Epigenetic Effects in Head and Neck Cancer and di-2-ethylhexyl phthalate (DEHP) Exposure

by

Siyu Liu

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Bioinformatics)
in the University of Michigan
2021

Doctoral Committee:

Associate Professor Maureen A. Sartor, Chair
Assistant Professor Justin Colacino
Associate Professor Ryan Mills
Associate Professor Arvind Rao
Associate Professor Laura Rozek

Siyu Liu

liusiyu@umich.edu

ORCID iD: 0000-0002-5681-7824

Dedication

To my grandfather, Shengming Liu, for his unconditional love and support throughout every chapter of my life.

Acknowledgements

Pursuing a doctorate degree is a long journey, but I have been fortunate through this whole experience to enjoy it to the fullest. First and foremost, I would like to extend my gratitude to my PhD advisor Maureen, who has helped me grow not only as a scientist but also as a person in the past four years. I would also like to thank my other committee members, Laura, Justin, Arvind and Ryan, my collaborators and both previous and current lab mates, for boosting my academic progress continuously. The Department of Bioinformatics has helped me tremendously ever since the beginning, which I am extremely grateful for.

In particular, I would like to thank my parents and my boyfriend, who trust me unconditionally and always come to rescue whenever needed. A special thanks to my friends both inside and outside the program for their support during the freezing winters of Michigan. Last, a special shoutout to my two pets, my dog Mochi and cat Miffy, who have prevented me from going out of my mind during the pandemic, and wish we have many more colorful years ahead to make more precious memories together.

Table of Contents

List of Tables

List of Figures

Abstract

Epigenetics refers to the study of heritable changes in gene expression that do not involve changes in the underlying DNA sequence. Typical investigated epigenetic marks include DNA methylation, modifications to histone proteins, and more recently DNA 5-hydroxymethylation (5hmC). Epigenetic profiles in carcinogenesis have been a topic of increasing interest, with several studies showing that epigenetic silencing of tumor suppressor genes and de-repression of heterochromatic intergenic regions via DNA methylation dysregulation play essential roles. Head and neck squamous cell carcinoma (HNSCC) is the 6th most common cancer worldwide, and accumulated scientific evidence suggests that epigenetic alterations, especially DNA methylation, are frequently involved in oral carcinogenesis, tumor progression, and resistance to therapy. In addition to cancer, other diseases and exposures can also modify the epigenome. For instance, environmental contaminants and toxicants such as the chemical DEHP, have been shown to alter epigenetic regulatory features such as DNA methylation and hydroxymethylation.

Infection with a high-risk strain of human papillomavirus (HPV) has been shown to be a risk factor for HNSCC development. Characteristics of HPV(+) HNSCC include higher risk of distant metastasis, poor differentiation, and more favorable prognosis, while HPV(-) HNSCC tends to progress locally and exhibit a higher level of differentiation and greater resistance to treatment. Previous studies have identified DNA methylation differences between these two classes of HNSCC, but we are the first to characterize 5hmC in HNSCC by HPV status (18 HPV(+); 18 HPV(-)). In Chapter 2, we showed significant hyper-hydroxymethylation in HPV(-)

tumors, characterized the heterogeneity in 5hmC in promoters and enhancers relative to tumor subtypes and clinical variables, and detected important cancer genes with differential 5hmC by HPV status and between subtypes.

In terms of clinical treatment in HNSCC, in Chapter 3, we tested the transcriptomic and epigenomic effects of a multi-component drug regimen called IRX-2 in a phase II clinical trial; IRX-2 is a human cell-derived biologic with multiple active cytokine components. The treatment regimen consists of 10 days of regional perilymphatic IRX-2 cytokine injections and daily oral indomethacin, zinc and omeprazole (Regimen 1) compared to the identical regimen without IRX-2 cytokines (Regimen 2). By comparing the IRX-2 versus control arms after 21 days of trial, we were able to characterize the significant change of more than 100 genes using NanoString immune gene expression data and explore the correlation between response (clinical and immunological) and immune signatures. The exploration of DNA methylation showed a slight overall increase in methylation after treatment in both regimens, especially for immune responders in the IRX-2 study arm. Although a small number of genes were identified that distinguished the IRX-2 from the control arm, most beneficial changes were common to both, suggesting that much of the benefit of the full IRX-2 treatment regimen is due to components other than the cytokine cocktail.

In Chapter 4, we focused on the epigenetic effect of early-life di-2-ethylhexyl phthalate (DEHP) exposure. DEHP is the most common member of the class of phthalates, and DEHP exposure during pregnancy has been shown to disrupt placental growth and development in mice, resulting in higher rates of low birthweight, premature birth, and fetal loss. We studied the effect of perinatal DEHP exposure on the DNA methylation profile in liver (a primary target tissue of DEHP) and blood (a common surrogate tissue) of both juvenile and adult mice. Despite

exposure ceasing at 3 weeks of age (PND21), we identified thousands of sex-specific differential DNA methylation events in 5-month-old mice, more than identified at PND21, both in blood and liver. However, only a small number of these differentially methylated cytosines overlapped between the time points, or between tissues (i.e. liver and blood), indicating changes by age and that blood may not be an appropriate surrogate tissue to estimate the effects of DEHP exposure on liver.

In summary, we facilitated the interpretation of epigenetic effects by exploring the 5hmC profile in head and neck cancer, characterizing the immune response gene expression and DNA methylation in oral cavity cancer immunotherapy, and capturing the effects of DEHP exposure on DNA methylation in mouse models.

Chapter 1 Introduction

1.1 Introduction to epigenetics

Epigenetics refers to the study of heritable chemical or physical changes that do not involve changes in the underlying DNA sequence. Typical investigated epigenetic marks include DNA methylation, modifications to histone proteins, and more recently DNA 5-hydroxymethylation (5hmC) (1). Epigenetic modifications have been shown to be correlated with many human diseases, including different cancers, autoimmune disorders, neurological disorders, aging and response to environmental exposures (2–5). Histone modification is a covalent post-translational modification to histone proteins, and many types of histone modification events include methylation, phosphorylation, acetylation, ubiquitylation, and sumoylation (6). These modifications regulate the chromatin structure either directly or by recruiting remodeling enzymes, and they can both positively and negatively regulate gene expression by changing the way in which histones bind to DNA (7).

DNA methylation indicates the process of adding a methyl group to bases in the DNA sequence, and the most commonly studied modification is the methylation of the C5 position on cytosine bases, which is called 5-methylcytosine (5mC) (**Figure 1.1**, modified from (8)). TET (ten-eleven translocation) proteins can oxidize 5mC to 5-hydroxymethylcytosine (5hmC) and other oxidative derivatives, and hydroxymethylation refers to the addition of hydroxymethyl group instead of methyl group. The conversion from 5mC to 5hmC results in a loss of transcriptional repression in promoters or enhancers, and is a common mechanism to activate

genes in differentiation and development (9). Genomic 5hmC is a relatively stable component of DNA, or can act as an intermediate in DNA demethylation by replication-dependent or replication-independent pathways, and it plays an essential role in the regulation of gene expression (10,11).

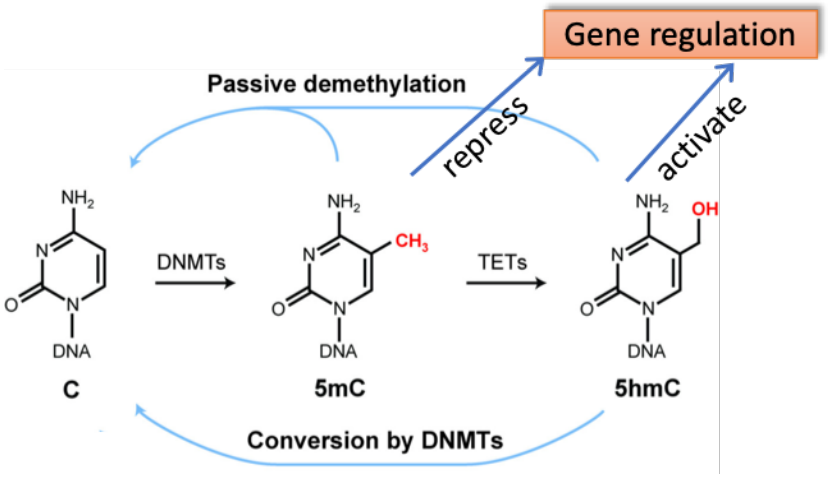


Figure 1.1 Chemical structures of 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC) and their potential associations with gene regulation

1.2 Methods to capture epigenetics

The rapid growth and lower cost of high-throughput sequencing methods and other molecular technologies have helped capture epigenetic profiles more accurately. Histone modifications can be detected using a variety of techniques including mass spectrometry and genomics approaches such as chromatin immunoprecipitation (ChIP), which can identify the location of a protein or modification of interest within the genome and measure its relative abundance at each location. There are two main methods to measure DNA methylation, which are affinity enrichment based and bisulfite-based methods (12). At a lower cost, the affinity enrichment strategies include methyl-DNA immunoprecipitation (MeDIP) and methyl-CpG binding domain protein (MBD)-based methods, which generally have lower resolution and a higher input requirement. The

bisulfite conversion methods include whole genome bisulfite sequencing (WGBS), which was considered the gold standard for several years, reduced representation bisulfite sequencing (RRBS), and Infinium (MethylationEPIC BeadChip, which interrogates over 850,000 methylation sites quantitatively), and though more labor and computationally intensive, they have low input requirement and single-nucleotide resolution (13). However, it is important to note that these standard bisulfite-based techniques cannot distinguish 5mC from other 5-carbon cytosine modifications, including 5hmC, whereas the affinity-based approaches are specific to one mark.

5hmC is the most prevalent intermediate on the oxidative DNA demethylation pathway, and multiple methods have been developed to capture the genomic 5hmC profile. Examples of affinity enrichment methods include hydroxymethyl-DNA immunoprecipitation sequencing (hMeDIP-seq), where a monoclonal 5-hmC antibody was used (14). Bisulfite conversion methods can be adjusted to characterize 5hmC, and the most common methods include OxBS-seq (Oxidative bisulfite sequencing, with an additional oxidative step to bisulfite to discriminate between 5mC and 5hmC) and TAB-seq (use of TET to distinguish 5mC and 5hmC) (15,16). Recently, a few newer and more efficient approaches to detect 5hmC at base resolution were introduced, such as Jump-seq (5hmC was labeled by the 6-N3-glucose moiety and connected to a hairpin DNA oligonucleotide) and TAPS-seq (TET-assisted pyridine borane sequencing; a bisulfite-free method to both 5mC and 5hmC detection) (17,18).

1.3 Epigenetics in head and neck cancer

In human cancers, alteration of DNA methylation patterns is frequently observed. On one hand, hyper-methylation patterns in the promoters of tumor suppressor genes could inhibit their

normal expression, resulting in changes in crucial biological processes, such as proliferation and cell cycle checkpoints, apoptosis, cell cycle, DNA repair, tumor invasion, and metastasis (19,20). On the other hand, global hypomethylation and hypomethylation at retrotransposon elements, centromeres and oncogenes can also induce genomic instability and contribute to cell transformation (21). In particular, lower methylation level in long interspersed elements (LINEs) or short interspersed elements (SINEs) is highly associated with carcinogenesis through genome destabilization (22). Global reduction in 5hmC has been observed in many cancer types studied to date (23–25), which are frequently associated with mutations in TET2 (26,27). More recent studies showed that the inhibition of TET activity by the onco-metabolite 2-hydroxyglutarate which accumulates through mutations of isocitrate dehydrogenases (IDH1/2) are also involved in the dynamic regulation of 5hmC levels (28,29).

Head and neck squamous cell carcinoma (HNSCC) is heterogeneous disease that involve multiple sites and cellular origins within the upper aerodigestive tract, which includes tumors in the oral cavity, larynx, oropharynx, and hypopharynx, and is the sixth most prevalent cancer worldwide(30–32). Globally, HNSCC affects approximately 680,000 patients and causes 330,000 deaths every year, with a five-year survival rate ranging from 37% to 62% (33,34),. While tobacco and alcohol consumption are long-recognized risk factors, high-risk strains of human papillomavirus (HPV), in particular HPV-16, account for an increasing number of cases (35). Epigenetic differences between normal and HNSCC tumor tissue are extensive, and there is increasing evidence showing that epigenetic modifications could be a potential cause of HNSCC, as shown by multiple genome-wide DNA methylation studies (32,36–39). Both hypermethylation and hypomethylation of promoter regions in HNSCC has been reported in many studies, and many genes are frequently analyzed such as p16, PTEN, DAPK, MGMT,

ECAD and RASSF1 (40). Another study re-evaluated the HNSCC DNA methylation markers, and hypermethylation of FAM135B and ZNF610 and hypomethylation of HOXA9 and DCC were validated from several HNSCC studies and 450K BeadChip data from The Cancer Genome Atlas (TCGA), and hypermethylation of FAM135B was significantly correlated with overall survival (41–43). The methylation alteration retrotransposon elements was also confirmed in HNSCC, where Arayataweegool et al. discovered hypomethylation of LINE-1 elements in coculture with cancer cells due to factors secreted by HNSCCs (44).

Compared with the relatively well studied DNA methylation events, only a few studies focus on hydroxymethylation in HNSCC. With the use of Immunohistochemistry methods, an early study by Jawert et al. discovered the loss of 5hmC and TET2 in oral squamous carcinoma, which is consistent with the observations in other cancer types (45). A more recent study using enzyme-linked immunosorbent assay and quantitative real-time PCR also low levels of 5hmC in HNSCC, and 5-hmC levels were significantly correlated with tumor stage and better disease-free survival (46).

1.4 HPV infection and head and neck cancer

During the last few decades, human papilloma virus (HPV) has emerged as a novel risk factor of HNSCC, especially for oropharyngeal squamous cell carcinoma (OPSCC) (47). HPV is a family of double-stranded DNA viruses of 8000 base pairs, and it can induce carcinogenic transformation of the infected mucosal epithelium by escaping cell-cycle checkpoints through degradation of p53 and Rb proteins, which is mediated by the accessory proteins E6 and E7, respectively (48,49). E6/E7 expression is frequently associated with integration of the viral genome into DNA regions of genomic instability, which is observed in the majority of invasive

cervical cancer and HPV(+) HNSCC (50). There are multiple methods to detect the presence of HPV, including E6/E7 HPV mRNA RT-PCR, HPV DNA in situ hybridization, p16in situ hybridization and measurement of HPV mRNA expression using high-throughput sequencing techniques (51,52).

HPV(+) HNSCC patients generally show better therapeutic response, improved prognosis and higher overall survival than their HPV(-) counterparts, with 5-year survival rates of 75–80% (53–55). At the molecular level, the gene CDKN2A (p16) is a marker for HPV etiology, due to its high expression level in HPV(+) tumors and common loss in HPV(-) tumors(56). HPV(+) HNSCC also have frequent activating mutations of genes involved in phosphoinositide 3-kinase (PI3K) pathway, affecting translation and transcription of multiple targets that are involved in various cellular properties such as proliferation, survival and motility (57,58). With tobacco consumption as the primary risk factor for development, HPV(-) HNSCC is often accompanied by p53 mutation, CDKN2A deletion or PIK3CA amplification (59). A few studies have examined molecular inter-tumor heterogeneity and identified subtypes of HNSCC (60–62), with the characteristic differences in global gene expression profiles between HPV(+) and HPV(-) tumors and among subtypes now established. Results point to differences by HPV status and tumor subtypes in several carcinogenic pathways, including basal epithelial-to-keratinocyte proliferation, immune response, cell adhesion and induction of DNA damage, that often correlate with clinical outcome (60,62).

During the last five years, multiple studies have discovered new evidence identifying different characteristics that define the main HPV(+) HNSCC subtypes. Based on gene expression profiling of a combination of 18 HPV+ HNSCC from University of Michigan Hospital and 66 TCGA HPV(+) HNSCC samples, Zhang et al. revealed two major HPV(+)

subtypes, named IMU and KRT. The IMU subtype was identified by a higher level of immune response and more mesenchymal differentiation, whereas the KRT subtype is identified by stronger keratinization and viral integration events (60). The study also linked the IMU subtype with stronger epithelial-to-mesenchymal transition (EMT) signatures and higher BCL2 gene expression, which could be contributing factors to IMU subtype's better prognosis.

The investigation of HPV integration status of the same cohort showed that the KRT subtype had more samples with HPV genic integration than the IMU subtype. The KRT group also showed lower expression of viral genes E2, E4, and E5. By analyzing SNP-array data from tumors and blood from the same cohort, the authors were able to identify more amplifications in KRT tumors than IMU, especially at chromosomal arms, and being compared with the results from gene expression suggested that the gain and loss of copy numbers can partially drive the expression differences between the two subgroups. In term of the differences in gene mutation frequencies between the two HPV(+) subtypes, the analysis on non-synonymous mutation from RNA-seq data and TCGA gene-level somatic mutation data displayed only one mutation with a difference of more than 20% between two groups, which occurred on oncogene PIK3CA in 37% of the KRT samples and only 16% of the IMU samples. By retrieving gene expression from 11 HNSCC studies, Locati et al. characterized three distinct HPV(+) subtypes, defined as Cl1 (immune-related), Cl2 (highly keratinized, epithelial mesenchymal transition-related), and Cl3 (highly keratinized, proliferation-related) (62). The Cl1 subtype was in agreement with the IMU subtype, while the KRT subtype was further stratified into Cl2 and Cl3 by their biological and prognostic characteristics.

A more recent study looked at the genetic variants profile of the same cohort of patients, with the use of a combination of mRNA-seq-based variant analysis and targeted gene sequencing on

matched tumor and blood samples (63). Although this study did not discuss any HPV(+) subtype specific mutation, it did identify a germline variant in a cancer related Fanconi Anemia (FA) pathway gene which was harbored by more than a third of patients.

1.5 Immunotherapy in head and neck cancer

The definition of cancer immunotherapy, or immuno-oncology, is a type of cancer treatment that uses components of the immune system to prevent, control, or eliminate cancer. There are many forms of immunotherapy, including checkpoint inhibitors, targeted antibodies, cancer vaccines, adoptive cell transfer, tumor-infecting viruses, and immune system modulators (64). Depending on the nature of treatment, cancer immunotherapy comes in forms of intravenous (IV), oral, topical, or intravesical. Checkpoint inhibitors are monoclonal antibodies that block the certain immune cells that need to be activated (65), and they work by blocking checkpoint proteins from binding with their partner proteins. Most widely used drugs target CTLA-4 and PD-1 (or partner protein PD-L1). As originally published by the National Cancer Institute, checkpoint inhibitors have been approved to treat a variety of cancer types, such as breast cancer, cervical cancer, colon cancer, Hodgkin lymphoma, liver cancer, lung cancer, and head and neck cancer, with response rate increasing from 6% to 13% during the last few years (66). Cytokines are proteins produced by white blood cells, and they are the most common form of immune modulators. The types of cytokines frequently used to treat cancer include interferons (INFs; inhibit B cell activation, enhance T cell activity, and increase the cellular-destruction capability of natural killer cells) and interleukins (ILs; regulate activated T cell proliferation and promote activation-induced cell death) (67). Other types of immune modulators include agonists, which

can activate pathways that promote adaptive immune responses, and adjuvants, which can activate pathways involved in the innate immune system.

Currently there are four approved monoclonal antibodies targeting PD-1 and PD-L1.The drugs nivolumab (opdivo) and pembrolizumab (keytruda) are anti-PD-1 antibodies, while avelumab (bavencio) and atezolizumab (tecentriq) are anti-PD-L1 antibodies (68,69). The first trial evaluating PD1/PDL1 inhibitor in HNSCC was published in 2016, where pembrolizumab was used for treatment of recurrent or metastatic HNSCC (70,71). The overall response rate was 18% during the Phase I trial regardless of HPV status, and the response rate stayed stable at 16% during the Phase II trial on HNSCC patients pre-treated with platinum and cetuximab (72). During the Phase III trial, the patients were allocated at a 2:1 ratio to receive either nivolumab or standard single-drug systemic therapy, and patients treated with nivolumab appeared to have longer overall survival, though lower than the pre-specified cutoff for survival improvement (73,74). In 2019, another Phase III clinical study using pembrolizumab in the treatment of relapsed or metastatic HNSCC reported superior treatment results in the pembrolizumab group (75). Based on the positive outcome, both nivolumab and pembrolizumab have been approved by the FDA for patients with relapsed or metastatic HNSCC who are cisplatin-resistant (76).

A number of trials have evaluated PD-L1 expression by IHC on HNSCC tumor cells alone or in combination with PD-L1 expression on tumor-infiltrating immune and stromal cells, so as to understand the association between certain gene expression and immunotherapy outcome. For instance, the aforementioned clinical trial showed that tumor PD-L1 expression generally correlates with improved efficacy with anti-PD-1/PD-L1 checkpoint inhibitors in HNSCC, with increased predictive value when including PD-L1 expression on tumor infiltrating immune cells (77). Apart from immune-related gene expression, epigenetics could also play an important role

in HNSCC immunotherapy. For instance, a Phase 2b clinical trial (NCT03019003) was conducted to test whether low dose of a demethylating agent (5'aza) is able to prime patient's immune response for immune checkpoint inhibitor therapy. In a few recent immune-related HNSCC studies, the characterization of DNA methylation profiles predicted a combination of hyper- and hypo-methylation markers in response to immune checkpoint inhibitors, and these methylation signatures could act as surrogate biomarkers to predict the objective response rate of PD-1/PD-L1 inhibition (78–81).

Immunomodulatory drugs are also known as biological response modifiers priming immune response in patients, and currently there are four drugs, namely, thalidomide (Thalomid), lenalidomide (Revlimid), pomalidomide (Pomalyst) and imiquimod (Aldara, Zyclara) in cancer treatment. Thalidomide, lenaliodomide, and pomalidomide are used to treat multiple myeloma, and they can boost the IL-2 release in cells, as well as stop tumors from forming new blood vessels (82,83). Imiquimod is a drug that is a cream applied to skin, and it can stimulate local immune response against skin cancer cells. Another form of immune modulation treatment is the use of Bacillus Calmette-Guérin (BCG), a weakened form of the bacteria that causes tuberculosis. BCG is a liquid put into the bladder through a catheter to treat early-stage bladder cancer, and as one of the earliest immunotherapies, it can also be used melanoma skin cancers (84). However, no immune modulation drugs have been approved in HNSCC to date.

1.6 Epigenetics in DEHP exposure

Environmental exposures compose of many forms, such as chemicals, metals, or particles, and can affect animal and human health negatively. For instance, various forms of exposures can alter the normal function of the endocrine system or harm the nervous, hematopoietic,

respiratory, or digestive systems by releasing toxins (85). Di(2-ethylhexyl) phthalate (DEHP) is a plasticizer (86) often found in polyvinyl chloride products, including medical equipment, car upholstery, food and beverage containers, and building materials (87). DEHP does not covalently bond to polymer chains in these products and is therefore likely to spread into the environment after repeated usage (88). Humans are exposed to DEHP through ingestion, inhalation and dermal exposure, and the exposure to DEHP is considered a universal and ubiquitous event (89,90). Due to the large quantity and wide variety of products containing DEHP that people interact with frequently, the estimated range of human exposure to DEHP is 3-30 μg/kg/day (91).

DEHP is a type of endocrine disruptor and a probable carcinogen, and has been detected in tissues such as blood, amniotic fluid, umbilical cord blood, and breast milk in humans (92–95), indicating that exposure starts as early as the fetal developmental stage, impacting generations that were not directly exposed to the chemical (96). The effect of DEHP can last across multiple generations during critical developmental window (97). The estimated human oral absorption varies from 25% to 50% depending on the dosage (98,99), while absorption in rodents can be as high as 58% of the oral dose, but 50% is assumed on average (100). The biological action of DEHP is very similar to a group of chemicals called peroxisome proliferators (PPs), and liver is a primary target organ for the effects of DEHP and other PPs (101). Prenatal DEHP exposure in rodents results in elevated fatty acid metabolism, as well as peroxisome proliferation and the accumulation of lipofuscin granules which are implicated in hepatocarcinogenesis (102,103). There are multiple modes of action of DEHP in hepatocytes, with the most well-known ones being activation of peroxisome proliferator-activated receptors (PPARs), induction of cell proliferation, suppression of apoptosis, oxidative DNA damage, and inhibition of gap junctional intercellular communication (101,104). PPARs can regulate histone deacetylation and DNA

methylation, and modulate a series of mechanistic pathways leading to increase in adipocyte formation and fat storage (105,106). Once absorbed, DEHP and its metabolites are distributed throughout the body via the blood promoting its endocrine disrupting properties.

Epigenetics, such as non-coding miRNA and DNA methylation, define and control cell and tissue development by regulating gene expression (107), and may mediate the effects of exposures. For example, Scarano et al. proposed that the long-term reproductive defects associated with phthalate exposure is exerted through non-coding miRNAs (108). DNA methylation, especially methylation patterns of imprinted genes is crucial for embryonic development (109). Previous animal and human studies have reported that perinatal exposure to DEHP is associated with altered DNA methylation (110–112), with multiple studies identifying DEHP-induced DNA hyper-methylation (113,114). A recent study in human cord blood showed significant DNA methylation changes in genes involved in androgen and estrogen responses, and spermatogenesis following prenatal DEHP exposure (115). Other studies of ovaries and oocytes showed prenatal DEHP exposure disrupted the expression of cell cycle regulators and changed the DNA methylation pattern of imprinted genes in germ cells (109,116). As the primary organ for filtering toxicants within the body, liver has also been identified as a primary target organ of DEHP. A recent study showed that DEHP exposed mice livers exhibited significant changes in global DNA methylation levels in all three subsequent generations, and the expression levels of DNA methyltransferase 1 (Dnmt1), which can establish and maintain DNA methylation patterns, were significantly changed in both the F1 and F2 generations (117,118). However, DNA methylation was only quantified using Nanodrop in this study, and the effect of prenatal DEHP exposure on genome-wide DNA methylation patterns in liver has not been assessed.

1.7 Dissertation overview

The epigenetic effects in both head and neck cancer and DEHP exposure have been described, but several questions remain to be answered in order to further understand the epigenetic mechanisms in the respective conditions. Compared with the abundance of literature in genetics, the field of epigenetics is not as well studied in head and neck cancer, especially in terms of the epigenetic differences within different HNSCC subtypes. With the increasing importance of HPV infection in HNSCC, it is essential to not only capture the epigenetic profiles of HPV(+) and HPV(-) tumors separately, but also investigate the epigenetic marks of subtypes within each tumor group. Even though immunotherapy has gained increasing usage in multiple cancer types, only a limited number of checkpoint inhibitor drugs are available for head and neck cancer. There is an urgent need to explore more types of immunotherapy in HNSCC, such as immune modulation drugs, as well as capture the immune profile of the patients before and after the treatment, both genetically and epigenetically. Multiple studies have shown the genetic effects of prenatal DEHP exposure, but the epigenetic changes associate with it has only been covered in a few studies. The exploration of DEHP's epigenetic effects, particularly in its target organs and surrogate organs such as blood, can cast light on the faster detection and better prognosis of DEHP exposure in human body. Based on the aforementioned observations and problems, in this dissertation I present three studies on the epigenetic effects in head and neck cancer and DEHP exposure.

In Chapter 2, we describe the genome-wide 5hmC profiles in HNSCC by HPV status and subtype in 18 HPV(+) and 18 HPV(-) well-characterized tumors. As described above, 5hmC is an oxidative derivative of DNA methylation and an important epigenetic marker in many cancer

types, but its heterogeneity among HNSCC subtypes, has not been studied. We show significant genome-wide hyper-5hmC in HPV(-) tumors, with both promoter and enhancer 5hmC able to distinguish meaningful tumor subgroups. We also identify specific genes whose differential expression by HPV status is driven by differential hydroxymethylation. Among the previously reported two HPV(+) subtypes, IMU (stronger immune response) and KRT (more keratinization), the IMU subtype revealed hyper-5hmC and up-regulation of genes in cell migration, and hypo-5hmC with down-regulation in keratinization and cell junctions. We experimentally validate our key prediction of higher secreted and intracellular protein levels of the invasion gene MMP2 in HPV(-) oral cavity cell lines.

In Chapter 3, we genetically and epigenetically characterize the immune response of oral cavity cancer patients from a randomized phase II trial conducted of the IRX-2 regimen, a homologous cell-derived complex containing a multi-cytokine biologic with multifaceted immune modulatory effects. The treatment regimen consists of 10 days of regional perilymphatic IRX-2 cytokine injections and daily oral indomethacin, zinc and omeprazole (Regimen 1) compared to the identical regimen without IRX-2 cytokines (Regimen 2). We elucidate changes before versus after the IRX-2 treatment in key immune response genes and explore the methylation profile change in both regimens during the treatment. Specific benefits in the tumor microenvironment of the cytokine cocktail in IRX-2 may include increased CD8+ T cell density, a slight global and repetitive element re-methylation of the genome, upregulation of the tumor suppressor DMBT1, and unchanged PD1/PDL1 for the subset of patients showing immune cell infiltration.

In Chapter 4, we use an established mouse model to study the effect of perinatal DEHP exposure on the DNA methylation profile in liver (a primary target tissue of DEHP) and blood (a

common surrogate tissue) of both juvenile and adult mice. Dysregulation of epigenetic modifications, including DNA methylation, have been shown to be an important mechanism for the pathogenic effects of prenatal exposures, including phthalates. Despite exposure ceasing at 3 weeks of age (PND21), we identify thousands of sex-specific differential DNA methylation events in 5-month-old mice, more than those identified at PND21, both in blood and liver. Only a small number of these differentially methylated cytosines (DMCs) are overlapped between the time points, or between tissues (i.e. liver and blood), indicating blood may not be an appropriate surrogate tissue to estimate the effects of DEHP exposure on liver DNA methylation. As part of the Toxicant Exposures and Responses by Genomic and Epigenomic Regulators of Transcription (TaRGET II) Consortium, this study implicates the usage of surrogate tissue instead of target tissue in human population-based studies and identifies epigenetic biomarkers for DEHP exposure.

Chapter 2

5-hydroxymethylation Highlights the Heterogeneity in Keratinization and Cell Junctions in Head and Neck Cancers

This work has been published as: Liu, S., de Medeiros, M. C., Fernandez, E. M., Zarins, K. R., Cavalcante, R. G., ... & Sartor, M. A. (2020). 5-hydroxymethylation highlights the heterogeneity in keratinization and cell junctions in head and neck cancers. *Clinical Epigenetics*, 12(1), 1-14.

2.1 Introduction

Head and neck squamous cell carcinoma (HNSCC) includes tumors in the oral cavity, larynx, and oropharynx, and is the sixth most prevalent cancer worldwide(30,31). Globally, HNSCC affects approximately 680,000 patients every year, with a five-year survival rate ranging from 37% to 62%(33). While tobacco and alcohol consumption are long-recognized risk factors, high-risk strains of human papillomavirus (HPV), in particular HPV-16, account for an increasing number of cases(35). HPV(+) HNSCC patients generally show better therapeutic response, improved prognosis and higher overall survival(53–55). At the molecular level, the gene CDKN2A (p16) is a marker for HPV etiology, due to its high expression level in HPV(+) tumors and common loss in HPV(-) tumors(56). A few studies have examined molecular inter-tumor heterogeneity and identified subtypes of HNSCC (60,61), with the characteristic

differences in global gene expression profiles between HPV(+) and HPV(-) tumors and among subtypes now established. Results point to differences by HPV status and tumor subtypes in several carcinogenic pathways, including basal epithelial-to-keratinocyte proliferation, immune response, cell adhesion and induction of DNA damage, that often correlate with clinical outcome.

