* Activation of LTBP3 gene by a long noncoding RNA (lncRNA) MALAT1 transcript in mesenchymal stem cells from multiple myeloma. J Biol Chem 2014;289(42):29365-75.

535. West JA, Davis CP, Sunwoo H, Simon MD, Sadreyev RI, *et al.* The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. Mol Cell 2014;55(5):791-802.

536. Han Y, Liu Y, Nie L, Gui Y, Cai Z. Inducing cell proliferation inhibition, apoptosis, and motility reduction by silencing long noncoding ribonucleic acid metastasis-associated lung adenocarcinoma transcript 1 in urothelial carcinoma of the bladder. Urology 2013;81(1):209 e1-7.

537. Hansen GM, Markesich DC, Burnett MB, Zhu Q, Dionne KM, *et al.* Large-scale gene trapping in C57BL/6N mouse embryonic stem cells. Genome Res 2008;18(10):1670-9.

538. Hogan B, Beddington R, Costantini F, Lacy E. Manipulating the mouse embryo: a laboratory manual. Plainview (USA): Cold Spring Harbor Laboratory Press; 1994. p 497.

539. Maruyama Y, Ono M, Kawahara A, Yokoyama T, Basaki Y, *et al.* Tumor growth suppression in pancreatic cancer by a putative metastasis suppressor gene Cap43/NDRG1/Drg-1 through modulation of angiogenesis. Cancer Res 2006;66(12):6233-42.

540. Hosoi F, Izumi H, Kawahara A, Murakami Y, Kinoshita H, *et al.* N-myc downstream regulated gene 1/Cap43 suppresses tumor growth and angiogenesis of pancreatic cancer through attenuation of inhibitor of kappaB kinase beta expression. Cancer Res 2009;69(12):4983-91.

541. Angst E, Dawson DW, Stroka D, Gloor B, Park J, *et al.* N-myc downstream regulated gene-1 expression correlates with reduced pancreatic cancer growth and increased apoptosis in vitro and in vivo. Surgery 2011;149(5):614-24.

542. Abdelrahim M, Baker CH, Abbruzzese JL, Safe S. Tolfenamic acid and pancreatic cancer growth, angiogenesis, and Sp protein degradation. J Natl Cancer Inst 2006;98(12):855-68.

543. Jutooru I, Chadalapaka G, Lei P, Safe S. Inhibition of NFkappaB and pancreatic cancer cell and tumor growth by curcumin is dependent on specificity protein down-regulation. J Biol Chem 2010;285(33):25332-44.

544. Jutooru I, Chadalapaka G, Abdelrahim M, Basha MR, Samudio I, *et al.* Methyl 2-cyano-3,12-dioxooleana-1,9-dien-28-oate decreases specificity protein transcription factors and inhibits pancreatic tumor growth: role of microRNA-27a. Mol Pharmacol 2010;78(2):226-36.

545. Jutooru I, Guthrie AS, Chadalapaka G, Pathi S, Kim K, *et al.* Mechanism of action of phenethylisothiocyanate and other reactive oxygen species-inducing anticancer agents. Mol Cell Biol 2014;34(13):2382-95.

546. Hedrick E, Cheng Y, Jin UH, Kim K, Safe S. Specificity protein (Sp) transcription factors Sp1, Sp3 and Sp4 are non-oncogene addiction genes in cancer cells. Oncotarget 2016;7(16):22245-56.

547. Guerra C, Barbacid M. Genetically engineered mouse models of pancreatic adenocarcinoma. Mol Oncol 2013;7(2):232-47.

548. Liby KT, Royce DB, Risingsong R, Williams CR, Maitra A, *et al.* Synthetic triterpenoids prolong survival in a transgenic mouse model of pancreatic cancer. Cancer Prev Res (Phila) 2010;3(11):1427-34.

549. Gu YZ, Hogenesch JB, Bradfield CA. The PAS superfamily: sensors of environmental and developmental signals. Annu Rev Pharmacol Toxicol 2000;40:519-61.

550. Nebert DW, Goujon FM, Gielen JE. Aryl hydrocarbon hydroxylase induction by polycyclic hydrocarbons: simple autosomal dominant trait in the mouse. Nat New Biol 1972;236(65):107-10.

551. Mimura J, Yamashita K, Nakamura K, Morita M, Takagi TN, *et al.* Loss of teratogenic response to 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) in mice lacking the Ah (dioxin) receptor. Genes Cells 1997;2(10):645-54.

552. Whitlock JP, Jr. Induction of cytochrome P4501A1. Annu Rev Pharmacol Toxicol 1999;39:103-25.

553. Hankinson O. The aryl hydrocarbon receptor complex. Annu Rev Pharmacol Toxicol 1995;35:307-40.

554. Kim DW, Gazourian L, Quadri SA, Romieu-Mourez R, Sherr DH, *et al.* The RelA NF-kappaB subunit and the aryl hydrocarbon receptor (AhR) cooperate to transactivate the c-myc promoter in mammary cells. Oncogene 2000;19(48):5498-506.

555. Vogel CF, Sciullo E, Li W, Wong P, Lazennec G, *et al.* RelB, a new partner of aryl hydrocarbon receptor-mediated transcription. Mol Endocrinol 2007;21(12):2941-55.

556. Singh NP, Nagarkatti M, Nagarkatti PS. Role of dioxin response element and nuclear factor-kappaB motifs in 2,3,7,8-tetrachlorodibenzo-p-dioxin-mediated regulation of Fas and Fas ligand expression. Mol Pharmacol 2007;71(1):145-57.

557. Vogel CF, Khan EM, Leung PS, Gershwin ME, Chang WL, *et al.* Cross-talk between aryl hydrocarbon receptor and the inflammatory response: a role for nuclear factor-kappaB. J Biol Chem 2014;289(3):1866-75.

558. Jackson DP, Li H, Mitchell KA, Joshi AD, Elferink CJ. Ah receptor-mediated suppression of liver regeneration through NC-XRE-driven p21Cip1 expression. Mol Pharmacol 2014;85(4):533-41.

559. Blankenship A, Matsumura F. 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) causes an Ah receptor-dependent and ARNT-independent increase in membrane levels and activity of p60(Src). Environ Toxicol Pharmacol 1997;3(3):211-20.

560. Li W, Matsumura F. Significance of the nongenomic, inflammatory pathway in mediating the toxic action of TCDD to induce rapid and long-term cellular responses in 3T3-L1 adipocytes. Biochemistry 2008;47(52):13997-4008.

561. Dong B, Matsumura F. The conversion of rapid TCCD nongenomic signals to persistent inflammatory effects via select protein kinases in MCF10A cells. Mol Endocrinol 2009;23(4):549-58.

562. Denison MS, Soshilov AA, He G, DeGroot DE, Zhao B. Exactly the same but different: promiscuity and diversity in the molecular mechanisms of action of the aryl hydrocarbon (dioxin) receptor. Toxicol Sci 2011;124(1):1-22.

563. Cheng Y, Jin UH, Allred CD, Jayaraman A, Chapkin RS, *et al.* Aryl hydrocarbon receptor activity of tryptophan metabolites in young adult mouse colonocytes. Drug Metab Dispos 2015;43(10):1536-43.

564. Hu W, Sorrentino C, Denison MS, Kolaja K, Fielden MR. Induction of cyp1a1 is a nonspecific biomarker of aryl hydrocarbon receptor activation: results of large scale screening of pharmaceuticals and toxicants in vivo and in vitro. Mol Pharmacol 2007;71(6):1475-86.

565. Denison MS, Seidel SD, Rogers WJ, Ziccardi MH, Winter GM, *et al.* Natural and synthetic ligands for the Ah receptor. In: Puga A, Kendall KB, editors. Molecular biology approaches to toxicology. London (UK): Taylor and Francis; 1998. p 3-33.

566. Jeuken A, Keser BJ, Khan E, Brouwer A, Koeman J, *et al.* Activation of the Ah receptor by extracts of dietary herbal supplements, vegetables, and fruits. J Agric Food Chem 2003;51(18):5478-87.

567. Safe S, Chadalapaka G, Jutooru I. AHR-reactive compounds in the human diet. In: Pohjanvirta R, editor. The Ah receptor in biology and toxicology. Hoboken (USA): John Wiler & Sons; 2012. p 331-42.

568. Henry EC, Bemis JC, Henry O, Kende AS, Gasiewicz TA. A potential endogenous ligand for the aryl hydrocarbon receptor has potent agonist activity in vitro and in vivo. Arch Biochem Biophys 2006;450(1):67-77.

569. Stevens EA, Mezrich JD, Bradfield CA. The aryl hydrocarbon receptor: a perspective on potential roles in the immune system. Immunology 2009;127(3):299-311.

570. Marshall NB, Kerkvliet NI. Dioxin and immune regulation: emerging role of aryl hydrocarbon receptor in the generation of regulatory T cells. Ann N Y Acad Sci 2010;1183:25-37.

571. Kerkvliet NI. AHR-mediated immunomodulation: the role of altered gene transcription. Biochem Pharmacol 2009;77(4):746-60.

572. Busbee PB, Rouse M, Nagarkatti M, Nagarkatti PS. Use of natural AhR ligands as potential therapeutic modalities against inflammatory disorders. Nutr Rev 2013;71(6):353-69.

573. Arsenescu R, Arsenescu V, Zhong J, Nasser M, Melinte R, *et al.* Role of the xenobiotic receptor in inflammatory bowel disease. Inflamm Bowel Dis 2011;17(5):1149-62.

