Epigenetically silenced class I human leukocyte antigen (HLA-I): a novel biomarker and therapeutic target in prostate cancer

By

Tamara Stephanie Rodems

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Cancer Biology)

at the

UNIVERSITY OF WISCONSIN-MADISON

2020

Date of final oral examination: October 12th, 2020

This dissertation is approved by the following members of the Final Oral Committee:

Joshua Lang, Associate Professor, Department of Medicine
William Sugden, Professor, Department of Oncology
David Beebe, Professor, Department of Biomedical Engineering
David Jarrard, Professor, Department of Urology
Elaine Alarid, Professor, Department of Oncology

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

The work represented by this thesis and my completed graduate studies could not have been done without the support of my family, friends, and colleagues. To Dr. Joshua Lang, thank you for choosing me as your first graduate student and giving me the opportunity to grow as an independent scientist. Your patient-focused philosophy on research and medicine has guided my graduate school research and inspired my future career path. I greatly appreciate your continual enthusiasm for my project and professional development. Thank you, especially, for your support and encouragement while I navigated graduate school as a new parent and all your help in transitioning out of graduate school and into an exciting start to my future career.

To the Lang Lab, past and present, thank you all for creating an awesome environment to do research. Thank you to Dr. Jamie Sperger for your continued support and advice on my graduate school projects and my future career. I appreciate you always being available to bounce experiment ideas off of and all your help and effort in pushing my project forward. Thank you as well to Dr. Erika Heninger for all your contributions to my project and thesis. Your expertise in epigenetics and immunology has made me a better and more knowledgeable scientist and was indispensable to my graduate school work. Thank you as well to all the Lang Lab research specialists for all the help processing samples and working with me to make sure my experiments were able to be completed successfully.

To the Cancer Biology program, thank you for putting students first and creating a wonderful graduate school experience. A special thank you to Dr. Elaine Alarid, Dr. Dan Loeb, Dr. Bill Sugden, and Jenny Schroeder for your support while I changed labs early

on. I will always appreciate the help, advice, and encouragement from each of you while making that transition. Thank you to my committee members for your support of my project and for all your advice on research and my future career. Thank you as well to all the Cancer Biology students during my time in graduate school. I will always be grateful for having a fantastic group of young scientists to experience graduate school with.

To my Madison friends, thank you for making these six years absolutely amazing. To Dalton, Jordan V., Mailee, Rob, Josh, Siera, Ed, Kyle, Jordan B., Halena, Jaye, Khoa, and Betsy, thank you all for being an awesome group of people and for all the support, advice, and fun times. The float trips, game nights, ultimate frisbee seasons, fantasy football leagues, badger game days, weddings, and cancer bio happy hours were absolutely essential parts of my graduate school experience. I will always cherish these six years and all the fun we had. Leaving the Midwest will be bittersweet specifically because of all of you. A special thank you to my cohort, Jordan Vellky and Dalton McLean. I am so grateful that I got to join grad school with two people who would become my lifelong friends. I can't imagine what this experience would have been like without the two of you. Thank you for all the badger tailgates in freezing parking lots or crammed in small apartments. Thank you for upkeeping cancer bio happy hour every Friday with me, leaving lab at 2pm to get a table at The Library so we could de-stress about failed experiments, frustrating meetings, and general grad school life. Thank you for the trips to Indianapolis, Michigan, Green Bay, and Door County. Thank you for your support during difficult times and making graduate school such a fantastic experience. Thank you for your continued support, advice, and excitement for my future. I appreciate you both and cannot wait to see how all three of us grow as scientists and people.

To my friends and family back home in San Diego, thank you for supporting me from afar. To Amanda, Ali, Jaymie, and Christine, thank you for your never ending support and encouragement and always believing in me. Thank you for your lifelong friendship and always being there when I would visit home. I can't wait to be closer to you all again. I appreciate you all always uplifting me and all your positive words of affirmation. It made being away from home easier to know that I had your love and support. Love you guys! To my siblings, Tyler and Bri, thank you for being supportive of me through all of my years of school. I'm so glad to be moving back close to you both again. I appreciate all the love and laughs that you both provide whenever we are together. I love you both. Thank you as well to my grandparents and late Oma, for your love and support in all my endeavors, past, present, and future. I love you all. To my parents, I cannot write a large enough thank you for everything you have done for me not just during grad school, but for my whole life. Your continued love and encouragement are the reason why I was able to get to this place in my life. I will forever appreciate your support, emotionally, financially, and otherwise. Thank you for always coming to visit, even when its -20 outside, and thank you for flying all of us out for every holiday. Visiting you guys and the rest of the family made being away for so long that much easier. Thank you for also being amazing grandparents to Elyse. Mom, I know you didn't want me to go to grad school so far away, but thank you for supporting me anyway! I am so grateful for your help and advice as I have balanced being a mom and a grad student. Dad, thank you for inspiring me to be a better scientist and all your support in starting my career. I appreciate all your advice on how to navigate grad school, a career in science, and life in general. Thank you, Mom and Dad, for everything you've done for me, Andy, and Elyse, we love you.

To Andy and Elyse, thank you for your daily love and support. A special thank you also to our dog, Spike, who passed away three weeks before my defense. Spike, thank you for being there with me for 15 years, through high school, undergrad, and grad school. You were a perpetual source of happiness and love and we will miss you forever, buddy. Elyse, thank you for your constant love and being the light of our lives. I can't imagine our lives without you. Your hugs and kisses after long days in lab were always what I needed to remember what really mattered. Thank you for always being joyful, laughing, and loving, even while I was constantly working! I am so excited to have so much more time and energy to spend with you. I love you, Lyser. Andy, thank you for being an amazing partner, especially through these last few years of navigating grad school and parenthood. I appreciate you always having words of positivity, always pushing me to keep going, and I especially appreciate you always encouraging me to take a step back and relax during stressful times. Your support during grad school and especially these last few months of writing my thesis, applying to jobs, interviewing, moving, and planning for our future, was essential to my success. I will forever appreciate your love and never ending support and can't image doing this with anyone else but you. I am so excited to see what the future holds for us. I love you and thank you for everything.

Epigenetically silenced class I human leukocyte antigen (HLA-I): a novel biomarker and therapeutic target in prostate cancer

Tamara S. Rodems

Under the supervision of Joshua M. Lang, MD, MS

at the University of Wisconsin – Madison

ABSTRACT

Downregulation of class I HLA (HLA-I) impairs immune recognition and surveillance in prostate cancer and is a mechanism of resistance to certain immunotherapies. However, the molecular mechanisms regulating HLA-I loss in prostate cancer have not been fully explored. Epigenetic changes are common in prostate cancer and have been proposed as drivers of prostate cancer progression. Here, we propose that epigenetic mechanisms 1) regulate HLA-I expression in prostate cancer, 2) are targetable by inhibition of epigenetic modifying proteins, and 3) have utility as potential biomarkers in prostate cancer circulating tumor cells (CTCs). We establish this through a comprehensive analysis of HLA-I genomic, epigenomic and gene expression alterations in primary and metastatic human prostate cancer. Genomic alterations were found to be extremely rare in the HLA-I genes in primary and metastatic prostate cancer and were not associated with HLA-I gene expression. Loss of expression of HLA-I genes was associated with increased DNA methylation and histone H3 lysine 27 tri-methylation as well as decreased chromatin accessibility and histone H3 lysine 27 acetylation. We found that epigenetic regulation of the HLA-I genes was targetable by inhibiting DNA methyltransferase (DNMT) and histone deacetylase (HDAC) protein families. DNMT and HDAC inhibition

decreased DNA methylation, increased H3 lysine 27 acetylation, and functionally re-expressed HLA-I on the surface of tumor cells. These results suggest the possibility for therapeutic use of epigenetic modifying agents to upregulate HLA-I on tumor cells to promote tumor clearance. Identifying patients who harbor epigenetically regulated HLA-I would allow for more personalized therapy decisions regarding epigenetic and immunotherapies. Described in this thesis is a method we developed for enrichment of methylated DNA from low-input samples, including CTCs. We validate the ability of this assay to detected HLA-I methylation in CTCs with low HLA-I expression, demonstrating the potential for methylated HLA-I as an epigenetic biomarker in prostate cancer.

ABBREVIATIONS

5AZA2	5-aza-2-deoxycytidine, aka decitabine
5hmC	5-hydorxymethylcytosine
5mC	5-methylcytosine
ac	Acetylation
ADT	Androgen deprivation therapy
APC	Antigen presenting cell
APC (GENE)	Adenomatous polyposis coli
APM	Antigen processing machinery
AR	Androgen receptor
ATAC-seq	Assay for Transposase-Accessible Chromatin using sequencing
AUC	Area under the curve
B2M	Beta-2-microglobulin
BRCA	BReast CAncer gene
CAF	Cancer associated fibroblast
CALR	Calreticulin
Cas9	CRISPR associated protein 9
CD11b	Cluster of differentiation 11b
CD14	Cluster of differentiation 14
CD27	Cluster of differentiation 27
CD34	Cluster of differentiation 34
CD44b	Cluster of differentiation 44b

CD45	Cluster of differentiation 45
cfDNA	Cell free DNA
ChIP	Chromatin immunoprecipitation
ChIP-seq	Chromatin immunoprecipitation followed by sequencing
CK	Cytokeratin
COMPARE-MS	Combination of methylated-DNA precipitation and methylation sensitive restriction enzymes
CpG	Cytosine guanine dinucleotide
CRISPR	Clustered regularly interspaced short palindromic repeats
Ct	Cycle threshold
CTA	cancer testis antigen
CTC	Circulating tumor cell
ctDNA	Circulating tumor DNA
dCas9	Catalytically dead Cas9
DMSO	Dimethyl sulfoxide
DNMT	DNA methyltransferase
EDTA	Ethylenediaminetetraacetic acid
ELR	Exclusive liquid repellency
EpCAM	Epithelial cell adhesion molecule
ER	Endoplasmic reticulum
ERG	ETS-related gene
ESP	Exclusion-based sample preparation
ETS	Erythroblast transformation-specific

EZH2	Enhancer of zeste homolog 2
FDA	Food and Drug Administration
FITC	Fluorescein isothiocyanate
FOLH1	Folate hydrolase 1; see also PSMA
GSTP1	Glutathione S-transferase pi
GTC	Guanidine thiocyanate buffer
H3K27	Histone H3 lysine 27
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
HER2	Human epidermal growth factor receptor 2
HLA	Human leukocyte antigen
HMT	Histone methyltransferase
KDM	Lysine demethylase
KLK2	Kallikrein related peptidase 2
LBH	Small molecule HDAC inhibitor LBH-589, aka Panobinostat
LiDS	Lithium dodecyl sulfate buffer
LINE1	Long interspersed nuclear element 1
LMP7	Low-molecular mass protein-7
LOH	Loss of heterozygosity
MBD	Methyl-CpG-binding domain
MBD-seq	Sequencing of DNA enriched with methyl-CpG-binding domain protein
MBD2	Methyl-CpG-binding domain protein 2

MBD2-MBD	Methyl-CpG-binding domain of methyl-CpG-binding domain protein 2
mCRPC	Metastatic castrate resistant prostate cancer
MCV	Max cycle value
MDSC	Myeloid derived suppressor cells
me3	Tri-methylation
MeCP2	Methyl-CpG-binding protein 2
MFI	Mean fluorescent intensity
MHC	Major histocompatibility complex
MHC-I	Class I major histocompatibility complex
MI	Methylation index
MIC	MHC class I polypeptide-related sequence
MS-PCR	Methylation specific polymerase chain reaction
MYOD1	Myoblast determination protein 1
NK	Natural killer cell
NKX3.1	NK3 homeobox 1
NSCLC	Non-small cell lung cancer
OT	Optimal threshold
PBMC	peripheral blood mononuclear cell
PBS	phosphate buffered saline
PBST	phosphate buffered saline + tween
PCR	polymerase chain reaction
PD-1	Programmed cell death protein 1

PD-L1	Programmed death-ligand 1
PDMS	Polydimethylsiloxane
PE	Phycoerythrin
PMP	Paramagnetic particle
POLR2A	RNA polymerase II subunit A
PRAD	Prostate Adenocarcinoma data set from TCGA
PRC2	Polycomb repressive complex 2
PSA	Prostate specific antigen
PSMA	Prostate specific membrane antigen; see also FOLH1
PSMB8	Proteasome 20S subunit beta 8
qPCR	Quantitative polymerase chain reaction
RARB	Retinoic acid receptor beta
RASSF1	Ras association domain family 1
RIPA	Radioimmunoprecipitation assay buffer
RNA-seq	RNA sequencing
ROC	Receiver operator characteristic
RPB1	Largest subunit of RNA polymerase II
RPL30	Large ribosomal subunit protein EL30
RRBS	Reduced representation bisulfite sequencing
SASCA	Semi-automated single cell aspiration
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEEMLIS	Semi-automated ESP-based methylated DNA enrichment of low-input samples

SEM	Standard error of the mean
SGI-110/SGI	Small molecule DNMT inhibitor, aka Guadecitabine
sgRNA	Small guide RNA
siRNA	Small interfering RNA
TAA	Tumor associated antigen
TAM	Tumor associated macrophage
TAP	Transporter associated with antigen processing
TCGA	The Cancer Genome Atlas
TCR	T cell receptor
TET	Ten-eleven translocation methylcytosine dioxygenase
TET1-CD	Catalytic domain of TET1
TMPRSS2	Transmembrane serine protease 2
TMPRSS2-ERG	Gene resulting from fusion of TMPRSS2 and ERG
Treg	Regulatory T cell
TSS	Transcription start site
ULBP	UL16 binding protein family
WBC	White blood cell
WGA	Whole genome amplification

LIST OF TABLES

Chapter 2

Chapter 3

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Appendix B

Appendix C

Chapter 1:

Introduction

Prostate Cancer Overview

Prostate cancer begins as a hormone-dependent disease, arising from cells in the prostate gland (1). Initial treatment for localized prostate cancer may include watchful waiting, removal of the prostate by radical prostatectomy, or radiation (2). The 5-year-survival rate for localized disease is 100%, but this drops to 30.5% once the cancer has metastasized to distant regions (3). Treatment for metastatic prostate cancer usually involves androgen deprivation therapy (ADT) because the growth of prostate cancer is initially dependent on androgen receptor (AR) signaling (1,2). The majority of men on ADT will eventually progress to metastatic castration-resistant prostate cancer (mCRPC) where treatment options are limited (4). Understanding the mechanisms that contribute to the transition of localized prostate cancer to metastatic disease is critical to providing better patient care and improving patient outcomes.

Immune Evasion in Cancer

In order for a tumor cell to travel from the primary tumor site to a site of metastasis, many molecular and microenvironmental changes need to occur including loss of cell-cell adhesion, resistance to cell cycle checkpoints, and changes to the microenvironment of metastatic seeding locations (5). Additionally, a tumor cell must survive encounters from various immune cells, whose function is to seek and destroy diseased cells (5,6). This avoidance of destruction by the immune system is termed immune escape or immune evasion. Multiple mechanisms of immune evasion have been proposed. These mechanisms involve many different cell types and cell signaling pathways stemming from both the tumor cell itself and cells within the tumor microenvironment (6).

The presence of various cell types within the tumor microenvironment can directly affect the ability of the immune system to find and eliminate tumor cells. Myeloid derived suppressor cells (MDSCs) are multifunctional when it comes to tumor immune evasion. MDSCs are involved in the formation of cancer associated fibroblasts (CAFs), the recruitment of T regulatory cells (T_{regs}), and their own differentiation into tumor associated macrophages (TAMs) (7). CAFs secrete multiple cytokines that modulate the composition of immune cells present in the microenvironment and induce changes to the structure of the microenvironment to physically block immune infiltration (8). Increased T_{regs} in the microenvironment can support survival of a tumor by promoting self-tolerance and effectively shutting down T-cell response to the tumor (9). TAMs support immune evasion through many mechanisms including suppression of CD8+ T-cell activation and recruitment of other immunosuppressive cell types (10).

Mechanisms of immune evasion that are intrinsic to tumor cells commonly involve ways of hiding from cells that are involved in both innate and adaptive immunity (6). Some of these mechanisms are highlighted in Figure 1.1. Innate immune responses by natural killer (NK) cells are subverted by tumor cells through downregulation or proteolytic cleavage of certain proteins that bind to NK cell receptors (11,12). These proteins include MICA, MICB, and the UL16-binding protein (ULBP) family (11,13). CD8+ cytotoxic T-cells are the main cell type responsible for adaptive immune responses. Successful recognition of tumor cells by CD8+ T-cells relies on the expression or lack of expression of certain proteins on both the tumor cell and the T-cell (14). The proteins involved in T-cell mediated recognition and lysis of tumor cells are numerous and include members of

antigen processing machinery (APM), negative co-stimulatory molecules, tumor associated antigens (TAAs), class I major histocompatibility complex (MHC-I) molecules, and programmed death ligand 1 (PD-L1) (14-16). Changes to the expression levels of these proteins in the tumor cell renders the T-cell response ineffective and leads to tumor cell survival (17,18).

Any or all of these mechanisms may be employed by a tumor cell to avoid destruction by the innate and adaptive immune system. However, the mechanisms involving alteration of CD8+ T-cell interactions with MHC-I are of particular interest because of the central role T-cells play in tumor elimination as well as the recent advances in immunotherapy, many of which rely on MHC-I expression. As a result, understanding these mechanisms and the regulation of MHC-I molecules in tumor cells can lead to improved treatment efficacy.

Class I Major Histocompatibility Complex (MHC-I) and the Class I Human Leukocyte Antigens (HLA-I)

The major histocompatibility complex (MHC) located on chromosome 6 consists of at least 250 protein-encoding and non-encoding genes, the majority of which are involved in regulating cellular immunity (19). Among these genes are the class I human leukocyte antigens. The class I human leukocyte antigens consist of three classical genes, *HLA-A*, *HLA-B*, and *HLA-C*, and three non-classical genes, *HLA-E*, *HLA-G*, and *HLA-F*. These genes code for proteins of the same names that serve important roles in the immune system. HLA-E, HLA-F, and HLA-G are each involved in regulation of NK cells through independent mechanisms (19). HLA-E and HLA-G also regulate certain subsets of T-cells

and HLA-G is able to inhibit B-cell proliferation (19). HLA-A, HLA-B, and HLA-C, referred to from here on as HLA-I, are by far the most researched and well understood of the human leukocyte antigens and are involved in regulation of CD8+ T-cells. HLA-I proteins are components of class I major histocompatibility complex molecules (MHC-I). Any of the three HLA-I proteins can be used in the formation of MHC-I (20,21). HLA-I genes have very similar sequences and the resulting proteins are similar in structure, which allow them to be interchangeable in the formation of MHC-I (22). However, polymorphisms in the genes allow for variations in peptide-binding and expression levels, which greatly increase the ability of the immune system to adapt to fight new infections and tumor development (20,23). MHC-I is expressed at the surface of virtually every nucleated cell in the human body (24). MHC-I expression is critical for T-cell mediated lysis of cells that have been infected with viruses or bacteria as well as tumor cells (25).

An overview of the formation of MHC-I molecules is shown in Figure 1.2. MHC-I molecules are assembled in the endoplasmic reticulum (ER) and trafficked through the Golgi apparatus to the cell surface (24). In the ER, the scaffolding proteins calreticulin, calnexin, and ERp57 aid in the folding and assembly of a heterodimer consisting of HLA-A, B, or C and beta-2-microglobulin (B2M) (24,26-28). In the cytoplasm, the immunoproteasome digests an intracellular protein into peptides that are transported into the ER by a complex called “transporter associated with antigen processing” (TAP) (24,29). The ER chaperone tapasin facilitates the loading of a peptide into the HLA-I and B2M dimer, resulting in a complete MHC-I molecule (24,30). The MHC-I molecule is then transported out of the ER and through the Golgi apparatus to the cell surface where it can interact with T-cells (24).

T-cells bind MHC-I through the T-cell receptor (TCR) and will also interact with co-stimulatory molecules and cell-cell adhesion molecules on the cell surface (25). Naïve T-cells “read” the peptide presented by MHC-I on professional antigen presenting cells (APCs) such as dendritic cells to determine if the peptide is self, non-self, or diseased-self (31). Naïve T-cells can also directly interact with tumor cells that have MHC-I present on the surface, however they are not considered strong activators of T-cells on their own (32). Recognition of a non-self or diseased-self antigen will activate the T-cell. Activated T-cells can then find other cells expressing this antigen in MHC-I complexes (31). Upon recognition, the T-cell will induce apoptosis by releasing perforin and granzyme B into the target cell and activating the caspase pathway (25). Disruption of any point in the MHC-I assembly, trafficking, or T-cell interaction processes, including the downregulation of HLA-I expression, can result in diseased cells surviving and proliferating despite the presence of cytotoxic T-cells (6,33-35).

HLA-I Downregulation in Prostate Cancer

Downregulation of HLA-I proteins has been widely described in multiple cancer types, including prostate, breast, colon, and cervical cancer (36,37). Previous studies have reported downregulation of HLA-I in approximately 70% of primary prostate tumors with complete loss in up to 34% of primary tumors and 80% of metastatic sites (37-39). Loss of expression of HLA-I at the cell surface eliminates cytotoxic T-cell response to tumor cells and renders many immunotherapies ineffective (36-38).

The interaction between T-cells and cancer cells through MHC-I is indispensable to the effectiveness of some immunotherapies, including certain methods of adoptive cell transfer, cancer vaccines, and PD-1/PD-L1 targeted therapies (40-42). When HLA-I is not expressed at the cell surface, T-cells are unable to bind through the TCR to elicit cytotoxic effects. When HLA-I is genetically downregulated such as though loss of heterozygosity (LOH) or mutational events, there is no way to re-express HLA-I apart from gene therapy (40,43). However, HLA-I that is transcriptionally downregulated has the potential to be re-expressed with drugs that target the source of downregulation (40,44). Distinguishing the mode of downregulation can allow the opportunity for patients with transcriptionally downregulated HLA-I to benefit from a treatment that causes HLA-I re-expression in combination with immunotherapy. Furthermore, re-expressing HLA-I even without the addition of immunotherapy may allow a patient's own immune system to be more efficient in tumor clearance.