Epigenetic differences between normal and HNSCC tumor tissue are extensive, as shown by genome-wide DNA methylation studies(36–38). HPV(+) status is associated with hypermethylation in the promoter of several specific genes(119), and HNSCC subtypes have been identified using DNA methylation data from The Cancer Genome Atlas (TCGA)(120). Previously, we showed that HNSCC DNA methylation profiles correlate with both patient diet and survival(119,121) and extensive genome-wide DNA hypomethylation in HPV(-) compared to HPV(+) squamous cell carcinoma (SCC) cell lines(122), which has since been validated by others in HNSCC tumors(123).

Most of the above epigenetic data relied on bisulfite treatment of DNA, which does not distinguish between methylation (5-methylcytosine or 5mC) and hydroxymethylation (5-hydroxymethylcytosine or 5hmC). TET (ten-eleven translocation) proteins can oxidize 5mC to 5hmC and other oxidative derivatives, with 5hmC being the most abundant form *in vivo*(124–126). This conversion results in a loss of transcriptional repression in promoters or enhancers, and is a common mechanism to activate genes in differentiation and development(127). Recent studies found that 5hmC is depleted in human cancers of many different origins(24,128,129), yet a recent study of oral cancers found that globally elevated 5hmC is positively associated with more aggressive tumors and worse survival(130). Genome-wide 5hmC profiles in HNSCC and

in specific tumor subtypes remain uncharacterized, and virtually nothing is known regarding the association of oncogenic viruses such as HPV with 5hmC levels.

Here, we capture genome-wide hydroxymethylation profiles and examine their heterogeneity among 18 HPV(+) and 18 HPV(-) previously well-characterized HNSCC tumors. We previously characterized two distinct HPV(+) HNSCC subtypes based on gene expression and copy number variation for these 36 tumors and those from TCGA(60). The IMU subtype is identified by a heightened immune response and more mesenchymal differentiation, whereas the KRT subtype is identified by more keratinization and viral integration events. Based on differential 5hmC profiles in other human cancers and the fundamental distinctions between HPV(+) and HPV(-) HNSCC, we reasoned that HPV infection would induce changes in hydroxymethylation, especially near differentiation and developmental genes, and corresponding genes differing by HPV status or tumor subtype. Specifically, since 5hmC levels are higher in more differentiated cells and lower in stem-like cells(131), we hypothesized an overall higher 5hmC level in HPV(-) tumors, since they tend to be more differentiated. Additionally, we predicted differential 5hmC to exist between the IMU and KRT subtypes.

We used hydroxymethylated DNA immunoprecipitation sequencing (hMeDIP-Seq) to assess 5hmC in our tumor cohort, and integrated results with previously generated RNA-seq data from the same tumors(132,133). Results pointed to extensive differential hydroxymethylation both by HPV status and HPV(+) subtype. The 5hmC levels at both promoter and enhancer regions distinguish meaningful tumor subgroups and associate with survival. We found a strong positive correlation between hydroxymethylation and gene expression. By integrating 5hmC with gene expression, we detected important pathways enriched in comparison based on HPV status and subtypes, including keratinization and cell junctions. Finally, we found that a much

higher portion of hyper-hydroxymethylated regions for HPV(-) samples fall in keratinocyte enhancer regions compared with HPV(+) samples. Since some of these enhancers can be linked to differentially expressed target genes, this result indicates that both promoter and enhancer hydryoxymethylation play important roles in HNSCC gene regulation. Our results partially explain different mechanisms responsible for previously noted subtype differences and suggest that 5hmC could be a potential epigenetic target in HNSCC based on HPV status and HPV(+) subtype.

2.2 Methods

2.2.1 Patient recruitment and hMeDIP-seq protocol

From 2011 to 2013, we identified 36 incident HNSCC patients with pre-treatment oropharynx or oral cavity squamous cell carcinoma at Michigan Medicine Hospital. HPV status was determined based on p16 staining and RNA-seq, as previously described(50,60). The details of tumor tissue acquisition can be found in Additional file 1: Supplementary Methods. After DNA extraction, the quality of the 36 DNA samples was measured by TapeStation genomic DNA kit (Agilent, Santa Clara, CA), followed by quantitation assessment by Qubit broad range dsDNA (ThermoFisher, Carlsbad, CA). Enzymes, PCR primers and indexed adaptors were supplied by New England BioLabs (Ipswich, MA) and Integrated DNA Technologies (Coralville, CA), respectively.

A total of 1μg of genomic DNA was used for shearing, blunt-end repair and phosphorylation process, and a single adenine nucleotide was then added to the 3' end of the resulting fragments for ligation preparation. DNA was cleaned by Qiagen's MinElute PCR purification columns (Qiagen, Germantown, MD). DNA samples were denatured and

resuspended in ice-cold immunoprecipitation buffer after the addition of DNA spike-ins for hMeDIP (Diagenode Denville, NJ). At this stage, 10% volume of the DNA solutions were kept as inputs, and immunoprecipitation overnight at 4°C with rotation was performed on the remaining solution, after adding a 5hmC-specific antibody (Cat # 39791, Active Motif, Carlsbad, CA). The 5hmC-enriched DNA fragments (IP) were released from the antibody and cleaned-up by Proteinase K (ThermoFisher, Carlsbad, CA) and AMPure XP beads (Beckman Coulter, Brea, CA), respectively. In order to evaluate the percent enrichment over input in the IP, qPCR with primers for spike-ins was conducted. For samples with good percent enrichment over input, PCR amplification was performed for library production, followed by cleaning with AMPure XP beads and quantification with the Qubit assay (ThermoFisher, Carlsbad, CA) and TapeStation High Sensitivity D1000 kit (Agilent, Santa Clara, CA). Each hMeDIP-seq sample with paired input was sequenced on a single lane of an Illumina HiSeq 2500, generating single-end, 50 bp reads.

2.2.2 hMeDIP-seq data analysis and peak finding

The main analysis steps were conducted using the Methylation Integration (*mint*) pipeline(134). Sample quality was first assessed with FastQC(135), then reads were aligned with bowtie2(136) after adapter and quality trimming with Trim Galore!. Peaks for each sample compared to input, i.e. the genome-wide regions of hydroxymethylation, were identified using MACS2.

Differential peaks, i.e. differentially hydroxymethylated regions (DhMRs) between HPV(+) and HPV(-) samples, or between HPV(+) subtypes, were identified using PePr(137). PePr takes replicates into account using a negative binomial model while improving variability estimates using information from neighboring sites. Differential peaks were called with false

discovery rate (FDR) < 0.05 and fold change (FC) > 2. Peaks were annotated using the R Bioconductor package *annotatr*(138). Peaks annotated to X and Y chromosomes were excluded to avoid confounding by sex. The overall 5hmC levels over gene bodies were calculated using MACS2 peaks with *metaGeneProfile* function in HOMER(139).

2.2.3 Principal component analysis (PCA) and singular value decomposition (SVD) analysis

PCA was performed using *prcomp* function in R, with the use of hMeDIP-seq counts in proximal promoters, gene bodies and custom enhancer regions. X and Y chromosome reads were removed to avoid sex bias. The *bedtools intersect* function was used to obtain 5hmC counts in promoter regions (1kb before to 1kb after TSS's) and gene bodies (from TSS to TES), followed by normalization by manual specification of library sizes in *DESeq2*, with the input values as covariate. The background was taken into account by calculating the log2 fold change for each region. SVD analysis was performed on the top principal components using the Bioconductor package *ChAMP*(140), to study correlation with variables of interest.

2.2.4 Tumor tissue acquisition

With written informed consents from patients and approval from University of Michigan Institutional Review Board, pretreated tumor tissues were collected for these patients, stored and flash frozen in liquid nitrogen at -80C. For each sample, H&E slides were prepared before passing the assessment criteria of at least 70% cellularity and less than 10% necrosis. Surface scrapings were then extracted from the tissue region where at least 70% tumor cellularity was confirmed while remaining frozen, followed by processing with Qiagen AllPrep DNA/RNA/Protein Mini Kit (Valencia, CA, USA) as per manufacturer protocol.

2.4.5 Determination of important genes in differential 5hmC analysis

The genes that are important in differential 5hmC analysis are those with highest number of DhMRs, those with the most significant DhMRs, those with DhMRs on promoters, and those that are related to cancer, especially HNSCC. A list of cancer-related and HNSCC-related genes were downloaded from COSMIC(141).

2.2.6 Generation of custom enhancer definitions

We defined a consensus set of human distal enhancers (>5kb from a TSS) from a combination of sources for enhancer locations and links to their target genes. The definition was based on enhancers from DNase hypersensitive sites (DHSs) from 125 cell types processed by ENCODE(142), distal and non-promoter DHS within 500kb of the correlated promoter DHSs from 32 cell types and FANTOM5(143). All enhancer regions were extended up to 1kb around their midpoint. To identify target genes, we used enhancer and gene interactions (<5kb from TSS) identified by ChIA-PET2(144) in 13 ChIA-PET datasets(145,146) from 6 cell lines.

2.2.7 RNA-seq analyses and association with 5hmC

RNA-seq data from the same 36 HNSCC samples as for hMeDIP-seq was downloaded from GEO (accession number GSE74956)(60). The raw sequences were aligned to hg19 using Tophat2 v2.0.11(147), and gene expression count data were generated with HTSeq(148). Differential expression levels were compared both by HPV status and by HPV(+) subtypes using edgeR-robust, as described in (149). HPV status was previously determined by viral gene expression; briefly, samples with more than 500 read pairs aligned to any high-risk HPV genome were classified as HPV(+)(60). When comparing gene expression with 5hmC levels, for each identified DhMR by HPV-status, we calculated the log fold change (logFC) of 5hmC level as log HPV(+) / HPV(-), so that one gene corresponds to one gene expression logFC value and one or multiple 5hmC logFC values. The odds ratio (OR) of 5hmC logFC and gene expression logFC

was calculated with a standard 2x2 contingency table containing the number of sites in each quadrant of Figure 3A.

2.2.8 Gene set enrichment testing on hmeDIP-seq and RNA-seq data

RNA-Enrich(150) was used for enrichment analysis on RNA-seq results, with the directional analysis option so that enrichment of gene sets in HPV(+) and HPV(-) samples could be distinguished.Gene Ontology (GO) term and pathway enrichment analysis for 5hmC peaks was conducted using the gene set enrichment program for genomic regions, ChIP-Enrich(151). The enrichment analysis on DhMR peaks was performed using the nearest TSS locus definition in ChIP-Enrich, and significant pathways were marked with an FDR < 0.1. Only "enriched" GO terms were selected for follow-up analyses. Broad GO terms with more than 1,500 genes were removed to keep the results more specific.

2.2.9 Keratinocyte enhancer regions download and analysis

ChromHMM(152) tracks of Primary Normal Human Epidermal Keratinocytes (NHEK) were downloaded from the UCSC Genome Browser, and the subset of regions overlapping between the ChromHMM track (strong enhancer, weak enhancer, active promoter and weak promoter regions only) and DhMRs by HPV status were identified using the GenomicRanges Bioconductor package. NHEK super-enhancers were defined using ROSE (Rank Ordering of Super-Enhancers)(153,154), and were downloaded from DbSUPER(155). NHEK ChIP-seq data for 9 histone marks (H3K27ac, H3K37me3, H3K36me3, H3K4me1, H3K4me2, H3K4m3, H3K9ac, H3K9me1, H4K20me1) were downloaded from ENCODE. The overall 5hmC levels over active enhancers and weak enhancers were calculated using MACS2 peaks in HOMER(139).

2.3 Results

2.3.1 Widespread differential hydroxymethylation between HPV(+) and HPV(-) tumors

The HNSCC cohort consisted of 18 HPV(+) and 18 HPV(-) patients, as previously determined based on RNA-seq alignment to 14 known high-risk HPV genomes. The cohort consisted of 26 males and 10 females, with an overall median age of 57 years. A total of 14 of the 18 HPV(+) patients were infected by subtype HPV16, and most were former or current smokers. hMeDIP-Seq was performed on these 36 HNSCC tumors to define their genome-wide hydroxymethylation signatures, examine how they differed by HPV status and tumor subtypes, and assess their relationship with clinical variables.

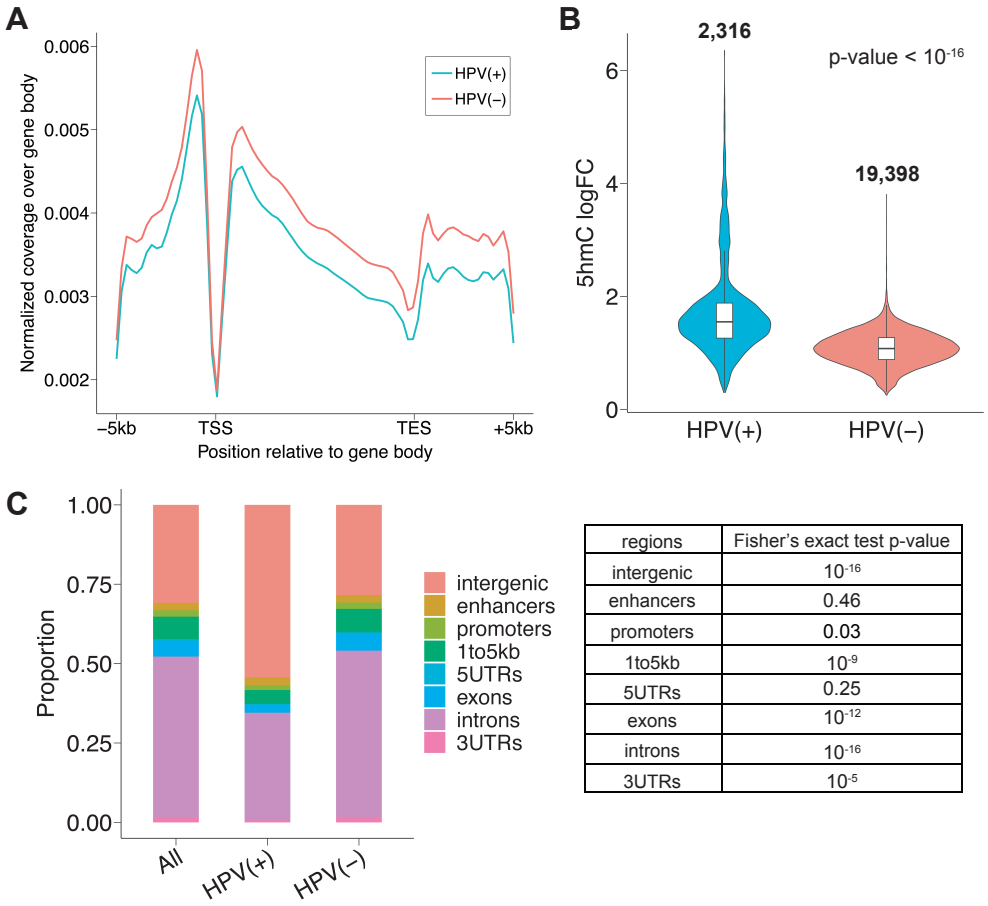
All 36 samples resulted in sufficient quality of data and hundreds of thousands of identified 5hmC peaks, with the sequencing depth ranging from 67.7 to 152.3 million reads mapped. The number of peaks detected generally ranged from 208,000 to 480,000. As expected, the number of peaks was positively correlated with the total number of reads mapped. There were no significant differences between HPV(+) and HPV(-) tumors based on peak numbers (p-value = 0.355, Wilcoxon signed-rank test), even after accounting for millions of reads mapped (p-value = 0.287, ANOVA test).

In general, a much higher level of hydroxymethylation (hyper-5hmC) was reported in HPV(-) HNSCC. Enrichment of 5hmC levels was plotted over gene bodies, and we observed a consistently higher level of 5hmC in gene bodies across the genome in HPV(-) tumors (**Figure 2.1A**). Consistent with previous studies, the average gene body profiles revealed a dip around the transcription start site (TSS) and transcription end site (TES) regions(156).

A total of 19,398 differentially hydroxymethylated regions (DhMRs) were detected as hyper-5hmC in HPV(-), as opposed to only 2,316 in HPV(+) tumors (p-value < 10^{-5}). Most

differential peak widths were narrow, between 100 - 200bp, and HPV(-) DhMRs were slightly longer than those for HPV. Although fewer peaks were hyper-5hmC in HPV(+) samples, they were in general stronger than the HPV(-) DhMRs, with larger fold change (**Figure 2.1B**).

Of the 2,316 DhMRs in HPV(+) tumors, about 46% were annotated to genes, with the majority in introns. Of the 19,398 DhMRs in HPV(-) tumors, more than 72% were annotated to genes, also with the majority in introns. By comparing the distribution of HPV(+) and HPV(-) DhMR annotations to the annotations of random genomic regions, we found a significantly higher proportion of DhMRs were in exons (5.34%) and introns (52.77%) in HPV(-) samples, as opposed to a smaller percentage in exons (2.32%) and introns (33.98%) for HPV(+) DhMRs (**Figure 2.1C**). Together, these results suggest that HPV positivity in HNSCC is linked to a reduced hydroxymethylation signature both in and around genes.



regions	Fisher's exact test p-value
intergenic	10^{-16}
enhancers	0.46
promoters	0.03
1to5kb	10^{-9}
5UTRs	0.25
exons	10^{-12}
introns	10^{-16}
3UTRs	10^{-5}

Figure 2.1 The level and distribution of hydroxymethylation varies by HPV status. (A) Global 5hmC distribution pattern over gene bodies in both HPV(+) and HPV(-) samples.
(B) Violin plot of 5hmC logFC in HPV(+) tumors (left) and HPV(-) tumors (right). The number on top indicates the total number of peaks being tested. Wilcoxon signed-rank test p-value < 10^{-16}. (C) The distribution of hyper-5hmC peaks from HPV(+) and HPV(-) samples, where first column represents the combination of HPV(+) and HPV(-) tumors. The table on the right displayed the p-values from Fisher's exact test between HPV(+) and HPV(-) HNSCC, where exons and introns showed a p-value of 10^{-12} and 10^{-16} respectively.

2.3.2 Genes and pathways with hyper-5hmC in HPV(+) tumors

A total of 623 genes were associated with hyper-hydroxymethylation in HPV(+) tumors. CDKN2A (p16), the most important biomarker of HPV status in clinical tests(157), had one of the highest number and also the most significant HPV(+) DhMRs. In total, 35 genes had hyper-

5hmC in promoter regions of HPV(+) tumors. As a prime example, 5hmC was enriched at the CDKN2A promoter in HPV(+) cases compared to HPV(-), a difference that was independent of copy number variations since this is controlled for by the use of input references. This is important, since loss of the CDKN2A locus is known to occur in HPV(-) cases. The raw coverage depth by sample and peak signal values of CDKN2A both showed great deviation in the promoter and along the gene body between HPV(+) and HPV(-) tumors.

Pathways enriched with higher 5hmC in HPV(+) HNSCC included desmosome, activation of NF-kappaB-inducing kinase activity, oxidoreductase activity, and mesenchymal cell differentiation (FDR<0.1). Mesenchymal development associated with epithelial to mesenchymal transition (EMT) is consistent with the higher risk of distant metastasis in HPV+ HNSCC. A total of 27 genes displayed higher 5hmC in HPV(+) tumors in mesenchymal cell differentiation and development, including the key EMT genes SNAI2, BMP2, SMAD2 and TGFB2, which are part of the TGFβ / Bone Morphogenic Protein (BMP) signaling pathway(158,159).

2.3.3 Genes and pathways with hyper-5hmC in HPV(-) tumors

A larger number of genes, 5,584, were found to be hyper-hydroxymethylated in HPV(-) HNSCC, of which 372 genes contained at least one promoter region DhMR. Some of the most important genes with promoter DhMRs were BCAR1, which plays crucial roles in metastasis and cell adhesion, and MMP2, which functions in EMT and immune response in multiple cancer types. We found 204 genes to harbor more than 10 HPV(-) hyper-5hmC regions. CDH13, a gene encoding a member of the cadherin superfamily that functions in cell-to-cell adhesion and is involved in several diseases, had 83 DhMRs. The peak signal values over each DhMR indeed demonstrated a higher level of hydroxymethylation in HPV(-) compared with HPV(+).

Pathway enrichment results identified cell morphogenesis, cell death, cell motility, and cytoskeletal rearrangement / cell-cell junction being among the significantly enriched (FDR<0.1). Several previously verified HNSCC-related genes were in a top enriched pathway. For example, frequently mutated HNSCC genes ERBB2, FGD1, NOTCH1, NR4A2, SEMA3E and ARAP3 had HPV(-) DhMRs in cell morphogenesis, and other head and neck-relevant genes such as CTGF, PKN2, TERT, TGFBR2 and TP63 in signal transduction had DhMRs.

2.3.4 Main sources of heterogeneity in hydroxymethylation in promoter and enhancer regions

We next sought to understand the sources of 5hmC heterogeneity in our cohort using principle component analysis (PCA). Interestingly, the greatest source of heterogeneity in promoter 5hmC profiles did not distinguish HPV(+) from HPV(-) tumors, but rather one subtype of the HPV(+) tumors (IMU) from all other tumors (see PC2 in **Figure 2.2A, 2.2C**). Consistent with the findings in Zhang, et al(60), the KRT subtype groups closer with HPV(-) HNSCC, partially due to the shared similarities of heightened keratinization. We thus sought to determine which, if any, known variables could explain the top PCs among HPV(+) tumors. Correlations between the top principle components and clinical, demographic and batch information, were calculated using singular value decomposition (SVD) analysis on proximal promoter regions for the 18 HPV(+) samples. PC1 was correlated with survival (p-value < 0.05), while both subtype (p-value < 0.01) and percentage of epithelial tissue (p-value < 0.05) were significant in PC2 (**Figure 2.2E**). TILs (tumor-infiltrating lymphocytes) score and batch effect were correlated with PC3, while survival and recurrence (p-value < 0.01) were correlated with PC4.

In terms of 5hmC heterogeneity in enhancer regions among the 36 tumors, a similar distinction between IMU and the other tumor samples was observed (see PC2 in **Figure 2.2B,**

2.2D). Correlating variables with the top enhancer PCs for the 18 HPV(+) samples, survival was again a significant factor in PC1 (p-value<0.05), and subtype (p-value < 0.01) and epithelial tissue (p-value < 0.05) were again both significant in PC2 (**Figure 2.2F**). However, the separation between the HPV(+) IMU subtype and the rest was not observed for 5hmC gene body. Instead, SVD analysis on gene body 5hmC showed other relevant clinical features associated with the top PCs, including survival, lymphocyte tissue and HPV integration status.

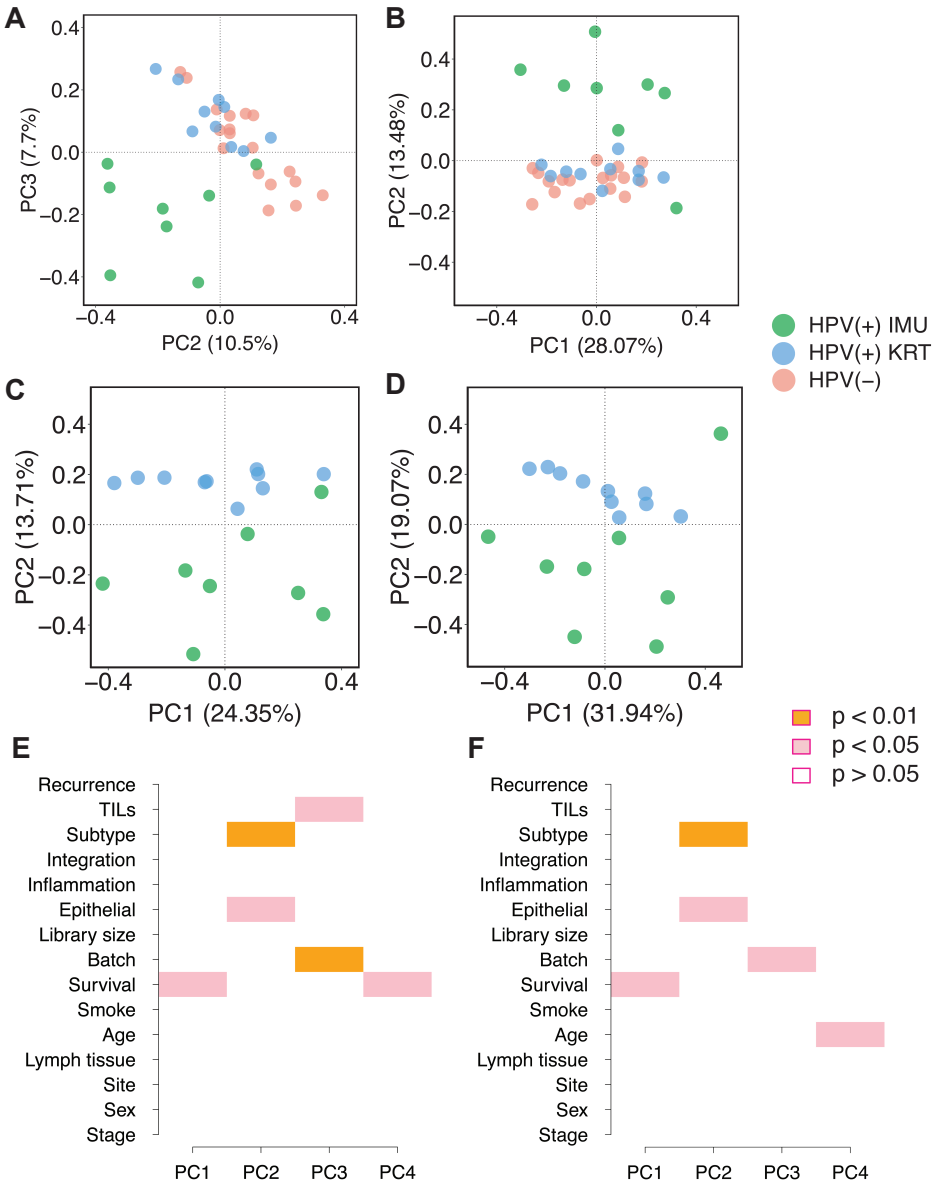


Figure 2.2 Principle component analysis illustrated sources of heterogeneity in 5hmC levels among HPV(+) HNSCC tumors.
(A) PC2 vs PC3 for proximal promoter regions. (B) PC1 vs PC2 for custom defined enhancers regions showed clear separation between the IMU subtype and the rest on PC2, which contributed 10.5% and 13.48% of the total variance respectively. (C-D) PC1 vs PC2 for both proximal promoters and custom defined enhancers showed clear separation by subtype within HPV(+) tumors. (E-F) The SVD analysis on proximal promoters and enhancers demonstrated several relevant clinical variables, such as survival, percentage of epithelial tissue and subtype, which significantly correlated with each principle component in 18 HPV(+) samples.

2.3.5 Differential hydroxymethylation and enriched pathways between the IMU and KRT tumor subtypes

Compared with the IMU subtype, an overall higher level of 5hmC was observed in the KRT subtype across the gene bodies, which was closer to the 5hmC levels of HPV(-) tumors. In terms of differential 5hmC between the two HPV(+) subtypes, there were significantly more instances of hyper-hydroxymethylation in the IMU subtype samples. A total of 63,859 hyper-5hmC regions were found in the IMU subtype, compared with only 1,833 hyper-5hmC regions in the KRT subtype. Only 838 (1.3%) of these regions were also among those found different by HPV status, out of which the majority (771 peaks) were hyperhydroxymethylated in HPV(+) and IMU tumors. Similar to the annotation of DhMRs based on HPV status, the majority of both IMU and KRT DhMRs were mapped to. Interestingly, cancer genes CDH13 and BCAR1 were found with multiple KRT DhMRs, and they were also important genes for hyper-5hmC in HPV(-) tumors, which is consistent with the previous finding that the KRT subtype shares more similarities with HPV(-) HNSCC.

The top enriched pathways marked by hyper-5hmC in the IMU subtype include cornification, epidermis development, keratinocyte differentiation, keratinization and cell differentiation. For the KRT subtype, cornification and keratinocyte differentiation also appear in the top enriched pathways. However, multiple pathways relevant to cytoskeleton organization or cell-cell junction, including cell adhesion and cytoskeleton structure, were found only in the KRT subtype, and were within the top 20 enriched terms. These terms were also significantly hyper-hydroxymethylated in HPV(-) compared to HPV(+), further explaining the similarity of the HPV(+) KRT subtype to HPV(-) HNSCC.

2.3.6 Hydroxymethylation is highly associated with gene expression in HNSCCs

RNA-seq data on the same 36 HNSCC samples were previously analyzed, resulting in 1887 up-regulated and 1644 down-regulated genes in HPV(+) samples (FDR<0.05 and absolute fold change>2) (60,160). A clear pattern of association can be observed between gene expression and 5hmC (Pearson's correlation coefficient = 0.62; odds ratio (OR) = 64.5) (**Figure 2.3A**), suggesting that 5hmC likely drives many of these observed differences. This positive correlation still holds when comparing gene expression with 5hmC logFC at enhancer, promoter, and gene body separately, with gene body regions showing the strongest correlation (Pearson's correlation coefficient = 0.53). The majority of genes (61.4%) are upregulated and hyper-hydroxymethylated in HPV(-) tumors, such as cell adhesion genes (including CDH13, CDH11, CDH2, CD44, GLI2, COL4A6), immune response genes (including TGFBR2, CD109, BCAR2, TIMP2, MMP2) and keratinization genes (CDH13, CD109, CDR, PALLD). In particular, TIMP2 and MMP2 also function in tumor invasion.

For CDKN2A, we determined that differential 5hmC was especially prominent in a 5kb region 5' of the second exon (chr9:21975000-21980000). To assess the association of this particular region with CDKN2A expression, we calculated 5hmC coverage per sample and found that 5hmC at this single region explained half of the variability in CDKN2A gene expression levels among the samples (Pearson's correlation coefficient = 0.7). (**Figure 2.3B**).

We next sought to determine the extent to which hydroxymethylation in promoter regions explained differences in gene expression. Twenty genes had at least one hyper-5hmC region in HPV(+) compared to HPV(-) in their promoter region, after excluding very low expressed genes. The expression of these genes was indeed significantly higher in HPV(+) IMU than HPV(-) samples (ANOVA p-value = 0.00066) (**Figure 2.3C**). The distinction between HPV(+) and HPV(-) samples was particularly clear for CDKN2A (**Figure 2.3D**). On the other hand, 296

genes had at least one hyper-5hmC region in HPV(-) in the promoter. However, there was no significant distinction found in the expression of the top 20 of these genes in any of the three comparisons (**Figure 2.3C**), leading us to hypothesize that the upregulation of genes due to hyper-5hmC in HPV(-) tumors is due to differences at enhancers rather than promoters. By using the genes with hyper-5hmC peaks on the promoter regions, we built networks using the shortest paths. The result showed that p16INK4 (CDKN2A) and p14ARF (alternate reading frame protein product of the CDKN2A locus) are the two center nodes for HPV(+) samples, while SMAD3, ABCC2 and IL32 are center nodes in the HPV(-) network.

Overall, we identified 35 GO terms both enriched with differential hydroxymethylation (whether proximal to the promoter or distal elements) and differential expression between HPV(+) and HPV(-) samples (**Figure 2.3E**). Twenty-five of these 35 GO terms were upregulated and hyper-hydroxymethylated in HPV(-) samples, including adherens junction, cell morphogenesis, chemotaxis and Ras signaling. This finding suggests that the higher expression of cell junction genes in HPV(-) tumors is at least partly regulated by hydroxymethylation. It also suggests that HPV infection could impact many cell junction biomarkers via the active demethylation process.