574. Singh NP, Singh UP, Singh B, Price RL, Nagarkatti M, *et al.* Activation of aryl hydrocarbon receptor (AhR) leads to reciprocal epigenetic regulation of FoxP3 and IL-17 expression and amelioration of experimental colitis. PLoS One 2011;6(8):e23522.

575. Furumatsu K, Nishiumi S, Kawano Y, Ooi M, Yoshie T, *et al.* A role of the aryl hydrocarbon receptor in attenuation of colitis. Dig Dis Sci 2011;56(9):2532-44.

576. Benson JM, Shepherd DM. Aryl hydrocarbon receptor activation by TCDD reduces inflammation associated with Crohn's disease. Toxicol Sci 2011;120(1):68-78.

577. Bansal T, Alaniz RC, Wood TK, Jayaraman A. The bacterial signal indole increases epithelial-cell tight-junction resistance and attenuates indicators of inflammation. Proc Natl Acad Sci U S A 2010;107(1):228-33.

578. Fukumoto S, Toshimitsu T, Matsuoka S, Maruyama A, Oh-Oka K, *et al.* Identification of a probiotic bacteria-derived activator of the aryl hydrocarbon receptor that inhibits colitis. Immunol Cell Biol 2014;92(5):460-5.

579. Li F, Hullar MA, Schwarz Y, Lampe JW. Human gut bacterial communities are altered by addition of cruciferous vegetables to a controlled fruit- and vegetable-free diet. J Nutr 2009;139(9):1685-91.

580. Venkatesh M, Mukherjee S, Wang H, Li H, Sun K, *et al.* Symbiotic bacterial metabolites regulate gastrointestinal barrier function via the xenobiotic sensor PXR and Toll-like receptor 4. Immunity 2014;41(2):296-310.

581. Jin UH, Lee SO, Sridharan G, Lee K, Davidson LA, *et al.* Microbiome-derived tryptophan metabolites and their aryl hydrocarbon receptor-dependent agonist and antagonist activities. Mol Pharmacol 2014;85(5):777-88.

582. D'Abaco GM, Whitehead RH, Burgess AW. Synergy between Apc min and an activated ras mutation is sufficient to induce colon carcinomas. Mol Cell Biol 1996;16(3):884-91.

583. Whitehead RH, VanEeden PE, Noble MD, Ataliotis P, Jat PS. Establishment of conditionally immortalized epithelial cell lines from both colon and small intestine of adult H-2Kb-tsA58 transgenic mice. Proc Natl Acad Sci U S A 1993;90(2):587-91.

584. Weige CC, Allred KF, Allred CD. Estradiol alters cell growth in nonmalignant colonocytes and reduces the formation of preneoplastic lesions in the colon. Cancer Res 2009;69(23):9118-24.

585. Turk HF, Kolar SS, Fan YY, Cozby CA, Lupton JR, *et al.* Linoleic acid and butyrate synergize to increase Bcl-2 levels in colonocytes. Int J Cancer 2011;128(1):63-71.

586. Kolar SS, Barhoumi R, Lupton JR, Chapkin RS. Docosahexaenoic acid and butyrate synergistically induce colonocyte apoptosis by enhancing mitochondrial Ca2+ accumulation. Cancer Res 2007;67(11):5561-8.

587. Baba T, Mimura J, Gradin K, Kuroiwa A, Watanabe T, *et al.* Structure and expression of the Ah receptor repressor gene. J Biol Chem 2001;276(35):33101-10.

588. Savas U, Bhattacharyya KK, Christou M, Alexander DL, Jefcoate CR. Mouse cytochrome P-450EF, representative of a new 1B subfamily of cytochrome P-

450s. Cloning, sequence determination, and tissue expression. J Biol Chem 1994;269(21):14905-11.

589. Diani-Moore S, Ram P, Li X, Mondal P, Youn DY, *et al.* Identification of the aryl hydrocarbon receptor target gene TiPARP as a mediator of suppression of hepatic gluconeogenesis by 2,3,7,8-tetrachlorodibenzo-p-dioxin and of nicotinamide as a corrective agent for this effect. J Biol Chem 2010;285(50):38801-10.

590. Lu YF, Santostefano M, Cunningham BD, Threadgill MD, Safe S. Substituted flavones as aryl hydrocarbon (Ah) receptor agonists and antagonists. Biochem Pharmacol 1996;51(8):1077-87.

591. Astroff B, Zacharewski T, Safe S, Arlotto MP, Parkinson A, *et al.* 6-Methyl-1,3,8-trichlorodibenzofuran as a 2,3,7,8- tetrachlorodibenzo-p-dioxin antagonist: inhibition of the induction of rat cytochrome P-450 isozymes and related monooxygenase activities. Mol Pharmacol 1988;33:231-6.

592. Zhou J, Gasiewicz TA. 3'-Methoxy-4'-nitroflavone, a reported aryl hydrocarbon receptor antagonist, enhances Cyp1a1 transcription by a dioxin responsive element-dependent mechanism. Arch Biochem Biophys 2003;416(1):68-80.

593. McDougal A, Wormke M, Calvin J, Safe S. Tamoxifen-induced antitumorigenic/antiestrogenic action synergized by a selective Ah receptor modulator. Cancer Res 2001;61:3901-7.

594. Jin UH, Lee SO, Safe S. Aryl hydrocarbon receptor (AHR)-active pharmaceuticals are selective AHR modulators in MDA-MB-468 and BT474 breast cancer cells. J Pharmacol Exp Ther 2012;343(2):333-41.

595. Choi EY, Lee H, Dingle RW, Kim KB, Swanson HI. Implications and development of AHR-based therapeutic agents. Mol Cell Pharmacol 2012;4(2):53-62.

596. Murray IA, Patterson AD, Perdew GH. Aryl hydrocarbon receptor ligands in cancer: friend and foe. Nat Rev Cancer 2014;14(12):801-14.

597. Cani PD, Delzenne NM. The gut microbiome as therapeutic target. Pharmacol Ther 2011;130(2):202-12.

598. Picard C, Fioramonti J, Francois A, Robinson T, Neant F, *et al.* Review article: bifidobacteria as probiotic agents -- physiological effects and clinical benefits. Aliment Pharmacol Ther 2005;22(6):495-512.

599. Kato K, Mizuno S, Umesaki Y, Ishii Y, Sugitani M, *et al.* Randomized placebo-controlled trial assessing the effect of bifidobacteria-fermented milk on active ulcerative colitis. Aliment Pharmacol Ther 2004;20(10):1133-41.

600. Furrie E, Macfarlane S, Kennedy A, Cummings JH, Walsh SV, *et al.* Synbiotic therapy *(Bifidobacterium longum*/Synergy 1) initiates resolution of inflammation in patients with active ulcerative colitis: a randomised controlled pilot trial. Gut 2005;54(2):242-9.

601. Ishikawa H, Matsumoto S, Ohashi Y, Imaoka A, Setoyama H, *et al.* Beneficial effects of probiotic bifidobacterium and galacto-oligosaccharide in patients with ulcerative colitis: a randomized controlled study. Digestion 2011;84(2):128-33.

602. Wildt S, Nordgaard I, Hansen U, Brockmann E, Rumessen JJ. A randomised double-blind placebo-controlled trial with *Lactobacillus acidophilus* La-5 and *Bifidobacterium animalis* subsp. lactis BB-12 for maintenance of remission in ulcerative colitis. J Crohns Colitis 2011;5(2):115-21.

603. Mori H, Sato Y, Taketomo N, Kamiyama T, Yoshiyama Y, *et al.* Isolation and structural identification of bifidogenic growth stimulator produced by *Propionibacterium freudenreichii*. J Dairy Sci 1997;80(9):1959-64.

604. Isawa K, Hojo K, Yoda N, Kamiyama T, Makino S, *et al.* Isolation and identification of a new bifidogenic growth stimulator produced by *Propionibacterium freudenreichii* ET-3. Biosci Biotechnol Biochem 2002;66(3):679-81.

605. Bentley R, Meganathan R. Biosynthesis of vitamin K (menaquinone) in bacteria. Microbiol Rev 1982;46(3):241-80.

606. Kang JE, Kim TJ, Moon GS. A novel *Lactobacillus casei* LP1 producing 1,4-dihydroxy-2-naphthoic acid, a bifidogenic growth stimulator. Prev Nutr Food Sci 2015;20(1):78-81.

607. Eom JE, Kwon SC, Moon GS. Detection of 1,4-dihydroxy-2-naphthoic acid from commercial *Makgeolli* products. Prev Nutr Food Sci 2012;17(1):83-6.

608. Okada Y, Tsuzuki Y, Miyazaki J, Matsuzaki K, Hokari R, *et al.* *Propionibacterium freudenreichii* component 1,4-dihydroxy-2-naphthoic acid (DHNA) attenuates dextran sodium sulphate induced colitis by modulation of bacterial flora and lymphocyte homing. Gut 2006;55(5):681-8.

609. Okada Y, Tsuzuki Y, Narimatsu K, Sato H, Ueda T, *et al.* 1,4-Dihydroxy-2-naphthoic acid from *Propionibacterium freudenreichii* reduces inflammation in interleukin-10-deficient mice with colitis by suppressing macrophage-derived proinflammatory cytokines. J Leukoc Biol 2013;94(3):473-80.

610. Nagata K, Inatsu S, Tanaka M, Sato H, Kouya T, *et al.* The bifidogenic growth stimulator inhibits the growth and respiration of *Helicobacter pylori*. Helicobacter 2010;15(5):422-9.