While the phenomenon of HLA-I downregulation has been well defined at the protein level in prostate cancer, the molecular mechanisms that contribute to HLA-I downregulation remain unknown. Downregulation of protein expression can be accomplished through multiple mechanisms including dysregulation of protein folding and trafficking, binding of transcriptional repressors, silencing of transcriptional activators, use of alternative promoters and enhancers, and epigenetics. Evidence from a growing body of literature suggests epigenetics may be at the heart of many of the mechanisms driving prostate cancer progression, including impaired antigen presentation.

Epigenetics Overview

Epigenetic alterations affect gene transcription without changing the sequence of the genome. Epigenetic alterations include DNA methylation, histone tail methylation, and histone tail acetylation. Other, more rare epigenetic alterations also exist, but are less well understood. The combination of DNA methylation and histone tail modifications allows genes to be reversibly silenced or activated depending on cellular needs both in the context of cellular differentiation and development and in response to environmental stimuli (45).

DNA methylation is the conversion of the DNA base, cytosine, to 5-methylcytosine by adding a methyl group at the fifth position of cytosine's pyrimidine ring (46). This base remains unchanged in its binding to guanine, thus leaving the sequence of the DNA unchanged through DNA replication and cell division. Methylation mainly occurs on cytosines that are located 5' of a guanine, termed a CpG dinucleotide (46). CpG's are often enriched at regulatory sites of genes, including gene promoters, and are referred to as CpG islands. Hypermethylation of CpG islands in the promoter regions of genes can silence gene transcription by preventing the binding of transcription factors and RNA polymerase (46,47). Actively transcribed genes have hypomethylated CpG islands in their promoters and some genes may lack them altogether, relying on other methods of transcriptional regulation to maintain expression levels (48).

DNA methylation is created and maintained in cells by the DNA methyltransferases (DNMTs) (49). DNMT1 is referred to as the maintenance methyltransferase. Its main function is to copy DNA methylation during DNA replication in mitosis to ensure methylation signatures are maintained from parent to daughter cell (49). DNMT3a and

DNMT3b are referred to as the de novo methyltransferases and are involved in adding new methylation to DNA in response to cellular signals (49,50). An additional DNA methyltransferase exists, DNMT3L, that lacks catalytic activity, but may aid in recruitment of the other epigenetic modifying proteins to DNA and is often found complexed with the de novo methyltransferases at sites of methylation (50-52).

Histones are a family of proteins that create chromatin structure in the nucleus through DNA interactions. DNA is wrapped around histone proteins forming nucleosomes (53). Certain histone proteins have tails or cores that can be posttranslationally modified at specific residues (54). These modifications are often referred to as the “histone code” and control histone to histone and histone to DNA interactions (53). Some modifications such as tri-methylation on lysine 27 of histone H3, are transcriptionally repressive because they promote condensation of chromatin by bringing histone complexes closer together, reducing the accessibility of transcriptional machinery in that location (53,55). Other modifications such as acetylation of the same lysine residue, are associated with regions of active transcription and promote open chromatin structure, allowing interaction between transcriptional machinery and DNA (56,57).

Histone tails are modified by several families of enzymes that catalyze the removal or addition of modifications of histone tails and cores. Among these are histone methyltransferases (HMTs), histone acetyltransferases (HATs), lysine demethylases (KDMs), and histone deacetylases (HDACs) (55,58-60). The interactions of all of these protein families and their contributions to chromatin regulation are an area of ongoing study.

Epigenetics in Prostate Cancer

Epigenetics is thought to be a driving force in the progression of prostate cancer (61). Prostate cancers generally do not have a large mutational burden, with a median somatic mutation frequency of 0.7 per megabase, one of the lowest in 21 cancers surveyed from the TCGA database (62). ETS fusions, such as the TMPRSS2-ERG fusion, and AR amplification are the most common genomic alterations in prostate cancer, but are found at frequencies of only 50% (63,64). Alternatively, multiple epigenetic alterations have been identified that occur in more than 90% of men (65,66). As such, the epigenetic landscape in prostate cancer has been the focus of many studies.

The epigenetic landscape in prostate cancer is summarized in Figure 1.3. The majority of prostate cancers have a global decrease in DNA methylation, as measured by methylation of transposable elements in the genome (67-69). Loss of DNA methylation is generally believed to contribute to genome instability and the activation of oncogenes (66,70,71). In contrast, there is an increase in DNA methylation at specific gene promoters, leading to gene silencing (65,66,72,73). Many of these genes are involved in tumor suppression, hormone regulation, DNA damage and cell cycle regulation, apoptosis, and cellular immunity (66,74). One of the most well characterized hypermethylated genes in prostate cancer is glutathione S-transferase pi (*GSTP1*). *GSTP1* methylation has been reported in numerous studies at frequencies ranging between 70% and 100% (65,66,75-78). Other genes like *APC* and *RASSF1*a have also been reported by certain groups to be methylated in more than 90% of samples tested (65,76,79,80). The significant commonality of these DNA methylation changes make them an attractive target for developing prostate cancer specific biomarkers.

The increase in DNA methylation at specific gene promoters is accompanied by changes in expression of the proteins that add and remove methyl groups to DNA. Protein expression of the DNMT family members is increased in prostate cancer tissues compared to normal tissue (81,82). Activity of these proteins is also increased in cell lines and ex vivo prostate tumor models (81,82). Expression of the ten-eleven translocation (TET) family of proteins, which initiate the de-methylation of DNA, have been found to be downregulated in a subset of prostate cancers and have been specifically implicated in the promotion of metastasis (83-86).

Histone modifications in prostate cancer have been largely studied at the global level. Gene specific histone modifications have not been well characterized because of the lack of techniques available that accommodate the low cell numbers obtained from biopsies (87,88). However, global studies of histone modifications have revealed changes in multiple histone modifications that correlate with silencing of tumor suppressor genes and activation of oncogenes (89). Global loss of heterochromatin in prostate cancer driven by a loss of histone tail acetylation has been reported in multiple studies (90). This global loss of heterochromatin results in expression of genes that are otherwise silenced or expressed in low levels in normal tissues, including genes involved in tumor growth and survival. Overexpression of HATs, including p300/CBP, have been reported in prostate cancer, supporting observations of increased histone acetylation (91). Inhibition of p300/CBP has also been proposed as a therapeutic strategy for mCRPC (92,93).

In contrast, levels of histone H3 lysine 27 tri-methylation (H3K27me3) have also been reported to be upregulated in prostate cancer (89). H3K27me3 has been found to

be enriched in the promoters of specific genes, including tumor suppressor genes *RASSF1* and *RARB* (89,94). In addition, the class I HDACs and some HMTs have been found to be upregulated in prostate tissues, supporting the shift towards chromatin-mediated gene silencing (95-99). The HMT that writes H3K27me3, EZH2, is now being investigated as a drug target to reverse the gene silencing caused by increased levels of H3K27me3 (100). These observations taken together point to an important role for epigenetic regulation in prostate cancer.

Epigenetic Regulation of MHC-I in Prostate Cancer

A small, but growing body of evidence has begun to implicate epigenetic mechanisms in the regulation of MHC-I and its processing machinery in prostate cancer. Our lab observed that inhibition of DNMT and HDAC proteins in prostate cancer cell lines and *ex vivo* cultures induces expression of cancer testes antigens, which are important components of effective tumor clearance by CTLs (101). HDAC inhibition was found to induce expression of various MHC-I and APM components in prostate cancer cell lines (39,102). Inhibition of BET bromodomain-containing proteins, which are readers of histone acetylation, lead to increased HLA-I protein expression and immunogenicity *in vivo* (103). However, the direct effect of inhibition of epigenetic modifying proteins on the epigenetic modifications of the HLA-I genes has not been explored. Furthermore, in depth characterization of the epigenetic landscape of the HLA-I genes in prostate cancer has not yet been performed. Understanding the prevalence and function of HLA-I epigenetic alterations can help solidify HLA-I as an important target of epigenetic therapy and identify the patient subsets that may have the most benefit from these therapies.

DNA Methylation as a Biomarker

New biomarkers are needed to aid clinicians in following the biology of patient disease and monitoring response to treatment. Genomic and transcriptomic biomarkers have been explored as biomarkers of disease presence and progression, but are either not present in the majority of patients or ineffective at predicting patient response to therapy. The TMPRSS2-ERG fusion and AR amplification are the most common genomic event in prostate cancer, but are found in only ~50% of men (63,104-106). AR splice variants, such as ARv7, have been reported in as many as 100% of metastatic tissues in various studies, but do not predict response to therapy unless found in circulation, where the detection rate is much lower, between 10-30% (107-109). Therefore, new classes of biomarkers have been an area of intense research. DNA methylation has been the subject of much of this research due to the notable high frequency of certain DNA methylation changes in prostate cancer. In addition to the high frequency of DNA methylation alterations, the inherent plasticity of epigenetic marks also makes this an attractive option for a biomarker of patient response to treatment.

The challenge in biomarker assessment in prostate cancer is the source of tumor material. Primary prostate biopsies are often a one-shot option due to the invasiveness of the procedure and the frequency of radical prostatectomy to treat the disease. Metastatic biopsies are also generally not repeated because prostate cancer most often metastasizes to the bone, making repeated biopsies invasive, painful, and expensive. Instead, liquid biopsies have been proposed as alternatives. Liquid biopsies include blood, lymph, and urine, and are minimally invasive, which allows for repeated and

longitudinal sampling (110,111). Sources for DNA to assess DNA methylation are cell free DNA (cfDNA) and circulating tumor cells (CTCs). However in both of these cases, the amount of tumor specific DNA that can be extracted from a single sample is incredibly low and often contains contaminating non-tumor DNA on the order of 10 to 1000 times the amount of tumor DNA (112,113). In order to use DNA methylation as a biomarker in prostate cancer liquid biopsies, a method must be developed to accommodate low cell inputs in the presence of high background.

Current methods for analyzing gene specific DNA methylation most often rely on bisulfite conversion of DNA, which is incredibly damaging and can lead to a loss of up to 90% of the starting material (114-116). Newer protocols have been developed for single cell technologies, but because of random DNA shearing or degradation during processing and low read coverage for sequencing based methods, retention of the target of interest cannot be guaranteed (117). There exists a need in the field for a better DNA methylation assay that can reliably extract DNA from low-input, heterogenous samples.

Aim of Thesis

Understanding the mechanisms that promote prostate cancer progression is key to improving patient outcomes and developing better therapies. One of these mechanisms is immune evasion, where cells employ various techniques to hide from immune cells. Downregulation of HLA-I has been reported as one of these techniques and is a common occurrence in prostate cancer. However, the molecular mechanisms regulating HLA-I loss in prostate cancer have not been fully explored. Epigenetic alterations in prostate cancer have been reported as molecular drivers of prostate cancer progression. Here, we

propose that epigenetic mechanisms contribute to HLA-I downregulation in prostate cancer. The goal of this thesis work is to explore the epigenetic landscape of HLA-I in prostate cancer, investigate the contribution of epigenetic alterations in the HLA-I genes to loss of HLA-I gene expression, and develop a method for analysis of low-input DNA methylation to evaluate methylated HLA-I as a liquid biomarker in CTCs.

Figure 1.1. Tumor-cell intrinsic mechanisms of immune evasion

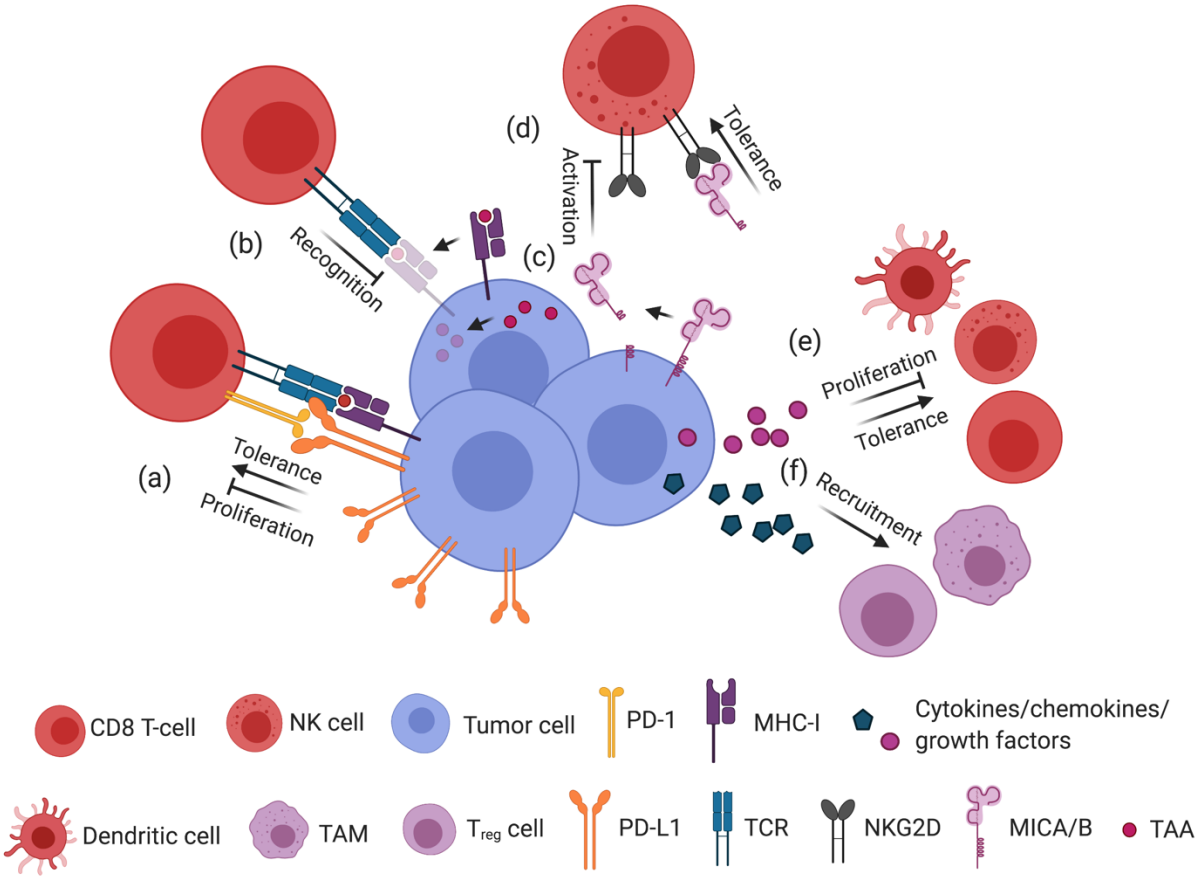


Figure 1.1. Tumor-cell intrinsic mechanisms of immune evasion. PD-L1 overexpression on the surface of tumor cells leads to T-cell tolerance and decreased proliferation (a). Loss or downregulation of MHC-I on tumor cells prevents T-cells from binding and reading antigens (b). Downregulation of TAAs in tumor cells can inhibit T-cell mediated recognition of tumor cells (c). Proteolytic cleavage of MICA/B proteins prevents direct activation of NK cells by the tumor cell, while cleaved proteins that bind to NK cells in the extracellular space promote NK cell tolerance (d). Various cytokines, chemokines, and growth factors secreted by tumor cells are able to prohibit proliferation of NK cells, dendritic cells, and T-cells and induce immune tolerance (e). Cytokines, chemokines, and growth factors secreted by tumor cells are also able to recruit TAMs and T_{reg} cells that support tumor growth and survival (f). Created with BioRender.com.

Figure 1.2. Overview of class I human leukocyte antigen (HLA-I) processing pathway

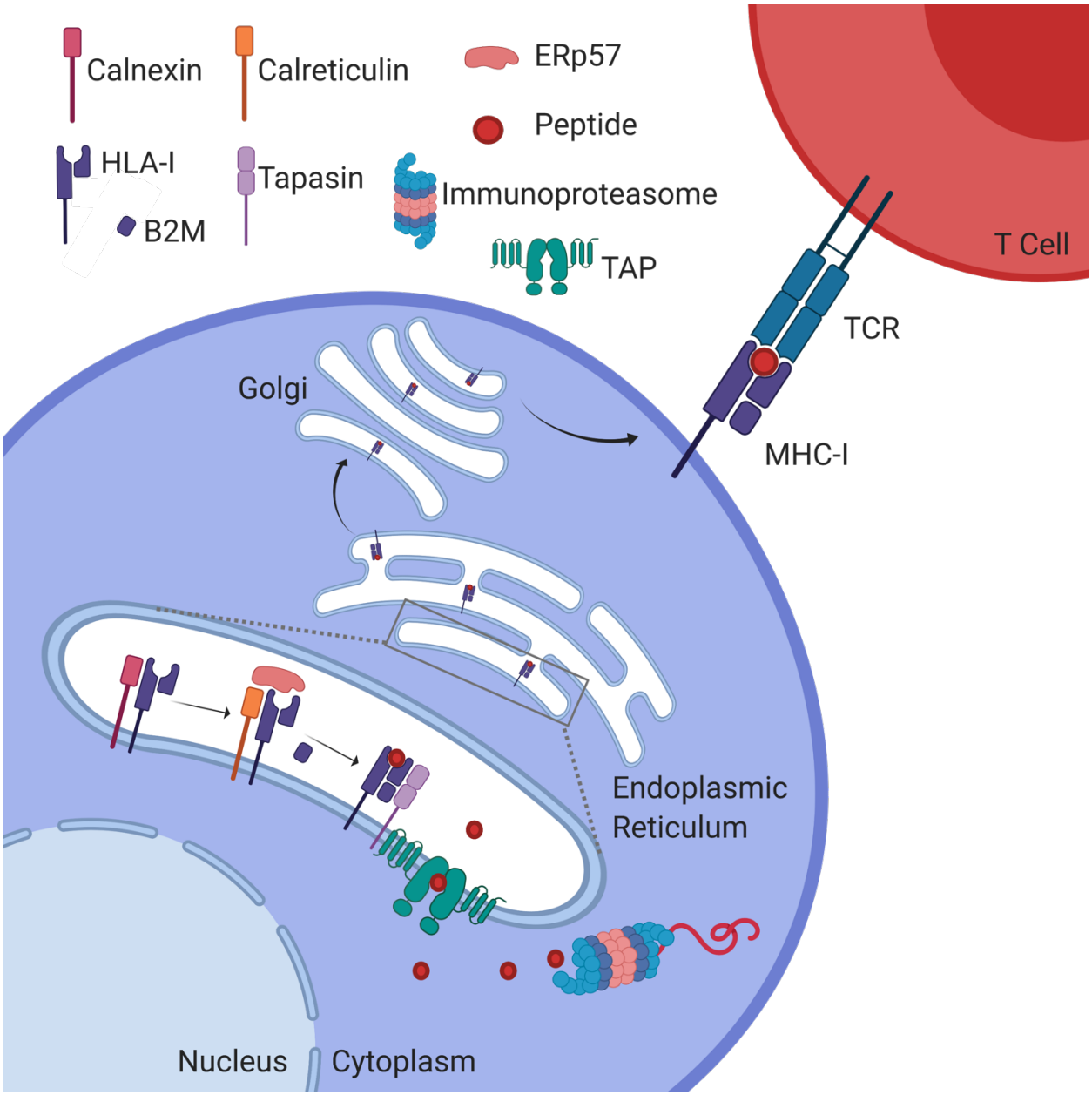


Figure 1.2. Overview of class I human leukocyte antigen (HLA-I) processing pathway. Dimerization of beta-2-microglubulin (B2M) and class I human leukocyte antigen (HLA-I) takes place in the endoplasmic reticulum with the help of chaperone proteins calnexin, calreticulin, and ERp57. Antigens are created by digestion of intercellular proteins by the immunoproteasome. Antigens are then transported into the endoplasmic reticulum by TAP and loaded into HLA-I:B2M complexes stabilized by tapasin. Antigen-loaded MHC-I molecules are trafficked to the cell surface through the Golgi apparatus. T cell receptors (TCR) bind functional MHC-I molecules at the cell surface to facilitate T cell mediated immunological responses. Created with BioRender.com.

Figure 1.3. Overview of the epigenetic landscape in prostate cancer

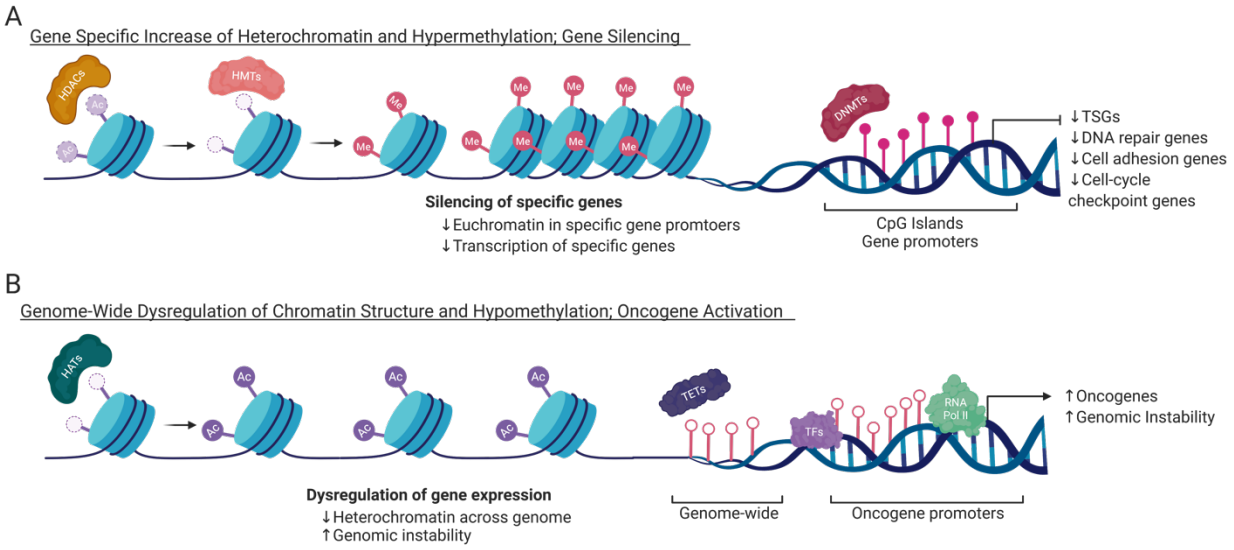


Figure 1.3. Overview of the epigenetic landscape in prostate cancer. A) The removal of histone marks associated with transcriptional activation and addition of marks associated with transcriptional repression takes place at specific gene promoters. Changes to histone signatures can be facilitated by various epigenetic readers, writers, and erasers, including HDACs and HMTs. Alteration of histone modifications is accompanied by hypermethylation, written by DNMTs, at specific gene promoters and CpG islands. The combination of increase in heterochromatin and hypermethylation leads to silencing of genes, including those involved in prevention of tumor growth and progression. B) In contrast, there is a genome-wide dysregulation of chromatin structure that is facilitated by epigenetic modifiers including HATs, which leads to an increase in histone marks associated with euchromatin and gene transcription. In addition, hypomethylation occurs genome-wide and in the promoters of oncogenes. Promoter hypomethylation allows for binding of transcriptional machinery, including TFs and RNA polymerase, leading to an increase in oncogene transcription. Overall, the dysregulation of chromatin structure and genome-wide hypomethylation causes increased genomic instability, which can promote tumor growth and survival. Created with BioRender.com.