Next, we explored the enriched pathways based on HPV(+) subtype, finding 38 GO terms with significant hyper-5hmC and up-regulation in the IMU subtype, including cell migration, phosphorylation, MAPK cascade and cytokine-mediated signaling pathways. Conversely, there were 11 GO terms with significant hyper-5hmC and up-regulation in the KRT subtype, including cell-cell junction, keratinization and epidermal cell differentiation, which is consistent with the more differentiated nature of the KRT subtype.

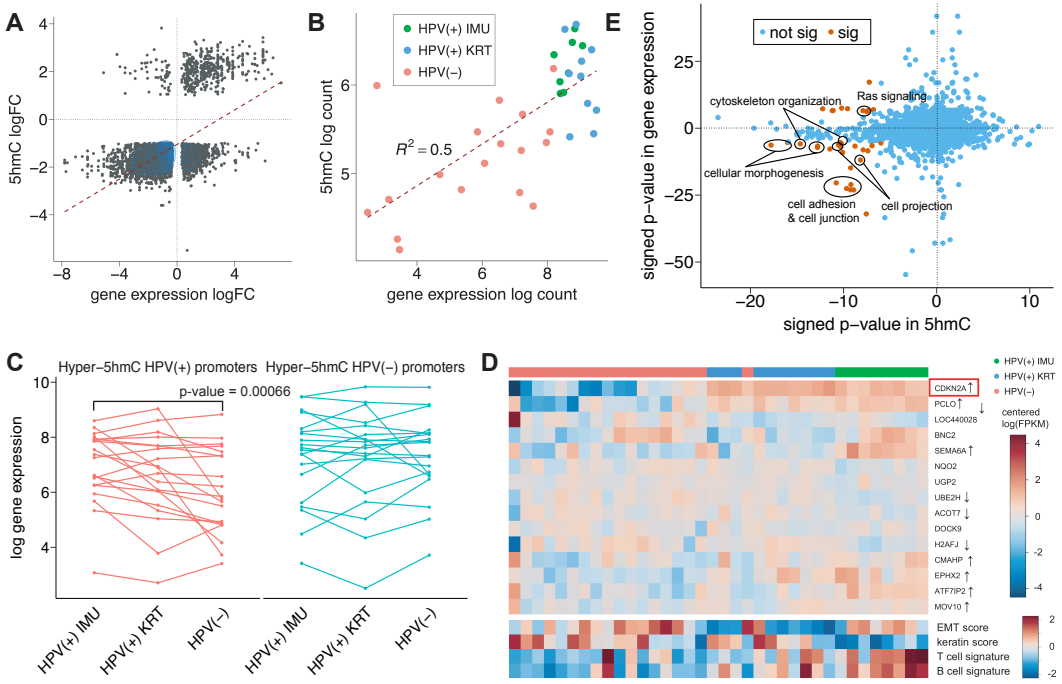


Figure 2.3 5hmC in HNSCC is highly correlated with gene expression.
(A) Scatterplot showing the positive correlation between gene expression and hydroxymethylation in HPV(+) and HPV(-) samples (Pearson correlation coefficient = 0.62). The top half represents genes that are significantly up-regulated in HPV(+) tumors, and the right half represents genes that are significantly hyper-hydroxymethylated in HPV(+) tumors. (B) Scatterplot showing a strong correlation between log gene expression and log 5hmC of 5kb intron region (chr9:21975000-21980000) of CDKN2A. HPV(+) samples were concentrated near the top right corner, indicating that both their gene expression and 5hmC coverage were higher compared with HPV(-) samples. (C) Spaghetti plot of log gene expression for top 20 genes with at least one DhMR in their promoter region for HPV(+) IMU, HPV(+) KRT and HPV(-) samples, respectively. (D) Heatmap showing the expression levels of sufficiently expressed genes with at least one HPV(+) DhMR in their promoter, which were mostly well clustered based on HPV status (marked with black and grey at the top). Most genes were significantly higher expressed in HPV(+) samples, such as CDKN2A. ↑ indicates genes that are also up-regulated in HPV(+) HNSCC, and ↓ indicates up-regulation in HPV(-) tumors. Keratin and EMT scores are measurements of keratinization level and EMT level, and T cell signature and B cell signature represent degree of immune response. Generally there was a higher level of keratinization and EMT in HPV(-) samples, while the immune response is more significant in HPV(+) samples. The detailed calculation can be found in Zhang et al. (60). (E) Enrichment analysis results for gene expression vs hydroxymethylation by HPV status. Each dot represents one GO term, and the color denotes the significance (yellow: significant; blue: not significant). Signed p-values are defined as >0 to indicate up-regulation in HPV(+) samples or hyper-hydroxymethylation in HPV(+) samples, and <0 to indicate upregulation in HPV(-) samples or hyper-hydroxymethylation in HPV(-) samples.

2.3.7 HPV(-) tumors are more hydroxymethylated in keratinocyte enhancer regions than HPV(+) tumors

Since 5hmC has been shown to be an important mark in enhancer regions(161), we revealed the distribution of hydroxymethylation across samples in relation to different chromatin states, including enhancers, in primary Normal Human Epidermal Keratinocyte (NHEK) cells. NHEK cells are isolated from the epidermis of juvenile foreskin or adult skin, and are similar to head and neck tissue both morphologically and physiologically(162). Around the center of NHEK active enhancers and weak enhancers, we observed much higher levels of 5hmC in HPV(-) tumors compared with both HPV(+) subtypes, although all of them showed the expected pattern of an increase around the enhancer centers. Enhancers also had more differences in hydroxymethylation than promoter regions, as indicated by strikingly more DhMRs in strong enhancers than in active promoters for both HPV(+) and HPV(-) tumors (**Figure 2.4A, Table 2.1**). Fisher's exact test showed an odds ratio (OR) of 2.56 (p-value < 10^{-16}) for strong enhancers and 2.43 (p-value < 10^{-16}) for weak enhancers comparing HPV(+) vs HPV(-) samples. This disparity between HPV(+) and HPV(-) samples is consistent with the fact that HPV(-) HNSCCs tend to be more differentiated than HPV(+) HNSCC.

	Strong enhancers	Weak enhancers	Active promoters	Weak / poised promoters	Super-enhancers
HPV(+) DhMRs (2,316)	94; 4.06%	110; 4.75%	17; 0.73%	29; 1.25%	34; 1.47%
HPV(-) DhMRs (19,398)	1909; 9.84%	2136; 11.01%	66; 0.34%	273; 1.41%	716; 3.45%
HPV(+) IMU DhMRs (63,859)	4971; 7.65%	6578; 10.30%	225; 0.35%	777; 1.22%	1374; 2.15%
HPV(+) KRT DhMRs (1,833)	230; 12.66%	197; 10.75%	8; 0.44%	20; 1.09%	117; 6.38%

Table 2.1 Number and percentage of HPV(+), HPV(-), HPV(+) IMU and HPV(+) KRT DhMRs that overlap with keratinocyte enhancers, promoters and super-enhancers.

We reconfirmed the higher keratinocyte (NHEK) enhancer 5hmC levels in HPV(-) tumors using ChIP-seq data for the histone mark H3K4me1, a mark for active and primed enhancers. Visualizing the H3K4me1 signals for the top 1000 hyper-5hmC regions for HPV(+) and HPV(-) tumors separately, showed the highest signal value within HPV(-) hyper-5hmC regions (**Figure 2.4B**). Similar to previous findings, this trend was not observed for H3K4me1 signals in HPV(+) tumors, indicating that HPV(-) samples have higher levels of 5hmC in keratinocyte enhancer regions.

The target genes of strong enhancers with at least one DhMR were determined using publicly available ChIA-PET data. In order to study the impact of enhancer hydroxymethylation on target gene expression, we specifically focused on target genes that were also differentially expressed. There were 5 and 66 hyper-hydroxymethylated enhancers associated with differentially up-regulated genes in HPV(+) and HPV(-) tumors, respectively. In particular, CLDN1, a cell-to-cell adhesion gene, was the target gene of a HPV(+) enhancer DhMR, and was also up-regulated in HPV(+) tumors. Conversely, differentially expressed genes CDH13, BCAR1 and TIMP3 not only displayed HPV(-) enhancer hyper-5hmC, but also contained multiple HPV(-) DhMRs in their exonic and intronic regions. Multiple strong enhancers showed a higher level of hydroxymethylation in HPV(-) samples in all three of these genes (**Figure 2.4C**).

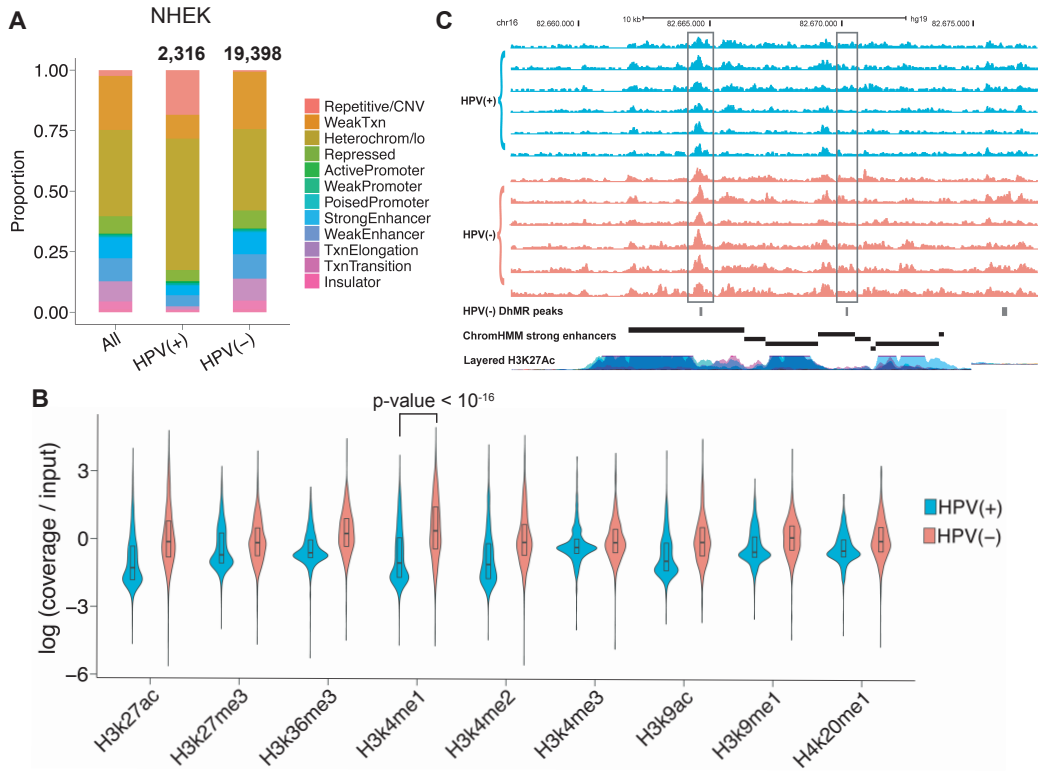


Figure 2.4 HPV(-) tumors have strong hyper-5hmC regions in epithelial and keratinocyte enhancer regions.
(A) HPV(-) HNSCCs showed much higher portion of DhMRs in strong and weak enhancers than both random regions and HPV(+) tumors. Regions were defined using the ChromHMM track for NHEK cells. The number on top indicates the total number of peaks being tested. (B) Violin plot showing ChIP-seq log(coverage/input) values for top 1000 HPV(+) and top 1000 HPV(-) hyper-5hmC peaks for 9 histone marks. Wilcoxon signed-rank test showed a p-value < 10^{-16} for H3k4me1 peaks. (C) UCSC Genome Browser view of NHEK strong enhancers near CDH13. Data shown are 5hmC profiles for 6 representative HPV(+) samples (upper 6 tracks) and 6 representative HPV(-) samples (lower 6 tracks), showing 3 regions of high 5hmC level for HPV(-).

2.3.8 Expression of invasion gene MMP2 in HPV(+) and HPV(-) cell lines

Our results show that many immune response genes, such as BCAR1, TIMP2 and MMP2, were both higher expressed and hyper-hydroxymethylated in the promoter and sometimes enhancer regions in HPV(-) tumors. In a previous study, 5hmC was shown to be positively correlated with depth of tumor invasion in colorectal cancer(163), and depletion of TET1, which could oxidize 5mC to 5hmC, could facilitate cell invasion(164). Along these same

lines, multiple studies reported that high levels of the MMP2 protein were linked with larger tumor size and more tumor invasion(165). Therefore, we hypothesized that hyper-hydroxymethylation and higher mRNA levels of MMP2 would result in higher MMP2 secreted protein, which can lead to stronger invasion in HPV(-) HNSCC.

First we reconfirmed that MMP2 in HPV(-) tumors showed both higher gene expression and overall higher 5hmC compared with HPV(+) tumors, and there is a positive correlation between the gene expression and 5hmC (**Figure 2.5A**). Next, we assessed the secreted and intracellular protein levels of MMP2 in two HPV(+) oropharynx, one HPV(+) oral cavity, and two HPV(-) oral cavity cell lines. The zymogram results showed higher levels of secreted MMP2 in HPV(-) cells compared to HPV(+) oropharynx cell lines, but not HPV(+) oral cavity (**Figure 2.5B**). Intracellular MMP2 was also higher in the HPV(-) cell lines than the HPV(+) oropharynx cells (**Figure 2.5C**). There is no clear distinction in the MMP2 mRNA levels between HPV(-) oropharynx and HPV(+) oral cavity cell lines (**Figure 2.5D**).

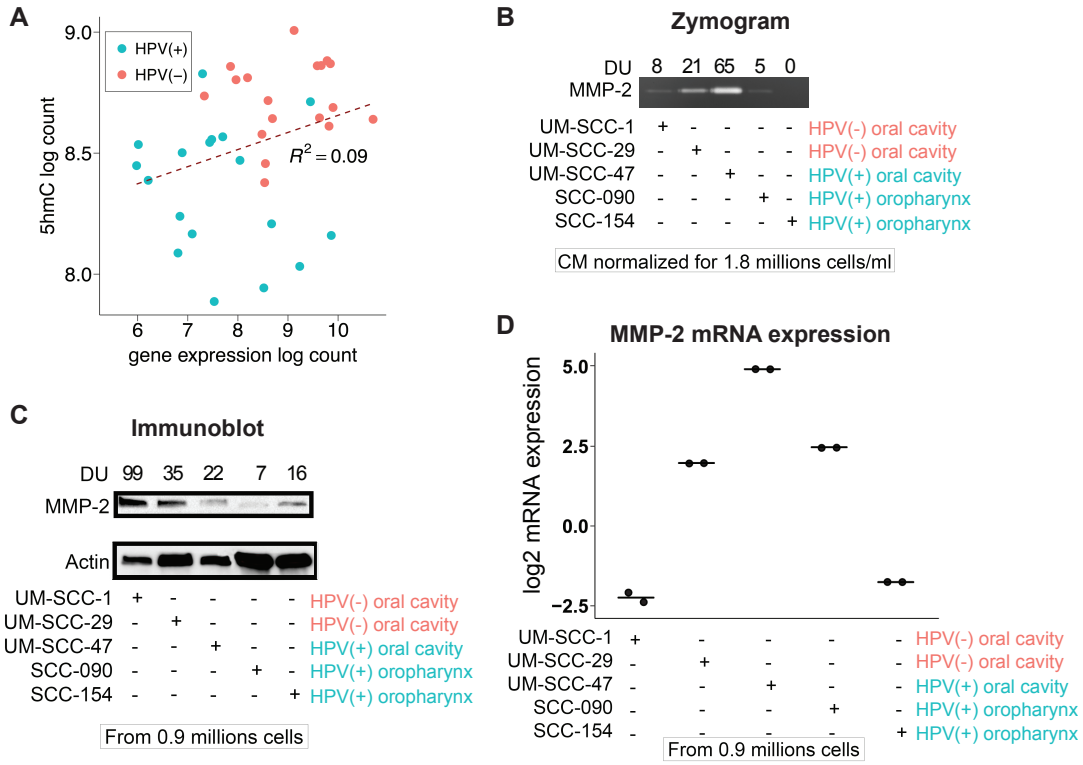


Figure 2.5 Protein expression level of MMP2 in 2 HPV(-) and 3 HPV(+) HNSCC cell lines.
(A) Scatterplot showing a correlation between log gene expression and log 5hmC of MMP2 gene. HPV(-) samples were concentrated near the top right corner, indicating that both their gene expression and 5hmC coverage were higher compared with HPV(+) samples. (B) Zymogram and (C) immunoblot indicated the secreted and intracellular level of MMP2 protein respectively. (D) Dot plot showing the mRNA level of MMP2 with two replicates. After removing the HPV(+) oral cavity cell line, which is an outlier, Wilcoxon signed-rank test showed a p-value = 0.34.

2.4 Discussion

5-hydroxymethylcytosine has been shown to be depleted in various human cancers, and is known to be more concentrated in differentiated cells(128,131). Stem cells, which are closest to the basal epithelial cells in HNSCCs, are known to have lower 5hmC levels, especially in the gene and enhancer regions required for differentiation(132,156). For our purposes, those differentiation-specific regions would be epithelial and keratinocyte-specific genic and enhancer regions. Differences between HPV(+) and HPV(-) HNSCC are extensive in terms of prognosis,

tumor recurrence patterns and survival(53,55,56). Similarly, molecular studies have shown marked differences in gene expression, DNA copy numbers, and DNA methylation profiles by HPV(+) status(60,120,122). For instance, multiple studies showed genome-wide DNA hypomethylation in HPV(-) HNSCC tumors(123), and the differential methylation of certain tumor suppressor genes could be potential markers for early HNSCC diagnosis(37).

The HPV lifecycle is tightly linked to epithelial cell differentiation, with HPV initially infecting the undifferentiated basal epithelial cells, and concluding its life cycle in differentiated keratinocytes. Upon HPV oncogene integration, heightened keratinization often occurs, potentially affecting metastatic risk(166). Our group has shown that patients with integrated HPV E6 & E7 had significantly worse overall survival(50). Unlike most cancers, evidence does not suggest that less differentiated HNSCCs are associated with worse survival; indeed, studies have suggested that more differentiated keratinocytes are associated with worse survival in oropharyngeal cancer(167). Due to the limited sample size of this study, we did not observe any significant survival difference between the two HPV(+) subtypes, where 2 and 0 deaths were reported after 36 months follow-up in the KRT and IMU subtype, respectively. Another recent publication on meta-analysis of HPV(+) OPSCC followed the observed trend of the IMU subtype having the best survival (labeled Cl1 in the paper) and the subtype with 100% HPV integration (similar to our KRT subtype; Cl2 in paper) having worse survival (62).

Although 5hmC tends to be overall lower in cancers, others have observed lower 5hmC in oral cancers to actually be indicative of better prognosis(130) (**Figure 2.6**). This could be due to the confounding effect of HPV; in this study we found lower 5hmC in HPV(+) patients, who have less differentiated tumors and better prognosis (**Figure 2.6**). However, mesenchymal differentiation may also lead to loco-regional or distant metastasis; thus the complete relationship

between differentiation and metastasis appears to be complex, as suggested by studies in oral cancer(168). Among our differential 5hmC genes, only CDK6 was identified as a clinically actionable target. CDK6 was hyper-hydroxymethylated in HPV(-) tumors at seven intronic regions, is targeted in the treatment of certain breast cancers(169), and recently showed response in the treatment of oral squamous cell carcinoma(170). CD20, which is regulated by the epigenetic markers NFκB and SMAD2/3(171), is targeted in chronic lymphocytic leukemia and follicular lymphoma(172). The NFκB pathway and SMAD2/3 genes were identified as important 5hmC markers in our study, suggesting a potential use of B cell markers in the immunotherapy of HNSCC.

Our study is the first to characterize genome-wide DNA hydroxymethylation in head and neck cancers. Among differentiation genes, we found strong hyper-5hmC in HPV(-) tumors especially concentrated in cell junction and cell adhesion pathways, which is consistent with previous findings that HPV(-) HNSCC is more differentiated compared with HPV(+) HNSCC. The overall higher level of hydroxymethylation in HPV(-) HNSCCs is also consistent with observed overall higher levels of DNA methylation in HPV(+) oropharyngeal cancer cases, considering the antagonizing effect of methylation and hydroxymethylation(123). Similar to DNA methylation, profiles of hydroxymethylation were highly affected by HPV status, particularly for p16. The great majority of hyper-hydroxymethylated genes were in HPV(-) HNSCC, many of which play important roles in cancer pathways. CDH13, a cell-cell adhesion gene and a member of the cadherin superfamily, was upregulated in HPV(-) tumors and also had the highest number of hyper-5hmC regions in HPV(-). Interestingly, other major cadherin family genes, (e.g., CDH1 and CDH11), are hypermethylated in HPV(+) HNSCC(122,173). Some genes known to have differential methylation by HPV status also displayed differential 5hmC in

the opposite direction, including cell adhesion genes COL4A6 and BCAR1, and tumor suppressor genes TIMP3 and SFRP4.

Even more than distinguishing HNSCC tumors by HPV status, 5hmC profiles distinguished the IMU HPV(+) subtype from the KRT HPV(+) subtype and HPV(-) tumors. The KRT subtype is more similar to HPV(-) HNSCC than the IMU subtype based on gene expression(60). Consistent with this, we found the same based on 5hmC, with a much higher level of 5hmC found in the KRT subtype, which is furthermore consistent with the more differentiated nature of this subtype. We also found cancer pathways such as cell migration enriched with hyper-5hmC in the IMU subtype, while cornification and keratinization were significantly enriched with hyper-5hmC in the KRT subtype.

A recent study on hydroxymethylation of pancreatic cancer showed positive correlation between 5hmC and open chromatin generated ATAC-seq in both cancer and control cells(174). While we did not have ATAC-seq data available, in our study we found a similar correlation with ChromHMM tracks of NHEK cells(152), and the especially strong enrichment of 5hmC on enhancers in HPV(-) HNSCC could partially be attributed to its more differentiated and/or malignant nature.

In summary, our comprehensive characterization of the genome-wide hydroxymethylation profiles in HNSCC revealed significant differential hydroxymethylation both by HPV status and between HPV(+) subtypes. We report the significance of CDKN2A hydroxymethylation by HPV status, as well as many other cancer-related genes, such as CDH1, TIMP3 and SFRP4. Overall, the results are closely in line with current knowledge of differences by HPV status, including differences in DNA methylation. We also discovered the important role

of the less reported gene CDH13 in HNSCC, and that the differential hydroxymethylation was especially concentrated in CDH13 enhancer regions.

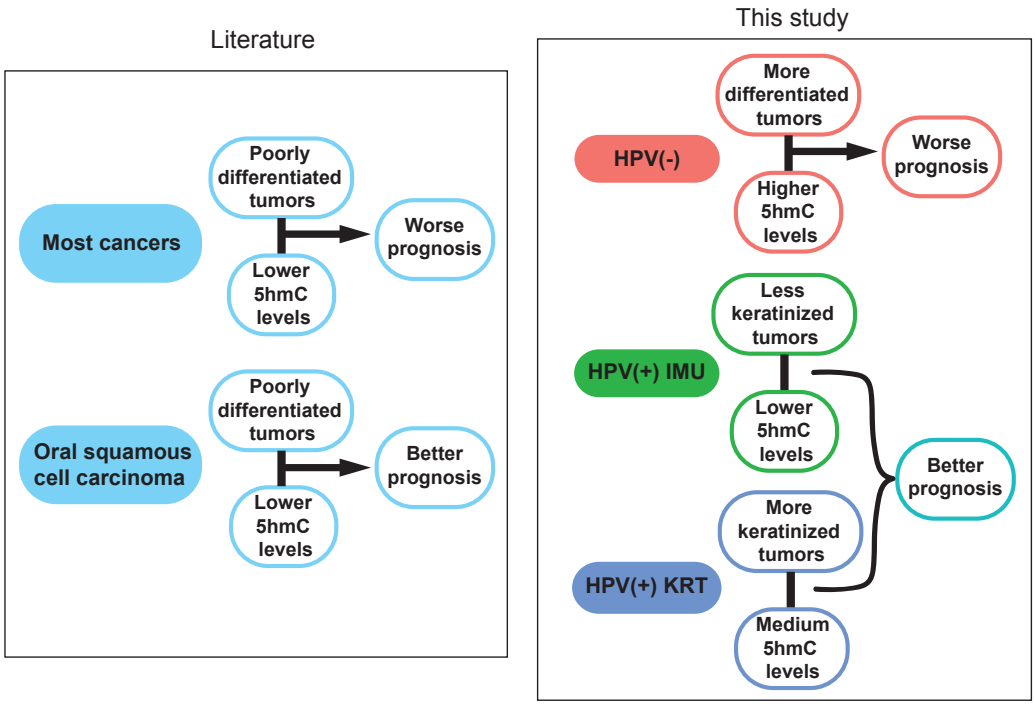


Figure 2.6 the correlation between 5hmC level and prognosis within different tumor types.
Schematic chart summarizing the correlation between 5hmC level and prognosis within different tumor types in published literature (left) and our study (right).

Chapter 3

Characterization of Immune Response in Patients with Cancer of the Oral Cavity after Neoadjuvant Immunotherapy with the IRX-2 Regimen

A manuscript covering this work is in preparation, with myself as first author,

3.1 Introduction

Head and neck squamous cell carcinoma (HNSCC) is the sixth most prevalent cancer worldwide, affecting approximately 680,000 patients every year(33,175). Traditional treatment of HNSCC involves a combination of surgery, radiotherapy and/or chemotherapy, and the five-year survival rate ranges from 55% to 66%(176). Immunotherapy with immune checkpoint inhibitors (ICI) has shown revolutionary progress in the treatment of multiple cancers. The PD-1/PD-L1 pathway is a key mechanism targeting T cell regulatory functions to enhance anti-tumor immune response(177), and PD-1 inhibitors pembrolizumab and nivolumab showed clinical advances in recurrent and metastatic HNSCC patients. However, the overall response rate for pembrolizumab is only 16% to 18% regardless of human papillomavirus (HPV) status(70,72), while patients treated with nivolumab have demonstrated significantly longer overall survival than those treated with standard single-drug systemic therapy(178). The challenge remains to discover how to increase these encouraging response rates and clinical results with novel immunotherapy approaches.

IRX-2 is a biologic product that contains multiple cytokines derived from normal donor phytohemagglutinin (PHA) stimulated mononuclear cells. The primary active components in IRX-2 are IL2, IL1β, IFNγ and TNFα. Previous studies showed that IRX-2 can protect T cells from activation-induced cell death and promote the cytolytic functions of natural killer (NK) cells(179). In addition to the direct effects of IRX-2 on antitumoral immunity, IRX-2 could potentially prime the tumor for positive response to immune checkpoint inhibitors. In a completed Phase 2a clinical trial in patients with oral carcinoma, IRX-2 immunotherapy was associated with increased immune infiltration and chemokine receptor expression using multiplex immunohistochemistry (IHC) and transcriptome analysis (NanoString Technologies) from 7 matched pre- and post-treatment tumor specimens(180).

Apart from the expression of immune response genes, methylation changes also play an important role in cancer immunotherapy. Hypermethylation at CpG islands and promoters as well as global hypomethylation have been documented in HNSCC in multiple studies (38,40,181). In particular, the hypomethylation of human retroelement long interspersed nucleotide element-1 (LINE-1) was associated with higher risk of oral cavity cancer (OCC) relapse, and is a potential predictive biomarker for OCC (182,183). In a few recent immune-related HNSCC studies, the characterization of DNA methylation profiles predicted a combination of hyper- and hypo-methylation markers in response to immune checkpoint inhibitors (ICI), and these methylation signatures could act as surrogate biomarkers to predict responsiveness to PD-1/PD-L1 inhibition (78–81). In addition, accompanied by genomics data, a methylation-based cell type deconvolution approach may distinguish between “immune hot” and “immune cold” HNSCCs (184).

Based on promising findings from a prior Phase 2a trial, we studied immune-related gene expression and genome-wide DNA methylation in tumor specimens from a larger scale randomized Phase 2b clinical trial. The trial was designed to determine whether IRX-2 cytokines induce increased lymphocyte infiltration into primary tumors in HNSCC and how these cytokines are active in patients with oral cavity cancer. The IRX-2 regimen started 3 weeks prior to surgery, consisting of an initial dose of cyclophosphamide followed by 10 days of regional perilymphatic IRX-2 cytokine injections and daily oral indomethacin, zinc and omeprazole (Regimen 1). The control regimen was identical to Regimen 1 without the IRX-2 cytokine injections (Regimen 2). A total of 96 patients were randomized 2:1 to Regimen 1 or Regimen 2 (64:32). To determine the effect of IRX-2 on the expression of immune genes, a paired transcriptome analysis (NanoString) was conducted on pre- and post-treatment tumor samples from 71 patients (45 in Regimen 1 and 26 in Regimen 2). In addition, DNA methylation EPIC BeadChip was conducted on a subset of patients (14 in Regimen 1 and 10 in Regimen 2) to explore the DNA methylation profiles of patients in both arms before versus after treatment.

3.2 Methods

3.2.1 Phase 2b clinical trial

From Jan 2016 to Jan 2018, a randomized Phase 2b clinical trial of neoadjuvant immunotherapy with the IRX-2 regimen was conducted on newly diagnosed Stage II, III or IVA oral cavity cancer patients across multiple centers in the USA and Brazil. A total of 96 patients were divided in two regimens, with 64 in the experimental regimen (Regimen 1) and 32 in the control regimen (Regimen 2). The treatment in Regimen 1 lasted 3 weeks, which included 10-day subcutaneous bilateral IRX-2 injections in the upper neck (2 injections of 1 mL each day), an

initial dose of cyclophosphamide, and a 3-week course of indomethacin (25mg; 3 times daily), zinc supplementation (65mg/day), and omeprazole (20mg/day) (179). The treatment in Regimen 2 was identical to that of Regimen 1 but excluded the IRX-2 (cytokine components) injections. A detailed study design can be found at http://www.brooklynitx.com/inspire-study/. All patients enrolled in this clinical trial underwent surgical resection 19 to 41 days post-randomization. This study was approved by the human experimentation Institutional Review Boards of the University of Michigan and other participating institutions. All patients have given written consent for the study, and the experimental methods comply with the Helsinki Declaration.

3.2.2 Tumor sample preparation

Formalin-fixed paraffin embedded blocks were requested from all enrolled subjects at each participating institution. 78 biopsy (pre-intervention) blocks and 84 surgical resection (post-intervention) blocks were requested. From these, blocks from 24 subjects with sufficient tissue in both biopsy and resection were selected for DNA methylation.

3.2.3 RNA extraction and Nanostring protocol

Whole tumor transcriptome analysis was commercially performed by the NanoString Technologies (Seatle, WA) central laboratory. Briefly, after being reviewed by a pathologist, formalin fixed, paraffin embedded (FFPE) slides were microdissected for RNA isolation from the tumor area of 5-µm slices using the High Pure microRNA FFPE Isolation Kit (Roche, Basel, Switzerland) or High Pure FFPE RNA Micro Kit (Roche) per manufacturer protocols. After being de-paraffinized with xylene and washed with ethanol, the tissues were lysed and treated with proteinase K for 3 hours at 55°C. After a washing step, lysates were applied onto spin columns and RNA was eluted once in 50 µL and twice in 40 µL elution buffer respectively. Next, RNA was purified and concentrated using the RNA Clean & Concentrator-5 Kit (Zymo

Research, Irvine, CA, USA). The yield of RNA was measured using NanoDrop 2000 (Implen GmbH, Munich, Germany) or Qubit RNA BR Assay Kit (Thermo Fisher Scientific) on the Qubit 3.0 Fluorometer (Thermo Fisher Scientific), and RNA quality was determined on a Lab-on-a-Chip 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). No sample was excluded solely based on RNA Integrity Number (RIN), since RNA from FFPE tissues generally have lower quality (RIN values <2).