611. Mok CF, Xie CM, Sham KW, Lin ZX, Cheng CH. 1,4-Dihydroxy-2-naphthoic acid Induces apoptosis in human keratinocyte: potential application for psoriasis treatment. Evid Based Complement Alternat Med 2013;2013:792840.

612. Cheng Y, Jin UH, Davidson L, Chapkin R, Jayaraman A, *et al.* Microbial-derived 1,4-dihydroxy-2-naphthoic acid and related compounds as aryl hydrocarbon receptor agonists/antagonists: structure-activity relationships and receptor modeling. Toxicol Sci (Accepted).

613. Van den Berg M, Birnbaum LS, Denison M, De Vito M, Farland W, *et al.* The 2005 World Health Organization reevaluation of human and Mammalian toxic equivalency factors for dioxins and dioxin-like compounds. Toxicol Sci 2006;93(2):223-41.

614. Hubbard TD, Murray IA, Perdew GH. Indole and tryptophan metabolism: endogenous and dietary routes to Ah receptor activation. Drug Metab Dispos 2015;43(10):1522-35.

615. Soshilov AA, Denison M. DNA binding (Gel Retardation Assay) analysis for identification of aryl hydrocarbon (Ah) receptor agonists and antagonists. In: Caldwell GW, Yan Z, editors. Optimization in drug discovery: in vitro methods. New York (USA): Springer Science; 2014. p 207-19.

616. Yang J, Zhang Y. I-TASSER server: new development for protein structure and function predictions. Nucleic Acids Res 2015;43(W1):W174-81.

617. Pandini A, Soshilov AA, Song Y, Zhao J, Bonati L, *et al.* Detection of the TCDD binding-fingerprint within the Ah receptor ligand binding domain by structurally driven mutagenesis and functional analysis. Biochemistry 2009;48(25):5972-83.

618. Motto I, Bordogna A, Soshilov AA, Denison MS, Bonati L. New aryl hydrocarbon receptor homology model targeted to improve docking reliability. J Chem Inf Model 2011;51(11):2868-81.

619. Xing Y, Nukaya M, Satyshur KA, Jiang L, Stanevich V, *et al.* Identification of the Ah-receptor structural determinants for ligand preferences. Toxicol Sci 2012;129(1):86-97.

620. Bisson WH, Koch DC, O'Donnell EF, Khalil SM, Kerkvliet NI, *et al.* Modeling of the aryl hydrocarbon receptor (AhR) ligand binding domain and its utility in virtual ligand screening to predict new AhR ligands. J Med Chem 2009;52(18):5635-41.

621. Wu D, Potluri N, Lu J, Kim Y, Rastinejad F. Structural integration in hypoxia-inducible factors. Nature 2015;524(7565):303-8.

622. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 2010;31(2):455-61.

623. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG. ZINC: a free tool to discover chemistry for biology. J Chem Inf Model 2012;52(7):1757-68.

624. Brooks BR, Brooks CL, 3rd, Mackerell AD, Jr., Nilsson L, Petrella RJ, *et al.* CHARMM: the biomolecular simulation program. J Comput Chem 2009;30(10):1545-614.

625. Best RB, Zhu X, Shim J, Lopes PE, Mittal J, *et al.* Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. J Chem Theory Comput 2012;8(9):3257-73.

626. Tamamis P, Lopez de Victoria A, Gorham RD, Jr., Bellows-Peterson ML, Pierou P, *et al.* Molecular dynamics in drug design: new generations of compstatin analogs. Chem Biol Drug Des 2012;79(5):703-18.

627. Sambuy Y, De Angelis I, Ranaldi G, Scarino ML, Stammati A, *et al.* The Caco-2 cell line as a model of the intestinal barrier: influence of cell and culture-related factors on Caco-2 cell functional characteristics. Cell Biol Toxicol 2005;21(1):1-26.

628. Carney MW, Erwin K, Hardman R, Yuen B, Volz DC, *et al.* Differential developmental toxicity of naphthoic acid isomers in medaka (*Oryzias latipes*) embryos. Mar Pollut Bull 2008;57(6-12):255-66.

629. Tamamis P, Kieslich CA, Nikiforovich GV, Woodruff TM, Morikis D, *et al.* Insights into the mechanism of C5aR inhibition by PMX53 via implicit solvent molecular dynamics simulations and docking. BMC Biophys 2014;7:5.

630. Tamamis P, Floudas CA. Molecular recognition of CCR5 by an HIV-1 gp120 V3 loop. PLoS One 2014;9(4):e95767.

631. Tamamis P, Floudas CA. Elucidating a key component of cancer metastasis: CXCL12 (SDF-1alpha) binding to CXCR4. J Chem Inf Model 2014;54(4):1174-88.

632. Tamamis P, Floudas CA. Elucidating a key anti-HIV-1 and cancer-associated axis: the structure of CCL5 (Rantes) in complex with CCR5. Sci Rep 2014;4:5447.

633. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph 1996;14(1):33-8, 27-8.

634. Eargle J, Wright D, Luthey-Schulten Z. Multiple Alignment of protein structures and sequences for VMD. Bioinformatics 2006;22(4):504-6.

635. Tamamis P, Morikis D, Floudas CA, Archontis G. Species specificity of the complement inhibitor compstatin investigated by all-atom molecular dynamics simulations. Proteins 2010;78(12):2655-67.

636. Tamamis P, Pierou P, Mytidou C, Floudas CA, Morikis D, *et al.* Design of a modified mouse protein with ligand binding properties of its human analog by molecular dynamics simulations: the case of C3 inhibition by compstatin. Proteins 2011;79(11):3166-79.

637. Engle MJ, Goetz GS, Alpers DH. Caco-2 cells express a combination of colonocyte and enterocyte phenotypes. J Cell Physiol 1998;174(3):362-9.

638. De Waard PW, De Kok TM, Maas LM, Peijnenburg AA, Hoogenboom RL, *et al.* Influence of TCDD and natural Ah receptor agonists on benzo[a]pyrene-DNA adduct formation in the Caco-2 human colon cell line. Mutagenesis 2008;23(1):67-73.

639. Till M, Riebniger D, Schmitz HJ, Schrenk D. Potency of various polycyclic aromatic hydrocarbons as inducers of CYP1A1 in rat hepatocyte cultures. Chem Biol Interact 1999;117(2):135-50.

640. Wilhelm M, Hardt J, Schulz C, Angerer J, Human Biomonitoring Commission of the German Federal Environment A. New reference value and the background exposure for the PAH metabolites 1-hydroxypyrene and 1- and 2-naphthol in urine of the general population in Germany: basis for validation of human biomonitoring data in environmental medicine. Int J Hyg Environ Health 2008;211(3-4):447-53.

641. Sul D, Ahn R, Im H, Oh E, Kim JH, *et al.* Korea National Survey for Environmental Pollutants in the human body 2008: 1-hydroxypyrene, 2-naphthol, and cotinine in urine of the Korean population. Environ Res 2012;118:25-30.

642. Jain RB. Association between polycyclic aromatic hydrocarbons and thyroid function among males and females: data from NHANES 2007-2008. Int J Environ Health Res 2016;26(4):405-19.

643. Nethery E, Wheeler AJ, Fisher M, Sjodin A, Li Z, *et al.* Urinary polycyclic aromatic hydrocarbons as a biomarker of exposure to PAHs in air: a pilot study among pregnant women. J Expo Sci Environ Epidemiol 2012;22(1):70-81.

644. Sudakin DL, Smit E, Cardenas A, Harding A. Naphthalene biomarkers and relationship with hemoglobin and hematocrit in White, Black, and Hispanic adults: results from the 2003-2004 National Health and Nutrition Examination Survey. J Med Toxicol 2013;9(2):133-8.

645. Soshilov AA, Denison MS. Ligand promiscuity of aryl hydrocarbon receptor agonists and antagonists revealed by site-directed mutagenesis. Mol Cell Biol 2014;34(9):1707-19.

646. Soshilov A, Denison MS. Ligand displaces heat shock protein 90 from overlapping binding sites within the aryl hydrocarbon receptor ligand-binding domain. J Biol Chem 2011;286(40):35275-82.

647. Cui XP, Qin CK, Zhang ZH, Su ZX, Liu X, *et al.* HOXA10 promotes cell invasion and MMP-3 expression via TGFbeta2-mediated activation of the p38 MAPK pathway in pancreatic cancer cells. Dig Dis Sci 2014;59(7):1442-51.

648. Sang Y, Zhou F, Wang D, Bi X, Liu X, *et al.* Up-regulation of long non-coding HOTTIP functions as an oncogene by regulating HOXA13 in non-small cell lung cancer. Am J Transl Res 2016;8(5):2022-32.

649. Wang J, Su L, Chen X, Li P, Cai Q, *et al.* MALAT1 promotes cell proliferation in gastric cancer by recruiting SF2/ASF. Biomed Pharmacother 2014;68(5):557-64.

650. Lu H, He Y, Lin L, Qi Z, Ma L, *et al.* Long non-coding RNA MALAT1 modulates radiosensitivity of HR-HPV+ cervical cancer via sponging miR-145. Tumour Biol 2016;37(2):1683-91.

651. Jin U-H, Lee S-O, Sridharan G, Lee K, Davidson LA, *et al.* Microbiome-drived tryptophan metabolites and their aryl hydrocarbon receptor-dependent agonist and antagonist activities. Mol Pharmacol 2014;85(5):777-88.