Chapter 2:

Targetable epigenetic alterations regulate class I HLA loss in prostate cancer

This chapter is adapted from the following manuscript in preparation: Rodems TS, Heninger E, Stahlfeld CN, Gilsdorf C, Carlson K, Kircher MR, Beebe DJ, McNeel DG, Haffner MC, Lang JM. Targetable epigenetic alterations regulate class I HLA loss in prostate cancer.

Contributions: Figure 2.7B was performed in collaboration with C.G. Figure 2.9 was performed by E.H. Figure 2.10 was performed by E.H. and K.C. in collaboration with D.G.M. All other experiments and analyses were performed by T.S.R.

ABSTRACT

Downregulation of class I HLA (HLA-I) impairs immune recognition and surveillance in prostate cancer and may underlie the ineffectiveness of checkpoint blockade. However, the molecular mechanisms regulating HLA-I loss in prostate cancer have not been fully explored. Here, we employed a comprehensive analysis of HLA-I genomic, epigenomic and gene expression alterations in primary and metastatic human prostate cancer. Loss of HLA-I gene expression was associated with increased promoter and CpG island DNA methylation, enrichment of tri-methylated lysine 27 on histone H3, and reduced chromatin accessibility. DNMT and HDAC inhibition decreased DNA methylation and increased H3 lysine 27 acetylation to re-express HLA-I on the surface of tumor cells. Re-expression of HLA-I on LNCaP cells by DNMT and HDAC inhibition increased activation of co-cultured $PSMA_{27-38}$-specific CD8+ T-cells. These results suggest the possibility for therapeutic use of epigenetic modifying agents to upregulate HLA-I on tumor cells to promote tumor clearance.

INTRODUCTION

Approximately 30,000 men die of metastatic prostate cancer per year in the US and the incidence of men presenting with metastatic disease is rising (3,118). There is a critical need to identify the molecular drivers that contribute to prostate cancer growth and metastasis. Immune evasion is one of the hallmarks of cancer pathogenesis and cancer be therapeutically targeted by immunotherapies that augment T-cell recognition and lysis of tumor cells. (5,119,120). However, display of a functional class I major histocompatibility complex (MHC-I) is required for recognition of tumor cells by cytotoxic T lymphocytes (CTLs). Lack of MHC-I display at the cell surface reduces tumor immunogenicity and drives resistance to immune checkpoint inhibitors (34,43,121,122). MHC-I is a multimeric protein composed of a class I human leukocyte antigen (HLA-I) protein (A, B, or C), beta-2-microglobulin (B2M), and a peptide derived from an intracellular protein (123). Downregulation of the MHC-I components has been proposed as a mechanism of immune evasion in numerous cancer types, including prostate cancer (6,34,37). with downregulation of HLA-I observed in ~70% of primary prostate tumors with complete loss in up to 34% of primary tumors and 80% of metastatic lesions (38,39). The molecular alterations that lead to HLA-I downregulation in prostate cancer remain largely unknown.

Recent findings have pointed to epigenetic mechanisms as drivers of prostate cancer progression (124,125). Overexpression of epigenetic modifying proteins, including the de novo methyltransferases, DNMT3A and DNMT3B, and class I histone deacetylases, HDAC1, HDAC2, and HDAC3, has been implicated in altering epigenetic programs and contributing to metastasis in prostate cancer (81,126). Studies in

esophageal squamous cell carcinoma and gastric cancer found that DNA methylation was a contributing mechanism to HLA-I downregulation, however epigenetic regulation of HLA-I has not yet been explored in prostate cancer (127,128). Inhibition of DNMT and HDAC proteins has been proposed as a therapeutic strategy in prostate cancer, though minimal clinical success has been observed in solid tumors (129). Despite this, there is promising evidence for the usefulness of epigenetic therapies in combination with immunotherapy (130). There remains a need to better understand the interplay between epigenetic and immune functions in cancer cells and for accessible biomarkers for both determining treatment benefit and monitoring treatment response (130). We have previously demonstrated that modulating the activity of DNMT and HDAC proteins in prostate cancer cell lines and *ex vivo* human prostate tissue induced the expression of cancer testis antigens (CTAs), which have been proposed as targets for tumor vaccine therapies (101). Taking advantage of the ability of epigenetic mechanisms to modulate aspects of the immune response, such as HLA-I expression, may improve the efficacy of certain immunotherapies. In this study, we demonstrate that transcriptional downregulation of HLA-I is coordinated by epigenetic silencing mechanisms, which can be reversed to functionally re-express HLA-I *in vivo* and restore HLA-I dependent T-cell activation.

RESULTS

Genomic alterations in HLA-I are rare in human prostate cancer

Previous studies have shown widespread HLA-I downregulation at the protein level in prostate cancer tissues (36-39). We sought to determine if genomic alterations in the HLA-I genes could be contributing to loss of HLA-I expression in prostate cancer. HLA-I copy number alterations and mutations were analyzed in primary and metastatic prostate adenocarcinoma samples in a data set from an analysis performed by Armenia et. al. of 6 independent studies (131). Genomic alterations in HLA-I were very low with these 3 genes individually exhibiting an alteration frequency of 6% or less in primary prostate cancer and 3% or less in metastatic prostate cancer (Figure 2.1A). Only 7/680 (1.0%) of primary adenocarcinoma samples and 6/333 (1.8%) of metastatic adenocarcinoma samples had a deletion or mutation event in any HLA-I gene and gene amplification was only seen in *HLA-C*. The majority of the identified alterations overall were the amplifications that occurred in *HLA-C*, which are not typically associated with downregulation of protein expression. The frequencies of deletion or mutation events had high study to study variation, but there were less than 5 deletions or mutations in any study.

We next analyzed HLA-I genomic gains and losses in relation to HLA-I mRNA expression in the TCGA PanCancer Prostate Adenocarcinoma (PRAD) data set to determine whether these alterations affect gene expression (Figure 2.1B). No significant differences in mRNA expression were observed between any of the groups. Groups with only one patient were not included in the analysis. Shallow deletions can be indicative of loss of heterozygosity (LOH), which has been reported to contribute to HLA-I

downregulation in breast and non-small cell lung cancer (132,133). While the shallow deletion groups did have a slightly lower mean mRNA expression level compared to the diploid group, this did not reach statistical significance and represented a relatively small subset of overall samples, suggesting LOH may not drive loss of HLA-I expression in prostate cancer. Overall, these analyses indicate the vast majority of HLA-I downregulation events in prostate cancer cannot be attributed to changes to the genome itself.

HLA-I gene expression is downregulated in subsets of human prostate cancer

We next analyzed changes in HLA-I gene expression in prostate cancer. We chose to analyze the PRAD data set for its abundance of samples (n=497), however this data set only includes primary site samples. In order to also compare metastatic gene expression to primary tumor gene expression, we also analyzed an additional data set from Taylor et al (134). The gene expression in the PRAD data set was determined by RNA-seq and the gene expression in the Taylor data set was determined by cDNA microarray. Data was converted into z-scores comparing tumor samples to the normal samples from the respective study to be able to compare between the two different experimental systems, however appropriate precaution should be used for any primary:primary comparisons due to the different methods of data acquisition.

We examined both the mean z-score in the population as well as the number and percentage of samples that were significantly up- and downregulated based on a confidence level of 95% (Figure 2.1C, Table 2.1). In the two primary data sets, only *HLA-A* showed noticeable negative shift in mean z-score, indicating a population level

downregulation of gene expression compared to normal samples. However, all three genes had significant negative shifts in mean z-score in the metastatic samples from the Taylor data set. The HLA-I genes were downregulated in 14% or less of samples in the primary data sets, but were downregulated in 58%-63% of the metastatic samples. Significant downregulation of HLA-I was also associated with decreased time to biochemical recurrence (two occurrences of PSA $\geq$ 0.2ng/mL) after radical prostatectomy when compared to all other patients or patients with significantly upregulated HLA-I expression (Figure 2.1D,E). These data show that HLA-I may be transcriptionally downregulated in a small subset of primary prostate cancers and the majority of metastatic prostate cancers and loss of HLA-I gene expression is a risk factor for biochemical recurrence.

Aberrantly expressed DNMT and class I HDAC genes are correlated to HLA-I in human prostate cancer

Previous studies have implicated the DNMTs and class I HDACs as a driving force in prostate cancer biology. We hypothesized that the DNMTs and class I HDACs could promote immune evasion by downregulation of HLA-I as well. We analyzed the expression of *DNMT1*, *DNMT3A*, *DNMT3B*, and the class I HDACs *HDAC1*, *HDAC2*, *HDAC3*, and *HDAC8* in the PRAD and Taylor data sets and their correlation to HLA-I expression in prostate cancer. Table 2.1 summarizes the number and percentage of patients in each study showing up- and downregulation of each of these genes. The *de novo* DNMTs and class I HDAC genes tended to be upregulated more often than downregulated with the exception of *HDAC8*, which was upregulated in 25% the samples

in the PRAD data set, but only in 1 sample from the Taylor data set. An analysis of the mean z-score in each data set revealed that *DNMT3A*, *DNMT3B* and *HDAC1/2/3* tended to have positively shifted z-scores indicating an increase in mean gene expression in the tumor populations (Figure 2.2A).

To investigate how the expression of these various genes associate with HLA-I expression, we calculated Pearson r values for the correlation between HLA-I and DNMT and HDAC genes in each data set (Figure 2.2B). Correlation patterns varied among the data sets, however we found that *DNMT3B* was highly negatively correlated to HLA-I gene expression in all three data sets and *HDAC2* was the most negatively correlated to HLA-I gene expression in the metastatic samples (Figure 2.2C). *DNMT3A* was also negatively correlated to HLA-I gene expression in the PRAD data set. *HDAC3* and *HDAC8* gene expression was positively correlated to HLA-I gene expression in the Taylor data sets, though showed weak to no correlation in the PRAD data set. Overall, the correlation of *DNMT3B* and *HDAC2* to HLA-I gene expression along with overexpression of these gene families, suggests a key role for epigenetic modification of DNA and histones in HLA-I downregulation, among other epigenetic pathways.

HLA-I CpG islands are methylated in primary prostate cancer

We next sought to investigate the DNA methylation signatures of the HLA-I genes in primary and metastatic prostate cancer biopsies. We analyzed the level of methylation at probes from the Illumina 450K methylation array within the HLA-I CpG islands in normal prostate tissue vs. prostate adenocarcinoma samples in the PRAD data set (Figure 2.3A). Every probe located within the *HLA-A* genomic region showed higher methylation levels

in prostate tumor samples compared to normal. There were also multiple probes located within *HLA-B* and *HLA-C* with significantly higher levels of methylation compared to normal. Notably, the probe located within 50bp leading up to the transcription start site for both *HLA-A* and *HLA-C* had the most significant increase in methylation in prostate tumors compared to normal samples (Figure 2.3B). *HLA-B* had significant methylation in the tumor samples at three sites within the gene body and within the promoter, 500bp upstream of the transcription start site.

We then explored whether the patients with increased levels of methylation in HLA-I had corresponding decreases in HLA-I gene expression. Correlations between matched patient gene expression from RNA-seq data and the methylation score at each probe were calculated (Table 2.2). Significant negative correlations were found in 12/12 HLA-A probes, 11/19 *HLA-B* probes, and 11/15 *HLA-C* probes in the tumor samples, and only 2/12, 5/19 and 3/15 for *HLA-A*, *HLA-B* and *HLA-C* normal samples respectively. There were also two probes, probe 1 and 3, in the *HLA-C* tumor samples that showed significant positive correlation between methylation and gene expression, which could indicate the presence of important regulatory elements in the *HLA-C* distal promoter and further upstream.

While many of these correlations have reached statistical significance, the Pearson r values are relatively weak, which may be due the smaller subset of patients in these cohorts that have significantly reduced HLA-I gene expression, as summarized in Table 2.1. To further investigate the association between gene expression and methylation in HLA-I genes, we separated the samples into three groups: high, medium, and low expression. High and low expression were defined by having a significant z-score using

a 95% confidence level relative to normal samples, with all non-significant samples placed into the medium category. We then compared methylation levels in the stratified samples to normal samples (Figure 2.3C). Methylation tended to be higher in the low HLA-I expression samples. Samples expressing medium and high levels of HLA-I were either not significantly different in methylation level, or high expressing samples had lower methylation levels than the medium expressing samples. We chose one probe from the promoter region of each gene to examine further and evaluate the statistical significance of the stratifications (Figure 2.3D). For each of these probes, the samples expressing low levels of HLA-I had significantly more methylation than the samples expressing medium or high levels of HLA-I. In *HLA-B* and *HLA-C*, the samples expressing high levels of HLA-I also had significantly lower methylation levels than the samples expressing medium levels of HLA-I. This analysis suggests a significant role for DNA methylation in HLA-I transcriptional downregulation in patients with prostate cancer.

Decreased chromatin accessibility is associated with HLA-I downregulation in primary prostate cancer

Studies on histone modifications in prostate cancer have been conducted largely at the global level with studies typically measuring overall levels of histone modification abundance (89,135). HLA-I specific histone modification signatures have not been explored in patients with prostate cancer. However, we can make inferences about epigenetic regulation of HLA-I in prostate cancer using ATAC-seq data from the TCGA PanCancer study. We analyzed the ATAC-seq signals in available primary prostate samples from the PRAD data set at two genomic locations: a proximal enhancer region

and the promoter. We then compared the changes in chromatin accessibility at these regions to matched HLA-I gene expression. While the sample size was small, those with high HLA-I gene expression tended to have higher scores for chromatin accessibility, especially in the promoter regions (Figure 2.4A). However, there were a few samples with low HLA-I gene expression and high chromatin accessibility. *HLA-A* gene expression was significantly positively correlated with ATAC-seq signal at the proximal enhancer region and *HLA-C* gene expression was significantly positively correlated with ATAC-seq signal at the promoter (Figure 2.4B). All other ATAC-seq signals were also positively correlated with HLA-I gene expression, but did not reach significance, likely due to the small sample size. This preliminary analysis suggests that HLA-I downregulation is associated with a decrease in accessible chromatin in key regulatory regions.

DNMT and HDAC inhibition induces HLA-I expression *in vitro*

The strong negative correlations between DNMT and HDAC family members and HLA-I expression in patient samples suggests that we may be able to re-express HLA-I by targeting the DNMT and HDAC families therapeutically. To investigate this, we utilized the prostate cancer cell lines LNCaP, PC3, 22rv1, and LAPC4 as well as a benign prostate cell line, RWPE1, as *in vitro* models. We first analyzed baseline HLA-I, DNMT, and HDAC expression in these cell lines. HLA-I protein expression was downregulated in LNCaP, 22rv1, PC3, and LAPC4 when compared to RWPE1 both at the cell surface as measured by immunofluorescence microscopy and in whole cell lysates visualized by western blot (Figure 2.5A). HLA-I gene expression was also significantly downregulated in these four prostate cancer cell lines when compared to RWPE1 gene expression levels

(Figure 2.5B). DNMT expression was measured by immunofluorescent microscopy in each cell line (Figure 2.5C). Increases in DNMT1, DNMT3a, and DNMT3b were seen in LNCaP, 22rv1, and LAPC4 cells compared to RWPE1. PC3 cells had increased levels of DNMT3a, but not DNMT 1 or DNMT3b. HDAC expression was measured by western blot (Figure 2.5D). Each cell line had increased protein expression of at least one HDAC family member. In contrast to what was observed in patient samples, DNMT and HDAC gene expression was not upregulated in the cancer cell lines compared to RWPE1 (Figure 2.5E). However, our findings are line with previous studies, which showed that while DNMT and HDAC gene expression is unchanged or downregulated in prostate cancer cell lines compared to RWPE1 cells, protein expression and activity is significantly upregulated (81,82,136).

We next pharmacologically inhibited DNMT and HDAC activity in the prostate cancer cell lines and RWPE1. We treated the cell lines with SGI-110 (SGI) to inhibit DNMTs or LBH-589 (LBH) to inhibit the class I HDACs either alone or in combination and measured HLA-I gene expression in response to the drugs. We found that the cancer cell lines all responded to at least one treatment condition (Figure 2.6). The combination treatment was the most effective to restore HLA-I expression in all cell lines tested. The cell lines had varying responses to the treatments when used as single agents. LNCaP and LAPC4 cells were responsive to both SGI treatment alone and LBH treatment alone. PC3 cells were also more sensitive to SGI treatment alone compared to LBH treatment alone. 22rv1 cells showed an opposite effect and were more responsive to LBH treatment alone compared to SGI treatment alone. Similar responses were also found in cell lines treated with a different DNMT inhibitor, 5-aza-2-deoxycytidine (5AZA2), alone and in

combination with LBH (Appendix A). Overall, these results support our hypothesis that loss of HLA-I expression is regulated by epigenetic mechanisms in prostate cancer cells.

HLA-I downregulation is associated with increased DNA methylation and H3K27 tri-methylation and loss of H3K27 acetylation prostate cancer cell lines

HLA-I induction by inhibition of epigenetic modifying proteins strongly suggests HLA-I is epigenetically silenced in prostate cancer cell lines. To confirm this, we analyzed epigenetic signatures in LNCaP, 22rv1, PC3, and LAPC4 compared to RWPE1. We designed a set of primers to target regions of the HLA-I genes that were identified as being differentially methylated in patient samples, including the distal and proximal promoter as well as two intragenic regions (IG) near Exon1/Intron1 (IG 1) and Exon2/Intron2 (IG 2) (Figure 2.7A). Methylation of HLA-I was evaluated using MBD2-based enrichment of methylated DNA followed by qPCR (Figure 2.7B). LAPC4 had the highest level of methylation in all three HLA-I genes. PC3 and LNCaP also had increased methylation in certain gene regions compared to RWPE1. 22rv1 had the lowest overall methylation of the cancer cell lines. Methylation in IG2 tended to be the most enriched in cancer cell lines and overall, methylation in the cancer cell lines was increased in at least 2 of the evaluated gene regions for each gene (Figure 2.7C). Overall, the methylation landscape in these four prostate cancer cell lines is comparable to the signatures found in patient samples.

Since HLA-I gene expression was associated with a less accessible chromatin state (Figure 2.4), we hypothesized that repressive histone signatures may also be present in the HLA-I genes in prostate cancer cell lines. Chromatin immunoprecipitation

(ChIP) was performed using antibodies targeting acetylated (H3K27ac), a marker of active transcription, and tri-methylated (H3K27me3) lysine 27 on histone H3, a marker of transcriptional repression. Primers were designed to locations near a peak in H3K27ac signature in GM12878, a lymphoblastoid cell line, determined by ChIP-seq from the ENCODE consortium (137) (Figure 2.7A). Overall, the prostate cancer cell lines showed significant decreases in H3K27ac and increases in H3K27me3 when compared to RWPE1 (Figure 2.7D). The cell lines 22rv1 and LAPC4 had particularly low levels of lysine 27 acetylation in HLA-I. LAPC4 and 22rv1 also had the highest overall level of tri-methylation. This strong repressive signature in 22rv1 may explain why HLA-I expression is so low in these cells, even though DNA methylation was not very high, suggesting that individual tumors may regulate HLA-I by different epigenetic mechanisms.

We generated correlation matrices for HLA-I protein expression, gene expression, methylation, and histone modifications to examine the relationships between each of these measures in the prostate cell lines (Figure 2.7E). We found that protein and gene expression were highly positively correlated in each gene. Methylation within each region tended to be positively correlated with methylation in other regions and with H3K27 tri-methylation, and negatively correlated with H3K27 acetylation, indicating co-occurrence of these epigenetic signatures. The pattern of correlation between methylation signatures in each gene region and HLA-I expression was heterogeneous. We found that methylation within the IG 1 and IG 2 region of *HLA-A* and *HLA-C* was negatively correlated with corresponding gene and protein expression. *HLA-B* IG 2 methylation was also negatively correlated with HLA-B gene and protein expression. The strongest correlations were found in the histone modification group. H3K27 acetylation was strongly

positively correlated with gene and protein expression and H3K27 tri-methylation was strongly negatively correlated with gene and protein expression for all three HLA-I genes. This analysis demonstrates that strong repressive epigenetic signatures are enriched at the HLA-I genes and correlated to HLA-I gene expression in prostate cancer cell lines.

DNMT and HDAC inhibition reverses repressive epigenetic signatures *in vitro*

To confirm that the changes in gene and protein expression in response to DNMT and HDAC inhibitors are accompanied by epigenetic changes in the genes themselves, we measured DNA methylation and histone modification changes in response to SGI, 5AZA2, and LBH. HLA-I methylation was reduced across the HLA-I genes in response to SGI and 5AZA2 (Figure 2.8A, Appendix A), regardless of whether gene expression was significantly induced from SGI treatment alone, suggesting DNA methylation loss may not be sufficient in all cases to re-express HLA-I. This may be due to histone modifications not changing from a repressive state even when DNA methylation has been removed. This is supported by the retention of H3K27me3 and H3K27ac levels in 22rv1 cells when treated with SGI (Figure 2.8B). This can also explain why cell lines tended to show the strongest response to combination treatments. To analyze changes in histone acetylation in response to HDAC inhibition by LBH, we performed ChIP analysis to look at H3K27ac in treated cell lines. H3K27ac was significantly increased in response to LBH in 22rv1 cells and a similar trend was observed in LNCaP and LAPC4 cells, though the results were not statistically significant (Figure 2.8C).