A minimum of approximately 50 ng of total RNA was used to generate whole tumor transcriptome profiles using the nCounter platform (NanoString Technologies) and the PanCancer Immune Profiling Panel from matched biopsy and resection samples. The NanoString captures the expression of 770 genes in total, including 730 immune response genes and 40 pan-cancer housekeeping genes. Out of 730 immune-related genes, 109 of them define 24 different immune cell types and populations, 30 encode known cancer/testis antigens, and more than 500 encode critical proteins in immune response pathways.

3.2.4 NanoString data analysis

A total of 71 patients out of 96 (45 in Regimen 1 and 26 in Regimen 2) had usable Nanostring data available both before and after treatment. A minimum of approximately 50 ng of total RNA was used to measure the expression of 730 immune-related genes and 40 housekeeping genes using the nCounter platform (NanoString Technologies) and the PanCancer Immune Profiling Panel. The expression data were code-count normalized, sample content normalized, and background corrected using the *NanoStringNorm* R Package(185). Paired tests with edgeR-QLF were used to conduct two types of differential analysis: before vs after treatment for each regimen separately, and a contrast between the change in Regimen 1 vs the

change in Regimen 2. Significantly differentially expressed genes had false discovery rate (FDR) < 0.05 and fold change (FC) > 1.5.

3.2.5 DNA extraction and EPIC BeadChip protocol

An H&E slide from each block was reviewed by an expert pathologist (Jon B. McHugh, MD) for areas of >70% tumor cellularity. The corresponding areas on unstained slides were microdissected for DNA isolation using the QIAamp DNA FFPE Tissue Kit (Qiagen, Hilden, Germany). All DNA samples were treated with RNAse using the QIAquick Nucleotide Removal Kit (Qiagen, Hilden, Germany). DNA quantity was assessed using the Qubit Broad Range Double Stranded DNA Assay (Thermo Fisher Scientific, Waltham, MA). Quality of these samples was evaluated using the Nanodrop Spectrophotometer (Thermo Fisher Scientific, Waltham, MA) and the Illumina FFPE QC Kit (Illumina, San Diego, CA). This kit evaluates quality of DNA extracted from FFPE tissues measuring through real-time PCR using the Bio-Rad CFX96 (Bio-Rad Laboratories, Hercules, CA). Controls (positive and no template) were included on each plate. Average Cq (quantification cycle) values for each sample were evaluated for deviance from the average Cq of the positive control (Delta Cq). As recommended by Illumina, samples selected for methylation measurement have Delta Cq values below 5 and have replicate Cq values that do not diverge by more than half a unit.

In total, 24-paired samples were selected based on quality and quantity for downstream processing. Concentrations ranged from 8.2 - 146.7 ng/µL, 260/280 ratios ranged from 1.85-2.13, and Delta Cq values ranged from -1.3 - 4.2 for final samples.

Bisulfite conversion was performed using 350 ng input FFPE DNA with the Zymo EZ DNA Methylation Kit following manufacturer recommendations. To restore bisulfite converted FFPE DNAs to an amplifiable form for the Infinium MethylationEPIC assay (Illumina), the

University of Michigan Advanced Genomics Core (UMAGC) performed the Infinium HD FFPE Restore protocol (15014614 RevC) according to manufacturer's instructions. The UMAGC generated Methylation Profiles of the restored FFPE samples using the Infinium MethylationEPIC Kit (Illumina) and the manufacturer's Infinium HD FFPE Methylation Assay Protocol (15027310 Rev.A). BeadChips were scanned and image data recorded by an iScan Reader with Methylation NXT settings. Methylation Profiles were generated using the GenomeStudio (v2011.1) Analysis Software with the Methylation Module and MethylationEPIC_v-1-0_B4 Manifest (Illumina).

3.2.6 EPIC BeadChip data analysis

MethylationEPIC BeadChip assay data (Illumina, San Diego, CA) were analyzed for 24 patients (14 in Regimen 1; 10 in Regimen 2) with both biopsy and resection samples. Raw data were pre-processed with functional normalization *funnorm*(186). A total of 17611 probes were removed due to failure of detection in at least 5% of samples. After dropping cross-reactive probes detected in 450K BeadChip and probes that map to X and Y chromosomes, a total of 777406 probes were used for the remainder of analysis. Linear regression and *eBayes* function in the *limma* R package were used to call DMPs (differential methylated probes, with methylation difference > 10% and FDR < 0.05) before versus after treatment in a paired analysis for each regimen, however no probes were significantly different. The R package *DMRcate*(187) was used to call potential DMRs (differentially methylated regions, with FDR < 0.05 and methylation change > 10%) from DMPs.

3.2.7 DNA methylation-based cell type deconvolution

Cell type deconvolution analysis has been shown to perform better using DNA methylation data compared to gene expression. We used MethylCIBERSORT(184) with inbuilt

head and neck cancer signatures on the beta values of the 24 patients with EPIC BeadChip data to generate a methylation matrix, which was then used in CIBERSORTx(188) for cell type deconvolution calculation. The cell types included were cancer, CD14 (monocytes / macrophages / myeloid-derived suppressor cells), CD19 (B cells), CD4_Eff (CD4+ effector T cells), CD56 (NK cells), CD8 (cytotoxic T cells), endothelial, eosinophils, fibroblast, neutrophils and regulatory T cells.

3.2.8 Determination of response to treatment

Two types of response to treatment were calculated for each patient having the relevant data: clinical tumor response and immune response (based on the change in T cell counts in the tumor microenvironment (**Table 3.1**). Clinical response categorized patients into partial responders (PR), stable disease (SD), or progressive disease (PD) based on RECIST (Response Evaluation Criteria in Solid Tumors) criteria, where the effect of the therapy was evaluated from unbiased centralized interpretation of CT and/or MRI scans. Immune responders (IR) were defined as having an increase in CD8+ TIL (tumor infiltrating lymphocytes) infiltrate score of at least 10 cells/mm^2, as previously published(189). The individual percent change in tumor size from clinician's tumor measurements was also available for all 96 patients, however this was not used as a measure of response due to potential subjectivity and variability in clinician estimates in this international, multi-institutional trial.

		Regimen 1	Regimen 2
Immune response (n=33)	Responder	15	4
	Non-responder	5	9
Clinical response (n=77)	PR	6	2
	SD	33	17
	PD	10	9

Table 3.1 The number of patients with immune response and clinical response status available in Regimen 1 and Regimen 2 respectively.

3.3 Results

3.3.1 Both regimens resulted in a differential pattern of immune gene expression before versus after treatment

From January 2016 to January 2018, we collected clinical data and tumor specimens from a total of 96 patients with newly diagnosed oral cavity squamous cell carcinoma who participated the neoadjuvant IRX-2 Phase 2b clinical trial with IRX-2 regimen. In order to determine the effect of IRX-2 on immune response, we have collected the expression level of 730 immune response genes with NanoString on 71 patients for paired pretreatment biopsy (3 weeks prior to surgery) and surgical resection specimens, out of which 45 were from patients randomized to Regimen 1 and 26 form Regimen 2. First, we sought to understand the degree of heterogeneity based on signatures in our cohort using principal component analysis (PCA). Although the samples did not separate by treatment regimen, a major source of heterogeneity was the distinction between biopsy and resection, indicating treatment induced immune response change in both regimens. By using the log2 fold change (FC) of resection versus biopsy samples, we did not observe any separation based on regimen, immune response, or clinical response.

By comparing the baseline timepoint (biopsy) to after 21 days of immunotherapy (resection), we elucidated changes due to the IRX-2 treatment with and without the cytokines in key immune response genes. A total of 51 and 79 immune response genes were found upregulated and downregulated respectively in the resection samples of Regimen 1, while 51 and 56 were found upregulated and downregulated respectively in the resection samples of Regimen 2. Among these differentially expressed genes, 69 overlapped between the two regimens, with 33 upregulated and 36 downregulated in resection (**Figure 3.1A**). Some of these overlapped upregulated genes have been previously linked with mutations in cancer, including CD63, BCL2,

MME1, DMBT1, and CXCL12, while the overlapped downregulated cancer genes include S100A7, HRAS, MST1R, CDH1, IL1B. Noticeably, CD8A, a gene marker for CD8 T cells, was only upregulated in Regimen 1.

To visualize the expression patterns of upregulated and downregulated genes in both regimens, we performed hierarchical clustering on the union of differentially expression genes in Regimen 1 and Regimen 2 (**Figure 3.1B**). The resulting heatmap separated upregulated genes from the downregulated genes, except for a few patients whose gene expression consistently decreased after treatment for nearly all genes. Interestingly, while the upregulated genes demonstrated a consistent up-regulation pattern across patients in both regimens, the downregulated genes were due to large decreases in expression in a subset of patients irrespective of regimen. No clear distinctions were observed between the two regimens or between the two immune response groups.

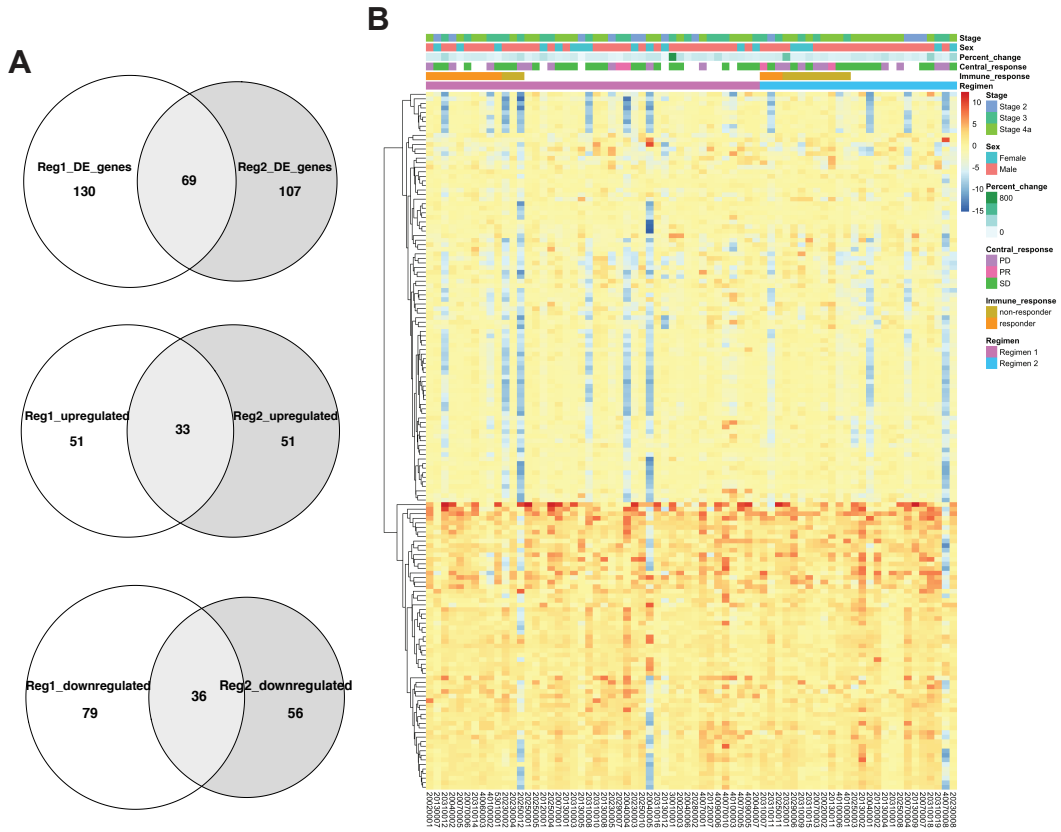


Figure 3.1 Immune response genes changed similarly in both Regimen 1 and Regimen 2.
(A) Venn diagrams showing the total number of differentially expressed genes, the number of upregulated genes, and the number of downregulated genes after treatment that are overlapped between Regimen 1 and Regimen 2, respectively. (B) Heatmap showing the $\log_2$ fold change of gene expression before versus after treatment (log2FC, resection/biopsy) of all 153 differentially expressed immune genes in either regimen. Samples are ordered in columns by regimen and immune responder status, and hierarchical clustering was performed on genes only, which are displayed as rows.

3.3.2 DMBT1 expression level showed stronger upregulation after treatment in Regimen 1 than Regimen 2

Comparing the changes observed after treatment between the regimens, we identified 9 genes which displayed significantly different changes between the two regimens (**Figure 3.2A**). Of these 9 genes, 6 of them (DMBT1, SEMG1, CD160, CCR9, CCL16, IL2) were associated

with a larger change between biopsy and resection samples in Regimen 1, while the changes in Regimen 2 were larger for KIR3DL2, IFNA7 and CTAG1B, and only DMBT1 was upregulated in both regimens. The only gene that showed previous mutations in HNSCC was DMBT1, a potential tumor suppressor gene functioning in calcium-dependent protein binding and signaling pattern recognition receptor activity(190). DMBT1 was found to be upregulated in resection samples in both regimens, but the extent of increases was more drastic in Regimen 1 (mean of 59.3 fold for Regimen 1 versus 19.43 fold for Regimen 2) (**Figure 3.2B**). Visualizing individuals' DMBT1 expression across response groups, we did not observe a strong association between change in DMBT1 expression level with either clinical tumor response or immune response (**Figure 3.2B**) in either regimen. However, this conclusion was limited because few patients with NanoString data were characterized as clinical responders after this brief immunotherapy regimen (n=4), and only a limited number of patients (n=26) had immune response information available, with only three immune responders in Regimen 2.

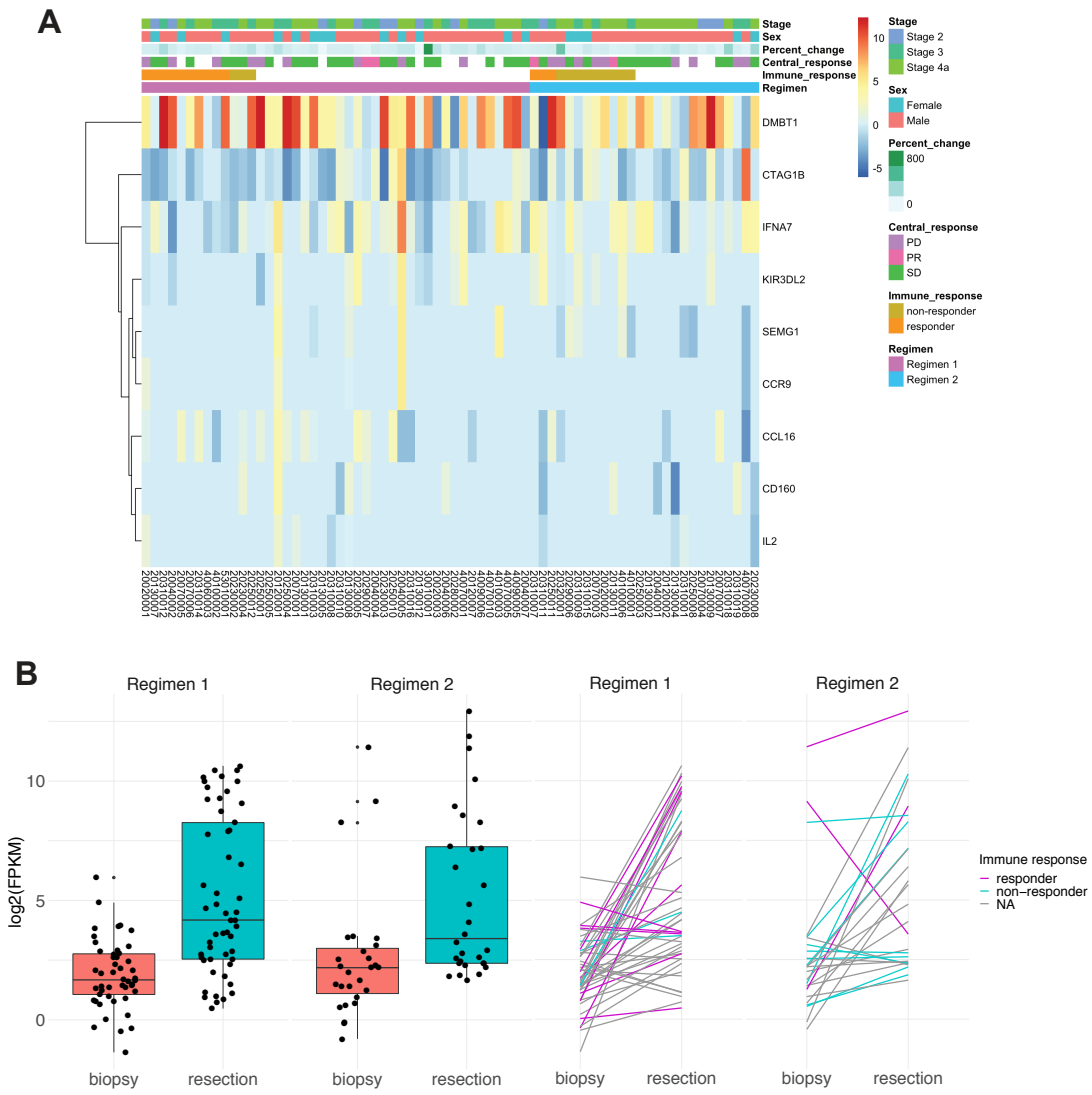


Figure 3.2 A total of 9 genes showed significantly different changes after treatment between the two regimens.
(A) Heatmap showing the $\log_2$ fold change (resection/biopsy) of the 9 genes whose difference before vs after treatment is significantly different between Regimen 1 and Regimen 2. (B) Box plot of the DMBT1 gene expression at biopsy and resection in the two regimens separately. Spaghetti plot showed the change of DMBT1 expression in patients of different immune response groups.

3.3.3 Change of different immune signatures after treatment in both regimens

To study the effect of IRX-2 treatment on key immune signatures, we first assessed the change in expression of PD1 and PDL1 in both regimens. Expression levels of both PDL1

(*CD274*) and PD1 (*PDCD1*) revealed large individual variability (**Figure 3.3**). The differential expression analysis between biopsy and resection described above revealed downregulation of PDL1 after treatment in regimen 2, and that this drop in expression was mainly concentrated in non-immune responders. Patients who were characterized as immune responders showed little to no change in either gene in either regimen.

Beyond PD1 and PDL1, we examined per patient changes in interferon, cytokine, antigen and inflammation gene groups, which were defined in Li et al.(191) and identified as being important immune responders in epithelial cancers. Overall, we observed a high level of variability among patients, yet 25% of the genes were differential expressed between biopsy and resection in both Regimen 1 and Regimen 2. (**Table 3.2**). After separating patients by their respective regimen and immune responder status, we discovered that the changes in interferon and antigen genes were more subtle in the immune responders of Regimen 1 than Regimen 2 and non-responders, but this trend was not as prominent in cytokine or inflammation genes. Although a definitive conclusion is difficult due to the existence of a few outlier patients with extremely large downregulation, we found that 62% and 75% of Regimen 1 immune responders had increased cytokine and inflammation gene expression; only 1 of 3 Regimen 2 immune responders had a similar increase.

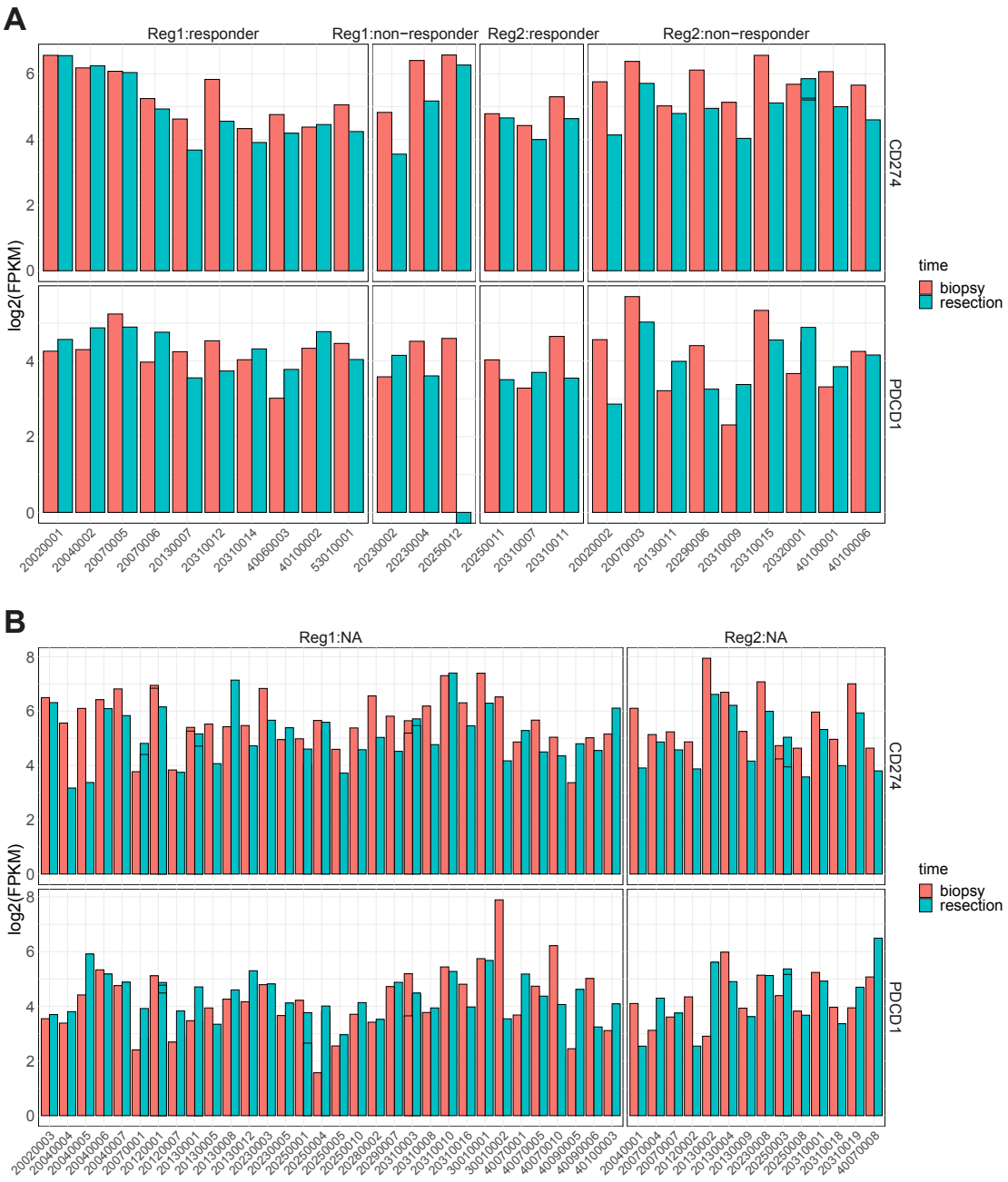


Figure 3.3 The expression levels of CD274 (PDL1) and PDCD1 (PD1) in both regimens.

The expression levels of CD274 (PDL1) and PDCD1 (PD1) in patients (A) with and (B) without immune response groups of the two regimens. The immune responder groups showed less change in both PDL1 and PD1 expression compared non-responders and patients without immune response information.

Immune gene types		interferon	cytokine	antigen	inflammation
Total # of genes with Nanostring data		17	30	5	18
Regimen 1	# upregulated in resection	2	1	0	1
	# downregulated in resection	5	6	0	2
Regimen 2	# upregulated in resection	2	1	0	1
	# downregulated in resection	6	7	0	2
Total # of differentially expressed genes		8	9	0	4

Table 3.2 The total number of interferon, cytokine, antigen and inflammation genes that are found in NanoString gene list and the number of genes in each category that are differentially expressed between biopsy and resection samples.

3.3.4 Higher methylation level in resection samples than baseline in both regimens

DNA methylation was assessed using the MethylationEPIC BeadChip Assay on 24 patients (14 in Regimen 1 and 10 in Regimen 2), to explore the changes in DNA methylation caused by IRX-2. The overall global methylation distribution results showed a slight overall increase in methylation after treatment in both regimens, and this increase was consistent across CpG islands, shores and shelves. The same trend was especially prominent on the LINE-1 and LINE-2 elements of Regimen 1 compared to Regimen 2, and it was mainly driven by the immune responders in both regimens (**Figure 3.4A**). Since the focus of this study was on the effect of IRX-2 on immune response, we also looked at the methylation level of the CpGs that fall on the promoter regions of immune response genes. In Regimen 1, the methylation level increase aftertreatment was more prominent in immune responders, which was not observed for immune responders in Regimen 2 (**Figure 3.4B**).

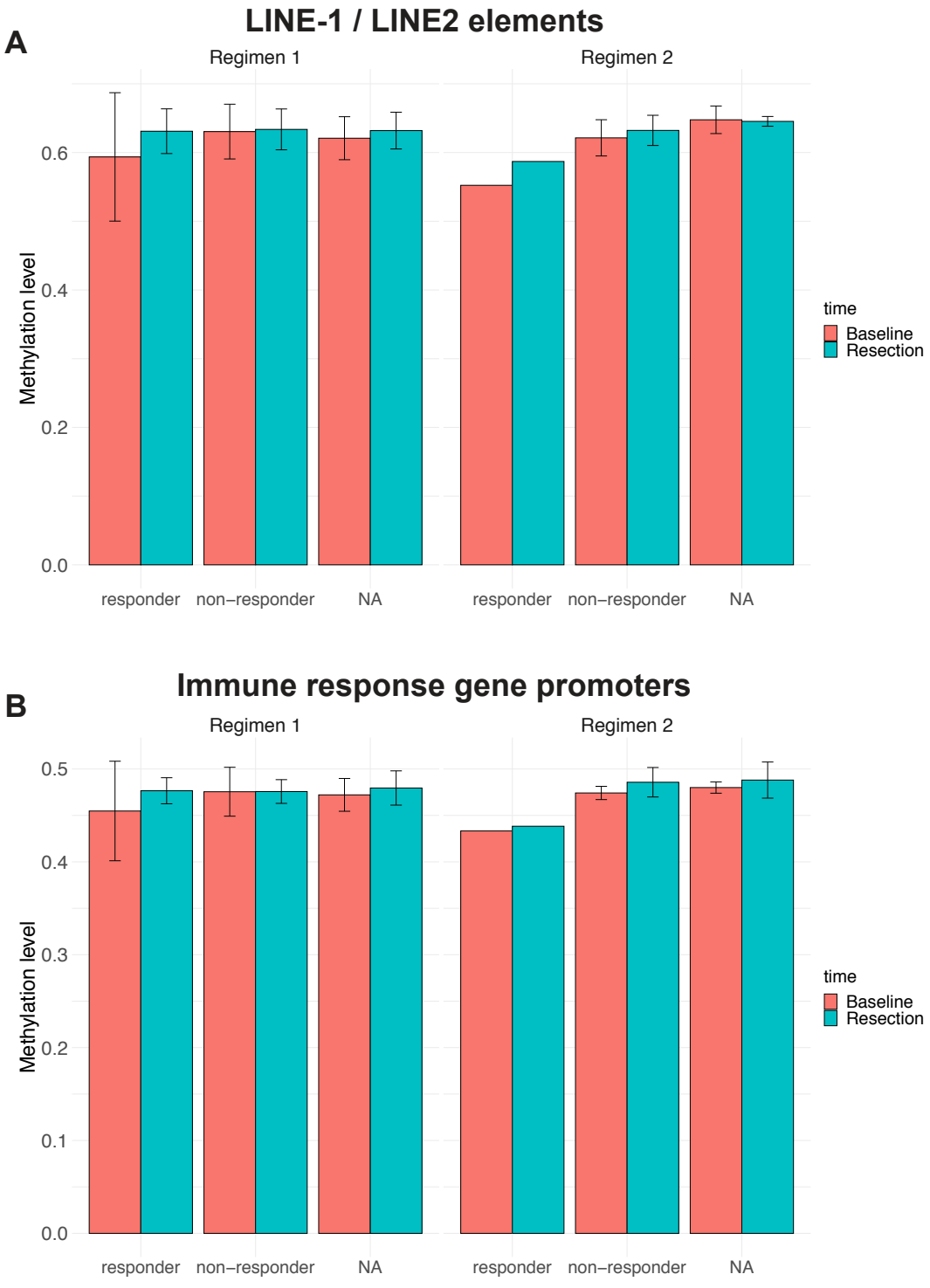


Figure 3.4 The average DNA methylation levels in both regimens.
The average DNA methylation levels in both regimens at (A) LINE-1 and LINE-2 elements and (B) immune response gene promoters. The resection samples showed higher DNA methylation levels than biopsy samples for almost all the comparisons, but the difference is more prominent in the immune responder group of Regimen 1.

3.3.5 Altered methylation of immune response and keratinization pathways after IRX-2 treatment

By comparing the methylation profile from EPIC BeadChip data before and after the treatment, we did not identify any differentially methylated probes (DMPs) in either regimen. Testing differentially methylated regions (DMRs) with FDR<5%, we identified 523 and 1 potential hyper- and hypo-methylated regions in the resection samples of Regimen 1 resepctively, including those mapped to immune response genes IFI27 and SPINK5 which were downregulated after treatment based on NanoString data (**Table 3.3**). Conversely, a total of 721 and 112 potential regions were found to be hyper- and hypo-methylated respectively in the resection samples of Regimen 2; of genes with potential hypermethylation, *CXCL14*, *CLEC5A*, *LAMP3*, *TOLLIP* and *IFITM1* displayed downregulation. A total of 21 potential DMRs were hyper-methylated in both regimens, including keratinocyte differentiation gene RIPK4, but none of the overlapped genes were among the immune response genes selected for NanoString. Pathway enrichment of the DMRs suggested cornification, keratinization and keratinocyte differentiation, and epidermic development and differentiation were enriched with hypermethylation in Regimen 1. Pathways including organ morphogenesis, epithelium development and embryonic development were most significant for Regimen 2 hypermethylation, but not keratinization or cornification.

Direction in resection	Regimen 1	Regimen 2
Hypermethylated	523	721
Hypomethylated	1	112
Total	524	833

Table 3.3 The number of potential differentially methylated regions (DMRs) derived from EPIC BeadChip data that are hyper- or hypo-methylated in Regimen 1 and Regimen 2 separately.

3.3.6 Cell deconvolution revealed lower percentage of cancer cells and higher percentage of T cells after treatment for immune responders

Cell type deconvolution analysis using bulk gene expression or DNA methylation data has become an important tool for interpreting changes in light of potential shifts in cell type proportions. Recently, it was shown that cell type deconvolution is more accurate using DNA methylation data than gene expression data (184). Due to the uncertainty in defining immune responders based on any one criterion, the cell type deconvolution results provide a complementary perspective. To determine the effect of IRX-2 treatment on cell type proportions in each regimen, we performed deconvolution with MethylCIBERSORT on the 48 samples (28 in Regimen 1 and 10 in Regimen 2) with DNA methylation data, estimating the proportions of 11 cell types. None of the cell types were significantly different either between the two regimens or between biopsy and resection samples. However, we observed decreased cancer cell proportions as well as increased CD8+ T cell proportions in immune responders of Regimen 1, which was not replicated in non-immune responders or Regimen 2 patients (**Figure 3.5A**). The cell type deconvolution results for CD8+ T cells were confirmed by the high correlation between the estimated CD8+ T cell proportion from MethylCIBERSORT and the CD8+ T cell density counts performed previously on tissue microarrays (189) (**Figure 3.5B**).