652. Vanommeslaeghe K, Hatcher E, Acharya C, Kundu S, Zhong S, *et al.* CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. J Comput Chem 2010;31(4):671-90.

653. Im W, Lee MS, Brooks CL, 3rd. Generalized born model with a simple smoothing function. J Comput Chem 2003;24(14):1691-702.

654. Jo S, Kim T, Iyer VG, Im W. CHARMM-GUI: a web-based graphical user interface for CHARMM. J Comput Chem 2008;29(11):1859-65.

655. Lee J, Cheng X, Swails JM, Yeom MS, Eastman PK, *et al.* CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. J Chem Theory Comput 2016;12(1):405-13.

656. Tamamis P, Floudas CA. Molecular recognition of CXCR4 by a dual tropic HIV-1 gp120 V3 loop. Biophys J 2013;105(6):1502-14.

657. Massova I, Kollman PA. Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. Perspect Drug Discov 2000;18(1):113-35.

658. Chen J, Im W, Brooks CL, 3rd. Balancing solvation and intramolecular interactions: toward a consistent generalized Born force field. J Am Chem Soc 2006;128(11):3728-36.

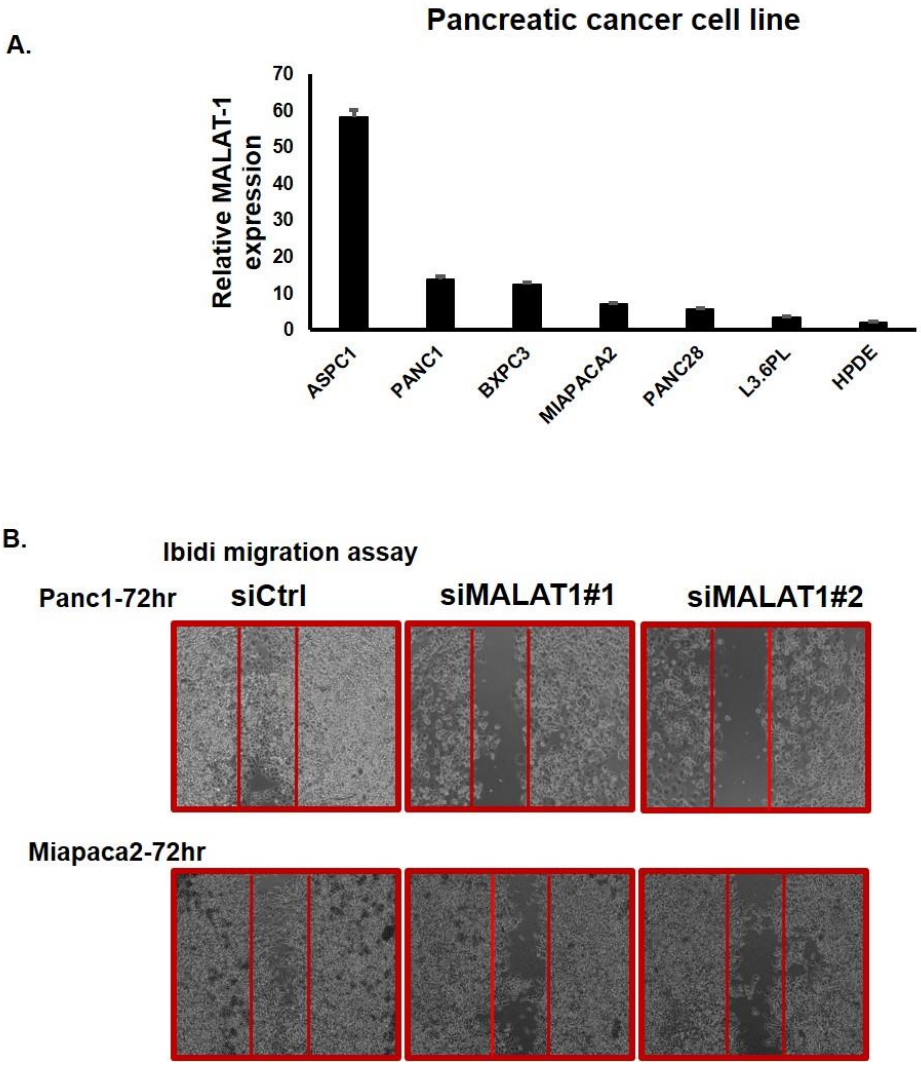
659. Gohlke H, Case DA. Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. J Comput Chem 2004;25(2):238-50.

660. Srinivasan J, Miller J, Kollman PA, Case DA. Continuum solvent studies of the stability of RNA hairpin loops and helices. J Biomol Struct Dyn 1998;16(3):671-82.

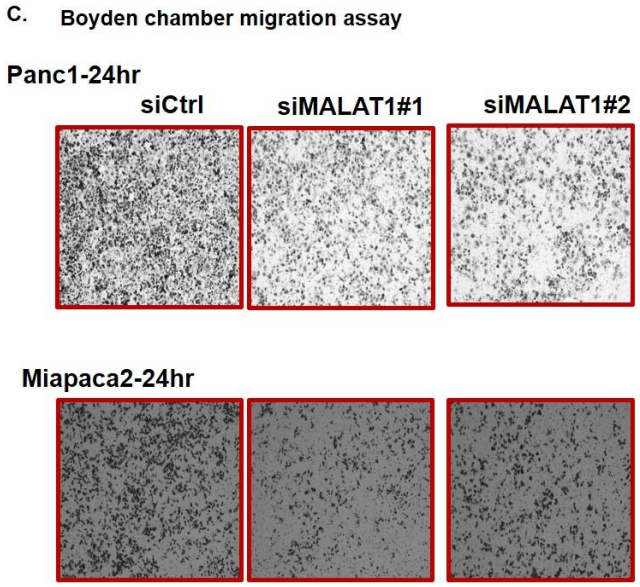
661. Hayes J, Archontis G. MM-GB(PB)SA calculations of protein-ligand binding free energies. In: Wang L, editor. Molecular dynamics - studies of synthetic and biological macromolecules. Rijeka (Croatia): InTech Press; 2012. p 171-90.

APPENDIX A

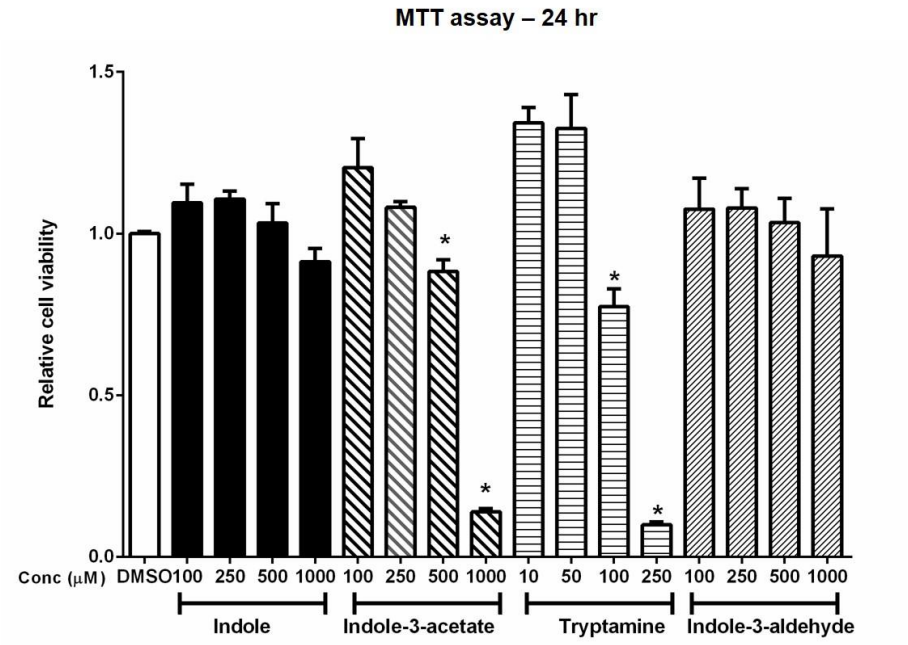
FIGURES



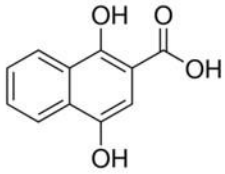
A-1. (A) The relative expression of MALAT1 in pancreatic cancer cell lines was determined by real time RT-PCR. MALAT1 knockdown reduced cell migration (B) and cell invasion (C) as determined by Boyden chamber assay (B) and Ibidi assay (C), respectively.



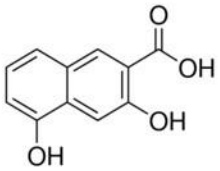
A-1. Continued.



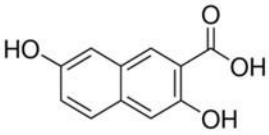
A-2. Cells were treated with different concentrations of the tryptophan metabolites for 24 hr and assayed for metabolic activity using the MTT assay. Significantly ($p<0.05$) decreased activity is indicated (*).