Enrichment of RNA polymerase II (RPB1) at the HLA-I gene promoters was increased in LNCaP cells in response to SGI treatment, showing that the increase in gene

expression is associated with increased transcriptional activity (Figure 2.8D). RBP1 enrichment at the HLA-I promoters was also increased in 22rv1 cells in response to LBH (Figure 2.8E). However, RPB1 enrichment not increased in LNCaP cells in response to LBH treatment. This difference in LBH induced RPB1 binding between LBH treated 22rv1 and LNCaP cells is corroborated by the differences in inducibility of HLA-I gene expression in response to LBH in these two cell lines, suggesting the epigenetic signatures are driving gene expression. While responses to DNMT and HDAC inhibition were varied in prostate cancer cell lines, the changes to the epigenetic landscape that accompany gene and protein induction suggest modulation of epigenetic proteins in prostate cancer may be useful to re-express epigenetically silenced HLA-I in patients.

DNMT and HDAC inhibition induces HLA-I gene expression *ex vivo*

The *in vitro* data suggest that DNMT and HDAC inhibition can significantly alter the expression of HLA-I. We next wanted to test whether this holds true in a more physiologically relevant system. Primary prostate tumor tissue was acquired from radical prostatectomy specimens and grown in an *ex vivo* culture. Expression of the HLA-I genes and three prostate specific genes, *AR*, *KLK3* (PSA), and *ACP3* (PAP), was measured to establish baseline HLA-I levels and confirm the presence of prostate cells in the culture (Figure 2.9A). Tissue in *ex vivo* culture was treated with DMSO, 5AZA2, or LBH either alone or in combination and induction of HLA-I expression was measured (Figure 2.9B). Similar to our *in vitro* cell culture studies, we found considerable heterogeneity regarding both the magnitude and pattern of HLA-I gene induction in response to either drug among these patient samples. 3 out of 5 patients showed robust induction in response to at least

one of the treatment regimens. Among the responders, 2 out of 3 responded to 5AZA2 and showed no response to LBH alone. 1 out of 3 responders showed strong induction to LBH and less response to 5AZA2. In 2 out of 5 patient samples we saw negligible response to 5AZA2 or LBH treatments. Overall, these *ex vivo* culture studies corroborate the cell line data and indicate a biologically relevant role for DNMT and HDAC inhibition in re-expression of HLA-I in prostate cancer. Moreover, these data again highlight the variability in response to inhibition of these two classes of epigenetic regulators and underlines the multiple epigenetic mechanisms that are at play.

HLA-I upregulation on tumor cells by DNMT and HDAC inhibition enhances activation of PSMA-specific T-cells

We next sought to assess the functional relevance of HLA-I upregulation by DNMT and HDAC inhibition. To do this, we measured the CD8+ T-cell response to LNCaP cells treated with 5AZA2, SGI, and LBH alone and in combination. $PSMA_{27-38}$-specific $CD8^+$ T-cells were raised by peptide vaccination in HHD mice expressing humanized HLA-A*02. HLA-A*02 expressing LNCaP cells were pre-treated with 5AZA2, SGI, or LBH or a combination of SGI or 5AZA2 and LBH and then co-cultured with splenocytes from vaccinated or unvaccinated mice (Figure 2.10A). As an additional control, an anti-HLA-I blocking antibody was used to block HLA-I at the LNCaP cell surface. After co-culture, T-cell activation markers were measured by flow cytometry in T-cells from each co-culture treatment condition (Figure 2.10B). We found that $PSMA_{27-38}$ tetramer-positive ($PSMA^+$) $CD8^+$ T-cells that were co-cultured with LNCaP cells treated with any DNMT or HDAC inhibitor increased in frequency and expressed increased levels of activation markers

CD69 and LFA-1 compared to those co-cultured with DMSO treated LNCaP cells. These T-cells also expressed increased levels of granzyme B and interferon-γ (IFN-γ), markers indicative of T-cell stimulation and differentiation into cytotoxic T-cells (CTLs). Additionally, the population of T-cells expressing CD107 (LAMP1), a marker of T-cell degranulation, was increased in the T-cells co-cultured with LNCaP cells treated with 5AZA2, SGI, and/or LBH.

The percent of $PSMA^+$ $CD8^+$ T-cells present after co-culture with LNCaP cells in each treatment condition is shown in Figure 2.10C. Treatment of LNCaP cells with SGI and LBH in combination significantly increased the percent of $PSMA^+CD8^+$ T-cells after co-culture with LNCaP cells at a 2:1 or 1:1 T-cell effector to tumor target (E:T) ratio SGI alone was able to significantly increase the percent of $PSMA^+CD8^+$ T-cells after co-culture with LNCaP cells at a 2:1 T-cell effector to tumor target ratio. The percent of $PSMA^+$ $CD8^+$ T-cells present increased after co-culture with LNCaP cells treated with 5AZA2 and LBH in combination, though this did not reach statistical significance. Treatment of LNCaP cells with HLA-I blocking antibody was able to ablate these effects. No significant changes in the percent of $PSMA^+$ T-cells where seen when LNCaP cells were co-cultured with T-cells from unvaccinated mice. This study confirms a clear functional role for HLA-I induction by DNMT and HDAC inhibition and suggests the utility of HLA-I re-expression for vaccine-based immunotherapies that rely on functional MHC-I expression on tumor cells.

DISCUSSION

The evidence presented here demonstrates that epigenetic mechanisms regulate expression of HLA-I genes in human prostate cancer. While it has been known for decades that HLA-I is downregulated at the protein level in prostate cancer, there have been limited investigations into the molecular underpinnings of this phenomenon, especially as it relates to epigenetic regulation of these genes. To address this lack of understanding, we have described the HLA-I methylome in prostate cancer patient samples and cell lines and confirmed the presence of repressive histone modifications in cell line models.

Our investigations into the regulation of class I HLA genes in prostate cancer revealed frequent HLA-I loss in metastatic tumors. The striking decrease in HLA-I expression in metastatic lesions identified in this study implies a possible role for loss of HLA-I expression in progression to metastasis. Previous findings showing that promoter DNA methylation increases during progression and that epigenetic mechanisms are important drivers prostate cancer progression (65,97,138). Whether the epigenetic alterations in HLA-I genes are a driver or passenger in the metastatic cascade will need to be further studied, but we hypothesize that immune evasion due to epigenetic loss of HLA-I plays a role in metastasis.

This study is in line with previous work showing HDAC inhibition upregulates HLA-I gene and proteins levels in LNCaP cells (39), and gene expression levels in PC3 and Du145 cells (39,102). Our work expands on these studies by examining the basal epigenetic signatures as well as changes in those signatures in response to HDAC and DNMT inhibition in multiple cell lines. We strengthen this idea further by showing

biological relevance for HLA-I induction by DNMT and HDAC inhibition in *ex vivo* tissue samples. Importantly, we also showed that HLA-I expression can be functionally induced on tumor cells by DNMT and HDAC inhibition, leading to increased activation of co-cultured T-cells from PSMA peptide vaccinated mice.

A previous study from our lab found DNMT and/or HDAC inhibition induces expression of CTAs (139) and another study found APM molecules were upregulated with HDAC inhibition (102). Our current study supports a wider role for epigenetics in regulating antigen presentation by also downregulating the HLA-I genes themselves. Inhibition of DNMT and HDACs is likely affecting many cellular pathways in addition to HLA-I genes, CTAs, and APM, leading to the phenotypes we observed in this study. A recent study found inhibition of BET bromodomain containing proteins, which are readers of histone acetylation, led to increased HLA-I protein expression and immunogenicity *in vivo*, supporting the important role we have found for histone modifications in HLA-I regulation (103). Further investigation into the contributions of other affected cellular pathways is needed to fully understand this phenomenon.

DNMT and HDAC inhibitors have been explored for their possible therapeutic efficacy and numerous clinical trials are ongoing for single or combination uses, including trials with 5AZA2, SGI, and LBH (140). Recently, the first epigenetic therapy for a solid tumor, a small molecule drug targeting EZH2, was approved by the FDA for use in epithelioid carcinoma (141,142). Previously, the only FDA approved uses for drugs targeting epigenetic modifying proteins were for hematologic malignancies. The limited success of epigenetic therapies in solid cancers may be due to the heterogeneity in the epigenetic signatures and responses to therapies as evidenced in our current study (129).

Additionally, failure of immunotherapies has been attributed to lack of immunogenicity of tumor cells as well as the inability to monitor treatment response in solid tumors (130). We anticipate that the monitoring HLA-I methylation status in patients as well as protein levels both before and during treatment may alleviate some of these challenges and we are currently working to develop epigenetically silenced HLA-I as a therapeutic biomarker. Overall, this study has implicated epigenetic mechanisms in the regulation of HLA-I in prostate cancer and points to epigenetic therapy as a promising option for enhancing the immune response in prostate tumors.

MATERIALS AND METHODS

Analysis of Public Data Sets

Genomic alteration data was accessed and analyzed in cBioPortal (RRID: SCR_014555) (143,144). TCGA-PRAD data was accessed and downloaded through UCSC Xena (145). Data from Taylor, et. al. (134) was accessed, analyzed, and downloaded through cBioPortal. ATAC-seq data from Corces, et al. was accessed, analyzed, and downloaded through UCSC Xena (145,146). Methylation beta values and matched gene expression values were accessed through Wanderer (147). Prism 8 (GraphPad Prism, RRID: SCR_002798) was used for correlation analyses. Z-scores for gene expression and ATAC-seq were calculated with the following formula:

$$Z=\frac{\chi-\mu}{\sigma}$$

where χ is the tumor or metastasis gene expression value, μ is the normal sample population mean, and σ is the normal sample population standard deviation. In the ATAC-seq data set, z-scores were calculated as compared to the tumor population, where χ is the tumor gene or ATAC-seq expression value, μ is the tumor sample population mean, and σ is the tumor sample population standard deviation.

Cell Lines and Cell Culture

LAPC4 (ATCC, Cat# CRL-13009, RRID: CVCL_4744) were maintained in DMEM Medium (Corning) supplemented with 20% fetal bovine serum (FBS) (Gibco, Cat# 10437028), 1% sodium pyruvate (Corning, Cat# MT25000CI), 0.5% beta-mercaptoethanol, and 1% penicillin-streptomycin (HyClone, Cat# SV30010). LAPC4 cells were cultured in poly-d-lysine coated flasks and/or plates (BioCoat flasks: Corning, Cat#

0877260; 6-well plates: Sigma-Aldrich, Cat# Z720798-20EA). RWPE1 (ATCC Cat# CRL-11609, RRID: CVCL_3791, LNCaP (RCB, Cat# RCB2144, RRID: CVCL_1379) – used for 5AZA/LBH epigenetic drug treatment gene expression experiments and epigenetic drug treatment protein expression experiments, 22Rv1 (ATCC, Cat# CRL-2505, RRID: CVCL_1045), and PC3 (ATCC, Cat# CRL-7934, RRID: CVCL_0035) cells were maintained in RPMI 1640 Medium (Corning, Cat# MT10040CV) supplemented with 10% FBS, 1% sodium pyruvate, 1% penicillin-streptomycin, 1% non-essential amino acids (HyClone, Cat# SH30238.01), and 0.1% beta-mercaptoethanol. LCL (HCC2218-BL, ATCC, Cat# CRL-2363, RRID: CVCL_1264) and LNCaP (ATCC, Cat# CRL-1740, RRID: CVCL_1379) – used for baseline gene expression, SGI/LBH epigenetic drug treatment experiments, DNA methylation, and chromatin immunoprecipitation experiments – cells were grown in RPMI 1640 Medium supplemented with 10% FBS and 1% penicillin-streptomycin (LCL grown in suspension). LCL, RWPE1, LNCaP, 22rv1, and PC3 were cultured in tissue culture treated flasks and/or plates (Flasks: Corning, Cat# 07202000; Plates: Thermo Fisher Scientific, Cat# 087721G).

Ex Vivo Culture of Prostate Tissue

Human prostate tissues were obtained from patients undergoing radical prostatectomy at the University of Wisconsin-Madison. All patients were consented under an Institutional Review Board (IRB) protocol #20130653. Absorbable gelatin sponges (Ethicon, Cat# 1973) were cut into pieces to fit in a 24-well tissue culture plate. Sponges were soaked in Ham's F-12 media (Fisher Scientific, Cat# SH3002601) supplemented with 0.25 units/ml regular insulin (Sigma-Aldrich, Cat# I9278-5ML), 1 µg/mL hydrocortisone (Sigma-Aldrich,

Cat# H0888-1g), 5 µg/mL human transferrin (Sigma-Aldrich, Cat# T8158-100mg), 2.7 mg/ml dextrose, 0.1 nM non-essential amino acids (HyClone, Cat# SH30238.01), 100 units/ml and 100 µg/mL Penicillin/Streptomycin respectively (HyClone, Cat# SV30010), 2 mM L-glutamine (Corning, Cat# 25-005-CI), 25 µg/mL bovine pituitary extract (Life Technologies, Cat# 13028014), and 1% fetal bovine serum (FBS) (Gibco, Cat# 10437028) until fully saturated. Each core was cut into ~1 mm^2 by 1 mm^2 cubes. Tissue was placed on the sponges and cultured for up to 4 days at 37° C at 5% CO_2 and 500µL media was replaced daily.

Immunoblotting

Whole cell lysates were collected from adherent cells by scraping into RIPA buffer after washing with cold PBS. Whole cell lysates were separated by SDS-PAGE and transferred onto nitrocellulose membrane. Membranes were blocked with SuperBlock blocking buffer (Thermo Scientific, Cat# 37515). Membranes were probed with primary antibodies diluted in 3% BSA in TBS plus 0.1% Tween-20 at 4°C overnight followed by incubation with HRP-linked secondary antibody (BioLegend, Cat# 405306, RRID:AB_315009) at RT for 1 hour and visualization by chemiluminescence. Primary antibodies: HLA-I clone W6/32 (BioLegend, Cat# 311412, RRID:AB_493132), α-tubulin (BioLegend, Cat# 627901, RRID:AB_439760).

Immunofluorescent staining in cell lines

For HLA-I staining, cells were seeded on coverslips in 6 well plates and fixed with 4% formaldehyde for 10 mins followed by washing with 1x PBS and blocking with 5% FBS in

1x PBS with 0.1% BSA. Fixed cells were stained with Hoechst 33342 (Thermo Fisher, Cat# PI62249) and anti-HLA-I conjugated to FITC (BioLegend, Clone W6/32, Cat# 311404) overnight at 4ºC in PBS with 0.1% BSA. Coverslips were mounted on slides with VECTASHIELD Antifade Mounting Medium (Vector Laboratories, Cat# H-1000-10). For DNMT staining, cells were seeded in 24 well plates and fixed with 4% formaldehyde for 10 minutes followed by washing with 1x PBS, permeabilization with 0.3% Triton X-100 for 5 minutes at RT, and blocking with 5% FBS in x PBS with 0.1% BSA. Cells were stained with Hoechst 33342 (Thermo Fisher, Cat# PI62249) and primary antibodies overnight in PBST with 0.1% Triton X-100 at 4ºC: DNMT1 (Cell Signaling, clone D63A6, cat# 5032, RRID:AB_10548197), DNMT3a (Thermo Fisher Scientific, clone 64B1446, Cat# MA5-16171, RRID:AB_2537690), or DNMT3b (Cell Signaling, clone E8A8A, cat# 57868, RRID:AB_.2799534). Cells were stained with secondary antibodies at RT for 1hr. Slides or plates were imaged at 10x using NIS Elements AR Microscope Imaging Software (NIS-Elements, RRID:SCR_014329) and analyzed using NIS Elements AR analysis software.

MBD2-MBD Enrichment of Methylated DNA from Cell Lines

Genomic DNA was isolated from cells using the AllPrep RNA/DNA Mini Kit (Qiagen, Cat# 80204) according to manufacturer's instructions. DNA was quantified by a NanoDrop 1000 spectrophotometer and 1µg DNA was sheared by sonication to an average size of around 200bp. Methylated DNA was enriched from sheared genomic DNA using the EpiXplore Methylated DNA Enrichment Kit (Takara Bio, Cat# 631962) according to manufacturer's instructions. Enrichment was measured by qPCR using primers designed

to various regions of *HLA-A*, *HLA-B*, and *HLA-C* (Integrated DNA Technologies). Primers are listed in Table 2.3.

Chromatin Immunoprecipitation

Chromatin immunoprecipitation (ChIP) was performed according to manufacturer's instructions using the SimpleChIP Enzymatic Chromatin IP Kit with Magnetic Beads (Cell Signaling Technology, Cat# 9003S). Immunoprecipitation was performed using the following antibodies from Cell Signaling Technology: Histone H3 (Clone D2B12, Cat# 4620S, RRID: AB_1904005), H3K27ac (Clone D5E4, Cat# 8173S, RRID: AB_10949503), H3K27me3 (Clone C36B11, Cat# 9733S, RRID: AB_2616029), and IgG (Cat# 2729S, RRID: AB_1031062). DNA was then analyzed by qPCR using primers designed to target *HLA-A*, *HLA-B*, and *HLA-C* (Integrated DNA Technologies). Primers are listed in Table 2.3. The H3K27ac signature in GM12878 determined by ChIP-seq from the ENCODE consortium (137) was accessed in the UCSC Genome Browser (RRID: SCR_005780) to aid in primer design.

Epigenetic Drug Treatments of Cell Lines and Ex Vivo Tissue

5-Aza-2′-deoxycytidine (5AZA2) (Sigma-Adrich, Cat# A3656-5MG), Panobinostat (LBH, LBH589) (Selleckchem, Cat# S1030), and SGI-110 (SGI) (Astex Pharmaceuticals) were dissolved in DMSO and stored at -80^{o}C in aliquots. Cells were treated with 10μM 5AZA2, 1μM SGI, or DMSO for 72 hours. 10nM or 100nM LBH was added for the last 24 hours after 48 hours of 5AZA2 or SGI treatment for combination treatments.

Gene Expression Analysis in Cell Lines

Total RNA was isolated from cells using the AllPrep RNA/DNA Mini Kit (Qiagen, Cat# 80204) according to manufacturer's instructions. RNA was quantified by a NanoDrop 1000 spectrophotometer and 1µg total RNA was reverse transcribed using the High Capacity RNA-to-cDNA kit (Thermo Fisher Scientific, Cat# 4388950). cDNA was diluted 10x and 5uL was used per reaction for qPCR. Pre-designed TaqMan probes (Thermo Fisher) for *HLA-A* (Hs01058806_g1), *HLA-B* (Hs00818803_g1), and *HLA-C* (Hs00740298), and *RPLP0* (4333761F) were used with iTaq Universal Probes Supermix (BioRad, Cat# 1725135). Gene expression was determined using the delta-delta-Ct method after normalization of each gene to housekeeping gene, *RPLP0 (P0)*.

Gene Expression Analysis in Epigenetic Drug Treated Ex Vivo Tissue and 5AZA2/LBH Treated Cell Lines

Total RNA was isolated from cells using RNeasy Mini Kit (Qiagen, Cat# 74106) according to manufacturer's protocol. Total RNA was isolated from ex vivo tissue using the Aurum™ Total RNA Fatty and Fibrous Tissue Kit (BioRad, Cat# 7326830) according to the manufacturer's protocol. RNA was quantified by a NanoDrop 1000 spectrophotometer and 1µg total RNA was reverse transcribed using iScript Reverse Transcription Supermix (Bio-Rad, Cat# 1708841). 1µL of the cDNA synthesis reaction was used to perform qPCR using SsoAdvanced Universal SYBR Green Supermix (BioRad, Cat# 1725274) according to the manufacturer's protocol. *HLA-A, HLA-B, HLA-C*, and *RPLP0* primers are listed in Table 2.3. Gene expression was determined using the delta-delta-Ct method after normalization of each gene to housekeeping gene, *RPLP0 (P0)*.

Peptide Vaccinations and T-cell Co-Culture

PSMA specific $CD8^+$ T-cells were generated by $PSMA_{27-38}$ peptide vaccination of HHD transgenic humanized mice expressing human HLA-I A*02. The HHD mice were a generous gift from Professor Francois Lemonnier at the Pasteur Institute, Paris (148). Mice were given once weekly subcutaneous injections of 100 ug synthetic PSMA peptide (VLAGGFFLL) (ProImmune, Oxford, UK) in 100uL CFA (Thermo Fisher Scientific, Cat# NC0916022) for the first injection or IFA vehicle (Sigma-Aldrich, Cat# AR002) for subsequent injections. Splenocytes were harvested 1 week after last immunization and the number of live $PSMA_{27-38}$-specific CD8+ splenocytes was determined by flow cytometry analysis following staining with GhostDye™ Violet 510 (Tonbo Biosciences, Cat# 13-0870), anti-mouse CD8 antibody (Tonbo Biosciences, Cat# 25-0081, RRID:AB_2621623) and Pro5® $PSMA_{27-38}$ A*02:01 MHC I pentamer (ProImmune, Oxford, UK). PSMA vaccinated splenocytes were then co-cultured with LNCaP cells that were pretreated with DMSO vehicle or 10uM 5AZA2 or 1uM SGI for 72 hours and/or 10nM LBH for the last 24 hours in RPMI media supplemented with 10% FBS. In control co-culture wells, LNCaP cells were treated with 5 μg of purified anti-HLA-A,B,C blocking antibody (clone W6/32) (BioLegend, Cat# 311412, RRID:AB_493132) prior to adding splenocytes. Cells were co-cultured for 72 hours and Golgi-stop (BD Biosciences, Cat# 554724) was added for the last 4 hours of culture, following the manufacturer's protocol. Cells were then harvested and subjected to labeling with GhostDye™ Violet 510 and surface markers: CD8 (Tonbo Biosciences, Cat# 25-0081, RRID:AB_2621623), Pro5® $PSMA_{27-38}$ A*02:01 MHC I pentamer (ProImmune), CD69 (BD Pharmigen, Cat# 551113,

RRID:AB_394051), LFA-I (BD Biosciences, Cat# 558191, RRID:AB_397055), and CD107 (BD Biosciences, Cat# 564347, RRID:AB_2738760) followed by fixation and permeabilization with BD Cytofix/Cytoperm (Thermo Fisher Scientific, Cat# 554714) according to the manufacturer's protocol and intracellular staining with antibodies against murine IFNγ (Tonbo Biosciences, Cat# 20-7311, RRID:AB_2621616) and Granzyme B (BD Biosciences, Cat# 560211, RRID:AB_1645488). Cells were then washed and acquired on an LSR II Fortessa or an Attune NxT instrument followed by data analysis by the FlowJo software v9.9.6 (FlowJo, RRID: SCR_008520). Gating controls included the fluorescent minus one (FMO) strategy.