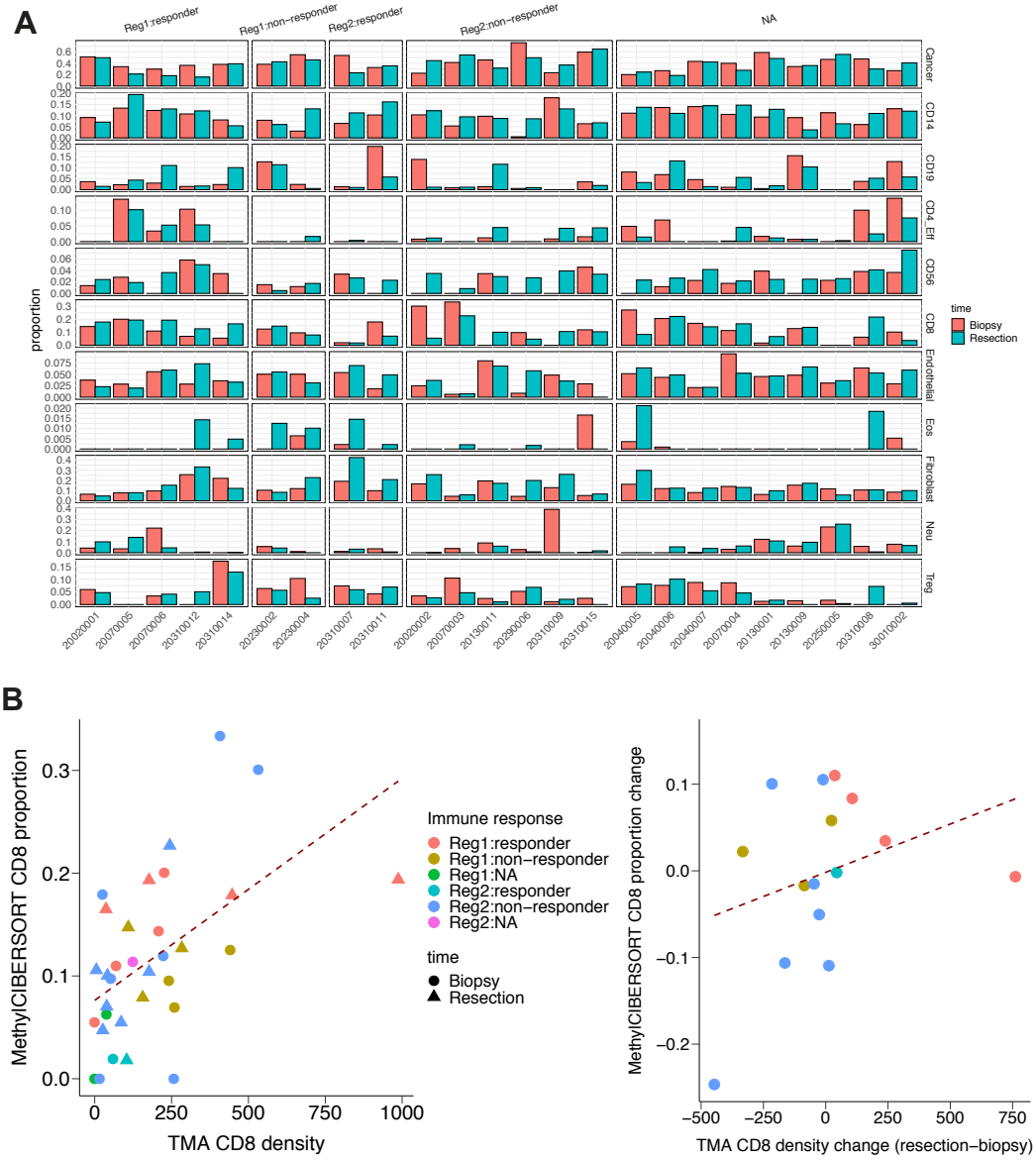


Figure 3.5 Cell deconvolution using DNA methylation data revealed higher CD8 in immune responders of Regimen 1.
(A) Bar plot showing the proportion of different cell types from cell deconvolution of 48 patients in different immune response groups. (B) Dot plot indicating the high correlation between MethylCIBERSORT CD8 proportion and TMA density (r = 0.51). The MethylCIBERSORT CD8 proportion change (resection - biopsy) was also correlated with TMA CD8 density change (r = 0.31) but not as strongly.

3.5 Discussion

Our study is the first to characterize immune-related gene expression and DNA methylation profiles after neoadjuvant immunotherapy with the IRX-2 regimen in a randomized clinical trial. We identified expression changes in immune response genes during the treatment with both IRX-2 and control regimens. Surprisingly, the differential expression pattern was similar in Regimen 1 and Regimen 2, with more than half of the differentially expressed genes in common between the two regimens, suggesting much of the action of the complete IRX-2 treatment protocol is due to elements other than the cocktail of cytokines. For example, one cytokine component in IRX-2, IL1B, was downregulated after treatment in both regimens, which indicated the cytokine injection did not significantly alter transcription of IL1B. However, the IL2 cytokine component, showed more significant change between biopsy and resection samples within Regimen 1 patients (**Figure 3.2A**), indicating a potential linkage between 21-day treatment of IRX-2 and higher activated CD4+ T cells and activated CD8+ T cells, which are major sources of IL2(192). DMBT1 (deleted in malignant brain tumor 1) is a potential tumor suppressor gene encoding a pattern recognition molecule that plays a key role within the innate immune system(193,194). Uniquely, DMBT1 was found to have more extensive upregulation in Regimen 1 than in Regimen 2. Previous studies confirmed the critical role DMBT1 plays in the interaction of tumor cells and the immune system, and downregulation of DMBT1 is thought to promote invasion in head and neck cancer(195,196), which could suggest modulation of tumor invasiveness after IRX-2 treatment in Regimen 1 in our study. However, follow-up studies are needed to further understand the downstream effects of DMBT1 upregulation in patients with oral squamous cell carcinoma.

Cyclophosphamide, a cytotoxic and immunoregulatory element of both regimens, slows cancer growth by inhibiting protein synthesis mainly by cross-linking strands of DNA and RNA,

and has also been shown to suppress regulatory T cells (Tregs) specifically with a single low dose(197). Given the complexity of the complete IRX-2 regimen and unknown interactions among its components, it is difficult to predict *a priori* the exact effect it could have on PD1 and PDL1 expression. On one hand, if active cytokine components in IRX-2 can prime patients' immune response, the expression of PD1 and PDL1 may remain stable or increase after treatment. Alternatively, if the IRX-2 regimen functions in a similar manner as a type of checkpoint inhibitor, we might expect more tumor cells to be targeted by infiltrating CD8+ T cells, accompanied by decreased PD1 and PDL1 expression. In our case, although we did not observe a clear distinction based on PD1 and PDL1 changes between the two regimens, the fact that PD1 and PDL1 level stayed relatively stable for immune responders in Regimen 1 showed that the effect of IRX-2 cytokines likely differs from checkpoint inhibitors and might be somewhere between restoring the immune response and stimulating T cell infiltration. Further, our measures did not specifically differentiate between tumor and immune cell expression of these biomarkers. However, both PD1 and PDL1 expression dropped in Regimen 2, particularly for non-responders and for non-responders in Regimen 1, which is consistent with an effect on PD-1 blockade(198) suggesting that such effects might not be clinically beneficial. In prior work we noted increases in CD68 tumor associated macrophages associated with IRX regimen 2 which would be consistent with such a negative effect(189). The multiplicity of interacting factors makes clear interpretations difficult, however the complexity of the physiologic concentration of cytokines in IRX-2 also make it an attractive approach to address the multiple immune deficits identified in HNSCC.

Utilizing the EPIC BeadChip DNA methylation data, we were able to identify an increase in global methylation level after treatment in both regimens, with a more prominent trend in

responders of Regimen 1 and in LINE1 & LINE2 repetitive elements. Since global and repetitive element hypomethylation is known to be common across cancer types, this increase is shifting the tumor's DNA methylation status closer to "normal". Other publications also report the potential importance of PD1 / PDL methylation in the prognosis of HNSCC patients(78,199), however we did not observe any significant change in methylation level in or near PD1 or PDL1. Pathway enrichment analysis with the DNA methylation data revealed pathways linked with hyper-methylation after treatment in Regimen 1, including cornification, keratinization and epidermic development, which could point to a change in the differentiation status of oral cavity carcinoma(200,201). Although our immune-focused NanoString data did not allow us to assess keratinization gene changes, this could be a counterpoint to the benefit of IRX-2 worth studying, since low keratinization has been associated with higher recurrence rates and a lower 5-year survival rates in oral cancer patients(202).

Previously, our group published the initial findings of this phase2b clinical trial that Regimen 1 resulted in higher levels of CD8+ T cell density post-treatment than did Regimen 2(189). Our study reconfirms this finding based on more significant CD8 markers expression (CD8A and IL2) in Regimen 1 only and on methylation-based cell type deconvolution. However, our analysis also suggested that most of the observed immune pattern change beyond this and the PD1/PDL1 findings, did not significantly differ by regimen. The main potential benefit of IRX-2 has been thought to be increased activation and migration of T cells from the regional lymph nodes into the tumor microenvironment. Direct effects of the cytokine preparation on the primary tumor would not necessarily be expected unless they are mediated by changes in TILs. Our current study focused on changes in the tumor microenvironment related to immune gene expression and epigenetic methylation rather than T cells. The lack of finding major changes in

gene expression or methylation is therefore not surprising. We found general changes associated with both immune modulating treatment regimens that probably reflect the non-cytokine components of the regimens but also encouraging findings that confirmed the T cell density counts previously reported with Regimen 1.

A main limitation of our study is the inconsistent availability of different types of datasets. A total of 96 patients were enrolled in the Phase 2b clinical trial, whose clinical characteristics and follow-up survival information were collected in a timely manner. However, only 71 patients had NanoString data sufficient for gene expression analysis, and only a subset of those (24 patients) were selected for DNA methylation analysis. A total of 26 patients had CD8+ T cell density measured for determining immune responder status, and clinical tumor response was only available in a subset of 77 patients, resulting in missing response data in a large percent of patients with gene expression or DNA methylation data. With gene expression analysis limited to NanoString data instead of full RNA-seq, our study also lacked the ability to study expression of other important pathways such as keratinization and differentiation status in general.

Chapter 4

Perinatal DEHP Exposure Induces Sex- and Tissue- Specific DNA Methylation Changes in both Juvenile and Adult Mice

This work has been published as: **Liu, S.**, Wang, K., Svoboda L. K., …& Sartor, M. A. (2021). Perinatal DEHP exposure induces sex- and tissue- specific DNA methylation changes in both juvenile and adult mice. *Environmental Epigenetics*, Volume 7, Issue 1, dvab004

4.1 Introduction

Di(2-ethylhexyl) phthalate (DEHP) is a plasticizer (86) often found in polyvinyl chloride products, including medical equipment, car upholstery, food and beverage containers, and building materials (87). DEHP does not covalently bond to polymer chains in these products and is therefore likely to spread into the environment after repeated usage (88). Due to the large quantity and wide variety of products containing DEHP that people interact with frequently, the estimated range of human exposure to DEHP is 3-30 μg/kg/day (91).

DEHP has been detected in tissues such as blood, amniotic fluid, umbilical cord blood, and breast milk in humans (92–95), indicating that exposure starts as early as the fetal developmental stage(96). As a lipophilic compound, DEHP can be absorbed through dermal exposure, inhalation, and oral ingestion (93,94,203). The estimated human oral absorption varies from 25% to 50% depending on the dosage (98,99), while absorption in rodents can be as high as

58% of the oral dose, but 50% is assumed on average (100). The biological action of DEHP is very similar to a group of chemicals called peroxisome proliferators (PPs), and liver is a primary target organ for the effects of DEHP and other PPs (101). Prenatal DEHP exposure in rodents results in elevated fatty acid metabolism, as well as peroxisome proliferation and the accumulation of lipofuscin granules which are implicated in hepatocarcinogenesis (102,103). There are multiple modes of action of DEHP in hepatocytes, including activation of peroxisome proliferator-activated receptors (PPARs), induction of cell proliferation, suppression of apoptosis, oxidative DNA damage, and inhibition of gap junctional intercellular communication (101). Once absorbed, DEHP and its metabolites are distributed throughout the body in the blood promoting its endocrine disrupting properties.

Epigenetic modifications, such as DNA methylation, define and control cell and tissue development by regulating gene expression (107). For example, DNA methylation patterning of imprinted genes is crucial for embryonic development (109). Previous animal and human studies have reported that perinatal exposure to DEHP is associated with altered DNA methylation (110–112), with multiple studies identifying DEHP-induced DNA hyper-methylation (113,114). A recent study in human cord blood showed significant DNA methylation changes in genes involved in androgen and estrogen responses, and spermatogenesis following prenatal DEHP exposure (115). Other studies of ovaries and oocytes showed prenatal DEHP exposure disrupted the expression of cell cycle regulators and changed the DNA methylation pattern of imprinted genes in germ cells (109,116). Although liver has been identified as a primary target organ of DEHP, the effect of prenatal DEHP exposure on genome-wide DNA methylation patterns in liver has not been assessed.

Contrary to *in vivo* environmental toxicology experiments for which target tissues are easily obtained, human environmental epigenetic epidemiology studies often make use of easily obtainable sources of DNA, such as blood and saliva, to serve as surrogate readouts of epigenetic alterations in target tissues, such as liver and brain (204). Therefore, it is essential to determine to what extent alterations measured in surrogate tissues are accurate reflections of the target tissues both epigenetically and transcriptionally, and whether exposure-associated alterations in genetic and epigenetic marks are consistent or different over the course of development. The Toxicant Exposures and Responses by Genomic and Epigenomic Regulators of Transcription (TaRGET II) Consortium was launched by the National Institute of Environmental Health Sciences (NIEHS) to partially address this issue of linking epigenetic alterations in surrogate and target tissues (204).

As part of the consortium, in this study we used an established mouse model of perinatal DEHP exposure to investigate the alterations in DNA methylation profiles in both liver (target tissue) and blood (surrogate tissue), at post-natal day (PND) 21 and 5 months of age in cohorts of male and female mice. We also studied the transcriptional effect of DEHP exposure by characterizing the gene expression profiles of the same cohort. Due to the rapid growth and physiologic changes mice go through from weanling (PND21) to adulthood (5-months), we hypothesized divergent DNA methylation patterning between the two time points, but to what extent the differential methylation marks overlap by sex and between liver and blood still remained unknown.

4.2 Methods

4.2.1 Animals and dosing paradigm

Mice utilized in this study were obtained from a colony maintained for over 230 generations with the *Avy* allele passed through the male line, resulting in forced heterozygosity on a genetically invariant background with 93% identity to C57BL/6J (205). Virgin a/a females (6–8 weeks old) were mated with virgin a/a males (7–9 weeks old), and randomly assigned to receive a control or DEHP-exposed diet through consumption of chow (206,207). DEHP was dissolved in corn oil and mixed with chow (AIN‐93G, TD.95092, Harlan Teklad) to achieve 7% corn oil and 25 mg DEHP per kg of chow. Control animals received the same chow with 7% corn oil. Dams began consumption of control or DEHP diets 2 weeks prior to mating, and exposure was continued during gestation and lactation.

DEHP was mixed into corn oil from Envigo to create a stock solution, and the stock solution was sent back to Envigo where it was mixed with the corn oil used to produce custom 7% corn oil chow in order to achieve uniform distribution of phthalates within the chow. The DEHP exposure level was selected based on a target maternal dose of 5mg/kg-day, assuming that pregnant and nursing female mice weigh approximately 25g and eat approximately 5g of chow per day. This target dose was selected based on literature demonstrating obesity-related phenotypes in offspring that were developmentally exposed to 5mg/kg-day of DEHP (208,209). The resulting exposure level is estimated to fall within the range of exposures experienced by humans (210). This is based on amniotic fluid levels of phthalates found in humans (ranging from <LOD to 100.6 ng/mL) and a study in rodents that orally ingested 11 mg/kg-day of phthalates resulting amniotic fluid levels of 68 ng/mL (211–218).

After weaning (PND21), pups were weighed, and all animals received DEHP-free chow for the remainder of the study (**Figure 4.1A**). Approximately 1–2 male and 1–2 female offspring per litter were followed until 5 months of age (the total number of animals per treatment group is

between 5 and 7, and the details can be found in **Figure 4.1B**). All animals had access to food and drinking water *ad libitum* throughout the experiment and were housed in polycarbonate-free cages. This study protocol was approved by the University of Michigan Institutional Animal Care and Use Committee (IACUC). All the experiments in this study were conducted according to procedures established by the NIEHS TaRGET II Consortium.

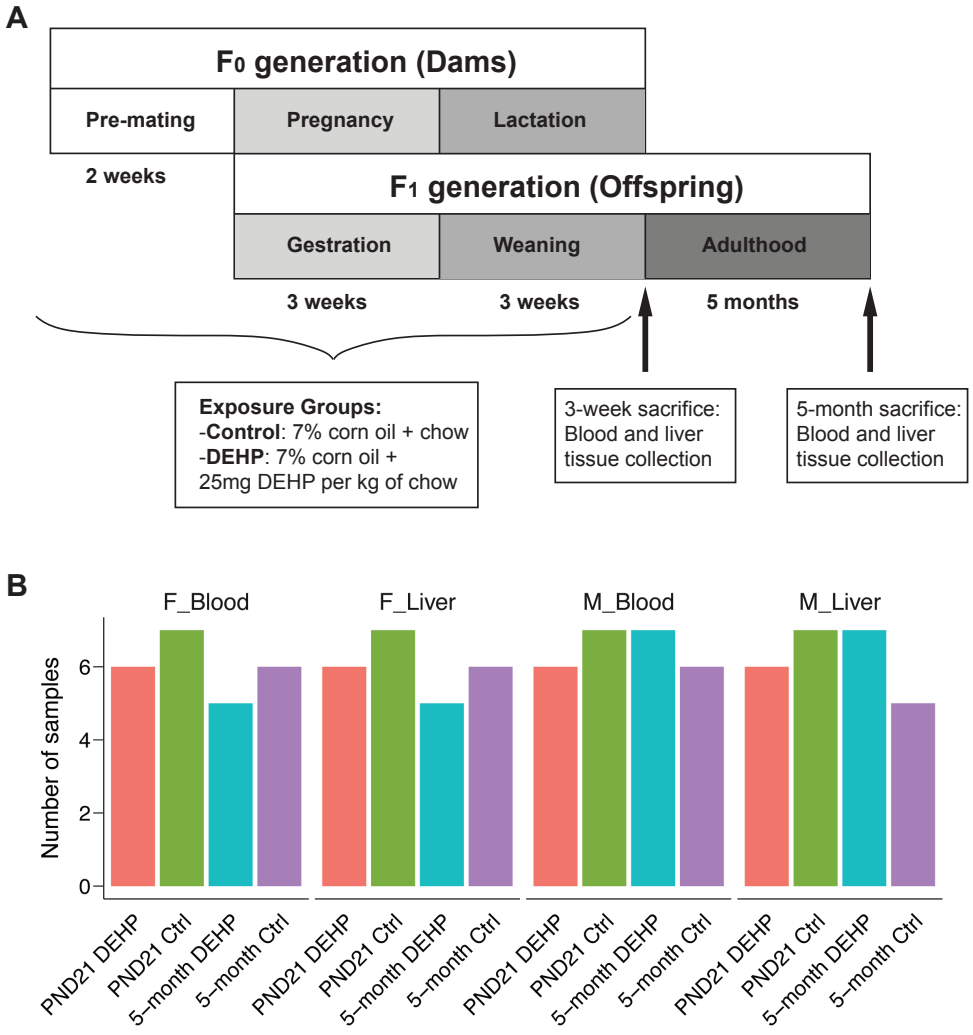


Figure 4.1 Overview of the perinatal DEHP exposure study design.
(A) DEHP exposure paradigm showing the duration of the exposure, details of the exposure, and times of tissue collection. (B) Bar plot showing the number of animals in each exposure group per tissue, age, and sex.

4.2.2 Tissue collection

Upon euthanasia at both PND21 and 5 months of age, blood and liver samples were collected according to protocols established by the TaRGET II Consortium (204). Briefly, blood was collected by cardiac puncture into tubes with EDTA anticoagulant, centrifuged, and plasma was removed. Liver tissue was weighed and then separated into individual lobes, and the left lobe of the liver was cryo-pulverized and suspended in Buffer RLT for nucleic acid extraction (Qiagen, cat # 1053393). Blood and liver DNA/RNA extraction was performed using the AllPrep DNA/RNA/miRNA Universal Kit (Qiagen #80224).

4.2.3 Enhanced reduced representation bisulfite sequencing (ERRBS)

A total of 50 ng of genomic DNA was digested using restriction enzyme MspI, and the digested DNA was then purified using phenol:chloroform extraction and ethanol precipitation in the presence of glycogen, before blunt-ending and phosphorylation. A single adenine nucleotide was next added to the 3′ end of the fragments in preparation for the ligation of the adapter duplex with a thymine overhang. The ligated fragments were cleaned, then processed for size selection on agarose gel. Selected fragments were treated with sodium bisulfite to convert unmethylated cytosines to uracils, which are then replaced with thymines during PCR amplification. These libraries were next cleaned up with AMPure XP beads (Product #A63880; Beckman Coulter), quantified using the Agilent TapeStation genomic DNA kit (Catalog #G2991AA; Agilent) and Qubit High Sensitivity dsDNA (Catalog #Q32850; Invitrogen). ERRBS samples were multiplexed, and sequencing was performed with Illumina HiSeq 4000, generating single-end, 50 bp reads for each library.

4.2.4 RNA-sequencing

After RNA extraction, library preparation was conducted in two batches using KAPA mRNA Hyper Prep Kit with Dual Indexing Adapters (PND21 samples) and Illumina TruSeq

stranded mRNA Library Prep Kit (5-month samples). These libraries were cleaned up and quantified using the Agilent 2200 TapeStation. Sequencing on these samples was carried out on the NovaSeq S2 flow cell and the Illumina HiSeq 4000 respectively for PND21 and 5-month samples, generating 50-bp paired end reads for each library.

4.2.5 ERRBS data analysis

Sequence quality per sample was first assessed with FastQC (v0.11.3), then Trim Galore! (v0.4.5) was used for adapter and quality trimming. More specifically, low-quality bases (quality score lower than 20), adapter sequences (required overlap of 6bp), and end-repair bases from the 3' end of reads were trimmed, and all reads shorter than 20bp after trimming were discarded. Bismark (219) was used for alignment and methylation calling, where reads were aligned to mm10 genome using Bowtie2 and methylation calls were reported for all nucleotides with a read depth of at least 5.

Differentially methylated cytosines (DMCs) and differentially methylated regions (DMRs) of 50 bp windows were identified using the methylSigDSS function of the methylSig R package (220,221), with significance cutoff of FDR <0.05 and an absolute difference in methylation of > 10%. The number of DMCs and DMRs were similar across conditions, so we used DMCs for all of the following results but enrichment analysis. Sufficient sequencing coverage for a minimum of 4 samples from each treatment group was required for the test. Batch effect was taken into consideration by using run number as a covariate, and sex was used as a covariate for combined analysis of both females and males. Sex stratified tests were also performed for each tissue and time point. The annotation was performed with the annotate_regions function of the annotatr R Bioconductor package (138). Gene set enrichment testing on DMRs were performed using the nearest TSS locus definition in the *ChIP-Enrich* R

Bioconductor package to find enriched pathways, and pathways with an FDR < 0.05 were considered significant (151).

4.2.6 Annotation of DMCs to CTD and imprinted genes

Genes having previous evidence of correlating with DEHP exposure were downloaded from the Comparative Toxicogenomics Database (CTD) (222), resulting in a list of 1186 genes from 1903 literature entries. A list of 316 mouse imprinted genes was generated by retrieving imprinted gene sets from two publications (223,224), followed by removal of the redundant genes.

4.2.7 RNA-seq data analysis

Data quality checks and gene feature summarization were conducted using the TaRGET II RNA-seq pipeline set by the Consortium Data Coordinating Center (DCC). Genes with >=1 CPM in at least 10 samples across all the consortium samples were used for normalization and differential analysis, which was conducted by the DCC using the RUVr function in *RUVSeq* (225) (k=3) separately for each institution's data and DESeq2 (226), respectively. As determined to be optimal by the DCC, differentially expressed genes were identified with absolute fold change > 1.5 and adjusted p-value < 0.001. Cell type deconvolution was conducted on the RNA-seq data of all blood samples at weanling (PND21) and adult (5-month) using CIBERSORTx (227). The differences of cell portions by treatment were tested with t-test.

4.3 Results

4.3.1 Genome-wide differential methylation analysis for PND21 and 5-month mice

To identify epigenetic changes in liver and blood induced by perinatal DEHP exposure in our mouse model, we focused on the effect of DNA methylation by conducting Enhanced

Reduced Representation Bisulfite Sequencing (ERRBS). All of the 99 samples (see **Figure 4.1B** for individual group sizes) resulted in sufficient quality of data covering 5% of all CpG sites across the mouse genome. The total alignment percentage ranged from 58.5% to 72.4%, and the average bisulfite conversion rate was 99.9%.

When comparing differential methylation between the DEHP exposed group and the control group at each time point, including sex as a covariate, we found fewer changes at PND21 (an average of 39 DMCs) than at 5 months (an average of 453 DMCs) of age for both blood and liver. The number of DMCs detected at 5 months was approximately an order of magnitude higher than those at PND21, despite DEHP exposure stopping at PND21 for the F1 offspring.

We next tested for sex-specific methylation changes, since endocrine disruptors often have sex-specific effects (228). The majority of the covered CpG sites (82.75% for female and 81.04% for male) are overlapped among the four respective groups (PND21 blood, PND21 liver, 5-month blood, 5-month liver), indicating the CpG sites were covered comprehensively among all comparisons. Larger numbers of DMCs were consistently identified across time and tissues using sex-specific analyses, again with more significant DMCs at 5 months than PND21 (**Figure 4.2**). Although the majority of differential CpG sites displayed methylation differences of less than 20%, we did identify some sites with more than 40% change and p-values $< 10^{-10}$. Compared to the overall, sex-combined analyses, we observed higher numbers of both hypo- and hyper- methylated sites with DEHP exposure, with similar numbers of hypermethylated and hypomethylated CpGs (**Figure 4.2**).

Annotating these sex-specific DMCs to genes and CpG islands revealed that at PND21, more DMCs occurred in CpG islands and shores in female blood compared with the other groups, and slightly higher proportion of DMCs were annotated to introns and exons in female

samples than male samples (**Figure 4.3A,B**). At 5-months, a similar but more significant trend was observed for higher rates of CpG island and shore DMCs in female blood (z-test p-value < 10^{-4}), although there was a higher proportion of DMCs annotated to introns in liver than blood in female samples (**Figure 4.3C,D**).

To examine how much inter-individual variation can be explained by methylation, we calculated the methylation level of the sex-specific DMCs identified above for all samples. The majority of individual animals could be clearly distinguished as PND21 or 5-months of age on the basis of methylation percentage of DMCs in both blood and liver (**Figure 4.3E,F**). More specifically, within blood samples, PND21 and 5-month animals separate clearly in PC1, with significantly more heterogeneity among the 5-month animals (**Figure 4.3E**). This separation is not as obvious in liver samples in the first three principal components, due to the even higher variation among 5-month livers (**Figure 4.3F**). The PND21 subjects do not separate well based on treatment either in blood or liver, while there is trend of separation between DEHP and control groups in blood samples at 5 months.

We identified a total of 1215 DMCs that mapped to gene promoters at PND21 (219 DMCs mapped to 149 genes) and 5 months (996 DMCs mapped to 715 genes). We further prioritized these DMCs by overlapping the promoter annotated genes with genes shown to be correlated with DEHP exposure, as indicated in the Comparative Toxicogenomics Database (CTD). Out of 2024 PND21 and 10312 5-month sex-specific DMCs, 74 (3.66%) and 358 (3.37%) were mapped on CTD genes respectively. Interestingly, DEHP was found to affect the mRNA expression of the *Gnas* gene (229), an imprinted gene in the G protein-coupled receptors and Ras signaling pathway. We identified promoter hyper-methylation of *Gnas* in the DEHP-exposed group in both male liver and blood at PND21, and both female liver and blood at 5

months. The prioritized gene that had the highest methylation change in female blood at PND21 was *Ucp2*, which encodes a mitochondrial uncoupling protein, and was found to be both hypermethylated (53.89% methylation difference) and DEHP-relevant in CTD with opposite direction of gene expression. We also identified 3 hypomethylated sites in the DEHP group in the promoter region of the *Esr1* (estrogen receptor 1) gene in female liver of 5 months.

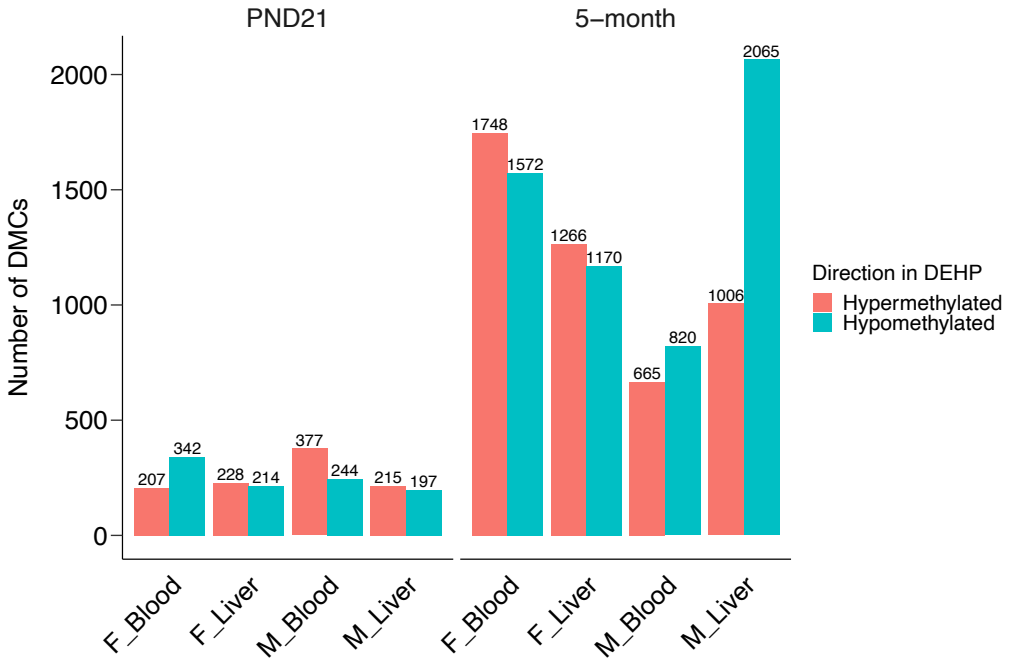


Figure 4.2 Overview of the number of DMCs between DEHP exposure and control groups.
The bar plot shows the number of hypermethylated and hypomethylated DMCs in DEHP separately for each comparison at both PND21 and 5 months.