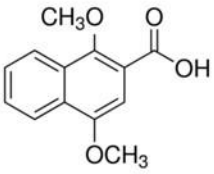
1.4-dihydroxy-2-naphthoic acid (1,4-DHNA)



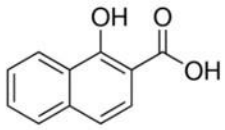
3,5-dihydroxy-2-naphthoic acid (3,5-DHNA)



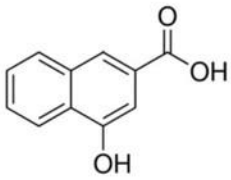
3,7-dihydroxy-2-naphthoic acid (3,7-DHNA)



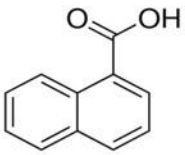
1,4-dimethoxy-2-naphthoic acid (1,4-DMNA)



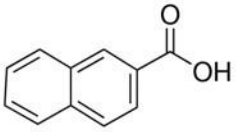
1-hydroxy-2-naphthoic acid (1-HNA)



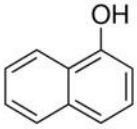
4-hydroxy-2-naphthoic acid (4-HNA)



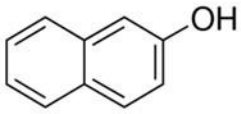
1-naphthoic acid (1-NA)



2-naphthoic acid (2-NA)

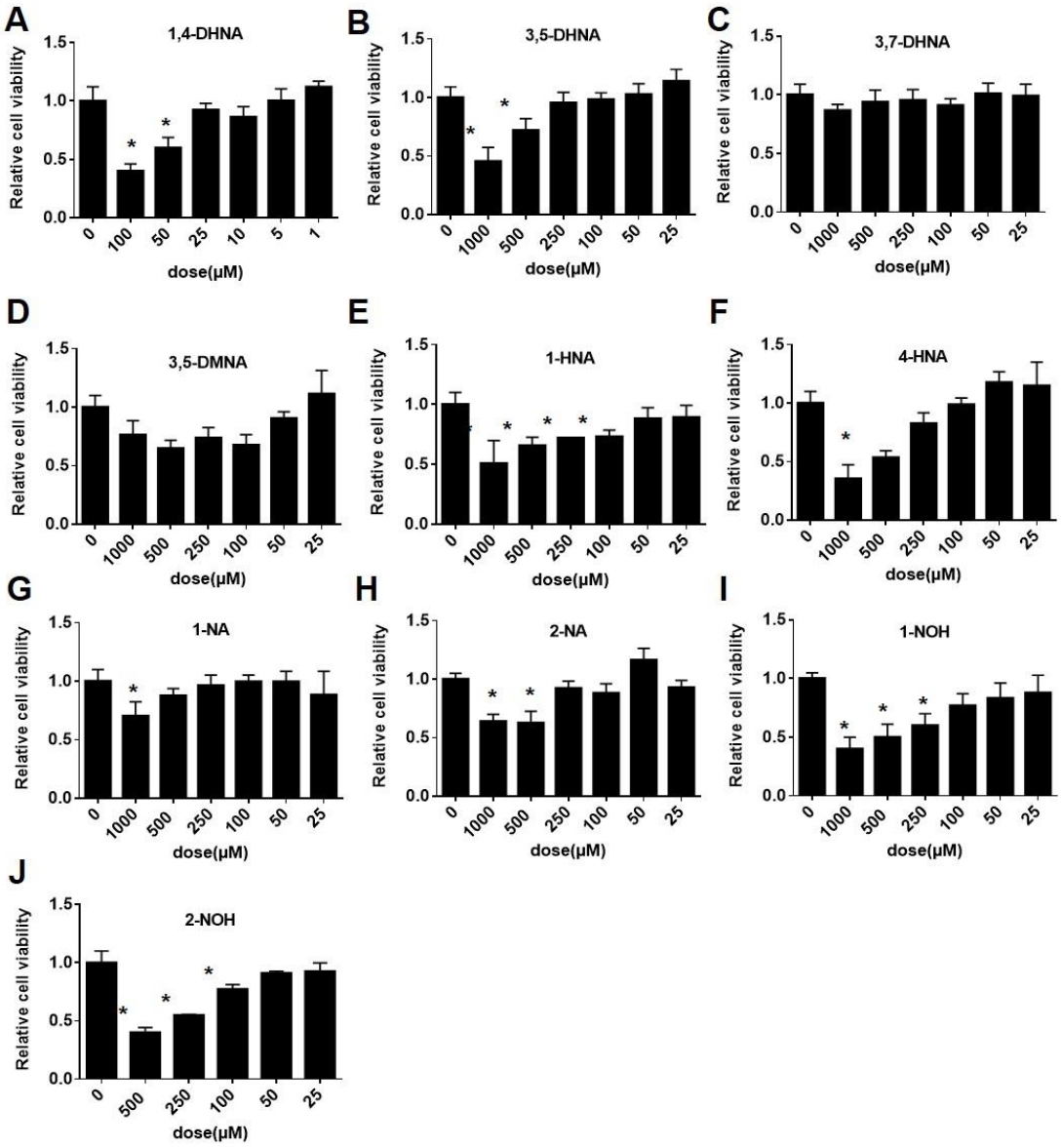


1-naphthol (1-NOH)

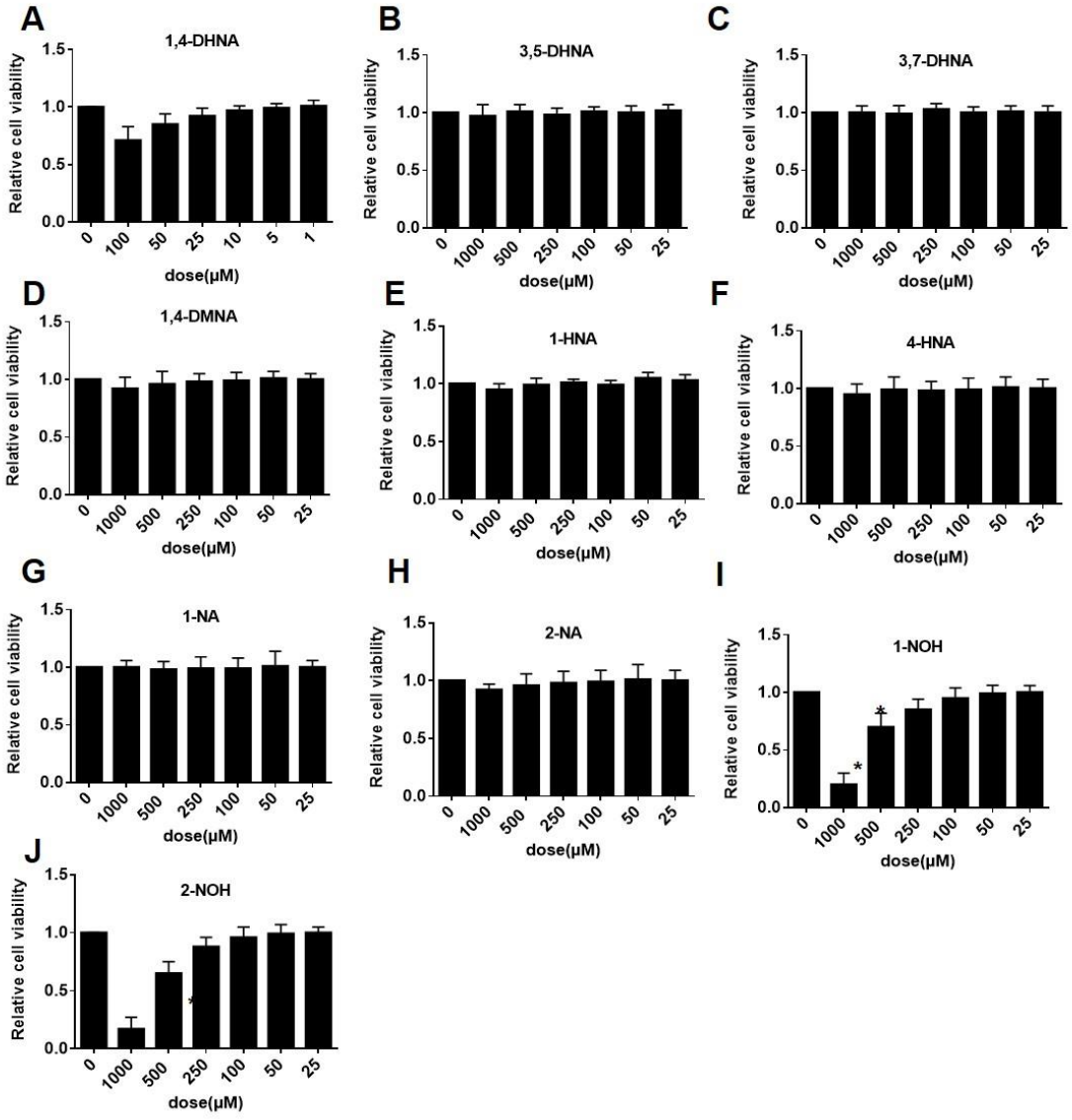


2-naphthol (2-NOH)