Statistical Analysis

Survival curve analysis was performed using the log-rank (Mantel-Cox) test. For the HLA-I genomic alteration and HLA-I protein expression studies, comparison between groups was made with an ordinary-way ANOVA followed by post hoc analysis with the Tukey test for correction of multiple comparisons. For baseline gene expression and ChIP experiments, comparisons between groups were made with 2-way ANOVA using the Dunnett method for correction of multiple comparisons. For drug treatment gene expression experiments, comparison between DMSO and treatment groups were made by t-test corrected for multiple comparisons by the Holm-Sidak method. Gene expression statistical analyses were performed on delta-Ct values. All statistical analyses were performed in Prism 8 (GraphPad Prism, RRID: SCR_002798).

ACKNOWLEDGEMENTS

We would like to thank the patients who donated samples for this study. The results using TCGA-PRAD data sets are based on data generated by the TCGA Research Network: https://www.cancer.gov.tcga.

Table 2.1. Up- and downregulation of HLA-I, DNMTs, and class I HDACs.

	Up (%)			Down (%)		
	Primary (TCGA)	Primary (Taylor)	Metastatic (Taylor)	Primary (TCGA)	Primary (Taylor)	Metastatic (Taylor)
	n=496	n=130	n=19	n=496	n=130	n=19
HLA-A	40 (8)	15 (12)	1 (5)	90 (18)	24 (18)	12 (63)
HLA-B	74 (15)	18 (14)	2 (11)	28 (6)	19 (15)	12 (63)
HLA-C	70 (14)	17 (13)	1 (5)	42 (8)	18 (14)	12 (63)
DNMT1	33 (7)	9 (7)	10 (53)	41 (8)	40 (31)	2 (11)
DNMT3A	179 (36)	28 (22)	9 (47)	6 (1)	11 (8)	1 (5)
DNMT3B	98 (20)	10 (8)	7 (37)	14 (3)	1 (1)	0 (0)
HDAC1	188 (38)	43 (33)	8 (42)	17 (3)	17 (13)	4 (21)
HDAC2	158 (32)	14 (11)	7 (37)	80 (16)	0 (0)	0 (0)
HDAC3	104 (21)	27 (21)	7 (37)	15 (3)	24 (18)	3 (16)
HDAC8	193 (39)	0 (0)	1 (5)	5 (1)	27 (21)	6 (32)

Table 2.2. Correlation of HLA-I gene expression to DNA methylation.

HLA-A

Probe	1	2	3	4	5	6	7	8	9	10	11	12
	Tumor											
Spearman r	-0.288	-0.340	-0.258	-0.202	-0.292	-0.257	-0.421	-0.422	-0.375	-0.381	-0.545	-0.371
P-value summary[a]	****	****	****	***	****	****	****	****	****	****	****	****
	Normal											
Spearman r	-0.122	-0.121	-0.095	-0.260	-0.129	-0.160	-0.421	-0.262	-0.213	-0.321	-0.245	-0.423
P-value summary[a]	ns	ns	ns	ns	ns	ns	*	ns	ns	ns	ns	*

HLA-B

Probe	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
	Tumor																		
Spearman r	-0.483	-0.199	-0.118	-0.095	-0.073	-0.091	-0.140	-0.005	-0.048	-0.057	-0.153	0.025	-0.027	-0.129	-0.237	-0.473	-0.339	-0.336	-0.297
P-value summary[a]	****	***	*	ns	ns	ns	*	ns	ns	ns	**	ns	ns	*	****	****	****	****	****
	Normal																		
Spearman r	-0.362	-0.245	-0.298	-0.420	0.139	0.257	0.027	-0.063	-0.291	-0.505	-0.161	-0.224	-0.423	-0.154	-0.100	-0.338	-0.296	-0.204	-0.392
P-value summary[a]	*	ns	ns	*	ns	ns	ns	ns	ns	**	ns	ns	*	ns	ns	ns	ns	ns	*

HLA-C

Probe	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	Tumor														
Spearman r	0.193	-0.094	0.232	-0.087	-0.125	-0.140	-0.131	-0.127	-0.309	-0.220	-0.303	-0.298	-0.308	-0.306	-0.452
P-value summary[a]	***	ns	****	ns	*	*	*	*	****	****	****	****	****	****	****
	Normal														
Spearman r	0.033	-0.134	0.028	-0.327	0.193	-0.049	-0.318	-0.291	-0.247	-0.339	-0.344	-0.311	-0.199	0.010	-0.357
P-value summary[a]	ns	ns	ns	ns	ns	ns	ns	ns	ns	*	*	ns	ns	ns	*

[a] ns: p>0.05; * 0.05>p>0.01; ** 0.01>p>0.001; *** 0.001>p>0.0001; **** p<0.0001

Table 2.3. Primers used for gene expression and epigenetic analysis of HLA-I.

	MBD2 Enrichment	Forward Primer (5'-3')	Reverse Primer (5'-3')
HLA-A	Distal Promoter	GCCAAGACTCAGGGAGACAT	AAACTGCGGAGTTGGGGAAT
	Proximal Promoter	CCAACTCCGCAGTTTCT	CACTGATTGGCTTCTCTGG
	Exon 1	CCAGAGAAGCCAATCAGTG	GAGTAGCAGGAGGAGGGTTC
	Exon 2	GACCAGGAGACACGGAATG	CTTCGGGGTGGATCTCGGA
HLA-B	Distal Promoter	GCCTTCAGAGAAAACTTCACCAGG	TCTTGTGTAGGGAAACTGAGCA
	Proximal Promoter	TTCCAGGATACTCGTGACGC	TCCCTCCCGACCCGC
	Exon 1	GGAGGGAAATGGCCTCTG	GGACACGGAGGTGTAGAAATAC
	Exon 2	CCGGAACACACAGATCTACAA	TCAGGGAGGCGGATCTC
HLA-C	Distal Promoter	GACTCTACACGTCCATTCCCAG	CTTTGCCTTACCTTACCTCACCT
	Proximal Promoter	CAGGGTCTCAGGCTCCAAGG	GGAAGAAGGACCCGACGCAG
	Exon 1	GAGACCTGGGCCTGTGA	GTCGAAATACCTCATGGAGTGG
	Exon 2	CTCCCACTCCATGAGGTATTTC	GCGCTTGTACTTCTGTGTCT
	ChIP		
	HLA-A	AACCCTCCTCCTGCTACTC	GGGACACGGATGTGAAGAAATA
	HLA-B	GGAGGGAAATGGCCTCTG	GGACACGGAGGTGTAGAAATAC
	HLA-C	GAGACCTGGGCCTGTGA	GTCGAAATACCTCATGGAGTGG
	Gene Expression (Figure 2.9)		
	HLA-A	GCGGCTACTACAACCAGAGC	CCAGGTAGGCTCTCAACT
	HLA-B	TCCTAGCAGTTGTGGTCATG	TCAAGCTGTGAGAGACATAT
	HLA-C	TCCTGGTTGTCCTAGTC	CAGGCTTTACAAGTGATGAG
	RPLP0	GACAATGGCAGCATCTACAAC	GCAGACAGACACTGGCAA

Figure 2.1. HLA-I genomic alterations and gene expression in primary and metastatic prostate cancer.

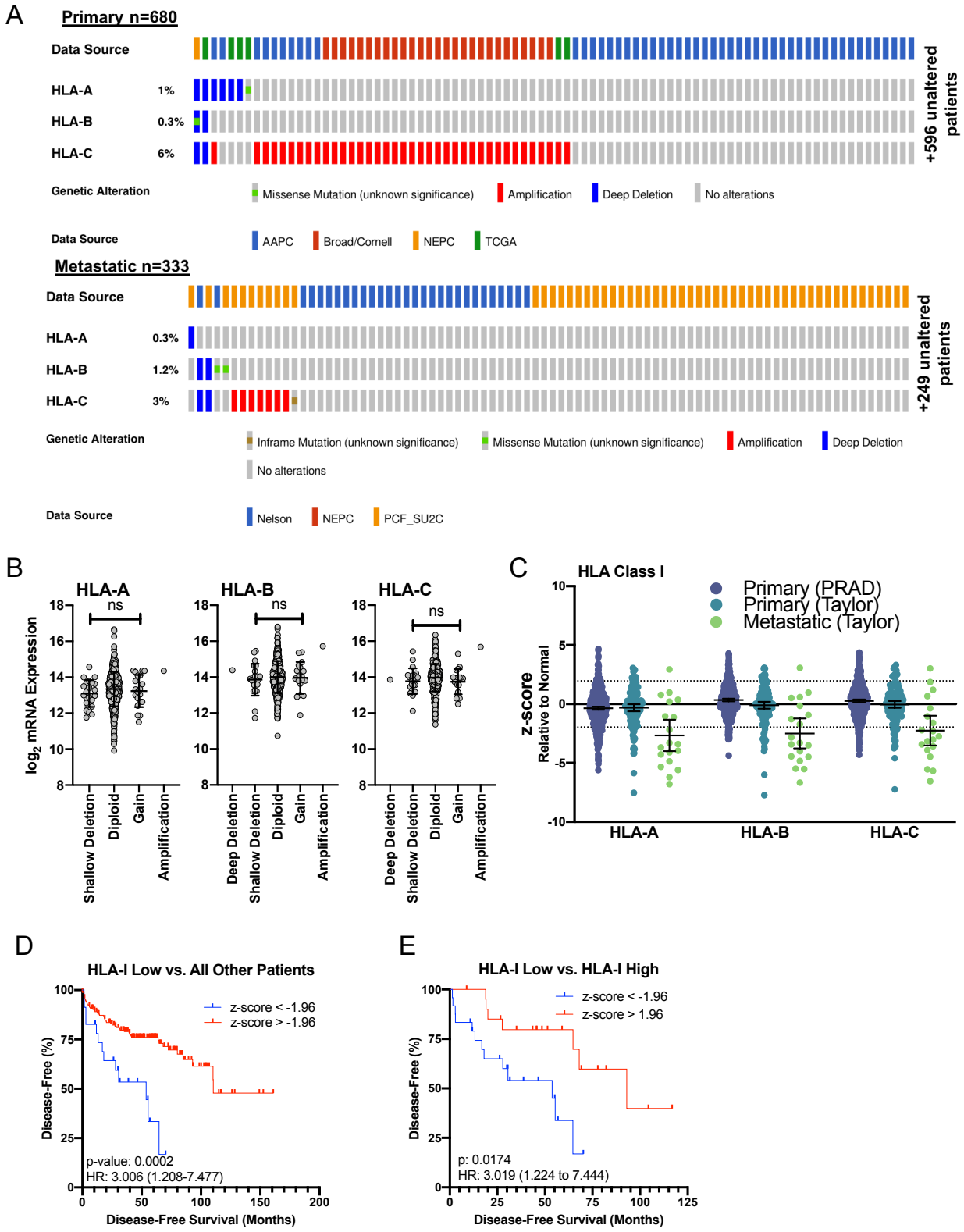


Figure 2.1. HLA-I genomic alterations and gene expression in primary and metastatic prostate cancer. A) Amplification, deletion, and mutations in *HLA-A*, *HLA-B*, and *HLA-C* in primary prostate cancer across 4 independent studies and metastatic prostate cancer across 3 independent studies. Overall percent altered for each gene is shown. All existing alterations are shown. B) HLA-I genomic alterations in relation to HLA-I gene expression in the PRAD dataset. C) *HLA-A*, *HLA-B*, and *HLA-C* expression in the PRAD, and Taylor data sets. Z-scores are relative to normal samples in the respective study. Dotted lines indicate z-score of +/- 1.96. Expression above or below dotted line is considered significantly up- or downregulated respectively. Line and error bars represent mean and 95% confidence interval. D) Kaplan-Meier curves showing disease-free survival time for patients in the Taylor data set with low expression (z-score < -1.96) vs. high expression (z-score > 1.96) of any HLA-I gene. E) Kaplan-Meier curves showing disease-free survival time for patients in the Taylor data set with low HLA-I expression in any HLA-I gene (z-score < -1.96) vs. all other patients (z-score > -1.96).

Figure 2.2. DNMT and HDAC gene expression and correlation to HLA-I in prostate cancer.

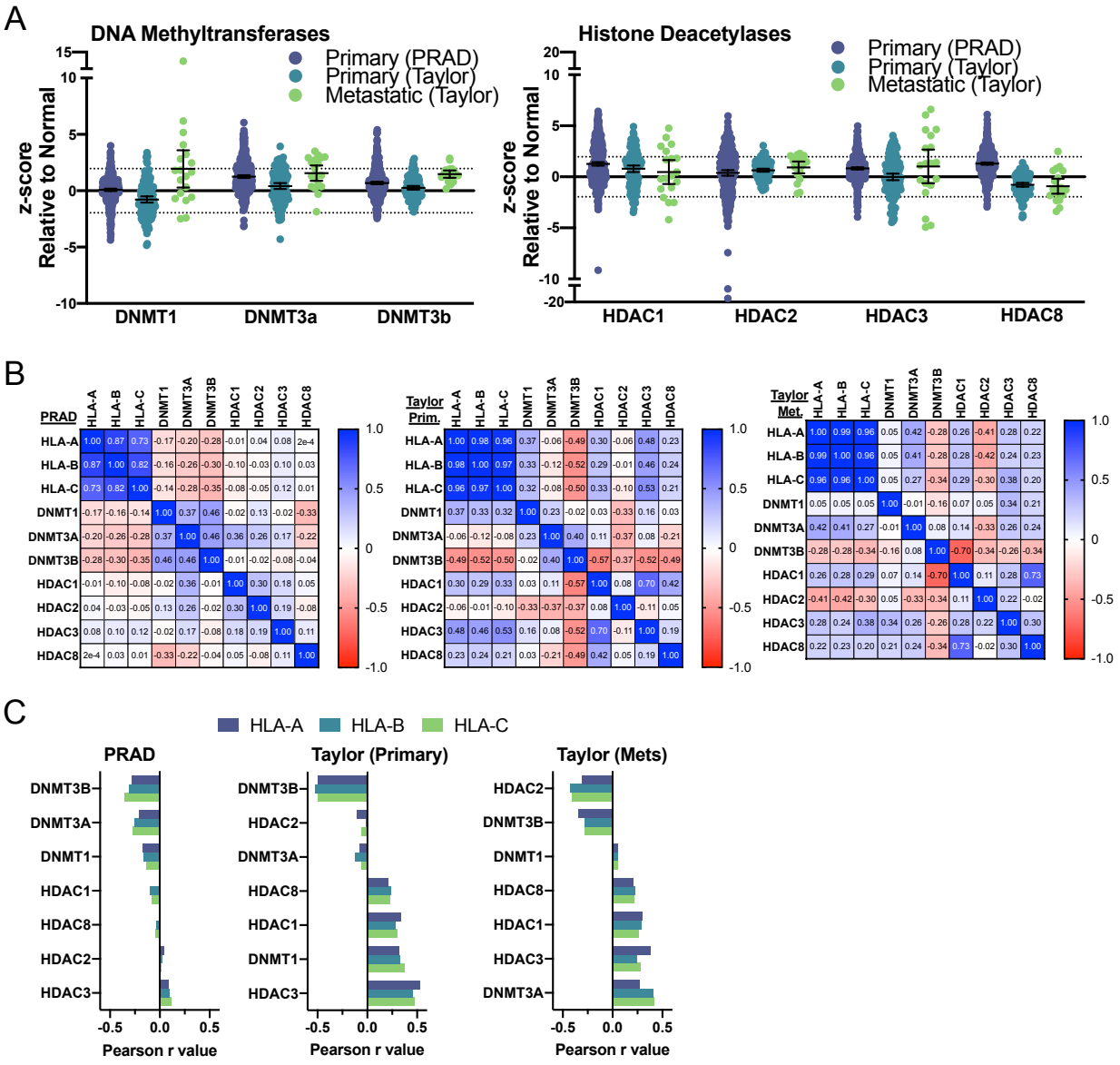


Figure 2.2 DNMT and HDAC gene expression and correlation to HLA-I in prostate cancer. A) Gene expression of the DNMTs and class I HDACs in the PRAD and Taylor data sets. Represented as z-scores relative to normal samples. Error bars represent SEM. Dotted lines indicate z-score of +/- 1.96. Expression above or below dotted line is considered significantly up- or downregulated respectively. Line and error bars represent mean and 95% confidence interval. B) Pearson r correlation matrix of gene expression of DNMTs and class I HDACs to HLA-I in each data set. C) Pearson r correlation values from (B) ordered from most negatively correlated to most positively correlated.

Figure 2.3. DNA methylation at HLA-I genes is associated with loss of HLA-I gene expression in prostate cancer.

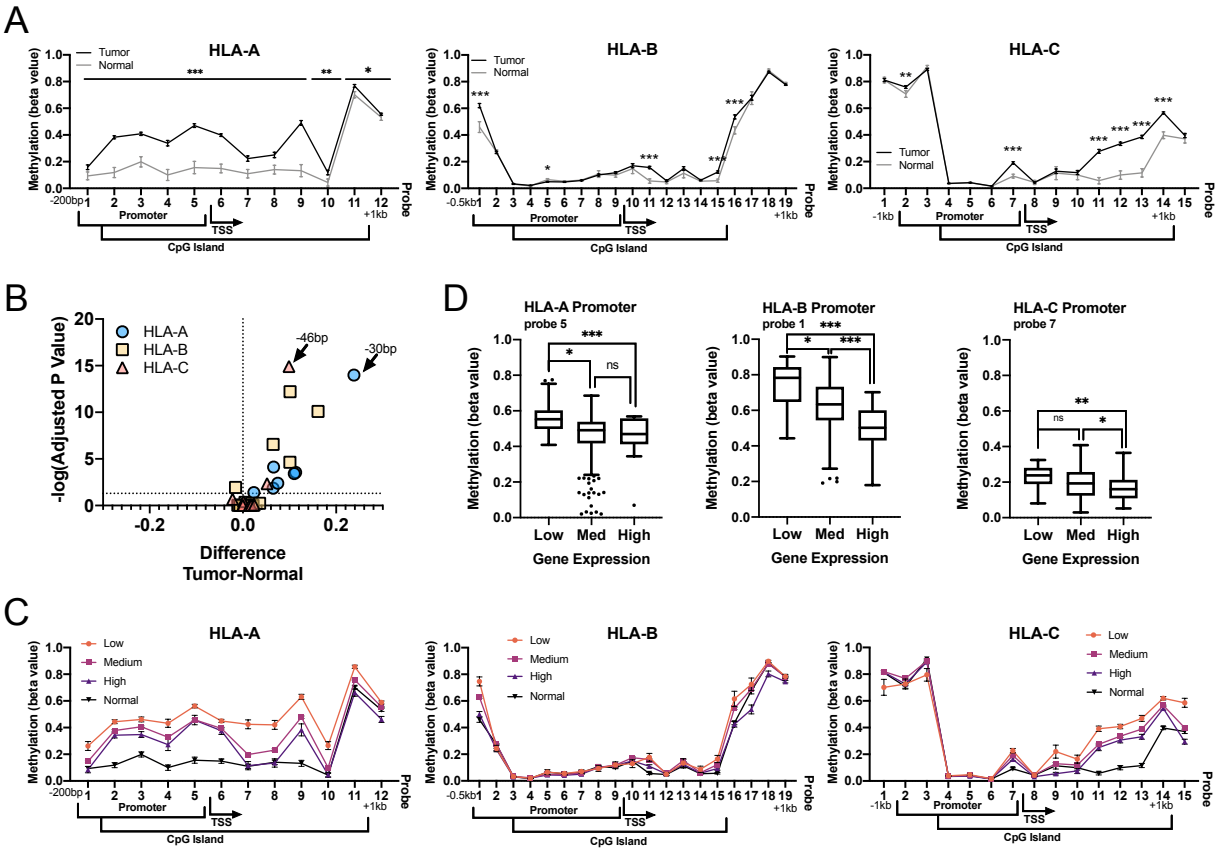


Figure 2.3. DNA methylation at HLA-I genes is associated with loss of HLA-I gene expression in prostate cancer. A) Methylation levels across the promoter and CpG islands of *HLA-A*, *HLA-B*, *HLA-C*. Gene regions where probes fall, TSS, and CpG island location are annotated. Distances between probes are not to scale. B) Volcano plot of statistical significance against difference in methylation levels between tumor and normal samples. Distance from the TSS for two points are indicated. C) Methylation beta values are shown at each probe across the HLA-I promoters and CpG islands for normal samples and tumor samples separated by gene expression level. D) Box plots of methylation levels of selected probes within the gene promoters separated by gene expression level. For C) and D) gene expression was categorized as: low expression: z-score < -1.96, med expression 1.96 < z-score > -1.96, high expression: z-score > 1.96.

Figure 2.4. Loss of chromatin accessibility at HLA-I genes is associated with loss of HLA-I gene expression in prostate cancer.