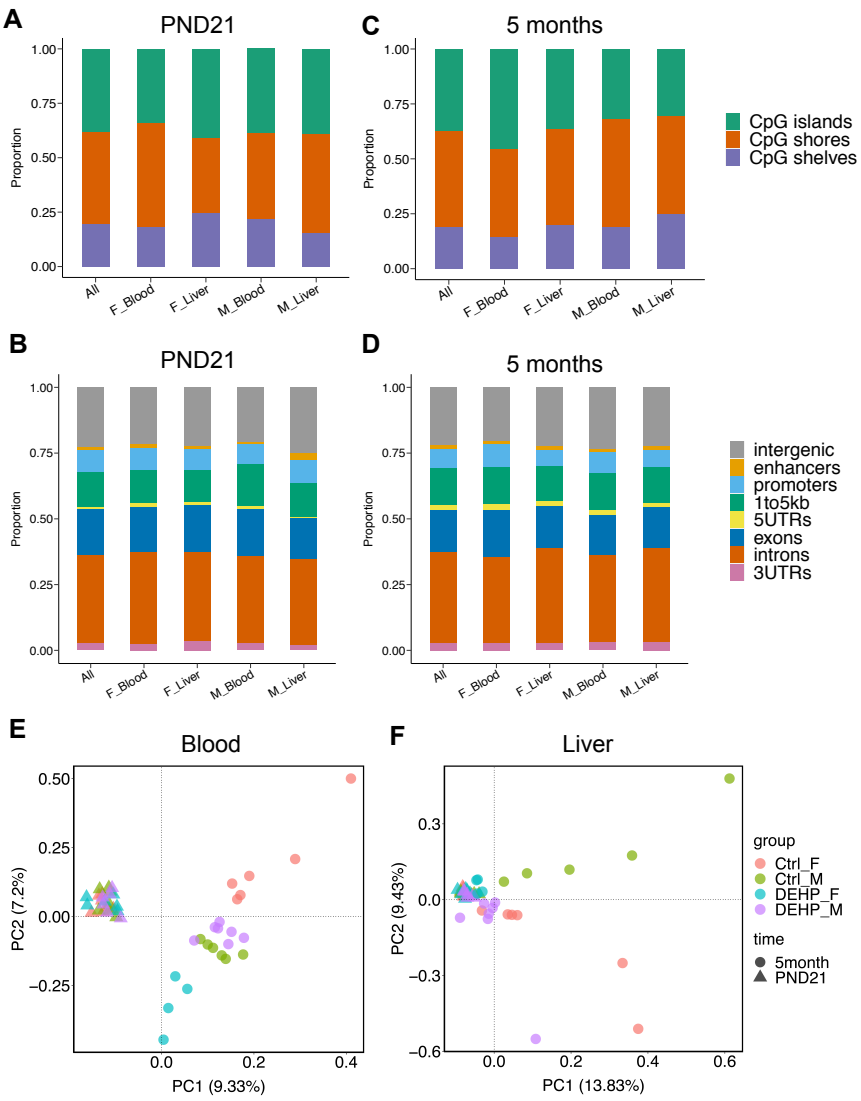


Figure 4.3 Annotated locations of the differentially methylated cytosines (DMCs) and heterogeneity among the animals in blood and liver.
(A) Stacked bar plot of 3-week (PND21) DMCs annotated to CpG islands, shores, shelves, and (B) regions relative to genes. (C) Stacked bar plot of 5-month DMCs annotated to CpG islands, shores, shelves, and (D) regions relative to genes. The "All" column indicates all the DMCs tested in each condition. (E) PCA plot showing the distribution of each individual mouse based on the methylation level of DMCs in at least one comparison in blood and (F) liver.

4.3.2 Pathways enriched with differential methylation in DEHP exposure

In addition to individual genes, we were interested in identifying sex-specific pathways enriched with differential methylation at PND21 and/or 5 months as a result of DEHP exposure.

We found more enriched GO terms for 5 months (n=408) than for PND21 (n=16). After combining related GO terms, organ growth and morphogenesis in thymus, muscle, eye, heart and embryos were enriched in female blood at 5 months, together with cell-cell adhesion. In female liver, cellular response to estrogen stimulus, DNA binding and intracellular transport were significant. For male-specific pathways, cell adhesion molecule binding was enriched in blood, while muscle contraction, cholesterol homeostasis and organ morphogenesis were enriched in liver.

4.3.3 Consistent differential methylation marks across tissues and time points

To address the TaRGET II consortium's aim of determining the extent to which a surrogate tissue (blood) reflects the changes observed in a target tissue (liver), we compared the DMCs from liver and blood at both PND21 and 5 months. We identified only one overlapped DMC in female for PND21 samples, which was near the promoter region of the *Eng* gene. *Eng* encodes the endoglin protein which plays an important role in the regulation of angiogenesis (**Table 4.1**). A total of 12 and 12 overlapped DMCs were detected for 5-month female and 5-month male respectively, but only 3 and 7 of these were changed in the same direction for both tissues (**Table 4.1**). The only gene showing prior correlation with DEHP exposure from CTD was a glycine transporter gene *Slc6a9*, and it was found with hypomethylation on exon 15 in both male blood and liver at 5 months. These findings overall represent a low level of consistency in DEHP-induced differential methylation between liver and blood, and a similar dearth of overlap between blood and liver was reported in our lead exposure study (230).

With the multiple differentially methylated sites and regions induced by DEHP exposure, we next determined which changes were consistent between PND21 and 5 months. The number of overlapped DMCs in the same direction between the two time points as well as their details

are listed in **Table 4.2**. Only one overlapped promoter DMC was reported, which was located in the *Ick* gene in female blood. Another overlapped DMC was identified in an intronic region of the same gene, and both of the CpGs were hypomethylated.

Group	# of DMCs overlapped	Direction in DEHP	location	gene	annotation	Methylation difference % (blood; liver)
PND21 Female	1	Hypermethylated	chr2:32672381	Eng	promoter; 1to5kb; exon	30.59; 32.98
PND21 Male	0					
5-month Female	3	Hypermethylated	chr4:153333737		intergenic	12.63; 16.47
		Hypermethylated	chr16:32173514		intergenic	17.33; 16.54
		Hypomethylated	chr9: 43151371	Pou2f3	intron	-13.50; -20.93
5-month Male	7	Hypermethylated	chr16:21402190		intergenic	10.85; 53.92
		Hypermethylated	chr6:114408371	Hrh1	intron	16.27; 11.60
		Hypomethylated	chrX:37950454	Rhox11-ps2-201	1to5kb	-10.05; -25.61
		Hypomethylated	chr6:116795937		intergenic	-12.63; -23.13
		Hypomethylated	chr13:53564366		intergenic	-12.34; -14.81
		Hypomethylated	chr4:117868149	**Slc6a9**; Ccdc24	exon; 3'UTR	-14.86; -17.66
		Hypomethylated	chr16:93981703	Cldn14	intron	-15.15; -14.91

Table 4.1 Sex-stratified differentially methylated cytosines (DMCs) consistent in both blood and liver at PND21 and 5 months.
Genomic locations and genic annotations are provided. The percentage of methylation difference was calculated as DEHP - Ctrl in blood and liver samples respectively. The gene marked in bold indicates prior correlation with DEHP exposure from CTD.

Group	# of DMCs overlapped	Direction in DEHP	position	gene	annotation	Methylation difference % (PND21; 5-month)
Female blood	4	Hypermethylated	chr5:117054318	Gm7478-201	1to5kb	27.34; 25.37
		Hypomethylated	chr9:78154861	Ick	promoter; intron	-18.20; -11.96
		Hypomethylated	chr9:118764757	Itga9	intron	-12.36; -18.38
		Hypomethylated	chr15:102955158	Hoxc11	exon	-17.39; -15.72
Female liver	1	Hypermethylated	chr7:40026066		intergenic	15.70; 12.93
Male blood	3	Hypermethylated	chr6:86098604	Add2	exon	39.24; 23.98
		Hypermethylated	chr5:146687971		intergenic	17.55; 15.07
		Hypomethylated	chr5:112269237	Crybb1	intron	-15.13; -14.46
Male liver	1	Hypomethylated	chr9:42520790	Grik4	exon; 3'UTR	-19.08; -33.01

Table 4.2 Sex- and tissue- specific differentially methylated cytosines (DMCs) consistent at both the 3-week and 5-month time points.
Genomic locations and genic annotations are provided. The percentage of methylation difference was calculated as DEHP - Ctrl in PND21 and 5-month samples respectively.

4.3.4 DNA methylation levels are higher in blood than liver at both PND21 and 5-month

DNA methylation differences between blood and liver are relevant for determining the extent to which blood can be used as a surrogate for biomarkers of exposure. We investigated the methylation level at CpG islands, shores and shelves for all the mice tested at PND21, and the result showed that levels were significantly higher in blood than liver for all these regions, especially in CpG shores and shelves (**Figure 4.4A**). No clear sex-specific difference was found. Upon replicating the comparison for 5-month mice, we found an even more drastic difference between the methylation levels of blood and liver (**Figure 4.4B**). In both cases, no clear overall difference in methylation levels was observed between DEHP exposure and the control group. These large baseline methylation differences in blood compared to liver may contribute to the lack of consistent changes observed due to DEHP exposure.

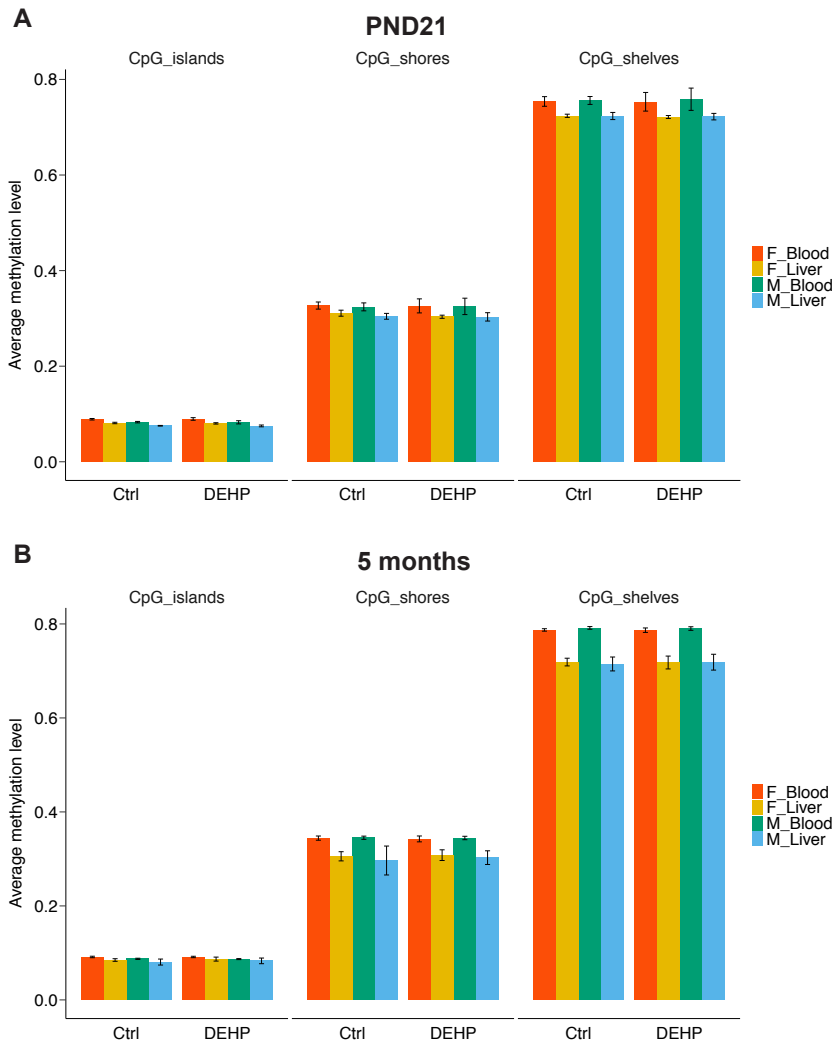


Figure 4.4 Sex- and tissue- stratified genome-wide DNA methylation levels at CpG islands, shores and shelves.
Sex- and tissue- stratified genome-wide DNA methylation levels at CpG islands, shores and shelves at (A) 3 weeks (PND21) and (B) 5 months. Blood shows higher DNA methylation levels than liver in CpG shores and shelves for both sexes, and with tissue differences increasing with mouse age.

4.3.5 DEHP exposure impacts DNA methylation of imprinted genes

Previous publications suggest that exposure to DEHP and other endocrine disruptors can profoundly alter epigenetic marks in imprinted genes (229,231,232), however the effect of perinatal DEHP in blood and liver specifically remains unknown. Here we compared the methylation level of mouse imprinted genes at CpG islands, shores and shelves, and discovered

that the effect of DEHP on imprinted genes is both sex and tissue specific (**Figure 4.5A,B**). For instance, male liver showed slight increased methylation at CpG shores with DEHP exposure at both PND21 and 5 months, while male blood had decreased methylation at CpG shelves at PND21. In general, changes in imprinted gene methylation levels by DEHP exposure were more obvious in liver than in blood, reinforcing the sensitivity of liver as a target tissue of DEHP exposure.

We found multiple mouse imprinted genes with DMCs in their promoter at both time points. For instance, the imprinted gene we found with the most DMCs was *Gnas*, which had 11 DMCs with DEHP hypermethylation in PND21 male blood and liver, as well as 5-month female blood, female liver and male liver. Another important imprinted gene, *Runx1*, which functions in hemopoiesis, was found to have promoter DEHP hypomethylation in female blood at PND21, but with DEHP hyper-methylation in female blood at 5 months. A similar pattern was found for *Zrsr1*, a maternally imprinted gene, which showed both DEHP hyper- and hypo-methylation in its promoter region at both time points in female liver. Despite lack of average methylation change on the gene body, *Igf2r* was found with several DMCs in both directions of female blood, female liver and male blood.

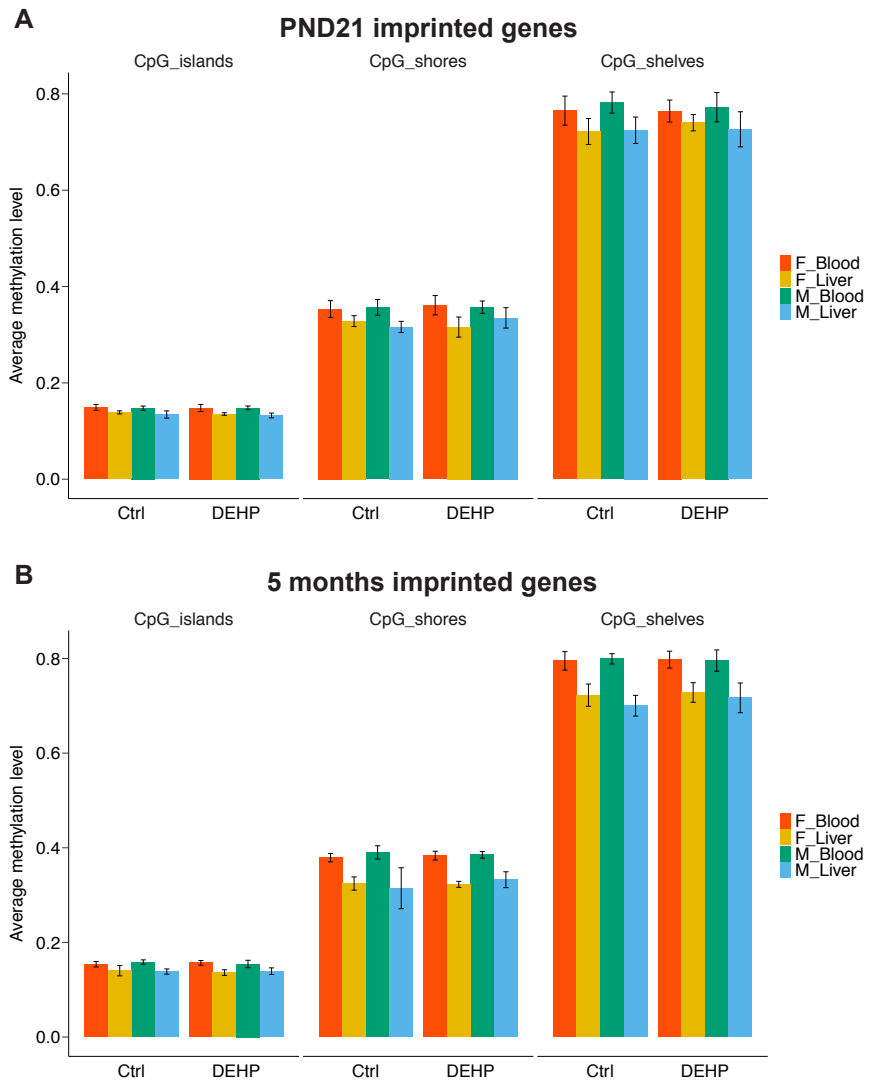


Figure 4.5 Sex- and tissue- stratified genome-wide DNA methylation levels of imprinted genes at CpG islands, shores and shelves.
Sex- and tissue- stratified genome-wide DNA methylation levels of imprinted genes at CpG islands, shores and shelves at (A) 3 weeks (PND21) and (B) 5 months. Blood shows higher DNA methylation levels than liver in CpG shores and shelves for both sexes, and slightly higher level of differences in liver were observed between DEHP and control group compared with the differences in blood.

4.3.6 DEHP exposure has greater effects on gene expression at PND21 than 5 months

To examine whether the differential methylation associated with DEHP exposure had an impact on gene expression at the same time points, RNA-seq data were generated using RNA

from the same animals. Results showed a subtle overall effect, with only a few differentially expressed genes identified at 5 months. At PND21, a total of 5 and 70 significant genes in male and female blood, respectively, and 52 and 706 in female and male liver (the total number of differentially expressed genes in each group are listed in **Figure 4.6A**). By comparing differential gene expression with previously identified DMCs, we did not observe any genes with both altered DNA methylation and change in gene expression at 5 months. However, *Zfhx3*, *Sh3pxd2a,* and *Plec* were all upregulated in the DEHP exposure group at PND21 in female liver, and these genes also harbored hypomethylation on the gene body in the respective group. On the other hand, downregulated gene *Col4a2* was found with hypermethylation on exon 23 in male blood. In PND21 male liver, a total of 20 differentially expressed genes were found to harbor 24 DMCs, with 1 of these genes being hypermethylated and downregulated, and 9 being hypomethylated and upregulated (**Figure 4.6B**). Two of the genes, *Dnmt3a* and *Abca2*, were DEHP-relevant genes in CTD with opposite direction between gene expression and methylation.

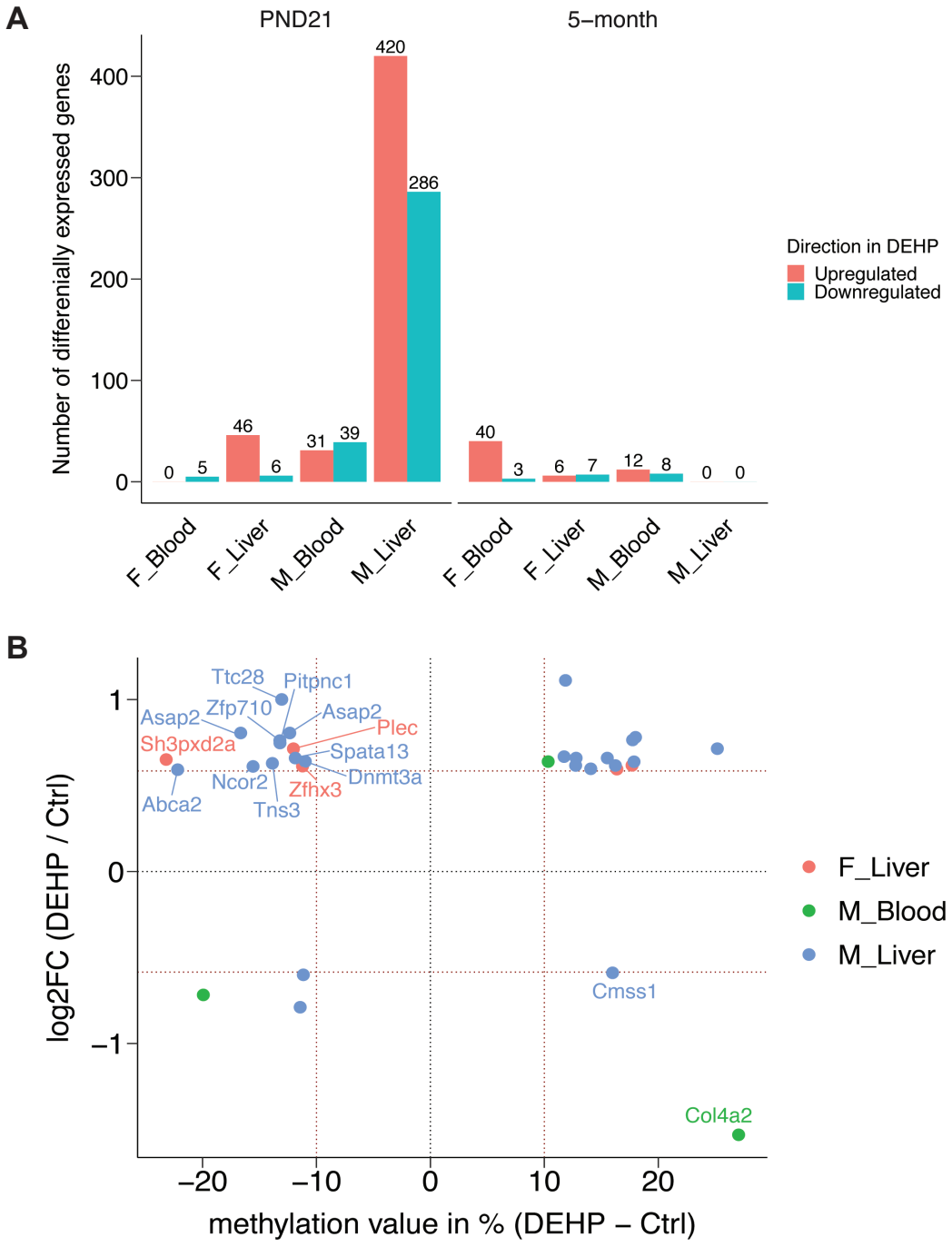


Figure 4.6 Differential gene expression analysis on the same cohort.
(A) Bar plot showing the number of up-regulated and down-regulated genes in blood and liver separately at the two time points. (B) Scatterplot showing all the DMCs that overlapped with upregulated or downregulated genes. The top half represents genes that are significantly upregulated in DEHP exposure groups, and the right half represents hypermethylated DMCs in DEHP exposure groups. Sites that showed opposite direction of DNA methylation and gene expression are labeled with gene symbols.

4.4 Discussion

In this study, we used a mouse model with pre- and peri-natal DEHP exposure to assess resultant genome-wide methylation responses at two ages (PND21 and 5 months) using ERRBS. We identified sex-specific DNA methylation changes in both juvenile and adult mouse liver and blood linked with the DEHP exposure. We found a few hundred genes harboring DMCs in their promoter regions, though only one (*Spata13*) showed significant gene expression change of the opposite direction from DNA methylation. For instance, the gene *Esr1* was found to harbor three hypomethylated sites upon DEHP exposure in the promoter region in livers of 5-month-old females. Interestingly, although previous studies demonstrated both positive and negative correlations between DEHP exposure and change in *Esr1* gene expression (233–236), the two studies of mouse liver both suggested that DEHP exposure is correlated with an increase in *Esr1* expression. Our finding of hypomethylation, which occurred in liver but not blood, is consistent with the fact that the previously-observed effects of DEHP on *Esr1* expression were liver-specific and not observed in other tissues (236).

Other interesting finding lies with the imprinted gene *Gnas*, which displayed promoter hyper-methylation in the DEHP group in both male liver and blood at PND21, and both female liver and blood at 5 months. As an important component in G protein-coupled receptors, Gnas plays a role in development of obesity and regulation of energy balance (237). Notably, Gnas was previously found to have differential methylation in response to many types of environmental exposures, including bisphenol A (238), lead (230), folic acid depletion (239), gases/fumes (240) and prenatal maternal stress (241), suggesting the Gnas imprinted cluster may be an especially labile region, sensitive to many environmental exposures. Although there was prior evidence from CTD showing the impact of DEHP on DNA methylation levels of imprinted

genes *Igf2r*, *Peg3,* and *H19* in mouse oocytes (109), we did not observe any significant changes in DNA methylation at these genes in either blood or liver.

Most of the differential methylation marks identified in this study were sex-specific, which is related to the fact that DEHP interacts with the androgen (AR), estrogen (ER), and PPARs, with a negative impact on testosterone production (104). Another explanation lies with the fact that reprogramming of post-fertilization methylation and in primordial germ cells occur with different kinetics and mechanisms between males and females (242). Multiple direct effects at the molecular level of both DEHP and its metabolites were observed in vitro by using a reporter assay with human cell lines on the estrogen receptor (*ESR1*) and on AR (97,243). In our study, *Esr1* demonstrated sex-specific differential methylation at its promoter region, while AR also showed sex-specific differential methylation, but in the gene body region. PPARs also regulate the expression of UCPs, including *Ucp2*, which harbored a hypermethylated site with 53.89% methylation difference, and previous research showed DEHP can result in decreased expression of *Ucp2* (244).

The differential methylation results also revealed that the number of DMCs at 5 months were almost 3-5 times the number of DMCs at PND21, indicating DEHP exposure early in life is correlated with more significant altered DNA methylation at a later stage of mouse development. Interestingly, significantly greater heterogeneity in methylation was observed among the 5-month animals than PND21 counterparts in both blood and liver. One explanation could be increased cell type methylation heterogeneity with age after DEHP exposure, which has been reported in mice of various ages (245–247), but we are not able to distinguish this hypothesis from the effect of a few outlier individuals without single-cell analysis. Another possibility lies with “environmental deflection”, a toxicant-mediated shift away from the baseline rates of age-

related methylation (248), where the deflection of DEHP could shift the methylation process during development. In support of DEHP causing more heterogeneity with age, we did observe slightly higher variance levels in the DEHP groups compared with control groups at 5 months for age-related CpGs (data not shown, average percentage increase of standard deviation = 5.65%).

In particular, we identified many enriched pathways at 5 months only, in either blood or liver. For instance, cellular response to estrogen stimulus and regulation of intracellular transport were significant in female liver, consistent with DEHP's effect on female hormones (249), while cholesterol homeostasis and channel activity were significant in male, and many of the aforementioned pathways are relevant to the action of DEHP in liver. Organ growth, morphogenesis, and cell-cell adhesion were enriched in female blood, suggesting the epigenetic effect of DEHP has potential effects on organ development and epithelial-mesenchymal transition (EMT) (250).

These methylation changes, however, did not translate to differential expression at the 5-month time point, which is in line with the finding of a recent study on endocrine disrupting chemicals showing that reprogrammed genes remained transcriptionally silent without any impact on metabolism until a later life stage (251). It may be that additional aging or an external stimulus leads to further differential gene expression. We did not identify many affected GO terms at PND21, possibly due to the lower degrees of differential methylation. We identified two genes, *Dnmt3a* and *Abca2* with DEHP hypomethylation and upregulation in male liver at PND21. *Dnmt3a* is an important epigenetic modifier in embryonic development and imprinting, and a few previous studies re-confirmed the DEHP exposure results in increased expression (252,253). Another study on low-dose pollutant mixture emphasized the association between

DEHP exposure and decreased expression of *Abca2* (254), which is a PPAR target gene and a member of the superfamily of ATP-binding cassette (ABC) transporters.

Our study did have limitations that may have resulted in missing some important findings. The use of ERRBS resulted in the analyzed CpG sites being biased towards CpG islands and shores, and other areas of higher-than-average CpG density. The technique of ERRBS does not distinguish 5hmC (5-hydroxymethylcytosine) from 5mC. Since we only focused on 5mC in this study, Oxy-RRBS would have been a more accurate measurement option. As the next step, we plan to analyze the methylation profile in the same cohort more comprehensively using whole genome bisulfite sequencing (WGBS), which captures CpG sites of the entire genome at single-base resolution. Due to limited resources, we did not measure the blood cell count before isolating DNA and RNA in our experiments, which could have ruled out the potential effects of blood cell type differences in the peripheral blood leukocytes between DEHP exposure and control groups (255). As an alternative to using the blood cell counts, we performed cell type deconvolution with the RNA-seq data on all the blood samples, and the results showed no significant difference between DEHP exposure and control groups regarding all cell types tested. Although this could indicate that the blood cell counts did not play a major role in the methylation level difference among each sex-specific group, we cannot rule out a slight overall shift in the percent of leukocytes. Our relatively small sample sizes when stratified by sex may also have resulted in missing many of the more subtle DNA methylation and gene expression differences, with small effect sizes at the given DEHP dosage.

As part of the TaRGET II consortium, one of the objectives of this study was to determine the extent to which surrogate tissues (blood) reflect the target tissues (liver) in terms of DNA methylation, and whether methylation marks are consistent or different ages of animals

(204,256). Although hundreds to thousands of DMCs and DMRs were detected in both liver and blood at both PND21 and 5 months, only a few DMCs overlapped among them. For example, the same promoter DMC near the *Eng* gene was detected in female blood and liver at PND21, and *Slc6a9* had the same hypomethylated mark between the two tissues in male at 5 months, whose upregulation was linked to DEHP exposure (257). Overlapped enriched pathways between liver and blood at 5 months included cell fate commitment, transcription factor activity and cyclic nucleotide metabolic process. Due to the small number of consistent marks between tissues in both juvenile and adult stages, we conclude that in general blood does not serve as a good surrogate tissue for liver, in terms of the DNA methylation profile changes due to our dose of perinatal DEHP exposure. However, the blood-specific methylation changes may be used as biomarkers for perinatal DEHP exposure. Indeed, we identified a small number of individual CpG sites affected in both tissues by DEHP in a sex-specific manner, which can serve as a starting point for targeted studies with larger

Chapter 5 Discussion

5.1 Conclusions

In this dissertation, I have described the epigenetic effects in carcinogenesis and response to environment, with a particular focus on regulatory roles of DNA methylation. Although genetic changes during DNA replication are one of the most significant risks of cancer development, environmental exposures to certain chemicals, such as tobacco smoke and radiation, can also cause DNA damage and epigenetic changes that contribute to carcinogenesis (258,259). For instance, both viral infections (including HPV) and tobacco consumption (first- and second-hand) are examples of environmental exposure that are known to increase the risk of multiple cancers, including HNSCC. The effects of some carcinogenic environmental exposures can occur later in life. Diet, for example, is a type of environmental exposure that influence the risk of development cancer, and numerous studies have confirmed multiple nutritional compounds have epigenetic targets in cancer cells (260). Thus, the interplay between carcinogenesis and environmental exposures has been essential genetically and epigenetically, as environmental factors continue to surface as potentially instrumental in explaining various pathways of carcinogens in different tumor types.

In Chapter 2, we described global 5hmC profiles and important genes with 5hmC changes in HNSCC both by HPV status and by HPV(+) subtype. Although 5hmC marks genes across a wide array of cellular processes, 5hmC profiles highlight genes turned on during differentiation, and can therefore be used for an in-depth characterization of the differentiation state of tumors. Thus, genome-wide 5hmC analysis is beneficial, especially to the extent that

differentiation state affects carcinogenic pathways, including cell junctions and adhesion, invasion, and migration in our study. Our data suggest that 5hmC profiles for the IMU subtype could be a useful biomarker in exploring the prognosis advantage of HPV(+) cancers and should be explored further. As the first study to characterize the genome-wide 5hmC profile in HNSCC, we identified significant genome-wide hyper-5hmC in HPV(-) tumors, with both promoter and enhancer 5hmC levels being clinically relevant and able to distinguish meaningful tumor subgroups. We also implicated 5hmC in key cancer-related processes that determine the likelihood of metastasis in head and neck cancer. Clinically, therapeutic de-escalation schedules are being introduced for HPV(+) patients, but the current challenge to such changes includes better identification of the small subset of HPV(+) cancer patients that have poor prognosis. Our study has important implications that 5hmC levels are crucial in defining tumor characteristics and potentially used to define clinically meaningful cancer patient subgroups for many cancer types.