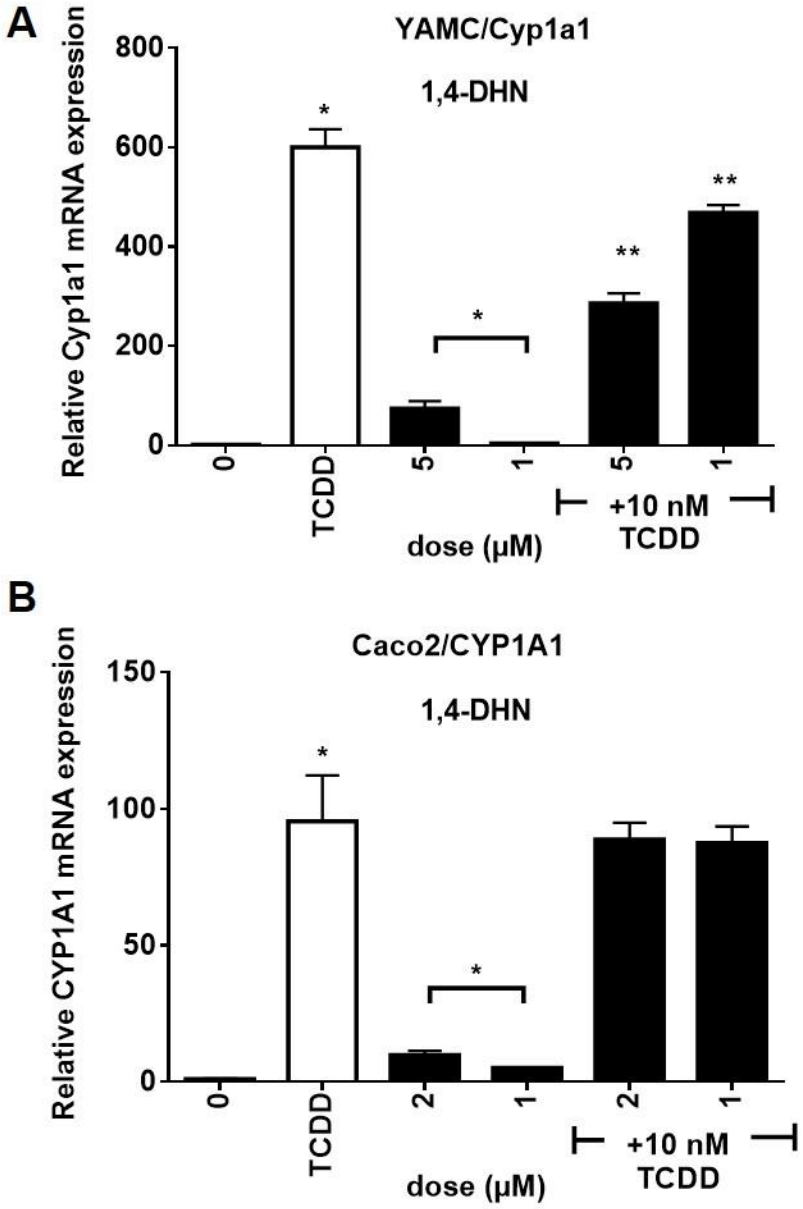
A-3. Structures, names and acronyms of the hydroxyl/carboxy naphthalene analogs used in this study.



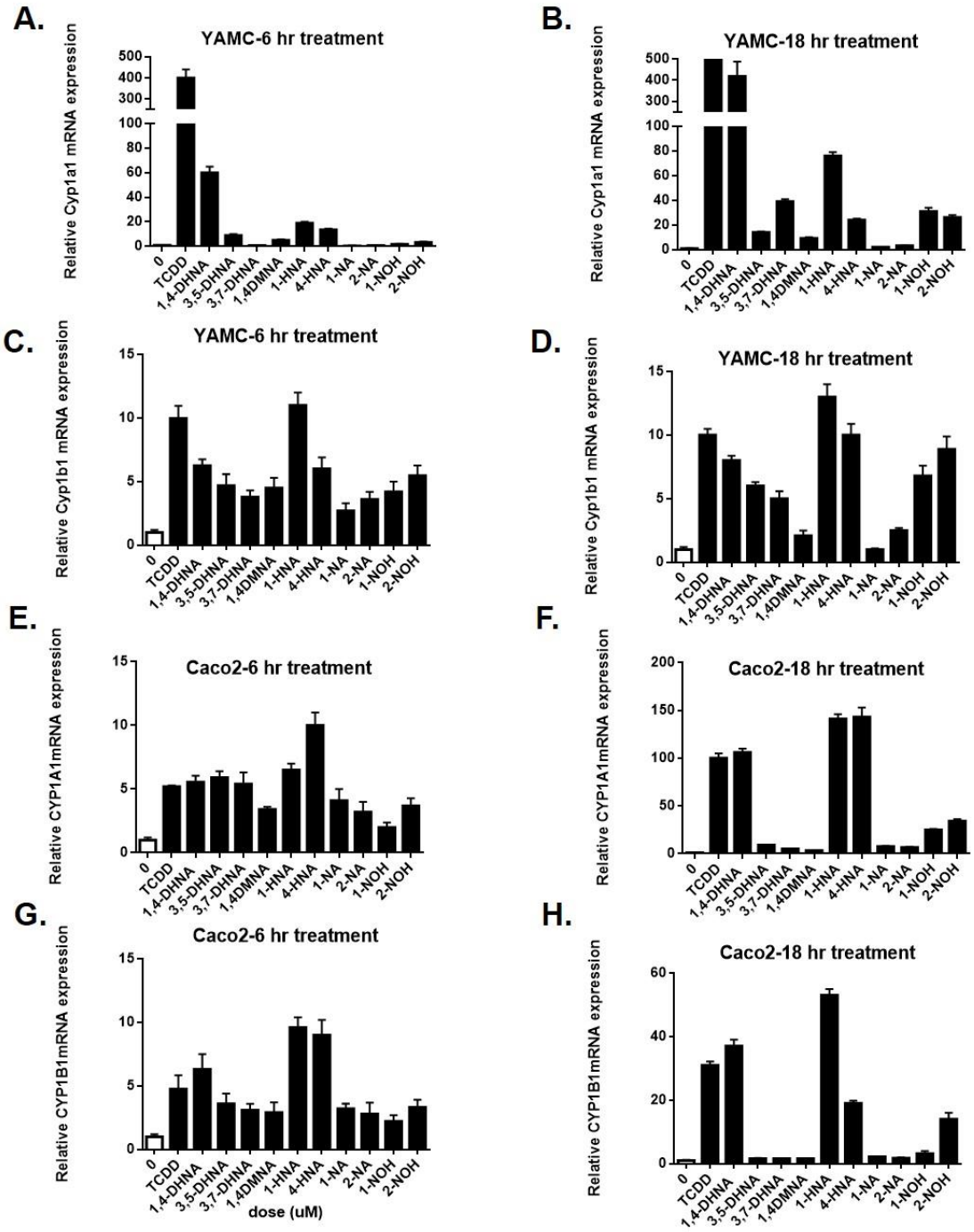
A-4. Cytotoxicity in YAMC cells. The effects of 1,4-DHNA and related compounds on cell viability were carried out as outlined in the Materials and Methods of Section V.



A-5. Cytotoxicity in Caco2 cells. The effects of 1,4-DHNA and related compounds on cell viability were carried out as outlined in the Materials and Methods of Section V.



A-6. 1,4-Dihydroxynaphthalene as an AhR agonist/antagonist. YAMC (A) or Caco2 (B) cells were treated with 1,4-dihydroxynaphthalene alone or in combination with 10 nM TCDD and *Cyp1a1*/*CYP1A1* mRNA levels were determined by real time PCR.



A-7. Comparative induction after 6 and 18 hr. Cells were treated with TCDD and maximal inducing concentrations of the naphthoic acid derivatives and *CYP1A1*/*CYP1B1* mRNA levels were determined by real time PCR as outlined in the Materials and Methods.

APPENDIX B

TABLES

B-1. The siRNA complexes used in this study.

Name	Sequence
siGL2	CGU ACG CGG AAU ACU UCG A
siHOTTIP	GCACAGAGAUAAUGGCAAAUU
siMLL	SASI_Hs01_00090459
siWDR5#1	SASI_Hs01_00046875
siWDR5#2	SASI_Hs01_00046876
siAURKA#1	SASI_Hs01_00241476
siAURKA#2	SASI_Hs01_00241477

B-2. Primers used for real time-PCR.

Name	Forward Primer	Reverse Primer
TBP	TGCACAGGAGCCAAGAGTGAA	CACATCACAGCTCCCCACCA
AHNAK	CTGAAGTGGTTCTGAGCGG	TCCACTCCATCTTCCGACTT
HOXA13	GGATATCAGCCACGACGAAT	ATTATCTGGGCAAAGCAACG
HOXA11	TGCCAAGTTGTACTTACTACGTC	GTTGGAGGAGTAGGAGTATGTCA
HOXB7	CGAGTTCCTTCAACATGCACT	TTTGCGGTCAGTTCCTGAGC
HOXA9	CTGTCCCACGCTTGACACTC	CTCCGCCGCTCTCATTCTC
HOXA10	TGGCTCACGGCAAAGAGTG	GCTGCGGCTAATCTCTAGGC
SMAD3	TGGACGCAGGTTCTCCAAAC	CCGGCTCGCAGTAGGTAAC
MMP3	GAGCTAAGTAAAGCCAGTGGA	GATATTTCTGAACAAGGTTCATCGT
HOTTIP	CCTAAAGCCACGCTTCTTTG	TGCAGGCTGGAGATCCTAGT
SGK1	GCGCTAACGTCTTTTCTGTCT	TGCTTCATGAAAGCGATGAG
GDF15	CTCCAGATTCCGAGAGTTGC	CACTTCTGGTGAGTATCC
MMP2	CCCACTGCGGTTTTCTCGAAT	CAAAGGGGTATCCATCGCCAT
CD44	TGCTACCAGAGACCAAGACA	CCCATGTGAGTGTCCATCTG
AURKA	GTCAAGTCCCCTGTCGGTT	AGTGAGACCCTCTAGCTGT

B-2. Continued.