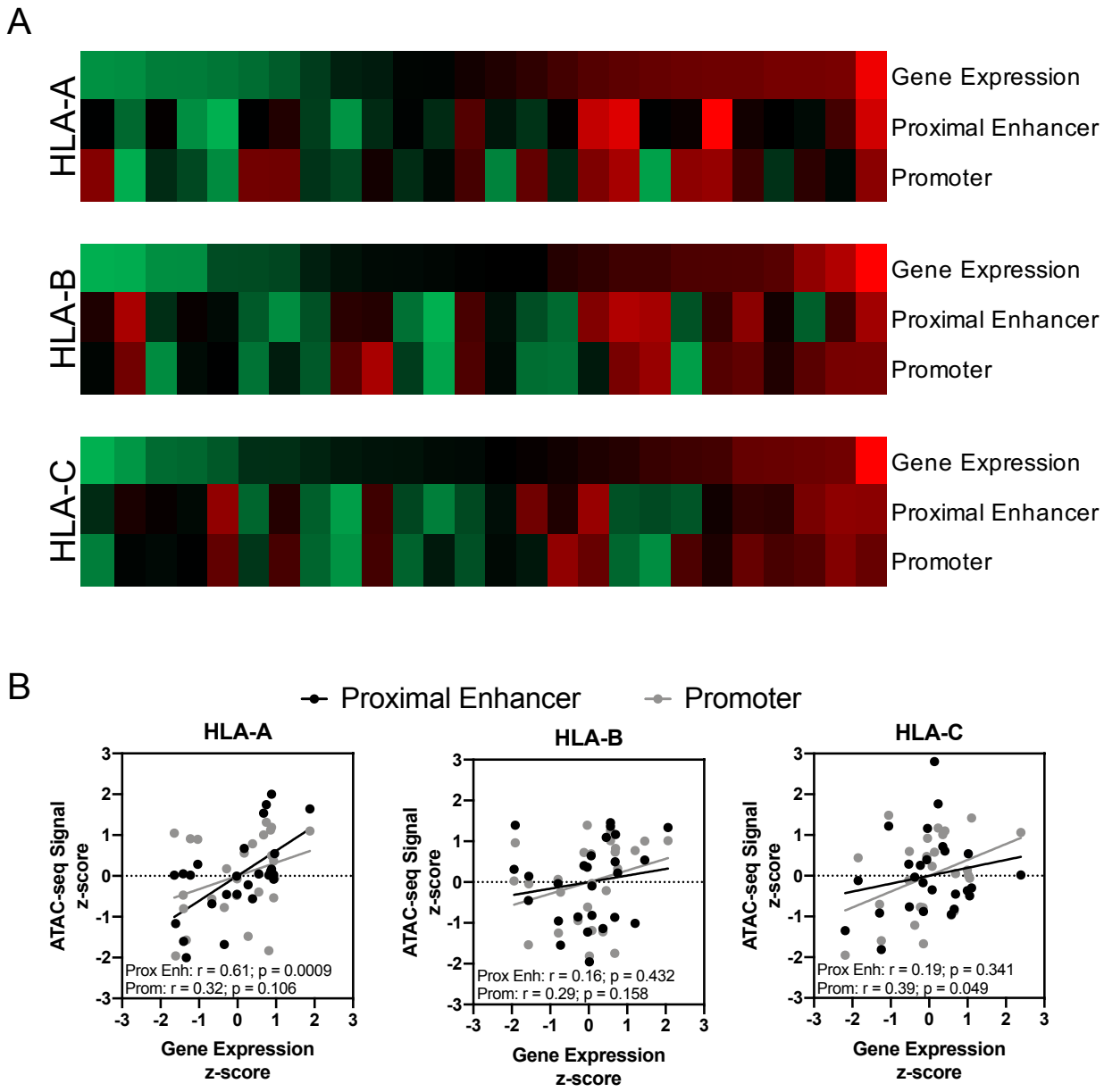


Figure 2.4. Loss of chromatin accessibility at HLA-I genes is associated with loss of HLA-I gene expression in prostate cancer. F) ATAC-seq from PRAD data set at a proximal enhancer and promoter in each HLA-I gene. G) Pearson r correlation of ATAC-seq signal from a proximal enhancer region and promoter to HLA-I gene expression.

Figure 2.5. Expression of HLA-I, DNMTs, and HDACs in prostate cancer cell lines.

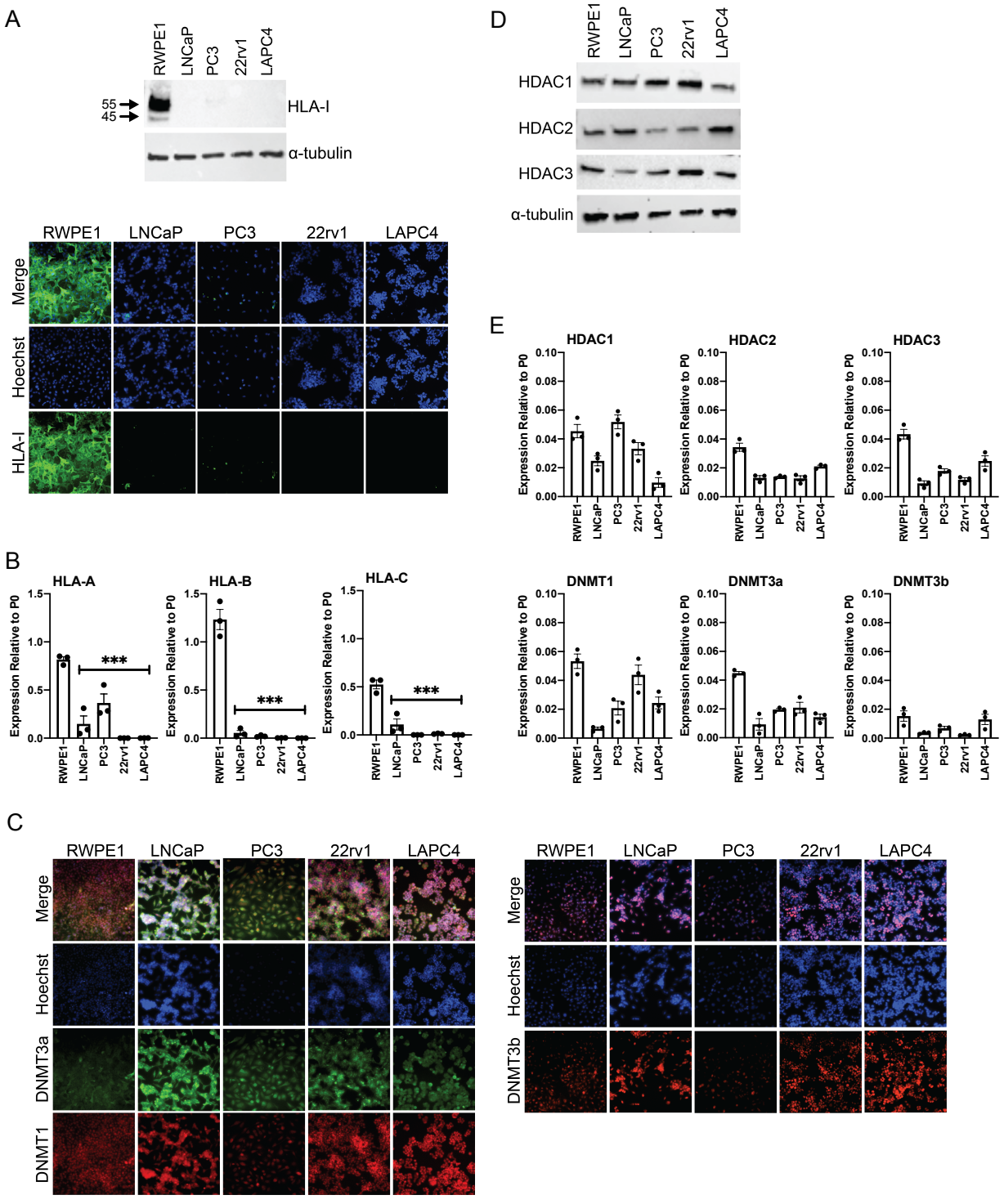


Figure 2.5. Expression of HLA-I, DNMTs, and HDACs in prostate cancer cell lines. A) HLA-I protein expression measured by western blot and immunofluorescence and (B) HLA-I gene expression in 4 prostate cancer cell lines (LNCaP, 22rv1, PC3, and LAPC4) compared to a normal prostate epithelial cell line, RWPE1. C) Immunofluorescent imaging of DNMT expression in prostate cancer cell lines and RWPE1. D) Western blot showing class I HDAC expression in prostate cancer cell lines and RWPE1. E) Gene expression of DNMTs and class I HDACs in prostate cancer cell lines and RWPE1.

Figure 2.6. Inhibition of DNMTs and HDACs induces HLA-I expression in prostate cancer cell lines.

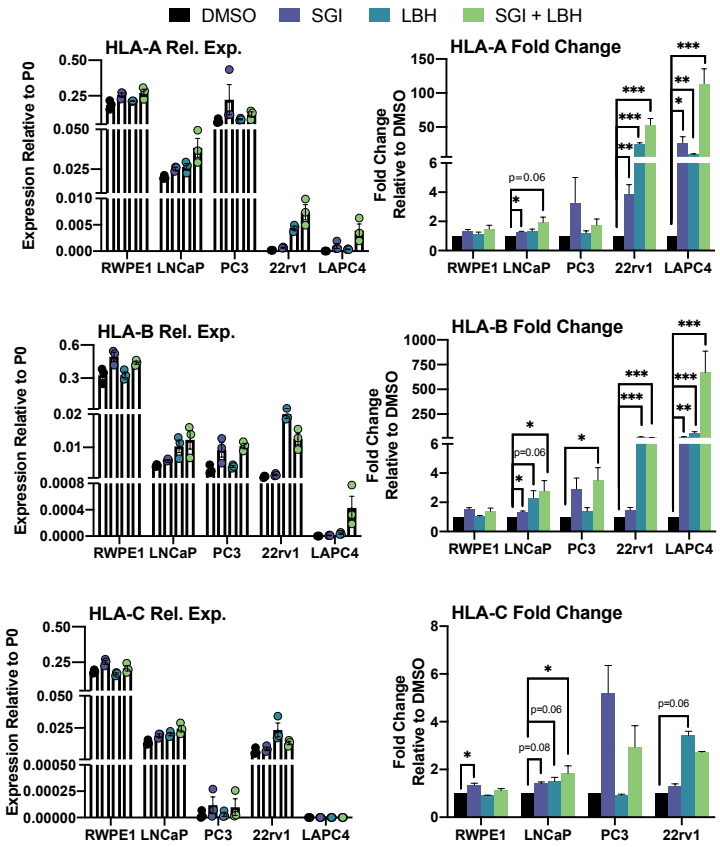


Figure 2.6. Inhibition of DNMTs and HDACs induces HLA-I expression in prostate cancer cell lines. HLA-I gene induction in cell lines treated with SGI or LBH alone and in combination. Data is represented as expression relative to housekeeping gene P0 and as a fold change relative to DMSO.

Figure 2.7. HLA-I downregulation is associated with a repressive epigenetic signature in prostate cancer cell lines.

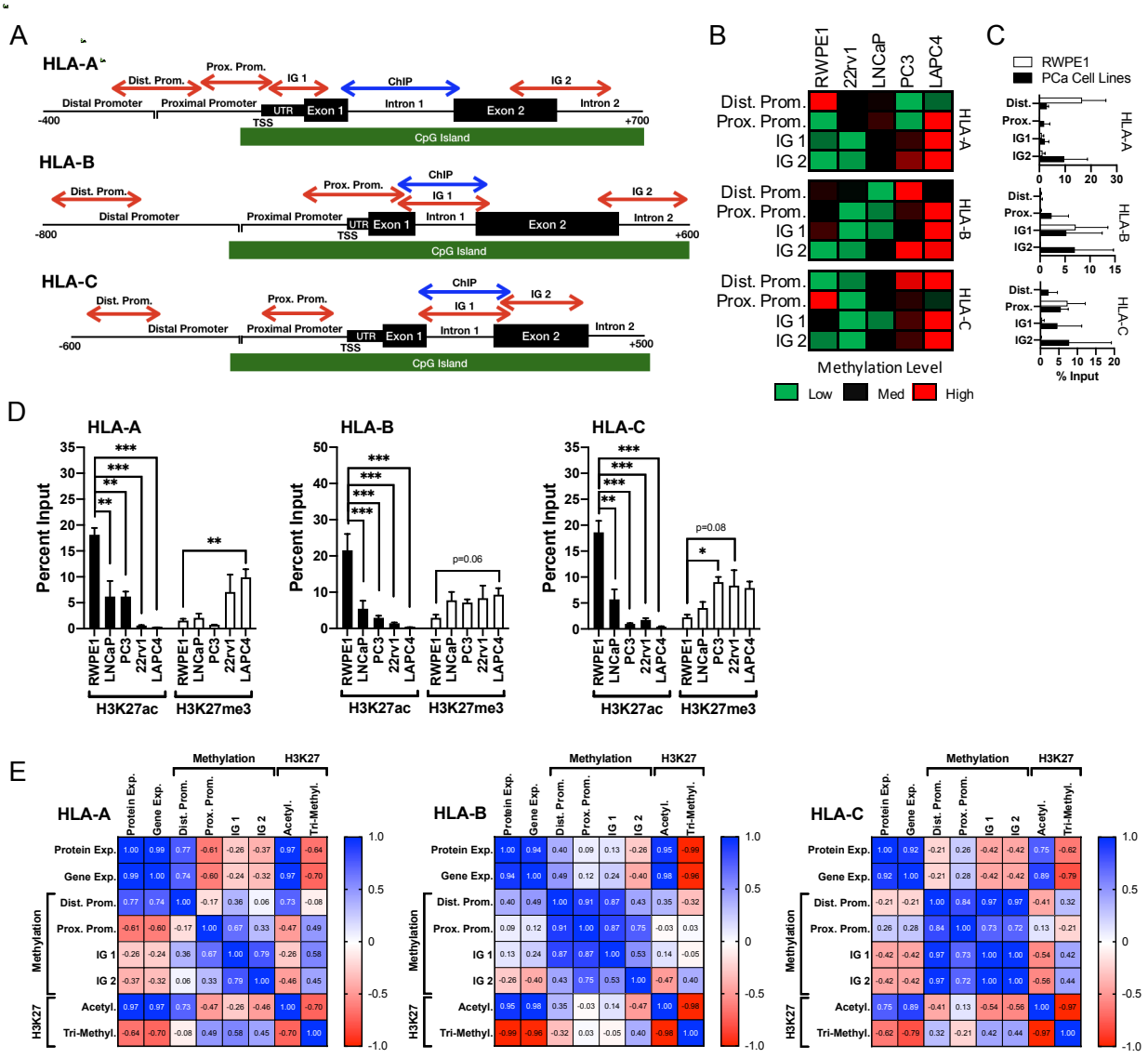


Figure 2.7. HLA-I downregulation is associated with a repressive epigenetic signature in prostate cancer cell lines. A) Primer locations for DNA methylation and histone modification analyses. B) DNA methylation in LNCaP, 22rv1, PC3, LAPC4, RWPE1, and a lymphoblastoid cell line (LCL) in four regions of the HLA-I genes. C) ChIP using antibodies to H3K27ac and H3K27me3 in LNCaP, 22rv1, PC3, and LAPC4 compared to RWPE1. D) Correlation matrix for HLA-I protein expression, gene expression, DNA methylation and H3K27 histone modifications. Pearson r values are indicated in each box.

Figure 2.8. Inhibition of DNMTs and HDACs alters epigenetic signatures in prostate cancer cell lines.

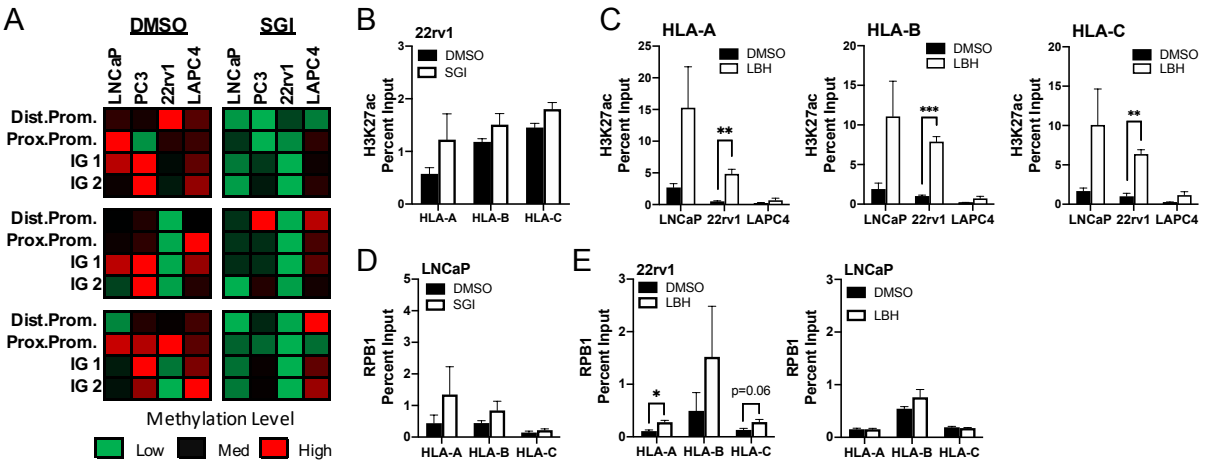


Figure 2.8. Inhibition of DNMTs and HDACs alters epigenetic signatures in prostate cancer cell lines. A) Heat map indicating the methylation level in cell lines treated with DMSO or SGI. B) ChIP using an antibody to H3K27ac in 22rv1 cells treated with DMSO or SGI. C) ChIP using antibodies to H3K27ac in cell lines treated with DMSO or LBH. D) ChIP using an antibody to Rpb1 in LNCaP cells treated with DMSO or SGI. E) ChIP using an antibody to Rpb1 in 22rv1 and LNCaP cells treated with DMSO or LBH.

Figure 2.9. Inhibition of DNMTs and HDACs induces HLA-I gene expression in primary prostate cancer *ex vivo*.

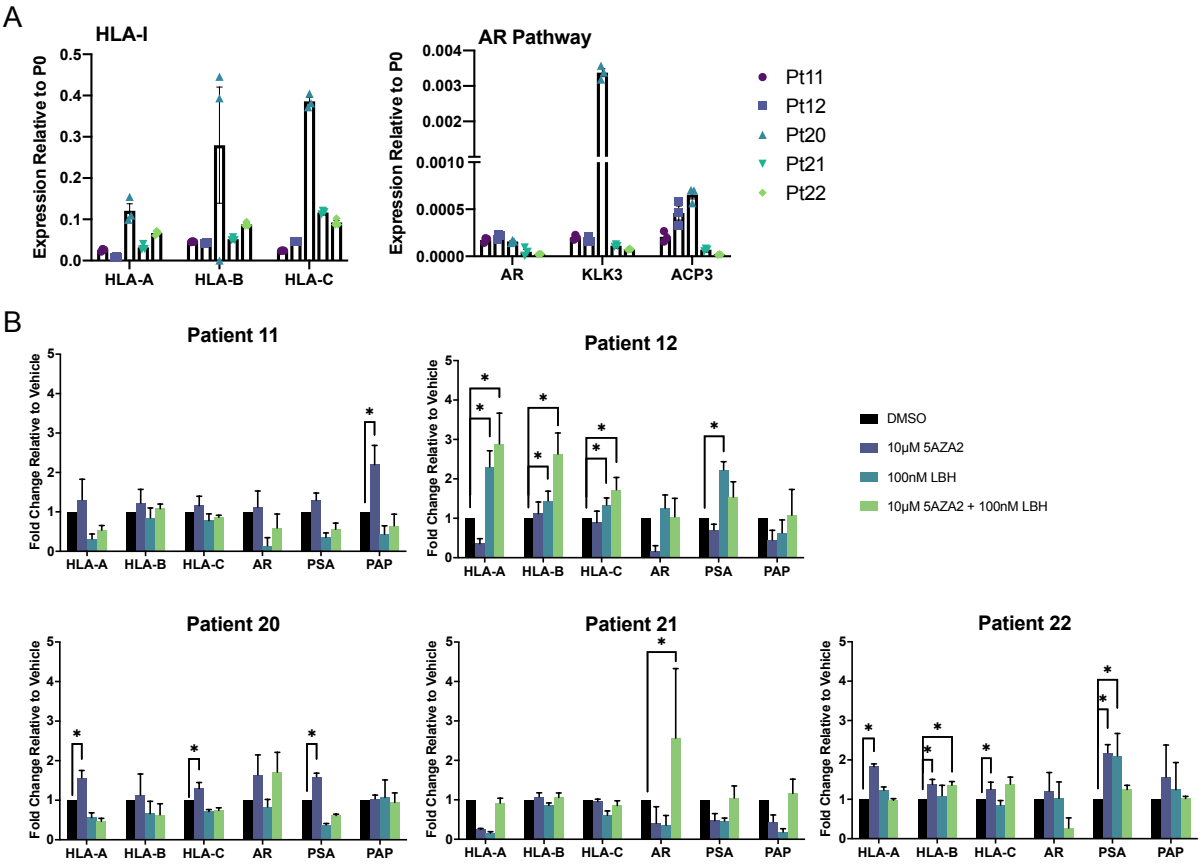


Figure 2.9. Inhibition of DNMTs and HDACs induces HLA-I gene expression in primary prostate cancer *ex vivo*. A) Gene expression analysis of the HLA-I genes and prostate specific elements of the AR pathway in prostate tumor tissue cultured *ex vivo*. B) Gene expression analysis of induction of HLA-I and AR pathway elements in ex vivo tissue treated with DMSO, 5AZA2, LBH, or 5AZA2+LBH. Error bars represent SD. $*$ $p<0.05$.

Figure 2.10. DNMT and HDAC inhibition in tumor cells increases co-cultured T-cell activation.