In Chapter 3, we captured the immune-related gene expression on a subset of oral cavity cancer patients during a Phase 2b clinical trial testing the effects of the cytokines in IRX-2. We were able to characterize the significant change of more than 100 immune-related genes in the control and/or treatment regimen and explore the correlation between response information and immune signatures. We conclude that the cytokine components of IRX-2 have definite but subtle effects on patient immune response, and that some of the effects of the complete IRX-2 treatment are due to the non-cytokine components. Specific benefits in the tumor microenvironment of the cytokine cocktail in IRX-2 may include increased CD8+ T cell density, a slight global and repetitive element re-methylation of the genome, upregulation of the tumor suppressor DMBT1, and unchanged PD1/PDL1 for the subset of patients showing immune cell infiltration. Future

studies with larger sample size and more complete data will be able to provide a more comprehensive picture and a better understanding of the effects of IRX-2 treatment and help in selecting patients most suitable for immunotherapy.

In Chapter 4, we used a mouse model with perinatal DEHP exposure to assess resultant genome-wide methylation responses at two ages (PND21 and 5 months) using ERRBS. We identified sex-specific DNA methylation changes in both juvenile and adult mouse liver and blood linked with the DEHP exposure. Although hundreds to thousands of DMCs and DMRs were detected in both liver and blood at both PND21 and 5 months, only a few DMCs overlapped among them, indicating blood does not serve as a good surrogate tissue for liver in terms of the DNA methylation profile changes at our dose of perinatal DEHP exposure. However, the blood-specific methylation changes may be used as biomarkers for perinatal DEHP exposure. Indeed, we identified a small number of individual CpG sites affected in both tissues by DEHP in a sex-specific manner, which can serve as a starting point for targeted studies with larger sample sizes.

5.2 Application of the important epigenetic marker 5hmC in other cancer types

In Chapter 2, we described the importance of 5hmC in HNSCC, especially in promoter and enhancer regions in distinguishing HPV(+) and HPV(-) tumors. We have collected whole genome bisulfite sequencing (WGBS) data on the same cohort of patients, which captures both 5mC and 5hmC. By combining WGBS and hMeDIP-seq data, we plan to investigate the quantitative relationships between gene expression and the two forms of DNA methylation at promoters, enhancers, and gene bodies (261). More specifically, even though the base-pair resolution is lower for hMeDIP-seq than WGBS, we can compute the average methylation and

hydroxymethylation by 150bp to 200bp windows. In addition, we can explore the correlation between DMRs and hyper-5hmC regions, and further study their regulatory role of gene expression by integrating them with RNA-seq data.

As part of the future directions, we are interested in the role of 5hmC in other cancer types, particularly the cancers that are associated with HPV infection, such as cervical cancer, where more than 95% of the cases are HPV(+). One preliminary study characterized the 5mC and 5hmC profiles in cervical cancer using MeDIP-seq and hMeDIP-seq, where the group identified downregulation of 5hmC in cancer compared with control samples (262). Due to the lack of control samples in our HNSCC study, we cannot compare our DMRs and DhMRs with those discovered in the cervical cancer study. However, we can use a similar approach to study 5mC and 5hmC together in HNSCC, and explore the trend of 5hmC distribution across the genome, especially in HPV(+) tumors. Just as we discovered the HPV(+) subtypes within HNSCC, a similar approach can be used on the gene expression and epigenetic profiles of cervical cancers to further define biologically meaningful subtypes. By conducting additional sequencing to capture 5hmC profiles in HPV(-) cervical cancers and comparing HPV(+) with HPV(-) tumors, we will be able to better understand the epigenetic effects of HPV infection in cervical cancers.

5.3 Improved characterization of immune response in IRX-2 immunotherapy

Based on the caveats discussed in Chapter 3, we propose a few ways to improve the IRX-2 immunotherapy study. As an immunomodulatory drug, the immunotherapy with IRX-2 regimen is expected to boost the host immune response, converting the tumor status from “cold” to “hot” by activating key elements of the tumor microenvironment. We did observe

upregulation of more than 50 immune-related genes after the treatment with IRX-2, but these changes are not specific to the cytokine components of the IRX-2 regimen. Although we did not identify any significant differential methylation between biopsy and resection samples in either regimen, we found a high concordance in estimated CD8+ T cells between cell type deconvolution and manual cell type counts (TMA). One of the main limitations of this study lies with the fact that NanoString technology only measures the expression of 730 immune response genes, but IRX-2 is involved in many biological processes not limited to immune response. If we can conduct RNA-seq on the same cohort, we will get a more comprehensive gene expression profile both before and after treatment, and better understand the functions of IRX-2 by conducting downstream pathway and network analyses. Second, we plan to perform manual cell type counts on more samples. The immune responder criterion used in our study is defined as patients having an increase of more than 10 CD8+ cells / mm^2, but this information is only available in 26 out of a total of 96 patients which greatly restricted the interpretation on the correlation between immune response genes and immune response in patients. Last, we also plan to expand the cohort to better study the DNA methylation shifts during IRX-2 treatment. So far, we only measured the methylation level in 24 patients using the EPIC BeadChip, and very few immune responder patients were included in the methylation analysis. The cell type deconvolution analysis on more patients would provide a useful confirmation of the impact of IRX-2 immunotherapy in changing different cell type proportions, especially in immune cells.

5.4 Effects of epigenetic modifications on other types of environmental exposures

The different types of environmental exposures TaRGET II consortium covers are arsenic, lead (Pb), bisphenol A (BPA), tributyltin (TBT), DEHP, the dioxin tetrachlorodibenzo-

p-dioxin (TCDD), and air pollution in the form of particulate matter <2.5μ ($PM_{2.5}$) (204). Aiming to explore the conservation of perturbations in epigenomic marks across target tissues/cells (adversely affected by environmental exposures) and surrogate tissues/cells (easily accessible and reflect the environmental exposures), the consortium has listed potential target tissues and surrogate tissues for each exposure separately in mouse models. In Chapter 4, we reported the effects of perinatal DEHP exposure on DNA methylation shift and gene expression profiles, and discovered that as a surrogate tissue, epigenetic changes in blood did not reflect those in liver, therefore not being a good surrogate tissue. In another study on epigenetics effects of perinatal lead exposure, a similar inconsistency was found between DNA methylation profile of liver and blood (230), indicating blood does not serve as a good surrogate tissue for liver in lead exposures either.

Now that we have a better understanding of the correlation between target and surrogate tissues based on DNA methylation, as the next step, we plan to explore the effects of other epigenetic modifications in DEHP, lead and other types of environmental exposures. For instance, previous studies linked 5hmC with different types of environmental exposures, such as particulate air pollution, Bisphenol A, hydroquinone and pentachlorophenol metabolites (263), but rarely were these studies conducted on a genome-wide scale. Based on the established association between DEHP exposures across three generations and significant changes in Tet1 and Tet2 expression in liver (118), we hypothesize that pre- and peri-natal DEHP exposure will shift the corresponding 5hmC profiles, especially in the target tissues of DEHP (brain, kidney and liver). Lead exposure can induce oxidative stress, and was previously linked with increased 5hmC and decreased 5hmC (264,265). The consortium has collected brain (cortex) tissues of mice of both sexes that are either exposed to lead or unexposed, and plan to quantify the

genome-wide 5hmC using hMeDIP-seq, so that we can further understand the sex-specific reprogramming of 5hmC with perinatal lead exposure. The epigenetic impact of BPA has been observed in many forms, including DNA methylation, hydroxymethylation, histone marks, and non-coding RNAs (238,266), but there is still a need to capture these epigenetic marks more systematically and at a genome-wide scale.

One of the biggest challenges in studying epigenetic effects of environmental exposures lies with the large variability within and between exposure groups, age groups, tissue types, and populations (259). Even though the use of animal models provides consistent and stable exposure regimes, gaps remain in the representation of different human populations and phenotypes that need to be addressed (266). Since health outcomes associated with both environmental exposures and epigenetic alterations can take years or generations to occur after the initial exposure, it is essential to have enough follow-up time and take measurements at multiple end points to study the long-term effects (267). The TaRGET II consortium not only enables improved understanding of the use of surrogate tissues in different exposure types, but also links epigenetics with transcriptomics at different time points, which casts light on understanding the regulatory mechanisms and aging effects of different epigenetic marks.

5.5 Closing statements

In this dissertation we captured and examined epigenetic signatures in head and neck cancer and DEHP exposure using high-throughput sequencing methods. Environment itself greatly impacts intergenerational and transgenerational epigenetic inheritance (268), and both cancers and exposures are examples of the environment. The interplay between epigenetics and DNA sequences is extremely crucial in the development of life sciences, affecting different

processes of the “Central Dogma”. The use of more advanced bioinformatics tools and pipelines greatly improve our understanding of the genetic and epigenetic mechanisms. In addition to investigating the genetics and epigenetics with new assays in new diseases, we should also focus on the interpretation and integration of existing data to further explore diseases from a more comprehensive point of view.

Bibliography

1. Egger G, Liang G, Aparicio A, Jones PA. Epigenetics in human disease and prospects for epigenetic therapy. Nature. 2004.
2. Moosavi A, Ardekani AM. Role of epigenetics in biology and human diseases. Iran. Biomed. J. 2016.
3. Feinberg AP. The Key Role of Epigenetics in Human Disease Prevention and Mitigation. N Engl J Med. 2018;
4. Pal S, Tyler JK. Epigenetics and aging. Sci. Adv. 2016.
5. Perera F, Herbstman J. Prenatal environmental exposures, epigenetics, and disease. Reprod. Toxicol. 2011.
6. Kulis M, Esteller M. DNA Methylation and Cancer. Adv Genet [Internet]. Academic Press; 2010 [cited 2018 Jan 9];70:27–56. Available from: http://www.sciencedirect.com/science/article/pii/B9780123808660600022?via%3Dihub
7. Kimura H. Histone modifications for human epigenome analysis. J. Hum. Genet. 2013.
8. Pfeifer GP, Kadam S, Jin SG. 5-hydroxymethylcytosine and its potential roles in development and cancer. Epigenetics and Chromatin. 2013.
9. Laird A, Thomson JP, Harrison DJ, Meehan RR. 5-hydroxymethylcytosine profiling as an indicator of cellular state. Epigenomics [Internet]. 2013;5:655–69. Available from: http://www.futuremedicine.com/doi/10.2217/epi.13.69
10. Hackett JA, Sengupta R, Zylicz JJ, Murakami K, Lee C, Down TA, et al. Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. Science (80-). 2013;
11. Ehrlich M, Ehrlich KC. DNA cytosine methylation and hydroxymethylation at the borders. Epigenomics. 2014.
12. Kurdyukov S, Bullock M. DNA methylation analysis: Choosing the right method. Biology (Basel). 2016.
13. Singer BD. A practical guide to the measurement and analysis of DNA methylation. Am. J. Respir. Cell Mol. Biol. 2019.
14. Jin C, Lu Y, Jelinek J, Liang S, Estecio MRH, Barton MC, et al. TET1 is a maintenance DNA demethylase that prevents methylation spreading in differentiated cells. Nucleic Acids Res. 2014;
15. Booth MJ, Branco MR, Ficz G, Oxley D, Krueger F, Reik W, et al. Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. Science (80-). 2012;
16. Yu M, Hon GC, Szulwach KE, Song CX, Zhang L, Kim A, et al. Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. Cell. 2012;
17. Hu L, Liu Y, Han S, Yang L, Cui X, Gao Y, et al. Jump-seq: Genome-Wide Capture and Amplification of 5-Hydroxymethylcytosine Sites. J Am Chem Soc. 2019;

18. Liu Y, Siejka-Zielińska P, Velikova G, Bi Y, Yuan F, Tomkova M, et al. Bisulfite-free direct detection of 5-methylcytosine and 5-hydroxymethylcytosine at base resolution. Nat Biotechnol. 2019;
19. Baylin SB, Jones PA. A decade of exploring the cancer epigenome-biological and translational implications. Nat. Rev. Cancer. 2011.
20. Skvortsova K, Stirzaker C, Taberlay P. The DNA methylation landscape in cancer. Essays Biochem. 2019.
21. Kulis M, Esteller M. DNA Methylation and Cancer. Adv. Genet. 2010.
22. Kitkumthorn N, Mutirangura A. Long interspersed nuclear element-1 hypomethylation in cancer: Biology and clinical applications. Clin. Epigenetics. 2011.
23. Jin SG, Jiang Y, Qiu R, Rauch TA, Wang Y, Schackert G, et al. 5-hydroxymethylcytosine is strongly depleted in human cancers but its levels do not correlate with IDH1 mutations. Cancer Res. 2011;
24. Yang H, Liu Y, Bai F, Zhang JY, Ma SH, Liu J, et al. Tumor development is associated with decrease of TET gene expression and 5-methylcytosine hydroxylation. Oncogene. 2013;32:663–9.
25. Bhattacharyya S, Yu Y, Suzuki M, Campbell N, Mazdo J, Vasanthakumar A, et al. Genome-wide hydroxymethylation tested using the HELP-GT assay shows redistribution in cancer. Nucleic Acids Res. 2013;
26. Konstandin N, Bultmann S, Szwagierczak A, Dufour A, Ksienzyk B, Schneider F, et al. Genomic 5-hydroxymethylcytosine levels correlate with TET2 mutations and a distinct global gene expression pattern in secondary acute myeloid leukemia. Leukemia. 2011.
27. Ko M, Huang Y, Jankowska AM, Pape UJ, Tahiliani M, Bandukwala HS, et al. Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. Nature. 2010;
28. Waha A, Müller T, Gessi M, Waha A, Isselstein LJ, Luxen D, et al. Nuclear exclusion of TET1 is associated with loss of 5- hydroxymethylcytosine in IDH1 wild-type gliomas. Am J Pathol. 2012;
29. Johnson KC, Houseman EA, King JE, Von Herrmann KM, Fadul CE, Christensen BC. 5-Hydroxymethylcytosine localizes to enhancer elements and is associated with survival in glioblastoma patients. Nat Commun. 2016;
30. Mendenhall WM, Mancuso a a, Amdur RJ, Stringer SP, Villaret DB, Cassisi NJ. Squamous cell carcinoma metastatic to the neck from an unknown head and neck primary site. Am J Otolaryngol. 2001;22:261–7.
31. Argiris A, Karamouzis M V., Raben D, Ferris RL. Head and neck cancer. Lancet. 2008.
32. Pai SI, Westra WH. Molecular pathology of head and neck cancer: Implications for diagnosis, prognosis, and treatment. Annu. Rev. Pathol. Mech. Dis. 2009.
33. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Piñeros M, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. Int. J. Cancer. 2019.
34. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;
35. Chaturvedi AK, Engels EA, Pfeiffer RM, Hernandez BY, Xiao W, Kim E, et al. Human papillomavirus and rising oropharyngeal cancer incidence in the United States. J Clin Oncol. 2011;
36. Baba S, Yamada Y, Hatano Y, Miyazaki Y, Mori H, Shibata T, et al. Global DNA

hypomethylation suppresses squamous carcinogenesis in the tongue and esophagus. Cancer Sci. 2009;
37. Lim Y, Sun CX, Tran P, Punyadeera C. Salivary epigenetic biomarkers in head and neck squamous cell carcinomas. Biomark. Med. 2016.
38. Castilho RM, Squarize CH, Almeida LO. Epigenetic modifications and head and neck cancer: Implications for tumor progression and resistance to therapy. Int J Mol Sci. 2017;18.
39. Zhou C, Ye M, Ni S, Li Q, Ye D, Li J, et al. DNA methylation biomarkers for head and neck squamous cell carcinoma. Epigenetics [Internet]. 2018 [cited 2019 Oct 9];13:398–409. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29927694
40. Gaździcka J, Gołąbek K, Strzelczyk JK, Ostrowska Z. Epigenetic Modifications in Head and Neck Cancer. Biochem. Genet. 2020.
41. Poage GM, Butler RA, Houseman EA, McClean MD, Nelson HH, Christensen BC, et al. Identification of an epigenetic profile classifier that is associated with survival in head and neck cancer. Cancer Res. 2012;
42. Stansfield JC, Rusay M, Shan R, Kelton C, Gaykalova DA, Fertig EJ, et al. Toward signaling-driven biomarkers immune to normal tissue contamination. Cancer Inform. 2015;
43. Zhou C, Ye M, Ni S, Li Q, Ye D, Li J, et al. DNA methylation biomarkers for head and neck squamous cell carcinoma. Epigenetics. 2018;
44. Arayataweegool A, Srisuttee R, Mahattanasakul P, Tangjaturonsasme N, Kerekhanjanarong V, Kitkumthorn N, et al. Head and neck squamous cell carcinoma drives long interspersed element-1 hypomethylation in the peripheral blood mononuclear cells. Oral Dis. 2019;
45. Jäwert F, Hasséus B, Kjeller G, Magnusson B, Sand L, Larsson L. Loss of 5-hydroxymethylcytosine and TET2 in oral squamous cell carcinoma. Anticancer Res. 2013;
46. Misawa K, Yamada S, Mima M, Nakagawa T, Kurokawa T, Imai A, et al. 5-Hydroxymethylcytosine and ten-eleven translocation dioxygenases in head and neck carcinoma. J Cancer. 2019;
47. Castellsagué X, Alemany L, Quer M, Halec G, Quirós B, Tous S, et al. HPV Involvement in Head and Neck Cancers: Comprehensive Assessment of Biomarkers in 3680 Patients. J Natl Cancer Inst. 2016;
48. Ghittoni R, Accardi R, Chiocca S, Tommasino M. Role of human papillomaviruses in carcinogenesis. Ecancermedicalscience. 2015.
49. Mirabello L, Clarke MA, Nelson CW, Dean M, Wentzensen N, Yeager M, et al. The intersection of HPV epidemiology, genomics and mechanistic studies of HPV-mediated carcinogenesis. Viruses. 2018;
50. Koneva LA, Zhang Y, Virani S, Hall PB, McHugh JB, Chepeha DB, et al. HPV Integration in HNSCC Correlates with Survival Outcomes, Immune Response Signatures, and Candidate Drivers. Mol Cancer Res. 2017;molcanres.0153.2017.
51. Taberna M, Mena M, Pavón MA, Alemany L, Gillison ML, Mesía R. Human papillomavirus-related oropharyngeal cancer. Ann. Oncol. 2017.
52. Khan A, Liu Q, Chen X, Stucky A, Sedghizadeh PP, Adelpour D, et al. Detection of human papillomavirus in cases of head and neck squamous cell carcinoma by RNA-seq and VirTect. Mol Oncol. 2019;
53. Blitzer GC, Smith MA, Harris SL, Kimple RJ. Review of the clinical and biologic aspects

of human papillomavirus- positive squamous cell carcinomas of the head and neck. Int. J. Radiat. Oncol. Biol. Phys. 2014.
54. Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DI, Nguyen-Tân PF, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. N Engl J Med. 2010;
55. Dok R, Nuyts S. HPV positive head and neck cancers: Molecular pathogenesis and evolving treatment strategies. Cancers (Basel). 2016.
56. Chung CH, Zhang Q, Kong CS, Harris J, Fertig EJ, Harari PM, et al. p16 protein expression and human papillomavirus status as prognostic biomarkers of nonoropharyngeal head and neck squamous cell carcinoma. J Clin Oncol. 2014;32:3930–8.
57. Tommasino M. The human papillomavirus family and its role in carcinogenesis. Semin. Cancer Biol. 2014.
58. Leemans CR, Snijders PJF, Brakenhoff RH. The molecular landscape of head and neck cancer. Nat. Rev. Cancer. 2018.
59. Johnson DE, Burtness B, Leemans CR, Lui VWY, Bauman JE, Grandis JR. Head and neck squamous cell carcinoma. Nat Rev Dis Prim. 2020;6.
60. Zhang Y, Koneva LA, Virani S, Arthur AE, Virani A, Hall PB, et al. Subtypes of HPV-positive head and neck cancers are associated with HPV characteristics, copy number alterations, PIK3CA mutation, and pathway signatures. Clin Cancer Res. 2016;22:4735–45.
61. Papillon-Cavanagh S, Lu C, Gayden T, Mikael LG, Bechet D, Karamboulas C, et al. Impaired H3K36 methylation defines a subset of head and neck squamous cell carcinomas. Nat Genet. 2017;
62. Locati LD, Serafini MS, Iannò MF, Carenzo A, Orlandi E, Resteghin C, et al. Mining of self-organizing map gene-expression portraits reveals prognostic stratification of HPV-positive head and neck squamous cell carcinoma. Cancers (Basel). 2019;
63. Qin T, Zhang Y, Zarins KR, Jones TR, Virani S, Peterson LA, et al. Expressed HNSCC variants by HPV-status in a well-characterized Michigan cohort. Sci Rep. 2018;
64. McCune JS. Immunotherapy to Treat Cancer. Clin. Pharmacol. Ther. 2016.
65. Lee L, Gupta M, Sahasranaman S. Immune Checkpoint inhibitors: An introduction to the next-generation cancer immunotherapy. J. Clin. Pharmacol. 2016.
66. Haslam A, Prasad V. Estimation of the percentage of us patients with cancer who are eligible for and respond to checkpoint inhibitor immunotherapy drugs. JAMA Netw. Open. 2019.
67. Ray A. Cytokines and their Role in Health and Disease: A Brief Overview. MOJ Immunol. 2016;
68. Balar AV, Weber JS. PD-1 and PD-L1 antibodies in cancer: current status and future directions. Cancer Immunol. Immunother. 2017.
69. Mei Z, Huang J, Qiao B, Lam AK yin. Immune checkpoint pathways in immunotherapy for head and neck squamous cell carcinoma. Int. J. Oral Sci. 2020.
70. Seiwert TY, Burtness B, Mehra R, Weiss J, Berger R, Eder JP, et al. Safety and clinical activity of pembrolizumab for treatment of recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-012): an open-label, multicentre, phase 1b trial. Lancet Oncol. 2016;
71. Chow LQM, Haddad R, Gupta S, Mahipal A, Mehra R, Tahara M, et al. Antitumor activity of pembrolizumab in biomarker-unselected patients with recurrent and/or

metastatic head and neck squamous cell carcinoma: Results from the phase Ib KEYNOTE-012 expansion cohort. J Clin Oncol. 2016;
72. Bauml J, Seiwert TY, Pfister DG, Worden F, Liu S V., Gilbert J, et al. Pembrolizumab for platinum- and cetuximab-refractory head and neck cancer: Results from a single-arm, phase II study. J Clin Oncol. 2017;
73. Ferris RL, Blumenschein G, Fayette J, Guigay J, Colevas AD, Licitra L, et al. Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck. N Engl J Med. 2016;
74. Harrington KJ, Ferris RL, Blumenschein G, Colevas AD, Fayette J, Licitra L, et al. Nivolumab versus standard, single-agent therapy of investigator's choice in recurrent or metastatic squamous cell carcinoma of the head and neck (CheckMate 141): health-related quality-of-life results from a randomised, phase 3 trial. Lancet Oncol. 2017;
75. Cohen EEW, Soulières D, Le Tourneau C, Dinis J, Licitra L, Ahn MJ, et al. Pembrolizumab versus methotrexate, docetaxel, or cetuximab for recurrent or metastatic head-and-neck squamous cell carcinoma (KEYNOTE-040): a randomised, open-label, phase 3 study. Lancet. 2019;
76. Polverini PJ, D'Silva NJ, Lei YL. Precision Therapy of Head and Neck Squamous Cell Carcinoma. J Dent Res. 2018;
77. Cohen EEW, Bell RB, Bifulco CB, Burtness B, Gillison ML, Harrington KJ, et al. The Society for Immunotherapy of Cancer consensus statement on immunotherapy for the treatment of squamous cell carcinoma of the head and neck (HNSCC). J Immunother Cancer. 2019;
78. Goltz D, Gevensleben H, Dietrich J, Schroeck F, de Vos L, Droege F, et al. PDCD1 (PD-1) promoter methylation predicts outcome in head and neck squamous cell carcinoma patients. Oncotarget. 2017;
79. Darvin P, Toor SM, Sasidharan Nair V, Elkord E. Immune checkpoint inhibitors: recent progress and potential biomarkers. Exp. Mol. Med. 2018.
80. Xue G, Cui Z-J, Zhou X-H, Zhu Y-X, Chen Y, Liang F-J, et al. DNA Methylation Biomarkers Predict Objective Responses to PD-1/PD-L1 Inhibition Blockade. Front Genet. Frontiers Media S.A.; 2019;10:724.
81. Starzer AM, Heller G, Tomasich E, Feldmann K, Hatziioannou T, Traint S, et al. DNA methylation profiling in patients with head and neck squamous cell carcinoma treated with immune checkpoint inhibitors. J Clin Oncol. American Society of Clinical Oncology (ASCO); 2020;38:e18527–e18527.
82. Chim CS, Kumar SK, Orlowski RZ, Cook G, Richardson PG, Gertz MA, et al. Management of relapsed and refractory multiple myeloma: Novel agents, antibodies, immunotherapies and beyond. Leukemia. 2018.
83. Matsushita M, Kawaguchi M. Immunomodulatory Effects of Drugs for Effective Cancer Immunotherapy. J. Oncol. 2018.
84. Joseph M, Enting D. Immune Responses in Bladder Cancer-Role of Immune Cell Populations, Prognostic Factors and Therapeutic Implications. Front. Oncol. 2019.
85. Machtinger R, Bollati V, Baccarelli AA. MiRNAs and lncRNAs as biomarkers of toxicant exposure. Toxicoepigenetics Core Princ Appl. 2018.
86. Services U. D of H and H. Toxicological profile: di(2-ethylhexyl)phthalate (DEHP). Agency Toxic Subst Dis Regist. 2002;
87. Kelley KE, Hernández-Díaz S, Chaplin EL, Hauser R, Mitchell AA. Identification of phthalates in medications and dietary supplement formulations in the United States and

Canada. Environ Health Perspect. 2012;
88. Bergé A, Cladière M, Gasperi J, Coursimault A, Tassin B, Moilleron R. Meta-analysis of environmental contamination by phthalates. Environ. Sci. Pollut. Res. 2013.
89. Shaz BH, Grima K, Hillyer CD. 2-(Diethylhexyl)phthalate in blood bags: Is this a public health issue? Transfusion. 2011.
90. Wen Y, Liu SD, Lei X, Ling YS, Luo Y, Liu Q. Association of PAEs with precocious puberty in children: A systematic review and meta-analysis. Int J Environ Res Public Health. 2015;
91. Kavlock R, Boekelheide K, Chapin R, Cunningham M, Faustman E, Foster P, et al. NTP center for the evaluation of risks to human reproduction: Phthalates expert panel report on the reproductive and developmental toxicity of butyl benzyl phthalate. Reprod Toxicol. 2002.
92. Högberg J, Hanberg A, Berglund M, Skerfving S, Remberger M, Calafat AM, et al. Phthalate diesters and their metabolites in human breast milk, blood or serum, and urine as biomarkers of exposure in vulnerable populations. Environ Health Perspect. 2008;
93. Kato K, Silva MJ, Reidy JA, Hurtz D, Malek NA, Needham LL, et al. Mono(2-ethyl-5-hydroxyhexyl) phthalate and mono-(2-ethyl-5-oxohexyl) phthalate as biomarkers for human exposure assessment to di-(2-ethylhexyl) phthalate. Environ Health Perspect. 2004;
94. Silva MJ, Reidy JA, Preau JL, Samandar E, Needham LL, Calafat AM. Measurement of eight urinary metabolites of di(2-ethylhexyl) phthalate as biomarkers for human exposure assessment. Biomarkers. 2006;
95. Latini G, De Felice C, Presta G, Del Vecchio A, Paris I, Ruggieri F, et al. Exposure to Di(2-ethylhexyl)phthalate in humans during pregnancy: A preliminary report. Biol Neonate. 2003;
96. Adibi JJ, Hauser R, Williams PL, Whyatt RM, Calafat AM, Nelson H, et al. Maternal urinary metabolites of Di-(2-ethylhexyl) phthalate in relation to the timing of labor in a US multicenter pregnancy cohort study. Am J Epidemiol. 2009;
97. Stenzid L, Rahban R, Pradosid J, Nef S, Giacobino AP. Genetic resistance to dehp-induced transgenerational endocrine disruption. PLoS One. 2019;
98. Schmid P, Schlatter C. Excretion and metabolism of di(2-ethylhexyl)-phthalate in man. Xenobiotica. 1985;
99. Anderson WAC, Castle L, Hird S, Jeffery J, Scotter MJ. A twenty-volunteer study using deuterium labelling to determine the kinetics and fractional excretion of primary and secondary urinary metabolites of di-2-ethylhexylphthalate and di-iso-nonylphthalate. Food Chem Toxicol. 2011;
100. Kurata Y, Shimamura N, Katoh M. Metabolite profiling and identification in human urine after single oral administration of DEHP. J Toxicol Sci. 2012;
101. Rusyn I, Peters J, Cunningham M. Modes of action and species-specific effects of di-(2-ethylhexyl)phthalate in the liver. Crit. Rev. Toxicol. 2006.
102. Zhang Y, Zhou L, Zhang Z, Xu Q, Han X, Zhao Y, et al. Effects of di (2-ethylhexyl) phthalate and high-fat diet on lipid metabolism in rats by JAK2/STAT5. Environ Sci Pollut Res. 2020;27:3837–48.
103. Marsman DS, Goldsworthy TL, Popp JA. Contrasting hepatocytic peroxisome proliferation, lipofuscin accumulation and cell turnover for the hepatocarcinogens wy-14,643 and clofibric acid. Carcinogenesis. 1992;

104. Rowdhwal SSS, Chen J. Toxic Effects of Di-2-ethylhexyl Phthalate: An Overview. Biomed Res. Int. 2018.
105. Stel J, Legler J. The role of epigenetics in the latent effects of early life exposure to obesogenic endocrine disrupting chemicals. Endocrinology. 2015.
106. Dutta S, Haggerty DK, Rappolee DA, Ruden DM. Phthalate Exposure and Long-Term Epigenomic Consequences: A Review. Front. Genet. 2020.
107. Zama AM, Uzumcu M. Epigenetic effects of endocrine-disrupting chemicals on female reproduction: An ovarian perspective. Front. Neuroendocrinol. 2010.
108. Scarano WR, Bedrat A, Alonso-Costa LG, Aquino AM, Fantinatti BEA, Justulin LA, et al. Exposure to an Environmentally Relevant Phthalate Mixture during Prostate Development Induces MicroRNA Upregulation and Transcriptome Modulation in Rats. Toxicol Sci. 2019;
109. Li L, Zhang T, Qin XS, Ge W, Ma HG, Sun LL, et al. Exposure to diethylhexyl phthalate (DEHP) results in a heritable modification of imprint genes DNA methylation in mouse oocytes. Mol Biol Rep. 2014;
110. Neier K, Cheatham D, Bedrosian LD, Dolinoy DC. Perinatal exposures to phthalates and phthalate mixtures result in sex-specific effects on body weight, organ weights and intracisternal A-particle (IAP) DNA methylation in weanling mice. J Dev Orig Health Dis. 2019;
111. Montrose L, Padmanabhan V, Goodrich JM, Domino SE, Treadwell MC, Meeker JD, et al. Maternal levels of endocrine disrupting chemicals in the first trimester of pregnancy are associated with infant cord blood DNA methylation. Epigenetics. 2018;
112. Bowman A, Peterson KE, Dolinoy DC, Meeker JD, Sánchez BN, Mercado-Garcia A, et al. Phthalate exposures, DNA methylation and adiposity in Mexican children through adolescence. Front Public Heal. 2019;
113. Wu S, Zhu J, Li Y, Lin T, Gan L, Yuan X, et al. Dynamic Epigenetic Changes Involved in Testicular Toxicity Induced by Di-2-(Ethylhexyl) Phthalate in Mice. Basic Clin Pharmacol Toxicol. 2010;
114. Abdel-Maksoud FM, Leasor KR, Butzen K, Braden TD, Akingbemi BT. Prenatal exposures of male rats to the environmental chemicals bisphenol A and di(2-ethylhexyl) phthalate impact the sexual differentiation process. Endocrinology. 2015;
115. Solomon O, Yousefi P, Huen K, Gunier RB, Escudero-Fung M, Barcellos LF, et al. Prenatal phthalate exposure and altered patterns of DNA methylation in cord blood. Environ Mol Mutagen. 2017;
116. Rattan S, Beers HK, Kannan A, Ramakrishnan A, Brehm E, Bagchi I, et al. Prenatal and ancestral exposure to di(2-ethylhexyl) phthalate alters gene expression and DNA methylation in mouse ovaries. Toxicol Appl Pharmacol. 2019;
117. Lyko F. The DNA methyltransferase family: A versatile toolkit for epigenetic regulation. Nat. Rev. Genet. 2018.
118. Wen Y, Rattan S, Flaws JA, Irudayaraj J. Multi and transgenerational epigenetic effects of di-(2-ethylhexyl) phthalate (DEHP) in liver. Toxicol Appl Pharmacol. 2020;
119. Degli Esposti D, Sklias A, Lima SC, Beghelli-de la Forest Divonne S, Cahais V, Fernandez-Jimenez N, et al. Unique DNA methylation signature in HPV-positive head and neck squamous cell carcinomas. Genome Med. 2017;
120. Lawrence MS, Sougnez C, Lichtenstein L, Cibulskis K, Lander E, Gabriel SB, et al. Comprehensive genomic characterization of head and neck squamous cell carcinomas.