Name	Forward Primer	Reverse Primer
HOXA1	TCCTGGAATACCCCATACTTAGC	GCACGACTGGAAAGTTGTAATCC
IGF2BP3	ATTTACAGTGGGAGGTGCTG	GCAGTTTCCGAGTCAGTGTT
TM4SF1	TGGTTCTTTTCTGGCATCGT	AGAAAGCATCGCACATCGTT
PAK2	GCCAAAGAATTATTACAGCATCC	TGCTTCTTTAGCTGCCATGA

B-3. Common genes overexpressed in pancreatic tumors (GSE16515) and downregulated in Panc1 cells transfected with siHOTTIP.

Gene	Entrez Gene ID	Gene Description	GSE16515 Pancreatic tumor _vs_ paired adjacent normal pancreatic tissue (Study: Pancreatic tumor compared to normal pancreatic tissue)	HOTTIP regulated gene list
TM4SF1	4071	transmembrane 4 L six family member 1	4.19	-5.21
SMS	6611	spermine synthase	2.1	-4.8
IGF2BP3	10643	insulin-like growth factor 2 mRNA binding protein 3	8.48	-4.38
NCEH1	57552	neutral cholesterol ester hydrolase 1	2.45	-4.06
UBASH3B	84959	ubiquitin associated and SH3 domain containing B	2.37	-4.02
PLAU	5328	plasminogen activator, urokinase	6.39	-3.69
NT5E	4907	5'-nucleotidase, ecto (CD73)	2.94	-3.65
PAK2	5062	p21 protein (Cdc42/Rac)-activated kinase 2	2.22	-3.62
AURKA	6790	aurora kinase A	2.52	-3.56
LDLR	3949	low density lipoprotein receptor	2.38	-3.43
F2RL1	2150	coagulation factor II (thrombin) receptor-like 1	2.29	-3.38

B-3. Continued.

Gene	Entrez Gene ID	Gene Description	GSE16515 Pancreatic tumor _vs_ paired adjacent normal pancreatic tissue (Study: Pancreatic tumor compared to normal pancreatic tissue)	HOTTIP regulated gene list
SLC44A1	23446	solute carrier family 44, member 1	2.32	-3.34
DORA2B	136	adenosine A2b receptor	3.45	-3.2
PNMA2	10687	paraneoplastic Ma antigen 2	2.02	-2.88
ALDOC	230	aldolase C, fructose-bisphosphate	2.27	-2.67
ERO1L	30001	ERO1-like (S. cerevisiae)	3.87	-2.52
ITGB1	3688	integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29 includes MDF2, MSK12)	2.11	-2.49
YWHAZ	7534	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide	2.37	-2.42
RAB31	11031	RAB31, member RAS oncogene family	3.39	-2.4
PALLD	23022	palladin, cytoskeletal associated protein	3.09	-2.39
ELF4	2000	E74-like factor 4 (ets domain transcription factor)	2.63	-2.38
KIF11	3832	kinesin family member 11	3.12	-2.36
DPY19L1	23333	dpy-19-like 1 (C. elegans)	2.69	-2.35
SEMA4B	10509	sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4B	2.48	-2.32
PHLDA1	22822	pleckstrin homology-like domain, family A, member 1	2.45	-2.32
CD97	976	CD97 molecule	2.32	-2.31

B-3. Continued.

Gene	Entrez Gene ID	Gene Description	GSE16515 Pancreatic tumor _vs_ paired adjacent normal pancreatic tissue (Study: Pancreatic tumor compared to normal pancreatic tissue)	HOTTIP regulated gene list
SLC16A3	9123	solute carrier family 16, member 3 (monocarboxylic acid transporter 4)	8.25	-2.29
ECT2	1894	epithelial cell transforming sequence 2 oncogene	5.83	-2.28
TMEM189	387521	transmembrane protein 189	2.05	-2.28
CALML4	91860	calmodulin-like 4	2.74	-2.24
SDC1	6382	syndecan 1	6.28	-2.18
ITGB5	3693	integrin, beta 5	2.34	-2.12
CCNB1	891	cyclin B1	4.14	-2.09
SLC6A10P	386757	solute carrier family 6 (neurotransmitter transporter, creatine), member 10, pseudogene	2.64	-2.09
NUAK1	9891	NUAK family, SNF1-like kinase, 1	2.74	-2.04
CLIC1	1192	chloride intracellular channel 1	2.7	-2.04
ENO2	2026	enolase 2 (gamma, neuronal)	3.91	-2.01
MALL	7851	mal, T-cell differentiation protein-like	6.77	-2
KIF2C	11004	kinesin family member 2C	2.65	-2

B-4. The comparison of differentially expressed gene in Panc1 cells and Malat1 transgenic mouse tumors.
Mouse 1251 genes differentially expressed ($p<0.05$, fold >1.5)
Panc1 890 genes differentially expressed ($p<0.05$, fold >1.5)
Common genes 50

Symbol	Exp Fold Change(mouse)	Exp Fold Change(human)
WISP2	-8.886	-1.653
TBX2	-7.193	-3.774
CPXM1	-6.869	-1.858
OLFM1	-6.733	-1.861
BCAS1	-6.492	-1.531
MDK	-5.778	-2.014
S1PR3	-4.863	-1.652
IGFBP5	-4.319	-2.924
TFF1	-4.250	-16.009
MYLIP	-4.073	-1.508
TSPAN6	-4.024	-1.903
FOXA1	-3.961	-6.948
PPM1H	-3.867	-1.786
FAM46A	-3.863	-3.324
TBC1D9	-3.717	-1.849
TACSTD2	-3.599	-13.771
ASS1	-3.285	-1.573
INHBB	-3.225	-1.646
LHFPL2	-3.173	-1.523
TSKU	-3.122	-4.661
NOTCH3	-3.117	-2.204
FAM174B	-3.022	-2.162
STAT1	-2.992	-1.514
LYN	-2.910	-1.530
UNC93B1	-2.897	-1.537
SLC7A2	-2.821	-3.742
PARP14	-2.759	-1.786
CEBPA	-2.744	-1.613
DKK3	-2.669	-2.378
ALDH1A3	-2.528	-1.624
SERPINA3	-2.508	-4.231
MAL2	-2.465	-4.240
PMEPA1	-2.448	-1.502
TAGLN	-2.432	-1.843
GPD1L	-2.223	-2.938

B-4. Continued.

Symbol	Exp Fold Change(mouse)	Exp Fold Change(human)
NCBP2	-2.210	-1.570
MYD88	-2.149	-1.572
TMEM64	-2.143	-3.618
RIOK3	2.143	1.573
PSPH	2.190	1.723
VPS37B	2.269	1.560
LMNA	2.473	1.836
PTBP2	2.509	1.583
NDRG1	2.883	2.662
SLC16A3	3.034	1.642
PHGDH	3.183	1.543
HIST1H1C	3.238	1.707
BAG2	3.335	1.860
CENPM	3.388	1.556

B-5. TCDD binding modes and their corresponding association free energies in complex with AhR.

TCDD Binding Mode	Association Free Energy (kcal/mol)
05h1	-42.50 ± 1.03
05h2	-44.84 ± 1.84
05h3	-46.53 ± 1.14
05q1	-47.88 ± 1.74
05q2	-49.69 ± 0.67
05q3	-43.08 ± 0.29
15h1	-48.04 ± 0.52
15h2	-47.10 ± 0.64
15h3	-45.55 ± 2.03
15q1	-46.58 ± 0.41
15q2	-47.13 ± 0.91
15q3	-47.34 ± 0.69
25h1	-45.78 ± 0.66
25h2	-46.27 ± 0.21
25h3	-44.90 ± 0.37
35h1	-48.20 ± 0.81
35h2	-46.63 ± 1.56
35h3	-46.24 ± 0.53

B-6. 1,4-DHNA binding modes and their corresponding association free energies in complex with AhR.

1,4 DHNA Binding Mode	Association Free Energy (kcal/mol)
05h1	-41.13 ± 0.44
05h2	-34.06 ± 2.35
05h3	-27.94 ± 1.05
05q1	-31.69 ± 0.39
05q2	-40.20 ± 2.27
05q3	-43.84 ± 0.39
15h1	-43.07 ± 1.22
15h2	-28.30 ± 0.99
15h3	-37.95 ± 3.83
15q1	-34.20 ± 3.30
15q2	-47.59 ± 0.56
15q3	-29.12 ± 2.08
25h1	-32.84 ± 1.15
25h2	-40.40 ± 1.27
25h3	-31.30 ± 1.54
35h1	-41.63 ± 0.30
35h2	-32.55 ± 2.18
35h3	-33.44 ± 2.06

APPENDIX C

SUPPLEMENTAL MATERIALS AND METHODS

C-1. Generation of docking poses

In all protocols, the MMFP module of CHARMM, version c39b2 (624), was used to constrain the ligands to the experimentally defined binding site. In four of the docking protocols, a harmonic potential of 10.0 kcal/(mol Å^2) was applied to the center of mass of the ligands to disallow structural deviations in the center of mass greater than either 0.5, 1.5, 2.5, or 3.5 Å from the initially determined center of mass determined by AutoDock Vina (622). In the remaining two docking protocols, a quartic spherical boundary potential was applied to the center of mass of the ligands, with the following form $E(r) = a \times (r - r_{\text{offset}})^2 \times \left((r - r_{\text{offset}})^2 - p\right)$, where $E(r)$ is the potential, a is the amplitude of the potential, r is the difference between the ligands' initial and new center of mass after each rotation, r_{offset} is a defined offset distance in the ligands' center of mass, and p is a parameter. With the quartic spherical boundary potential, an energetic well is created away from the initial position of the ligands, which may enhance the sampling of the binding modes. Here, a was set to 25.0 kcal/(mol Å^2), p was set to 1, and r_{offset} was set to either 0.5 or 1.5 Å. The molecules were parameterized using CGENFF (652).