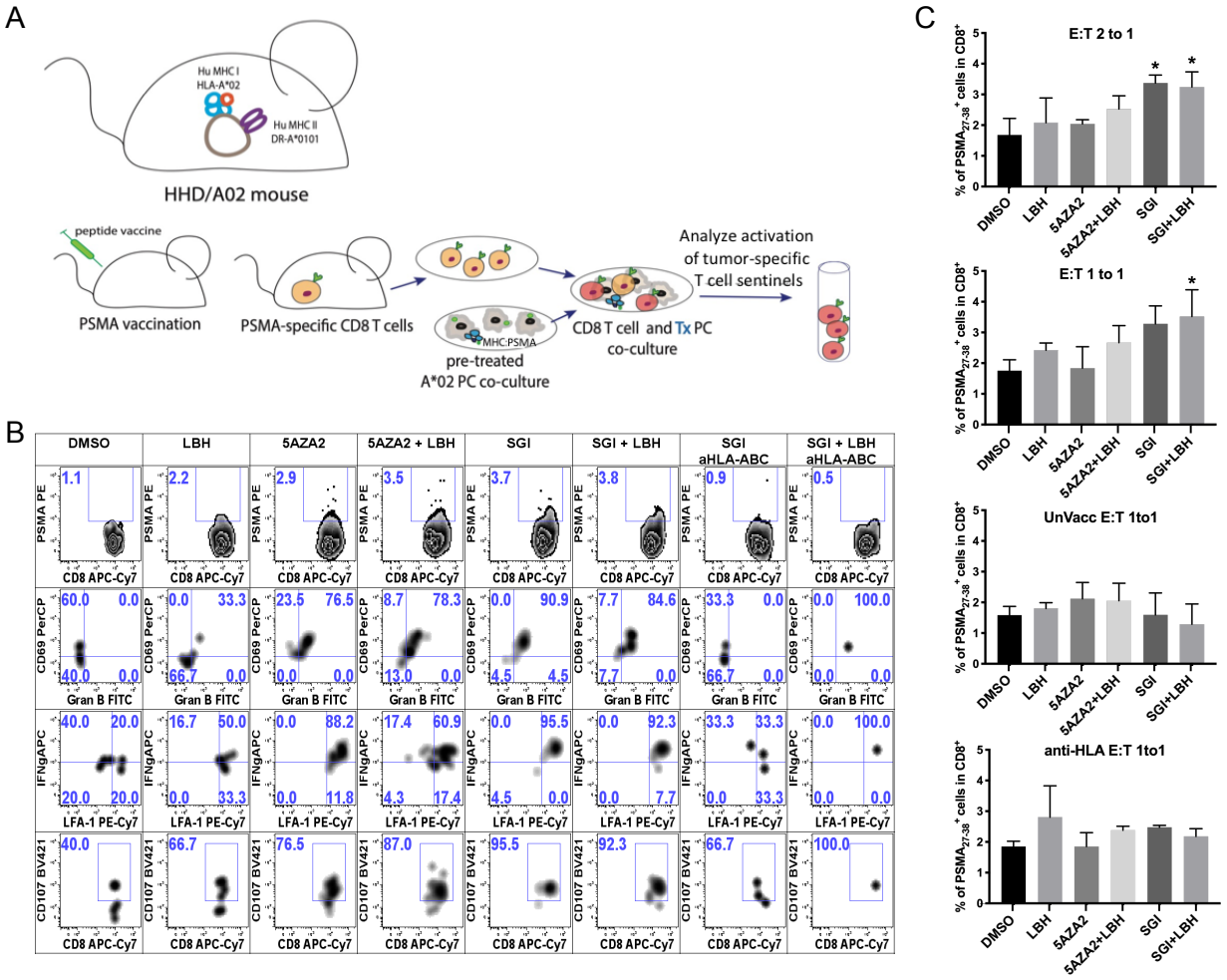


Figure 2.10. DNMT and HDAC inhibition in tumor cells increases co-cultured T-cell activation. A) Vaccination and co-culture scheme to analyze $PSMA_{27-38}$-specific T-cell response to LNCaP cells treated with epigenetic modifying agents. B) Plots represent the rare-event cell analysis of $PSMA_{27-38}$-specific T-cells to assess the cytolytic features. Top row shows $PSMA_{27-38}$ tetramer binding on the total $CD8^+$ splenocyte subset. $PSMA_{27-38}$ $tetramer^+$ cells are projected in rows 2-4 and gate frequencies are expressed as a percentage of total $PSMA_{27-38}$ $CD8^+$ T-cells. Plots represent expression of activation markers CD69, LFA-1, IFNγ, granzyme B, and CD107 (LAMP1). C) Graphs represent the frequency of $PSMA_{27-38}$ pentamer positive cells within the total $CD8^+$ T-cell population after co-culture with LNCaP cells treated with epigenetic modifying agents. Splenocytes were co-cultured with LNCaPs in ratios of 2:1 and 1:1 $PSMA_{27-38}$ $tetramer^+$ $CD8^+$ T-cell Effector to tumor Target (E:T). In control conditions, splenocytes from unvaccinated mice were co-cultured with LNCaPs treated with epigenetic modifying agents or HLA-I as blocked by an anti-HLA-A,B,C blocking antibody. Error bars represent SD. $*$ $p<0.05$.

Chapter 3:

SEEMLIS: A Flexible and Semi-Automated Method for Enrichment of Methylated DNA from Low-input Samples

This chapter is adapted from the following publication under review: Rodems TS, Juang DS, Stahlfeld CN, Gilsdorf C, Krueger TE, Heninger E, Zhao SG, Sperger JM, Beebe DJ, Haffner MC, Lang JM. SEEMLIS: A Flexible Semi-Automated Method for Enrichment of Methylated DNA from Low-input Samples.

Contributions: Figure 3.5A was performed in collaboration with J.M.S. Figure 3.9 was performed in collaboration with C.N.S. and C.G. All other experiments and analyses were performed by T.S.R.

ABSTRACT

DNA methylation alterations have emerged as hallmarks of cancer and have been proposed as screening, prognostic, and predictive biomarkers. Traditional approaches for methylation analysis have relied on bisulfite conversion of DNA, which can damage DNA and is not suitable for targeted gene analysis in low-input samples. Here we have adapted methyl-CpG binding domain protein 2 (MBD2)-based DNA enrichment for use on a semi-automated exclusion-based sample preparation (ESP) platform for robust and scalable enrichment of methylated DNA from low-input samples, called SEEMLIS. We show that combining methylation-sensitive enzyme digestion with ESP-based, MBD2 enrichment allows single gene analysis with high sensitivity for *GSTP1* in highly impure, heterogenous samples. We also show that ESP-based, MBD2 enrichment coupled with targeted pre-amplification allows analysis of multiple genes with sensitivities approaching the single cell level in pure samples for *GSTP1* and *RASSF1* and sensitivity down to 14 cells for these genes in highly impure samples. We demonstrate the potential utility of SEEMLIS using circulating tumor cells from patients with prostate cancer. In summary, this novel assay provides a platform for determining methylation signatures in rare cell populations with broad implications for research and clinical applications.

INTRODUCTION

Epigenetic modifications to DNA are fundamental to human biology, including histone tail modifications, changes in chromatin structure, and DNA methylation. The ability of epigenetic modifications to alter gene expression without changing the sequence of the genome is essential to human development and disease (149). DNA methylation in particular has been widely studied for its contributions to human development and disease (150). DNA methylation refers to the addition of a methyl group to the fifth carbon of the nucleotide cytosine. In humans, cytosine methylation mainly occurs at cytosines that are 5' to a guanine, termed CpG (46). CpG methylation contributes to gene regulation directly by blocking binding sites of transcription factors or RNA polymerase and indirectly by recruiting other epigenetic modifiers that promote chromatin reorganization (46,47,151). Importantly, these changes are heritable and conserved, but can also be plastic in nature (152,153), making them attractive targets for studying disease progression and developing biomarkers and therapies.

DNA methylation changes are a hallmark of almost all malignancies (154). In prostate cancer for instance, alterations in DNA methylation have been extensively studied and have been shown to exhibit exquisite biomarker properties for early detection and disease monitoring (77,154-158). This approach is particularly relevant since there is a lack of recurrent genomic alterations in prostate cancer (62), with the most common genomic alterations only occurring in ~50% of patients (64,104-106). The lack of genomic alterations in prostate cancer generates a need for non-genomic prognostic and predictive biomarkers. During tumor progression, specific gene promoters and CpG islands are hypermethylated, leading to silencing of genes including tumor suppressor

genes such as *APC* and *RASSF1* (65,66,159). Importantly, many of these genes are methylated in more than 75% of patients, with some genes such as *GSTP1* being methylated in over 95% of patients with prostate cancer (65). Therefore, DNA methylation changes, especially in specific gene promoters, have the potential to be powerful prognostic and predictive biomarkers.

An increasing number of recent studies have investigated the feasibility of using DNA methylation signatures in DNA from liquid biopsies such as circulating tumor cells (CTCs) and cell-free DNA (cfDNA) as prognostic and predictive biomarkers. Studies in prostate cancer cfDNA and breast cancer cfDNA and CTCs, for example, have identified multiple candidate biomarkers that may be useful for early detection and diagnosis (160-163). However, robustly assessing DNA methylation changes from small amounts of DNA remains a major technical challenge. The overwhelming majority of these studies utilize bisulfite conversion based approaches such as methylation specific PCR (MS-PCR) or reduced representation bisulfite sequencing (RRBS). Bisulfite conversion has been shown to extensively damage DNA and can result in a loss of up to 90% of DNA yield (114). Due to the limited amount of tumor DNA in liquid biopsies, bisulfite conversion-based approaches are suboptimal and may not enable clinical diagnostic use.

To circumvent the issues that arise from bisulfite-based approaches, multiple studies have investigated affinity enrichment-based approaches using either antibodies or proteins that specifically bind methylated DNA. For instance, the use of 5-methylcytosine (5mC) or 5-hydroxymethylcytosine (5hmC) antibodies has been shown to enrich for methylated DNA from as little as 0.5ng of starting material (164,165). The methyl-CpG binding domain (MBD) of methyl-CpG binding proteins such as MeCP2 and

MBD2 have also been used to enrich methylated DNA (166-168). An assay termed combination of methylated-DNA precipitation and methylation-sensitive restriction enzymes (COMPARE-MS) employs the methyl-CpG binding domain of MBD2 (MBD2-MBD) coupled with DNA digestion with methylation-sensitive restriction enzymes to enrich methylated DNA in heterogenous samples. This assay was used to profile multiple genes from 20ng of DNA from prostate cancer biopsy samples (168). Here we sought to develop a robust and scalable enrichment platform for targeted analysis of methylated DNA from low-input samples. We combined the strength of the COMPARE-MS assay with semi-automated exclusion-based sample preparation (ESP) to enrich methylated DNA from low-input samples, in an assay called SEEMLIS. We developed SEEMLIS to enable methylated DNA analysis in low-input samples such as CTC-derived DNA.

Our lab already has a well-established platform for CTC capture, protein staining, and nucleic acid extraction both with a handheld and a semi-automated microfluidic ESP system (139,169,170). Both of these systems harness the rapid, gentle, and low-loss physical characteristics of the ESP system to enable extraction of high quality DNA from low-input samples, which can be used for virtually any DNA-based experimental endpoint. In this study, we show that DNA extracted from this system can be used directly for DNA methylation analysis by SEEMLIS. We validated the performance of SEEMLIS for single gene analysis using CTCs from patients with prostate cancer as our source of low-input target DNA. Additionally, we show that SEEMLIS followed by a pre-amplification step, can allow methylation analysis of multiple gene targets from low cell inputs, with sensitivity down to a single cell for certain targets in samples with low non-tumor cell contamination.

SEEMLIS was designed for seamless integration into our CTC capture, imaging, and nucleic acid extraction ESP-based system. However, this assay is applicable for use with any genomic DNA source where methylation marks and genomic integrity are sufficiently preserved. The clinical utility of this assay is far reaching, including the ability to monitor responses to epigenetic therapies and to identify patients with specific methylation signatures that are prognostic of disease outcomes or render them resistant or susceptible to certain therapies. Importantly, while we have developed this system in the context of prostate cancer, this assay is readily transferable to other types of cancer and diseases where DNA methylation changes are of importance

RESULTS

Range of detection of *GSTP1* promoter in LNCaP DNA after digestion with a methylation-sensitive restriction enzyme and MBD2-MBD enrichment

An overview of the SEEMLIS workflow in shown in Figure 3.1. SEEMLIS was designed for seamless integration into our existing ESP-based CTC isolation platforms. In this report, we will use these techniques for sample isolation, DNA extraction, and methylated DNA enrichment where applicable. However, SEEMLIS may be performed from step 2 of the workflow (MBD2-MBD Enrichment) with DNA isolated from samples of interest in any way that preserves methylation and does not significantly damage or shear DNA (168). In this study, DNA was extracted from lysed cells by magnetic silica coated beads using ESP-based approaches, which our lab has used extensively for nucleic acid isolation (139,170,171). To fragment DNA, a combination of methylation-sensitive and in-sensitive restriction enzymes was used as described previously (168). The addition of a methylation-sensitive restriction enzyme that only cuts if the restriction site is unmethylated further reduces unspecific background signal. Tagged MBD2-MBD protein was immobilized on cobalt coated magnetic beads to capture methylated DNA. Bead-bound DNA is purified by automated magnetic transfer via ESP through a wash buffer and into water for elution.

To test the range of detection for a single, methylated gene, we chose *GSTP1* as a benchmark. *GSTP1* promoter hypermethylation is present in more than 95% of prostate cancers and has previously been used in various biomarker studies (65,75,77,172). Methylation of the repetitive element, *LINE1*, was used as a positive control for successful enrichment of methylated DNA since *LINE1* is highly abundant in the human genome and

known to be methylated in virtually all human cells (173,174). We chose to use the LNCaP cell line as our positive sample. Bisulfite sequencing has previously been performed on the LNCaP cell line which shows that the promoter and CpG island of *GSTP1* is heavily methylated (168,175,176). White blood cells (WBC) from healthy donors and patients with prostate cancer were used to represent the WBC population in circulation as our negative control. It has been previously shown that *GSTP1* is hypomethylated in white blood cells (65,177). *GSTP1* primer location, genomic context, and methylation levels for LNCaP (176,178) and WBCs (177,178) are shown in Figure 3.2. For all LNCaP and WBC validation experiments, DNA was extracted by semi-automated ESP as described in the methods unless otherwise indicated.

To determine the sensitivity, specificity, and range of SEEMLIS for *GSTP1* in LNCaP cells, we generated serial dilutions of LNCaP DNA at 1000, 100, 10, and 1 cell(s) and used 1000 WBCs as a negative control. Enrichment was performed as described in the methods and Figure 3.1. DNA was digested with AluI and HhaI restriction enzymes prior to enrichment of methylated DNA with MBD2-MBD coated magnetic beads. Quantitative PCR (qPCR) was performed on enriched DNA using primers for *GSTP1* and *LINE1*. Raw Ct values were used to create ROC curves and calculate *GSTP1* and *LINE1* methylation index (MI) for each dilution of LNCaP cells and WBC samples. An ROC curve was created using all LNCaP and WBC values, which had an area under the curve (AUC) of 0.86 (Figure 3.3A). Youden's J statistic was used to find the optimal threshold (OT), which resulted in 81.25% sensitivity and 85.71% specificity for methylated *GSTP1* to distinguish LNCaP from WBC. This threshold was applied to the methylation index values calculated for each dilution of LNCaP and WBC samples (Figure 3.3B). We were able to

detect methylated *GSTP1* from LNCaP cells above the OT 8/8 times at the 1000 and 100 cell dilutions, 7/8 times at the 10 cell dilution and 3/8 times at the single cell dilution. GSPT1 was found above the OT 4/28 times in any WBC sample. *LINE1* methylation was detected in all samples in an input dependent manner, indicating successful enrichment of methylated DNA from all sample types (Figure 3.3B).

We also created an ROC curve using only LNCaP samples with greater than or equal to 10 cells to demonstrate the increased sensitivity when not working with single cell samples (Figure 3.3A). This curve results in an AUC very close to 1.0 (0.98). A similar threshold under these conditions results in an improvement in sensitivity (95.83%), but no improvement to specificity. The threshold determined to be optimal by Youden's J statistic for this sample set is more restrictive, but with higher sensitivity and specificity of 87.50% and 96.43% respectively. The ROC curve generated with data from LNCaP cell inputs of greater than or equal to 100 cells has an AUC of 1.0 with 100% sensitivity and specificity at the calculated optimal threshold (Figure 3.3A).

To determine the efficiency of SEEMLIS and calculate a detection limit, we plotted the *GSTP1* Ct values against *LINE1* Ct values and cell input number. *GSTP1* and *LINE1* Ct values were positively correlated to each other across the various dilutions of LNCaP cells (Figure 3.3C). Similar relative enrichment of *GSTP1* and *LINE1* in individual enrichment reactions suggests stochastic variations in how much input DNA is added from the serial dilutions is responsible for differences in *GSTP1* enrichment from the same dilution, rather than inefficient capture of methylated DNA. Importantly, *GSTP1* is detected in MBD2-MBD enriched LNCaP DNA in an input dependent manner, where each 10-fold dilution of starting cells results in a gain of 3.42 Ct values as determined by the

slope of the best fit line when plotting raw Ct values vs. log cell input, resulting in an assay efficiency of 96.06% (Figure 3.3D). The equation of the best fit line was used to calculate a detection limit for GSTP1 in this assay based on the OT from Figure 3.3B. The detection limit based on these values is 3 cells.

Range of detection of *GSTP1* promoter in heterogenous samples after methylation-sensitive restriction enzyme digestion and MBD2-MBD enrichment

Next, we wanted to determine range of detection of *GSTP1* in samples representative of the purity of DNA isolated from circulating tumor cells (CTCs). DNA collected form CTC samples is often contaminated with large amounts of WBC DNA due to persistence of WBCs in the sample even after CTC enrichment. To test the ability of SEEMLIS to detect *GSTP1* in impure, heterogenous populations, we generated contrived samples to be representative of the purity levels we may obtain from CTC samples. We used LNCaP cells to represent CTCs and spiked them into patient-derived WBCs. We spiked approximately 1000, 100, 10, and 1 LNCaP cells by serial dilution into 1000 WBCs. These spike-ins represent a range of purity levels from CTCs being 50% of the population to “worst case scenario” purity levels, where CTCs are less than 1% of the cell population. The input amounts are representative of the range of CTC numbers we may get from patient samples, although we also observe higher numbers of CTCs and greater purity in certain patients. We were able to detect methylated *GSTP1* from the contrived sample sets at similar levels as LNCaP cells alone at each dilution (Figure 3.3E). Methylated *GSTP1* was detected from LNCaP cells spiked into 1000 WBCs above the OT 4/4 times at the 1000 LNCaP cell dilution, 3/4 times at the 10 LNCaP cell dilution and 2/4 times at

the single LNCaP cell dilution. Detection of *LINE1* methylation corresponded to the total cell number present in the sample for each of these spike-in experiments.

We also wanted to confirm that different levels of WBCs would not introduce more false positive *GSTP1* signals compared to 1000 WBC inputs, since circulating tumor cell samples may have varying levels of WBCs. We measured *GSTP1* and *LINE1* detection in WBC inputs of 5000, 2000, 100, 10, and 1 cells to confirm that *GSTP1* signal was consistently low in various WBC input amounts (Figure 3.4). The *GSTP1* signal in WBCs did not increase significantly at inputs greater than 1000 cells. *GSPT1* signal was below the detection limit 100% of the time in dilutions lower than 1000 cells. *LINE1* was detected relative to input amount up to 2000 cells, where saturation of *LINE1* signal occurred. Therefore, in these assay conditions, interpretation of *LINE1* as a readout of total cellular level is only applicable to inputs of 2000 cells or less.

Detection of methylated *GSTP1* in prostate cancer CTCs

A cohort of patients was selected to test the performance of SEEMLIS on patient CTC samples. The samples in this cohort were collected by EpCAM positive selection following CD45+ depletion using a handheld ESP device called the VERSA where RNA was extracted prior to DNA extraction (171). The VERSA technology integrates multiple analysis steps into one microfluidic device, including capture, staining, and imaging of CTCs followed by nucleic acid extraction from the same cells. The RNA from the samples in this cohort was used to assess gene expression of various prostate specific markers to ensure the presence of CTCs in the sample. All samples had gene signatures that support the presence of prostate epithelial cells by expression of multiple prostate specific

genes (Figure 3.5A). Estimates of CTC and WBC numbers and sample purity were made from images taken prior to nucleic acid extraction, where CTCs were considered EpCAM and Hoechst positive and negative for WBC exclusion markers CD45, CD14, and CD44b (Table 3.1). Sample 274 had a purity greater than 50%. Sample 411 did not have an image available. All other samples ranged in purity from 1.0%-10.2%. CTC number ranged from 7 to 237.

DNA obtained from the patient samples was enriched for methylated DNA by MBD2-MBD and qPCR was performed for *GSTP1* and *LINE1* using the enriched DNA. Methylation index was calculated for each sample and the OT determined for *GSTP1* in LNCaP cells was applied. We were able to detect methylated *GSTP1* above the OT in 5/6 (83%) samples (Figure 3.5B). *LINE1* methylation was detected in all samples indicating methylated DNA was captured even in sample 578 where methylated *GSTP1* was not detected. *GSTP1* methylation index was significantly correlated to the CTC number determined by imaging (Figure 3.5C). However, *LINE1* methylation index did not significantly correlate to total cell number. This may be due to inaccuracy in determining total cell number from live cell images or loss of *LINE1* methylation seen during prostate cancer progression leading to *LINE1* methylation levels not correlating with cell number (67). This pilot study shows that SEEMLIS can successfully detect methylated *GSTP1* from a range of input target cell numbers and sample purities for seamless integration of targeted methylated DNA analysis into CTC capture, imaging, and gene expression analysis to allow comprehensive biomarker evaluation.

SEEMLIS performance without the use of a methylation-sensitive enzyme to facilitate multiple target detection

While detection of one gene is useful in many contexts, there are myriad reasons for wanting to detect methylation in multiple gene targets from the same low-input sample. Splitting the sample into separate reactions for each gene is not ideal when working with heterogenous, low-input samples. Multiplexing primers in a pre-amplification reaction would allow analysis of multiple genes from the same limited sample, while still allowing a single methylated DNA enrichment per sample. However, due to the complexity of methylation patterns in each different gene region that we may want to include, the use of a methylation-sensitive enzyme in the assay poses a problem.

A methylation-sensitive enzyme will cut at its restriction site only if the DNA is unmethylated. Designing primers that contain this restriction site can improve background from non-specifically captured DNA by preventing amplification during qPCR. However, in the context of evaluating methylation at multiple targets in low-input samples, each restriction site for this enzyme is required to be methylated in each gene in the list of targets to be analyzed. While the use of the methylation-sensitive enzyme is beneficial for reducing the amount of background unmethylated target DNA, finding a methylation-sensitive restriction enzyme that is compatible with each region of interest becomes prohibitively difficult as the number of targets increases. Therefore, depending on the chosen genes, it may be necessary to use only non-methylation-sensitive restriction enzymes for detection in the same low-input sample.