Nature. 2015;
121. Choudhury JH, Ghosh SK. Promoter hypermethylation profiling identifies subtypes of head and neck cancer with distinct viral, environmental, genetic and survival characteristics. PLoS One. 2015;10.
122. Sartor MA, Dolinoy DC, Jones TR, Colacino JA, Prince MEP, Carey TE, et al. Genome-wide methylation and expression differences in HPV(+) and HPV(-) squamous cell carcinoma cell lines are consistent with divergent mechanisms of carcinogenesis. Epigenetics. 2011;6:777–87.
123. Lleras RA, Smith R V., Adrien LR, Schlecht NF, Burk RD, Harris TM, et al. Unique DNA methylation loci distinguish anatomic site and HPV status in head and neck squamous cell carcinoma. Clin Cancer Res. 2013;19:5444–55.
124. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science (80-). 2009;324:930–5.
125. Ruzov A, Tsenkina Y, Serio A, Dudnakova T, Fletcher J, Bai Y, et al. Lineage-specific distribution of high levels of genomic 5-hydroxymethylcytosine in mammalian development. Cell Res. 2011;21:1332–42.
126. Tan L, Xiong L, Xu W, Wu F, Huang N, Xu Y, et al. Genome-wide comparison of DNA hydroxymethylation in mouse embryonic stem cells and neural progenitor cells by a new comparative hMeDIP-seq method. Nucleic Acids Res. 2013;
127. Laird A, Thomson JP, Harrison DJ, Meehan RR. 5-hydroxymethylcytosine profiling as an indicator of cellular state. Epigenomics. 2013.
128. Jin SG, Jiang Y, Qiu R, Rauch TA, Wang Y, Schackert G, et al. 5-hydroxymethylcytosine is strongly depleted in human cancers but its levels do not correlate with IDH1 mutations. Cancer Res. 2011;71:7360–5.
129. Jäwert F, Hasséus B, Kjeller G, Magnusson B, Sand L, Larsson L. Loss of 5-hydroxymethylcytosine and TET2 in oral squamous cell carcinoma. Anticancer Res. International Institute of Anticancer Research; 2013;33:4325–8.
130. Wang Y, Hu H, Wang Q, Li Z, Zhu Y, Zhang W, et al. The level and clinical significance of 5-hydroxymethylcytosine in oral squamous cell carcinoma: An immunohistochemical study in 95 patients. Pathol Res Pract. 2017;
131. Ficz G, Gribben JG. Loss of 5-hydroxymethylcytosine in cancer: Cause or consequence? Genomics. 2014.
132. Ficz G, Branco MR, Seisenberger S, Santos F, Krueger F, Hore TA, et al. Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature. 2011;473:398–404.
133. Wu H, D'Alessio AC, Ito S, Wang Z, Cui K, Zhao K, et al. Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. Genes Dev. 2011;
134. Cavalcante RG, Patil S, Park Y, Rozek LS, Sartor MA. Integrating DNA methylation and hydroxymethylation data with the mint pipeline. Cancer Res. 2017;
135. Andrews S, Krueger F, Seconds-Pichon A, Biggins F, Wingett S. FastQC. A quality control tool for high throughput sequence data. Babraham Bioinformatics. Babraham Inst. 2015.
136. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;

137. Zhang Y, Lin YH, Johnson TD, Rozek LS, Sartor MA. PePr: A peak-calling prioritization pipeline to identify consistent or differential peaks from replicated ChIP-Seq data. Bioinformatics. 2014;30:2568–75.
138. Cavalcante RG, Sartor MA. Annotatr: Genomic regions in context. Bioinformatics. 2017;33:2381–3.
139. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. Mol Cell. 2010;38:576–89.
140. Tian Y, Morris TJ, Webster AP, Yang Z, Beck S, Feber A, et al. ChAMP: Updated methylation analysis pipeline for Illumina BeadChips. Bioinformatics. 2017;
141. Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, et al. COSMIC: The Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Res. 2019;
142. Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, et al. The accessible chromatin landscape of the human genome. Nature. 2012;489:75–82.
143. Lizio M, Harshbarger J, Shimoji H, Severin J, Kasukawa T, Sahin S, et al. Gateways to the FANTOM5 promoter level mammalian expression atlas. Genome Biol. 2015;
144. Li G, Chen Y, Snyder MP, Zhang MQ. ChIA-PET2: A versatile and flexible pipeline for ChIA-PET data analysis. Nucleic Acids Res. 2017;
145. Dowen JM, Fan ZP, Hnisz D, Ren G, Abraham BJ, Zhang LN, et al. Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. Cell. 2014;
146. Tang Z, Luo OJ, Li X, Zheng M, Zhu JJ, Szalaj P, et al. CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. Cell. 2015;
147. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 2013;14.
148. Anders S, Pyl PT, Huber W. HTSeq-A Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31:166–9.
149. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics [Internet]. 2010;26:139–40. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19910308%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2796818
150. Lee C, Patil S, Sartor MA. RNA-Enrich: A cut-off free functional enrichment testing method for RNA-seq with improved detection power. Bioinformatics. 2016;32:1100–2.
151. Welch RP, Lee C, Imbriano PM, Patil S, Weymouth TE, Smith RA, et al. ChIP-Enrich: Gene set enrichment testing for ChIP-seq data. Nucleic Acids Res. 2014;42.
152. Ernst J, Kellis M. ChromHMM: Automating chromatin-state discovery and characterization. Nat. Methods. 2012. page 215–6.
153. Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell. 2013;
154. Lovén J, Hoke HA, Lin CY, Lau A, Orlando DA, Vakoc CR, et al. Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell. 2013;
155. Khan A, Zhang X. DbSUPER: A database of Super-enhancers in mouse and human genome. Nucleic Acids Res. 2016;
156. Wernig-Zorc S, Yadav MP, Kopparapu PK, Bemark M, Kristjansdottir HL, Andersson

PO, et al. Global distribution of DNA hydroxymethylation and DNA methylation in chronic lymphocytic leukemia. Epigenetics and Chromatin. 2019;
157. Li J, Poi MJ, Tsai MD. Regulatory mechanisms of tumor suppressor P16INK4A and their relevance to cancer. Biochemistry. 2011;
158. Fenouille N, Tichet M, Dufies M, Pottier A, Mogha A, Soo JK, et al. The epithelial-mesenchymal transition (EMT) regulatory factor SLUG (SNAI2) is a downstream target of SPARC and AKT in promoting melanoma cell invasion. PLoS One. 2012;7.
159. Rahman MS, Akhtar N, Jamil HM, Banik RS, Asaduzzaman SM. TGF-β/BMP signaling and other molecular events: Regulation of osteoblastogenesis and bone formation. Bone Res. 2015.
160. Robinson MD, McCarthy DJ, Smyth GK. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2009;
161. Li J, Wu X, Zhou Y, Lee M, Guo L, Han W, et al. Decoding the dynamic DNA methylation and hydroxymethylation landscapes in endodermal lineage intermediates during pancreatic differentiation of hESC. Nucleic Acids Res. 2018;
162. Cohen EEW, Lingen MW, Zhu B, Zhu H, Straza MW, Pierce C, et al. Protein kinase Cζ mediates epidermal growth factor-induced growth of head and neck tumor cells by regulating mitogen-activated protein kinase. Cancer Res. 2006;
163. Tian Y ping, Lin A fen, Gan M fu, Wang H, Yu D, Lai C, et al. Global changes of 5-hydroxymethylcytosine and 5-methylcytosine from normal to tumor tissues are associated with carcinogenesis and prognosis in colorectal cancer. J Zhejiang Univ Sci B. 2017;
164. Hsu CH, Peng KL, Kang ML, Chen YR, Yang YC, Tsai CH, et al. TET1 Suppresses Cancer Invasion by Activating the Tissue Inhibitors of Metalloproteinases. Cell Rep. 2012;
165. Dong W, Li H, Zhang Y, Yang H, Guo M, Li L, et al. Matrix metalloproteinase 2 promotes cell growth and invasion in colorectal cancer. Acta Biochim Biophys Sin (Shanghai). 2011;
166. Gillison ML, Trotti AM, Harris J, Eisbruch A, Harari PM, Adelstein DJ, et al. Radiotherapy plus cetuximab or cisplatin in human papillomavirus-positive oropharyngeal cancer (NRG Oncology RTOG 1016): a randomised, multicentre, non-inferiority trial. Lancet. 2019;
167. Cooper T, Biron VL, Adam B, Klimowicz AC, Puttagunta L, Seikaly H. Association of keratinization with 5-year disease-specific survival in oropharyngeal squamous cell carcinoma. JAMA Otolaryngol - Head Neck Surg. 2015;
168. Mohanta S, Siddappa G, Valiyaveedan SG, Ramanjanappa RDT, Das D, Pandian R, et al. Cancer stem cell markers in patterning differentiation and in prognosis of oral squamous cell carcinoma. Tumor Biol. 2017;
169. Rocca A, Farolfi A, Bravaccini S, Schirone A, Amadori D. Palbociclib (PD 0332991): Targeting the cell cycle machinery in breast cancer. Expert Opin Pharmacother. 2014;
170. Zainal NS, Lee BKB, Wong ZW, Chin IS, Yee PS, Gan CP, et al. Effects of palbociclib in oral squamous cell carcinoma and the role of PIK3CA in conferring resistance. Cancer Biol Med. 2019;
171. Pavlasova G, Mraz M. The regulation and function of CD20: An “enigma” of B-cell biology and targeted therapy. Haematologica. 2020.
172. Lin TS. Ofatumumab: A novel monoclonal anti-CD20 antibody. Pharmgenomics. Pers. Med. 2010.

173. Colacino JA, Dolinoy DC, Duffy SA, Sartor MA, Chepeha DB, Bradford CR, et al. Comprehensive Analysis of DNA Methylation in Head and Neck Squamous Cell Carcinoma Indicates Differences by Survival and Clinicopathologic Characteristics. PLoS One. 2013;
174. Bhattacharyya S, Pradhan K, Campbell N, Mazdo J, Vasantkumar A, Maqbool S, et al. Altered hydroxymethylation is seen at regulatory regions in pancreatic cancer and regulates oncogenic pathways. Genome Res. 2017;
175. Argiris A, Karamouzis M V, Raben D, Ferris RL. Head and neck cancer. Lancet (London, England) [Internet]. 2008;371:1695–709. Available from: http://www.sciencedirect.com/science/article/pii/S014067360860728X
176. Pulte D, Brenner H. Changes in Survival in Head and Neck Cancers in the Late 20th and Early 21st Century: A Period Analysis. Oncologist. 2010;
177. Cramer JD, Burtness B, Ferris RL. Immunotherapy for head and neck cancer: Recent advances and future directions. Oral Oncol. 2019.
178. Ferris RL, Blumenschein G, Fayette J, Guigay J, Colevas AD, Licitra L, et al. Nivolumab for recurrent squamous-cell carcinoma of the head and neck. N Engl J Med. 2016;
179. Wolf GT, Moyer JS, Kaplan MJ, Newman JG, Egan JE, Berinstein NL, et al. IRX-2 natural cytokine biologic for immunotherapy in patients with head and neck cancers. Onco. Targets. Ther. 2018.
180. Berinstein NL, McNamara M, Nguyen A, Egan J, Wolf GT. Increased immune infiltration and chemokine receptor expression in head and neck epithelial tumors after neoadjuvant immunotherapy with the IRX-2 regimen. Oncoimmunology. 2018;
181. Smith IM, Mydlarz WK, Mithani SK, Califano JA. DNA global hypomethylation in squamous cell head and neck cancer associated with smoking, alcohol consumption and stage. Int J Cancer. 2007;
182. Jung H, Kim HS, Kim JY, Sun JM, Ahn JS, Ahn MJ, et al. DNA methylation loss promotes immune evasion of tumours with high mutation and copy number load. Nat Commun. 2019;
183. Misawa K, Yamada S, Mima M, Nakagawa T, Kurokawa T, Imai A, et al. Long interspersed nuclear element 1 hypomethylation has novel prognostic value and potential utility in liquid biopsy for oral cavity cancer. Biomark Res. 2020;
184. Chakravarthy A, Furness A, Joshi K, Ghorani E, Ford K, Ward MJ, et al. Pan-cancer deconvolution of tumour composition using DNA methylation. Nat Commun. 2018;
185. Waggott D, Chu K, Yin S, Wouters BG, Liu FF, Boutros PC. NanoStringNorm: An extensible R package for the pre-processing of nanostring mRNA and miRNA data. Bioinformatics. 2012;
186. Fortin JP, Labbe A, Lemire M, Zanke BW, Hudson TJ, Fertig EJ, et al. Functional normalization of 450k methylation array data improves replication in large cancer studies. Genome Biol. 2014;
187. Peters TJ, Buckley MJ, Statham AL, Pidsley R, Samaras K, V Lord R, et al. De novo identification of differentially methylated regions in the human genome. Epigenetics and Chromatin. 2015;
188. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. Nat Biotechnol. 2019;
189. Wolf GT, Liu S, Bellile E, Sartor M, Rozek L, Thomas D, et al. Tumor infiltrating

lymphocytes after neoadjuvant IRX-2 immunotherapy in oral squamous cell carcinoma: Interim findings from the INSPIRE trial. Oral Oncol. 2020;
190. Mollenhauer J, Wiemann S, Scheurlen W, Korn B, Hayashi Y, Wilgenbus KK, et al. DMBT1, a new member of the SRCR superfamily, on chromosome 10q25.3-26.1 is deleted in malignant brain tumours. Nat Genet. 1997;
191. Li H, Chiappinelli KB, Guzzetta AA, Easwaran H, Yen RWC, Vatapalli R, et al. Immune regulation by low doses of the DNA methyltransferase inhibitor 5-azacitidine in common human epithelial cancers. Oncotarget. 2014;
192. Liao W, Lin JX, Leonard WJ. IL-2 family cytokines: New insights into the complex roles of IL-2 as a broad regulator of T helper cell differentiation. Curr. Opin. Immunol. 2011.
193. End C, Bikker F, Renner M, Bergmann G, Lyer S, Blaich S, et al. DMBT1 functions as pattern-recognition molecule for poly-sulfated and poly-phosphorylated ligands. Eur J Immunol. 2009;
194. Tuttolomondo M, Casella C, Hansen PL, Polo E, Herda LM, Dawson KA, et al. Human DMBT1-Derived Cell-Penetrating Peptides for Intracellular siRNA Delivery. Mol Ther - Nucleic Acids. 2017;
195. PIAO S, PIAO S, BANERJEE R, LIU M, INGLEHART RC, D'SILVA N. DOWNREGULATION OF DMBT1 PROMOTES INVASION IN HEAD AND NECK CANCER. Oral Surg Oral Med Oral Pathol Oral Radiol. 2017;
196. Singh P, Banerjee R, Piao S, Costa de Medeiros M, Bellile E, Liu M, et al. Squamous cell carcinoma subverts adjacent histologically normal epithelium to promote lateral invasion. J Exp Med. 2021;
197. Ahlmann M, Hempel G. The effect of cyclophosphamide on the immune system: implications for clinical cancer therapy. Cancer Chemother. Pharmacol. 2016.
198. Chen PL, Roh W, Reuben A, Cooper ZA, Spencer CN, Prieto PA, et al. Analysis of immune signatures in longitudinal tumor samples yields insight into biomarkers of response and mechanisms of resistance to immune checkpoint blockade. Cancer Discov. 2016;
199. Bais M V. Impact of Epigenetic Regulation on Head and Neck Squamous Cell Carcinoma. J. Dent. Res. 2019.
200. Outh-Gauer S, Alt M, Le Tourneau C, Augustin J, Broudin C, Gasne C, et al. Immunotherapy in head and neck cancers: A new challenge for immunologists, pathologists and clinicians. Cancer Treat. Rev. 2018.
201. Canning M, Guo G, Yu M, Myint C, Groves MW, Byrd JK, et al. Heterogeneity of the head and neck squamous cell carcinoma immune landscape and its impact on immunotherapy. Front. Cell Dev. Biol. 2019.
202. Wolfer S, Elstner S, Schultze-Mosgau S. Degree of Keratinization Is an Independent Prognostic Factor in Oral Squamous Cell Carcinoma. J Oral Maxillofac Surg. 2018;
203. Wittassek M, Koch HM, Angerer J, Brüning T. Assessing exposure to phthalates - The human biomonitoring approach. Mol. Nutr. Food Res. 2011.
204. Wang T, Pehrsson EC, Purushotham D, Li D, Zhuo X, Zhang B, et al. The NIEHS TaRGET II Consortium and environmental epigenomics. Nat. Biotechnol. 2018.
205. Navarro SJ, Trinh T, Lucas CA, Ross AJ, Waymire KG, MacGregor GR. The C57BL/6J mouse strain background modifies the effect of a mutation in Bcl2l2. G3 Genes, Genomes, Genet. Genetics Society of America; 2012;2:99–102.
206. Waterland RA, Jirtle RL. Transposable Elements: Targets for Early Nutritional Effects on

Epigenetic Gene Regulation. Mol Cell Biol. 2003;
207. Weinhouse C, Anderson OS, Bergin IL, Vandenbergh DJ, Gyekis JP, Dingman MA, et al. Dose-dependent incidence of hepatic tumors in adult mice following perinatal exposure to bisphenol A. Environ Health Perspect. 2014;
208. Schmidt J-S, Schaedlich K, Fiandanese N, Pocar P, Fischer B. Effects of di(2-ethylhexyl) phthalate (DEHP) on female fertility and adipogenesis in C3H/N mice. Environ Health Perspect. 2012;120:1123–9.
209. Lin Y, Wei J, Li Y, Chen J, Zhou Z, Song L, et al. Developmental exposure to di(2-ethylhexyl) phthalate impairs endocrine pancreas and leads to long-term adverse effects on glucose homeostasis in the rat. Am J Physiol Metab. 2011;301:E527–38.
210. Neier K, Cheatham D, Bedrosian LD, Dolinoy DC. Perinatal exposures to phthalates and phthalate mixtures result in sex-specific effects on body weight, organ weights and intracisternal A-particle (IAP) DNA methylation in weanling mice. J Dev Orig Health Dis. 2018;1–12.
211. Lorber M, Calafat AM. Dose reconstruction of di(2-ethylhexyl) phthalate using a simple pharmacokinetic model. Environ Health Perspect. 2012;120:1705–10.
212. Wittassek M, Angerer J, Kolossa-Gehring M, Schäfer SD, Klockenbusch W, Dobler L, et al. Fetal exposure to phthalates--a pilot study. Int J Hyg Environ Health. 2009;212:492–8.
213. Wittassek M, Angerer J. Phthalates: metabolism and exposure. Int J Androl. 2008;31:131–8.
214. Chang J-W, Lee C-C, Pan W-H, Chou W-C, Huang H-B, Chiang H-C, et al. Estimated Daily Intake and Cumulative Risk Assessment of Phthalates in the General Taiwanese after the 2011 DEHP Food Scandal. Sci Rep. 2017;7:45009.
215. Silva MJ, Reidy JA, Herbert AR, Preau JL, Needham LL, Calafat AM. Detection of phthalate metabolites in human amniotic fluid. Bull Environ Contam Toxicol. 2004;72:1226–31.
216. Huang P-C, Tsai C-H, Liang W-Y, Li S-S, Huang H-B, Kuo P-L. Early Phthalates Exposure in Pregnant Women Is Associated with Alteration of Thyroid Hormones. Gonzalez-Bulnes A, editor. PLoS One. 2016;11:e0159398.
217. HUANG P, KUO P, CHOU Y, LIN S, LEE C. Association between prenatal exposure to phthalates and the health of newborns☆. Environ Int. 2009;35:14–20.
218. Calafat AM, Brock JW, Silva MJ, Gray LE, Reidy JA, Barr DB, et al. Urinary and amniotic fluid levels of phthalate monoesters in rats after the oral administration of di(2-ethylhexyl) phthalate and di-n-butyl phthalate. Toxicology. 2006;217:22–30.
219. Krueger F, Andrews SR. Bismark: A flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics. 2011;
220. Park Y, Figueroa ME, Rozek LS, Sartor MA. MethylSig: A whole genome DNA methylation analysis pipeline. Bioinformatics. 2014;
221. Park Y, Wu H. Differential methylation analysis for BS-seq data under general experimental design. Bioinformatics. 2016;
222. Davis AP, Grondin CJ, Johnson RJ, Sciaky D, Wiegers J, Wiegers TC, et al. Comparative Toxicogenomics Database (CTD): update 2021. Nucleic Acids Res. 2020;
223. Williamson C, Blake A, Thomas S, Beechey C V, Hancock J, Cattanach BM, et al. Mouse Imprinting Data and References. MRC Harwell, Oxfordsh. 2010.
224. Tucci V, Isles AR, Kelsey G, Ferguson-Smith AC, Bartolomei MS, Benvenisty N, et al. Genomic Imprinting and Physiological Processes in Mammals. Cell. 2019.

225. Risso D, Ngai J, Speed TP, Dudoit S. Normalization of RNA-seq data using factor analysis of control genes or samples. Nat Biotechnol. 2014;
226. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;
227. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;
228. Buckley JP, Doherty BT, Keil AP, Engel SM. Statistical approaches for estimating sex-specific effects in endocrine disruptors research. Environ Health Perspect. 2017;
229. Kang ER, Iqbal K, Tran DA, Rivas GE, Singh P, Pfeifer GP, et al. Effects of endocrine disruptors on imprinted gene expression in the mouse embryo. Epigenetics. 2011;
230. Svoboda LK, Neier K, Wang K, Cavalcante RG, Rygiel CA, Tsai Z, et al. Tissue and sex-specific programming of DNA methylation by perinatal lead exposure: implications for environmental epigenetics studies. Epigenetics. 2020;
231. Tindula G, Murphy SK, Grenier C, Huang Z, Huen K, Escudero-Fung M, et al. DNA methylation of imprinted genes in Mexican-American newborn children with prenatal phthalate exposure. Epigenomics. 2018;
232. Grindler NM, Vanderlinden L, Karthikraj R, Kannan K, Teal S, Polotsky AJ, et al. Exposure to Phthalate, an Endocrine Disrupting Chemical, Alters the First Trimester Placental Methylome and Transcriptome in Women. Sci Rep. Nature Publishing Group; 2018;8:1–9.
233. Kawano M, Qin XY, Yoshida M, Fukuda T, Nansai H, Hayashi Y, et al. Peroxisome proliferator-activated receptor α mediates di-(2-ethylhexyl) phthalate transgenerational repression of ovarian Esr1 expression in female mice. Toxicol Lett. 2014;
234. Zhou C, Flaws JA. Effects of an environmentally relevant phthalate mixture on cultured mouse antral follicles. Toxicol Sci. 2017;
235. Julien B, Pinteur C, Vega N, Labaronne E, Vidal H, Naville D, et al. Evidence for estrogeno-mimetic effects of a mixture of low-dose pollutants in a model of ovariectomized mice. Environ Toxicol Pharmacol. 2018;
236. Naville D, Pinteur C, Vega N, Menade Y, Vigier M, Le Bourdais A, et al. Low-dose food contaminants trigger sex-specific, hepatic metabolic changes in the progeny of obese mice. FASEB J. 2013;
237. Weinstein LS, Xie T, Qasem A, Wang J, Chen M. The role of GNAS and other imprinted genes in the development of obesity. Int. J. Obes. 2010.
238. Kochmanski JJ, Marchlewicz EH, Cavalcante RG, Perera BPU, Sartor MA, Dolinoy DC. Longitudinal effects of developmental bisphenol a exposure on epigenome-wide DNA hydroxymethylation at imprinted loci in mouse blood. Environ Health Perspect. 2018;
239. Wang L, Chang S, Wang Z, Wang S, Huo J, Ding G, et al. Altered GNAS imprinting due to folic acid deficiency contributes to poor embryo development and may lead to neural tube defects. Oncotarget. 2017;
240. Van Der Plaat DA, Vonk JM, Terzikhan N, De Jong K, De Vries M, La Bastide-Van Gemert S, et al. Occupational exposure to gases/fumes and mineral dust affect DNA methylation levels of genes regulating expression. Hum Mol Genet. 2019;
241. Vangeel EB, Izzi B, Hompes T, Vansteelandt K, Lambrechts D, Freson K, et al. DNA methylation in imprinted genes IGF2 and GNASXL is associated with prenatal maternal stress. Genes, Brain Behav. 2015;
242. Messerschmidt DM, Knowles BB, Solter D. DNA methylation dynamics during

epigenetic reprogramming in the germline and preimplantation embryos. Genes Dev. 2014.
243. Engel A, Buhrke T, Imber F, Jessel S, Seidel A, Völkel W, et al. Agonistic and antagonistic effects of phthalates and their urinary metabolites on the steroid hormone receptors ERα, ERβ, and AR. Toxicol Lett. 2017;
244. Eveillard A, Lasserre F, de Tayrac M, Polizzi A, Claus S, Canlet C, et al. Identification of potential mechanisms of toxicity after di-(2-ethylhexyl)-phthalate (DEHP) adult exposure in the liver using a systems biology approach. Toxicol Appl Pharmacol. Academic Press; 2009;236:282–92.
245. Gravina S, Dong X, Yu B, Vijg J. Single-cell genome-wide bisulfite sequencing uncovers extensive heterogeneity in the mouse liver methylome. Genome Biol. 2016;
246. Petkovich DA, Podolskiy DI, Lobanov A V., Lee SG, Miller RA, Gladyshev VN. Using DNA Methylation Profiling to Evaluate Biological Age and Longevity Interventions. Cell Metab. 2017;
247. Bell CG, Lowe R, Adams PD, Baccarelli AA, Beck S, Bell JT, et al. DNA methylation aging clocks: Challenges and recommendations. Genome Biol. 2019.
248. Kochmanski J, Montrose L, Goodrich JM, Dolinoy DC. Environmental deflection: The impact of toxicant exposures on the aging epigenome. Toxicol. Sci. 2017.
249. Park C, Lee J, Kong B, Park J, Song H, Choi KO, et al. The effects of bisphenol A, benzyl butyl phthalate, and di(2-ethylhexyl) phthalate on estrogen receptor alpha in estrogen receptor-positive cells under hypoxia. Environ Pollut. 2019;
250. Oral D, Erkekoglu P, Kocer-Gumusel B, Chao MW. Epithelial-mesenchymal transition: A special focus on phthalates and bisphenol A. J Environ Pathol Toxicol Oncol. 2016;
251. Treviño LS, Dong J, Kaushal A, Katz TA, Jangid RK, Robertson MJ, et al. Epigenome environment interactions accelerate epigenomic aging and unlock metabolically restricted epigenetic reprogramming in adulthood. Nat Commun. 2020;
252. Prados J, Stenz L, Somm E, Stouder C, Dayer A, Paoloni-Giacobino A. Prenatal exposure to DEHP affects spermatogenesis and sperm DNA methylation in a strain-dependent manner. PLoS One. 2015;
253. Schaedlich K, Schmidt JS, Kwong WY, Sinclair KD, Kurz R, Jahnke HG, et al. Impact of di-ethylhexylphthalate exposure on metabolic programming in P19 ECC-derived cardiomyocytes. J Appl Toxicol. 2015;
254. Labaronne E, Pinteur C, Vega N, Pesenti S, Julien B, Meugnier-Fouilloux E, et al. Low-dose pollutant mixture triggers metabolic disturbances in female mice leading to common and specific features as compared to a high-fat diet. J Nutr Biochem. 2017;
255. Campanella G, Gunter MJ, Polidoro S, Krogh V, Palli D, Panico S, et al. Epigenome-wide association study of adiposity and future risk of obesity-related diseases. Int J Obes. 2018;
256. Perera BPU, Svoboda LK, Dolinoy DC. Genomic tools for environmental epigenetics and implications for public health. Curr. Opin. Toxicol. 2019.
257. Ren H, Aleksunes LM, Wood C, Vallanat B, George MH, Klaassen CD, et al. Characterization of peroxisome proliferator-Activated receptor α-Independent effects of PPARα activators in the rodent liver: Di-(2-ethylhexyl) phthalate also activates the constitutive-activated receptor. Toxicol Sci. 2009;
258. Baccarelli A, Bollati V. Epigenetics and environmental chemicals [Internet]. Curr. Opin. Pediatr. NIH Public Access; 2009 [cited 2021 Jun 9]. page 243–51. Available from: /pmc/articles/PMC3035853/

259. Marsit CJ. Influence of environmental exposure on human epigenetic regulation. J. Exp. Biol. 2015.
260. Hardy TM, Tollefsbol TO. Epigenetic diet: Impact on the epigenome and cancer. Epigenomics. 2011.
261. Li L, Gao Y, Wu Q, Cheng ASL, Yip KY. New guidelines for DNA methylome studies regarding 5-hydroxymethylcytosine for understanding transcriptional regulation. Genome Res. 2019;
262. Wang J, Su Y, Tian Y, Ding Y, Wang X. Characterization of DNA hydroxymethylation profile in cervical cancer. Artif Cells, Nanomedicine Biotechnol. 2019;
263. Efimova OA, Koltsova AS, Krapivin MI, Tikhonov A V., Pendina AA. Environmental epigenetics and genome flexibility: Focus on 5-hydroxymethylcytosine. Int. J. Mol. Sci. 2020.
264. Sen A, Heredia N, Senut MC, Hess M, Land S, Qu W, et al. Early life lead exposure causes gender-specific changes in the DNA methylation profile of DNA extracted from dried blood spots. Epigenomics. 2015;
265. Sen A, Cingolani P, Senut MC, Land S, Mercado-Garcia A, Tellez-Rojo MM, et al. Lead exposure induces changes in 5-hydroxymethylcytosine clusters in CpG islands in human embryonic stem cells and umbilical cord blood. Epigenetics. 2015;
266. Chung FFL, Herceg Z. The promises and challenges of toxico-epigenomics: Environmental chemicals and their impacts on the epigenome. Environ. Health Perspect. 2020.
267. Barouki R, Melén E, Herceg Z, Beckers J, Chen J, Karagas M, et al. Epigenetics as a mechanism linking developmental exposures to long-term toxicity. Environ. Int. 2018.
268. Cavalli G, Heard E. Advances in epigenetics link genetics to the environment and disease. Nature. 2019.