In each of the six docking simulation protocols, twenty docking molecular dynamics (MD) simulation runs were performed in implicit solvent (653) using CHARMM (624). To preserve the non-binding conformation of AhR during the docking simulations, a

harmonic constraint of 1.0 kcal/(mol*Å^2) was applied to AhR residues, excluding residues with backbone atoms within 8.5 Å of TCDD in its initial placement. These residues were freed from constraints as to facilitate the docking procedure and allow flexibility in the docking site. Every two picoseconds, a rotation of 120 degrees was attempted on the docked ligand around a randomized axis with the final two picosecond structure minimized under fifty steps of steepest descent and saved for further analysis. The rotations aimed at allowing the ligands to explore different poses within the binding sites. Two-hundred steps of two picosecond MD runs were performed in each of the twenty docking MD simulation runs. Thus, four thousand snapshots of each ligand in the binding site of AhR per protocol were produced. As an initial screening, from each docking simulation protocol, out of the four thousand complex structures produced, we extracted the three complex structures with the lowest interaction energy for further analysis. As a result, 18 docking conformations of TCDD in complex with AhR and 18 docking conformations of 1,4-DHNA in complex with AhR.

The extracted binding modes were named based on which protocol they originated from. For example, binding mode 05h1 is the mode with the lowest interaction energy from the protocol using a harmonic potential allowing a deviation of 0.5 Å in the ligand's center of mass. Likewise, binding mode 15q2 is the mode with the second lowest interaction energy from the protocol using a quartic potential with an r_{offset} of 1.5 Å.

C-2. Molecular dynamics simulations of selected TCDD:AhR and 1,4-DHNA:AhR complexes

Each complex was solvated in a 105 Å truncated octahedral explicit water box. Potassium chloride ions were introduced to each water box resulting in an ion concentration of 0.15 M. The charge of the systems was neutralized by adding additional potassium and chloride ions as needed. The ions were initially placed using 2,000 steps of Monte Carlo simulations (654,655). Subsequently, 50 steps of steepest descent minimization followed by 50 steps of Adopted Basis Newton-Raphson minimization was performed on the solvent molecules. An additional 50 steps of steepest descent minimization and 50 steps of Adopted Basis Newton-Raphson minimization were performed in each system with heavy atoms of both ligands and all AhR backbone atoms constrained under 1.0 kcal/(mol*Å^2) harmonic constraints and AhR heavy side-chain atoms under 0.1 kcal/(mol*Å^2) harmonic constraints.

The systems were equilibrated in five consecutive stages, of 0.2 ns duration each. The total duration of equilibration was equal in the first of the five stages, the heavy atoms of the ligand, backbone atoms of the protein, and heavy side-chain atoms of the protein were constrained with a 5.0 kcal/(mol*Å^2), 5.0 kcal/(mol*Å^2), and 0.5 kcal/(mol*Å^2) harmonic force, respectively. In the subsequent four stages, the harmonic constraints on the ligand heavy atoms, protein backbone atoms, and protein heavy side-chain atoms, respectively, were reduced to 4.0 kcal/(mol*Å^2), 4.0 kcal/(mol*Å^2), and 0.4 kcal/(mol*Å^2) in the second stage, 3.0 kcal/(mol*Å^2), 3.0 kcal/(mol*Å^2), and 0.3

kcal/(mol*Å^2) in the third stage, 2.0 kcal/(mol*Å^2), 2.0 kcal/(mol*Å^2), and 0.2 kcal/(mol*Å^2) in the fourth stage, and finally 1.0 kcal/(mol*Å^2), 1.0 kcal/(mol*Å^2), and 0.1 kcal/(mol*Å^2) in the fifth stage. These constraints were then released for the production stage of the simulations.

In the production stage, each complex was simulated for 10 ns with frames extracted every 20 ps. The temperature for all simulations was set at 300 K using the Hoover thermostat, and the pressure for all simulations was set at 1.0 atm. The simulations were executed using the Leap-frog Verlet algorithm of CHARMM, version c39b2 (624). Light (0.1 kcal/(mol*Å^2)) harmonic constraints were introduced to the backbone atoms of AhR residues 241 through 264 in order to preserve the structure in that domain with respect to the rest of the receptor. After the MD simulation runs reached 10 ns, the trajectories were stripped of all solvent atoms and the 10 ns simulation snapshots of each complex were analyzed for their energetic favorability as described below. Simulation snapshots were extracted every 20 ps, and thus 500 snapshots were analyzed per system.

MM GBSA association free energy calculations. According to the MM GBSA approximation, the association free energy, ΔG, is determined by subtracting the total free energy of the free protein, G_P, and the free ligand, G_L, from the total free energy of the complex, G_{PL}, as shown in the equation below (630,635,656):

$$\Delta G = G_{PL} - G_P - G_L \qquad \text{Eq. 1}$$

The individual free energies were estimated using the following equation (657).

$$G = E^{\text{Bonded}} + E^{\text{Elec}} + E^{\text{GB}} + E^{\text{vdW}} + g \times \text{SASA} \quad \text{Eq. 2}$$

Where E^{Bonded}, E^{Elec}, E^{GB}, E^{vdW}, and SASA are the bonded energy, electrostatic interaction energy, generalized-Born energy, van der Waals energy, and solvent-accessible surface area of the system, respectively. The polar contribution to the total MM GBSA association free energy is represented by the sum of electrostatic interaction energy and generalized-Born energy terms. The nonpolar contribution to the total MM GBSA association free energy is represented by the sum of the van der Waals energy and solvent-accessible surface area terms.

In this calculation, we used the one-trajectory approximation (658,659), as in refs. (629,630,635,656), in which the free state of the ligand and receptor protein have the same conformation as in the bound state and entropy contribution from the ligand and protein are assumed to be negligible. Entropy contributions from the solvent are taken into account implicitly (660). The generalized Born with a simple switching (GBSW) model was introduced to calculate the polar and non-polar solvation free energies (653). The non-polar surface tension coefficient used for these calculations was set to the default value of 0.03 kcal/(mol Å^2).

The association free energies of each set of conformations for both TCDD and 1,4-DHNA were averaged over four segments. The 10 ns production runs were divided into four equal segments of 2.5 ns. The average association free energies of each segment, for the entire 2.5 ns trajectory, were taken to be the measurement for the 2.5 ns interval. The

average association free energies and standard deviation values, reported in Appendix B-5 for TCDD binding modes and Appendix B-6 for 1,4-DHNA binding modes, were calculated using the four measurements for each binding mode.

C-3. Per AhR residue interaction free energy calculations

The production runs were also divided into four, 2.5 ns segments for the per residue interaction free energy calculations. For each 2.5 ns segment, the per AhR residue interaction free energy between a residue *R* and *L* (ligand) was calculated using the following equation (626,635,636).

$$\Delta G_{RL}^{\text{inte}} = \frac{1}{f}\sum_{k\in f}\left(\sum_{i\in R}\sum_{j\in L}\left(E_{ij}^{\text{Elec}} + E_{ij}^{\text{GB}}\right) + \sum_{i\in R}\sum_{j\in L} E_{ij}^{\text{vdW}} + \gamma\sum_{i\in R,L}\Delta\left(\text{SASA}_i\right)\right)$$

The polar, Van der Waals (vdW), and non-polar solvation interaction free energy between a residue *R* and *L* are represented by the first, second, and third components of equation above respectively. For our study, *R* corresponds to a given residue in the AhR protein, and *L* corresponds to either ligand, TCDD or 1,4-DHNA, in the binding pocket of AhR. The sum of the per residue interaction free energies across the 2.5 ns segment is averaged over *f (=125)*, the number of snapshots used in the calculation. The average per AhR residue interaction free energies and their standard deviations, presented in Figure 8, between *R* and *L* were calculated using the per residue interaction free energy values of the four 2.5 ns segments.

The polar component of the total interaction free energy, consisting of electrostatic interaction $\left(E_{ij}^{\text{Elec}}\right)$and generalized-Born $\left(E_{ij}^{\text{GB}}\right)$ energy contributions between residues *R* and *L*, represents the interaction between residues *R* and *L* and the interaction between residue *R* and the solvent polarization potential induced by *L*. The non-polar component, consisting of the van der Waals interactions $\left(E_{ij}^{\text{vdW}}\right)$ between the two residues and the change in the non-polar solvation free energy due to binding ($\gamma \cdot \Delta\text{SASA}_i$), represents the non-polar interactions with the surrounding solvent and cavity contributions.

The solvation terms were determined using the GBSW generalized-Born model (653). These calculations were performed using a non-polar surface tension coefficient, γ, of 0.03 kcal/(mol*Å^2). Atoms surrounding residues *R* and *L* affect the generalized-Born energy contribution $\left(E_{ij}^{\text{GB}}\right)$ and solvent accessible surface area (ΔSASA_i) terms (661). For the $\left(E_{ij}^{\text{GB}}\right)$ term, all atoms were included, and the charges of atoms outside the groups *RL*, *L*, and *R* were set to zero. The ΔSASA_i term corresponds to the difference in solvent accessible surface areas of the bound and unbound states of residues *R* and *L*.