We tested the effect of not using a methylation-sensitive enzyme on *GSTP1* detection in LNCaP and WBCs pre- and post-MBD2-MBD enrichment (Figure 3.6). Samples were digested with the methylation-sensitive enzyme, HhaI or the non-methylation-sensitive enzyme, HpyCH4V. All samples were also digested with the non-methylation-sensitive enzyme AluI. We chose to add an additional non-methylation-sensitive enzyme to replace HhaI to increase the number of cut sites in our target regions. This ensures that the fragments are small enough to mitigate false positives from enrichment of methylated regions far away from the primer site. We found that while the use of the methylation-sensitive enzyme reduces signal from WBCs, the signal without the enzyme is still low enough to warrant use in heterogenous populations, if we apply a higher limit of detection to limit false positives. The combination of enzyme types should be empirically determined by the researcher for each application of this assay based on the desired target regions.

Detection of multiple targets by pre-amplification of MBD2-MBD enriched DNA

In order to detect methylation at multiple genes from low cell input samples, we tested the addition of a pre-amplification step prior to qPCR. We used the TaqMan system for targeted pre-amplification, which amplifies pre-selected targets with the same probes used in subsequent qPCR-based analysis. We chose this targeted method for pre-amplification rather than a traditional whole genome amplification (WGA) method for its speed, reduced hands-on time, and flexibility. Additionally, because we are analyzing a known subset of genes, amplifying the entire genome increases the risk of unpredictable biased amplification, which may be exacerbated by the uneven fragment lengths

produced by restriction enzyme digestion. We have published multiple studies using the TaqMan system of pre-amplification for gene expression and have done extensive testing in this context to ensure pre-amplification is not affecting the results of our experiments (139,171,179). In order to determine that the pattern and amount of pre-amplification of our DNA primers is similar across our genes of interest, we compared Ct values for each primer set using 5ng and 0.5ng of pre-amplified and unamplified LNCaP DNA (Figure 3.7A). The amplification pattern is similar across all primer sets after pre-amplification and at different starting concentrations. *GSTP1* gained approximately 10 Ct values. *RASSF1*, *APC*, and *RARB* gained approximately 12 Ct values.

To test the performance of SEEMLIS with the pre-amplification step included, we generated serial dilutions of 1000, 100, 10, and 1 LNCaP cell(s) and used 1000 and 100 patient-derived WBCs as a negative control. We also spiked 1000, 100, 10, and 1 LNCaP cell(s) into 1000 patient-derived WBCs to mimic CTC samples. The samples were enriched for methylated DNA by MBD2-MBD capture. The enriched DNA was then placed directly into a pre-amplification reaction. We included primers for four genes in the pre-amplification pool (*RASSF1*, *APC*, *RARB* and *GSTP1*) which have previously been identified as being methylated in a large percentage of prostate cancers (65). Primer locations, genomic context, and methylation level determined by whole genome bisulfite sequencing in LNCaP (176) and WBCs (177) are shown in Figure 3.2. It is important to note that *RARB* is methylated at a low level in WBCs, which is likely to result in a more restrictive detection limit for this gene. *LINE1* was left out of the pre-amplification pool because it is abundant enough to be detected in samples that have been diluted after the pre-amplification without pre-amplification of *LINE1* itself. Pre-amplified DNA was then

diluted 1:5 and qPCR was performed for *GSTP1*, *RASSF1*, *APC*, *RARB*, and *LINE1*. Ct values from all LNCaP and WBC samples, excluding spike in samples, were plotted together for each gene to determine cut off Ct values to be used for methylation index calculations, as described in more detail in the Methods (Figure 3.7B). Ct values from the 1000 cell WBC samples and LNCaP samples of 10 cells or greater were used to create ROC curves for each gene (Figure 3.8A). The 1 cell LNCaP dilution samples were not included in the ROC curves because of the increased background we expected from not using a methylation-sensitive enzyme for digestion. The AUCs for *GSTP1*, *RASSF1*, *APC*, and *RARB* were 0.88, 0.84, 0.83, and 0.67 respectively. Youden's J statistic was used to find the OT for each gene, which are listed in Figure 3.8A.

Methylation index was calculated for each gene as described in the Methods section for pre-amplified samples and the OTs calculated for each gene in Figure 3A were applied (Figure 3.8B). We were able to detect methylated *GSTP1* and *RASSF1* from LNCaP cells above the OT 4/4 times at the 1000 and 100 cell dilutions and 2/4 times at the 10 cell dilution. We were able to detect methylated *APC* above the OT 4/4 times at the 1000 and 100 cell dilutions and 1/4 times at the 10 cell dilution. *RARB* was only detected above the optimal threshold 4/4 times at the 1000 cell dilution and 2/4 times at the 100 cell dilution. *GSTP1*, *RASSF1*, and *APC* were detected above the OT 4/4 times at the 1000 and 100 cell dilutions for the spike in samples. *RARB* was detected above the OT 4/4 times at the 1000 cell dilution and 3/4 times at the 100 cell dilution. Only *RASSF1* was ever detected above the optimal threshold at the 10 cell dilution for the spike-in samples (1/4 times). As expected, we were not able to detect any gene above the optimal thresholds at the 1 cell dilution for LNCaP alone or the spike-in samples. *LINE1*

methylation was detected at all dilutions in all samples in a dilution dependent manner. *LINE1* Ct values were significantly positively correlated with Ct values for each gene (Figure 3.8C). Methylated DNA was enriched in an input dependent manner for inputs of 1000, 100, and 10 cells for each gene, with efficiencies for *GSTP1*, *RASSF1*, *APC*, and *RARB* of 87%, 101%, 92%, and 122% respectively (Figure 3.8D). Undetected samples were excluded from efficiency analyses. Detection limits listed in Figure 3.8A were calculated based on the OT using the equation of the best fit line determined by plotting Ct values vs. cell input (Figure 3.8D). *GSTP1* and *RASSF1* had a detection limit of 14 cells while *APC* and *RARB* had detection limits of 24 and 152 cells respectively.

We also performed SEEMLIS on 100 WBCs in addition to 1000 WBCs to see if lowering the amount of background cells would reduce background signal (Figure 3.8B). The *GSTP1* and *RASSF1* signals from 100 WBCs were reduced to effectively 0. *APC* signal was reduced to the equivalent level of a single LNCaP cell. *RARB* signal was not reduced between 1000 and 100 WBCs. These results indicate that reducing the amount of background WBCs in the sample may lower the threshold enough to approach single cell sensitivity for certain genes.

Detection of multiple targets in pre-amplified MBD2-MBD enriched DNA from cells isolated by single cell aspiration

These data suggest that the lack of a methylation-sensitive enzyme and addition of a pre-amplification step can allow us to look at multiple gene targets in CTC samples where WBCs are present, but only in certain high purity or high CTC-burden samples. In order to look at multiple gene targets in samples where purity or CTC-burden is low, further

purification steps are warranted. One way to achieve 100% CTC purity is to use a single-cell aspirator to select only CTCs or only WBCs. We tested the feasibility of this by using a semi-automated single cell aspirator (SASCA) developed in our lab (180). We tested the feasibility of this in 3 prostate cancer CTC samples that were first enriched for EpCAM positive cells as well as LNCaPs and WBCs as positive and negative controls respectively. An example image of a single WBC in the single cell aspirator microarray is shown in Figure 3.10A. We aspirated 2-3 groups of approximately 10-15 CTCs, LNCaP cells,, and WBCs. We then performed SEEMLIS without the use of a methylation-sensitive enzyme with the pre-amplification step included on each of these samples. Methylation index was calculated for each gene in each group (Figure 3.9B). We were able to detect *LINE1* from each sample at comparable levels to each other and to our 10 cell dilution samples in Figure 3.8. We were able to detect *GSTP1* from 1 LNCaP sample and in 1 CTC group from Pt. 1. *RASSF1* was detected in both LNCaP control groups as well as 2/2 CTC groups from patient 1 and 2/3 CTC groups from patient 2. *APC* was detected in 1/2 LNCaP groups, but was not detected in any CTC groups. *RARB* was detected in all CTC groups from Patients 2 and 3 and in 1/2 CTC groups from patient 3. None of the genes, apart from *LINE1*, were detected in the aspirated WBC group. Generating pure samples by single cell aspiration allowed us to perform a multiplexed analysis of methylation signatures in prostate cancer CTCs. These data demonstrate the utility of this assay for targeted methylation analysis in CTCs and other low-input samples.

DISCUSSION

Analyzing epigenetic signatures in low-input samples such as DNA derived from liquid biopsies has been an on-going challenge. Most techniques that measure DNA methylation rely on bisulfite conversion of DNA, which can lead to extensive damage of input material. This problem necessitates large input amounts, or reduces sensitivity such that targeted analysis of specific regions is not reliable. This loss of sensitivity is especially true in heterogenous samples due to the increased need for assay sensitivity to distinguish target DNA from background DNA. The assay presented here uses MBD2-MBD enrichment of methylated DNA in a semi-automated ESP based system to support targeted DNA methylation analysis in low-input samples without chemical DNA alteration or lengthy protocols.

While we have developed SEEMLIS in the context of prostate cancer and specifically with CTC biomarker development in mind, this assay is readily transferable for targeted methylation analysis and biomarker development in any disease state, including blood and solid tumors and non-cancer diseases. This assay also provides a method for analysis of single cell heterogeneity, which may be useful not only for research in disease progression, but also for non-disease states such as development and cellular differentiation.

SEEMLIS also lends itself to further downstream applications such as sequencing the MBD2-MBD captured DNA (MBD-seq) and we are currently exploring this endpoint. There have been multiple studies that have performed reduced representation bisulfite sequencing (RRBS) from low-input samples including single cells, but this necessitates using sequencing methods that even further reduce sequence representation than when

performed with thousands of cells (181,182). Furthermore, because the damage caused by bisulfite conversion is a random event, analysis of specific gene loci at the single cell level is reliant on chance that the site of interest remains intact, is converted efficiently, and is able to be mapped successfully post-sequencing. This assay provides a way to alleviate some of these challenges by enabling enrichment of methylated DNA from low-input samples without chemically altering or damaging the DNA. While previous studies have demonstrated success in MBD-seq on samples inputs as low as 5ng (183), we anticipate that our semi-automated capture assay will be able to lower the minimum input even further based on the results in this study, especially in pure samples. As such, the clinical and research utility of this technique is incredibly far-reaching.

A limitation of capture-based assays such as this one is the lack of single nucleotide resolution. As of now, bisulfite or enzymatic based conversion techniques must be used in order to achieve this level of resolution. However, resolution at the single base level is often not needed for clinical tests or to understand the contribution of DNA methylation to disease progression or heterogeneity. SEEMLIS is not meant to reveal the transcriptional nuances that may arise from alterations in methylation at specific CpG sites, but rather is designed to sensitively and specifically identify regions of DNA methylation in low-input samples. Identification of differences in methylated regions by this assay may then lead to more in depth, higher resolution studies of DNA from larger samples.

This adaptability of SEEMLIS to many research-based and clinical needs, including highly sensitive, single gene analysis and analysis of multiple genes from single or multiple cells, makes it an exciting tool for methylation analysis in various settings. The

ability to quickly and reliably analyze epigenetic biomarkers is important for researchers and clinicians due to revelations over the last few decades on epigenetic influence on development and disease.

MATERIALS AND METHODS

Cells for Assay Validation

LNCaP cells (ATCC) were a gift from Dr. David Jarrard and were cultured in RPMI medium (Corning) supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin (HyClone). LNCaP cells were harvested when confluent and frozen in aliquots in growth medium plus 10% DMSO (Fisher Scientific) at -80^{o} C and quick thawed for use in assay validations experiments. White blood cells (WBCs) for assay validation experiments were derived from healthy donor blood or from the blood of a patient with prostate cancer. WBCs were selected on CD45 positivity using magnetic LS MACS columns (Miltenyi). WBCs were frozen in aliquots in PBS plus 10% DMSO (Fisher Scientific) at -80^{o} C and quick thawed for use in assay validations experiments.

DNA Extraction

Semi-automated DNA extraction was performed on a Gilson PIPETMAX liquid handling robot enabled for exclusion-based sample preparation (ESP), termed EXTRACTMAX (184). LiDS buffer (90mM Tris-HCL, 500mM lithium chloride, 1% Igepal CA-630, 10mM EDTA, 1mM dithiothreitol) and MagneSil Paramagnetic Particles (PMPs) (Promega) resuspended in GTC buffer (10 mM Tris-HCl, 6 M guanidinium thiocyanate, 0.1 % Igepal CA-630, pH 7.5) are added to the EXTRACTMAX extraction microplate (Gilson) by the robot. Cells were added to the microplate well containing LiDS, GTC, and MagneSil beads and mixed by the robot. Cells were allowed to lyse and DNA was allowed to bind to MagneSil PMPs for 5 minutes. The robot then transferred the MagneSil PMPs with bound DNA by exclusive liquid repellency (ELR) through one PBST (PBS containing 0.1%

Tween-20) wash, one PBS wash, and into water for elution. Beads were manually resuspended in the elution well and allowed to elute for 2 minutes. The MagneSil PMPs are magnetically transferred out of the elution well, leaving eluted DNA in water. LNCaP and WBC DNA for restriction enzyme and primer pre-amplification validation experiments was extracted using the AllPrep DNA/RNA mini kit (Qiagen) according to manufacturer's instructions.

Restriction Enzyme Digestion

DNA is digested using 1 µL of each chosen restriction enzyme (AluI at 10 units/µL, HhaI at 20 units/µL, and/or HpyCH4V at 5 units/µL; NEB) in 20 µL reactions containing 1x CutSmart Buffer (NEB) for 15 minutes at 37 ºC followed by enzyme inactivation for 20 minutes at 80 ºC.

Methylated DNA Enrichment

25 µL of TALON magnetic beads (Takara) were washed 3x with 100 µL 1x Binding Buffer (BB) (4% glycerol, 1 mM $MgCl_2$, 0.5 mM EDTA, 120 mM NaCl, 2 mM Tris-HCl pH 7.4, 0.2% Tween-20, and 0.5mM DTT). Washed beads were resuspended in 100 µL MBD2-MBD Coupling Buffer (1x BB, 1x Halt protease inhibitor cocktail (Thermo Scientific), 500 ng Unmethylated Lambda DNA (Promega), 5 µL tagged MBD2-MBD (EpiXplore Kit, Takara)) and placed on shaker at RT for 1 hour to bind MBD2-MBD to the TALON beads. MBD2-MBD bound beads were washed 3x with 100 µL 1x BB and resuspended in 88 µL 1x BB with 1x Halt protease inhibitor cocktail and added to 20 µL restriction enzyme digested DNA in 200 µL PCR tubes. This reaction was placed on a shaker at RT for 3

hours to bind methylated DNA to MBD2-MBD conjugated TALON beads. PCR tubes were placed onto the Gilson PIPETMAX liquid handling robot (EXTRACTMAN system enabled for ESP as previously described (184)) for washing and elution steps. The robot transferred the whole volume from the PCR tubes onto the EXTRACTMAX extraction microplate (Gilson) and then magnetically transferred the TALON beads through a wash containing 1x BB with 1x Halt protease inhibitor cocktail and into water for elution. The whole elution volume including beads was manually pipetted into new 200 µL PCR tubes and placed in a thermocycler at 95ºC for 15 minutes to ensure complete elution of methylated DNA. If pre-amplification of captured DNA was being performed, the elution volume was manually pipetted into new 200 µL PCR tubes containing the pre-amplification reaction mix and placed directly into the thermocycler under pre-amplification cycling conditions. The resulting eluate was used in downstream applications. Volumes indicated are per reaction.

Quantitative Real Time PCR and Pre-amplification

Quantitative PCR was performed using TaqMan hydrolysis probes (Applied Biosystems) and iTaq Universal Probes Supermix (Bio-Rad). Custom primer pairs and FAM dye-labeled TaqMan MGB probes were used for methylation analysis. Primer and probe sequences are listed in Table 3.2. Commercially available probes from Applied Biosystems were used for gene expression analysis for *AR*_Total (Hs00907242_m1), *TMPRSS2* (Hs01120965_m1), *KLK2* (Hs_00428383_m1), *FOLH1* (Hs00379515_m1), *NKX3.1*(Hs00171834_m1), *POLR2A* (Hs00172187_m1), and *RPLP0* (4333761F). Cycling conditions: 5 minutes at 95 ºC for initial denaturation and enzyme activation

followed by 45 amplification cycles of 5 seconds at 95 ºC and 30 seconds at 60 ºC. Pre-amplification was performed using custom hydrolysis probes and TaqMan PreAmp Master Mix (Applied Biosystems) when indicated according to manufacturer specifications. Cycling conditions: 10 minutes at 95 ºC for enzyme activation followed by 14 cycles of 95 ºC for 15 seconds and 60 ºC for 4 minutes. Pre-amplified samples were diluted 1:5 with TE buffer.

Single Cell Aspiration

For LNCaP and WBC aspiration, cells were first diluted to 100 cells/µl with PBS. For CTC aspiration, CTCs were first enriched using the VERSA platform as described above were stained in the VERSA with Hoechst 33342 (Thermo Fisher) and anti-bodies to EpCAM conjugated to PE (Abcam) and exclusion markers: CD27, CD45, CD34, and CD11b each conjugated to AlexaFluor 647 (BioLegend). Single cell aspiration was performed using a custom *semiautomated single-cell aspirator* (SASCA) platform as previously described (180). PDMS microarrays were prepared as previously described and adhered to a cleaned glass microscope slide. For diluted LNCaP and WBC samples, 6 µl cells were seeded for a total of ~600 cells per microarray. Stained CTC samples were seeded directly from the VERSA into the microarray. The microarray was imaged on a Nikon Ti-E Eclipse inverted fluorescent microscope, and target cells were identified by phenotypic staining analysis for CTCs or brightfield imaging for LNCaP and WBC control groups. CTCs were identified as EpCAM positive, exclusion (CD45/CD34/CD11b/CD27) negative cells, whereas WBCs were classified as EpCAM negative, exclusion positive cells. Groups of 10-15 CTCs, LNCaPs, and WBCs were aspirated from microwells and

dispensed directly into 10 µL PBS in the extraction plate for DNA extraction. Images of the microarray were analyzed using NIS Elements AR Microscope Imaging Software (NIS-Elements, RRID:SCR_014329) to obtain HLA-I mean fluorescent intensity (MFI) values.

Whole Blood Processing and CTC Capture

Blood is collected and processed as previously described (171,185). Briefly, whole blood collected by venipuncture into EDTA tubes was separated by centrifugation with Ficoll-Paque PLUS (Fisher Scientific). The layer containing mononucleated cells was depleted of CD45+ cells by magnetic LS MACS columns (Miltenyi). CTCs were isolated using an anti-EpCAM goat polyclonal antibody (R&D Systems). RNA was isolated using oligo (dT) Dynabeads (Invitrogen) and DNA was isolated using MagneSil Paramagnetic Particles (PMPs) (Promega) as previously described (170).

Methylation Index Calculation

Methylation Index (MI) was calculated by the delta Ct method using the max cycle value (MCV) as the “control” Ct value:

$$\text{Methylation Index} = 2^{-(Ct-MCV)}$$

MCV for *GSTP1* (no pre-amplification) is 45. Unamplified wells were given the MCV as their Ct value for analysis. MI of 1 is then interpreted as no enrichment of methylated DNA for the target. MCV for each gene for pre-amplified samples was assigned for each gene by determining a Ct cutoff where serially diluted replicates are no longer reliably detected or where clustering of negative controls is seen (Figure 3.7B). MCV for pre-amplified

GSTP1 was set at 35. MCV for pre-amplified *RASSF1* was set at 33, *APC at 28*, and *RARB* at 30. MCV for *LINE1* in either condition is 45.

Statistical Analysis

Receiver operator characteristic (ROC) curves were generated for each gene using Prism 8 (GraphPad) by plotting sensitivity vs. 100-specificity for the raw Ct values of LNCaP (true positive) and WBC (true negative). Optimal Threshold (OT) values were determined using Youden's J Statistic which is defined as the maximum value achieved from subtracting 100 from the sum of the sensitivity and 100-specificity values (in percentages). The associated Ct value was then converted into a methylation index (MI) using the MCV for each gene as described in the methods. Area under the curve (AUC) with 95% confidence intervals were found and reported, which indicate the probability that a randomly selected true positive sample will have a greater MI value than a randomly selected true negative sample. Simple linear regression and semi-log non-linear fit analyses were performed in Prism 8 and r, R^2 and slope are reported where relevant to data interpretation. Assay efficiencies (E) were calculated using slopes of best fit lines when comparing raw Ct values to log cell input as follows:

$$E = -1+10^{-\frac{1}{slope}}$$

All error bars represent SEM.

ACKNOWLEDGEMENTS

We would like to thank the patients who donated samples for this study. Figure 1 image was created using BioRender.com. Data for methylation levels in LNCaP and PBMCs and CpG island location is available on the GRCh37/hg19 genome assembly online at the UCSC Genome Browser.

Table 3.1. CTC, WBC, and Total Cell Counts for Patient Samples

Patient #	CTCs	WBCs	Total Cells	Purity (%)
274	237	242	479	49.5
408	72	1355	1427	5.0
581	7	238	245	2.9
411	NA	NA	NA	NA
501	28	2783	2811	1.0
578	9	79	88	10.2

Table 3.2. Custom TaqMan probe sequences for DNA methylation analysis

Gene	Primer	Sequence (5' - 3')
GSTP1	Fwd	TTCGCTGCGCACACTTC
	Probe	CGGTCCTCTTCCTGCTGTCTGTTT
	Rev	CTTTCCCTCTTTCCCAGGTC
RASSF1	Fwd	CCTCCAGAAACACGGGTA
	Probe	TTTGCGGTCGCCGTCGTTGT
	Rev	CTTCCTTCCCTCCTTCGTC
APC	Fwd	TTATTACTCTCCCTCCCACCTC
	Probe	TCTTGTGCTAATCCTTCTGCCCTGC
	Rev	TGGCAGTTGACACGCATAG
RARB	Fwd	GAAGGAGAACTTGGGATCTT
	Probe	CTAACCGGCTCGTTCGGACCTTT
	Rev	AGCCTGTAATTGATCCAAATGA
LINE1	Fwd	CGCAGGCCAGTGTGTGT
	Probe	CCGTGCGCAAGCCGA
	Rev	TCCCAGGTGAGGCAATGC

Figure 3.1. SEEMLIS workflow.

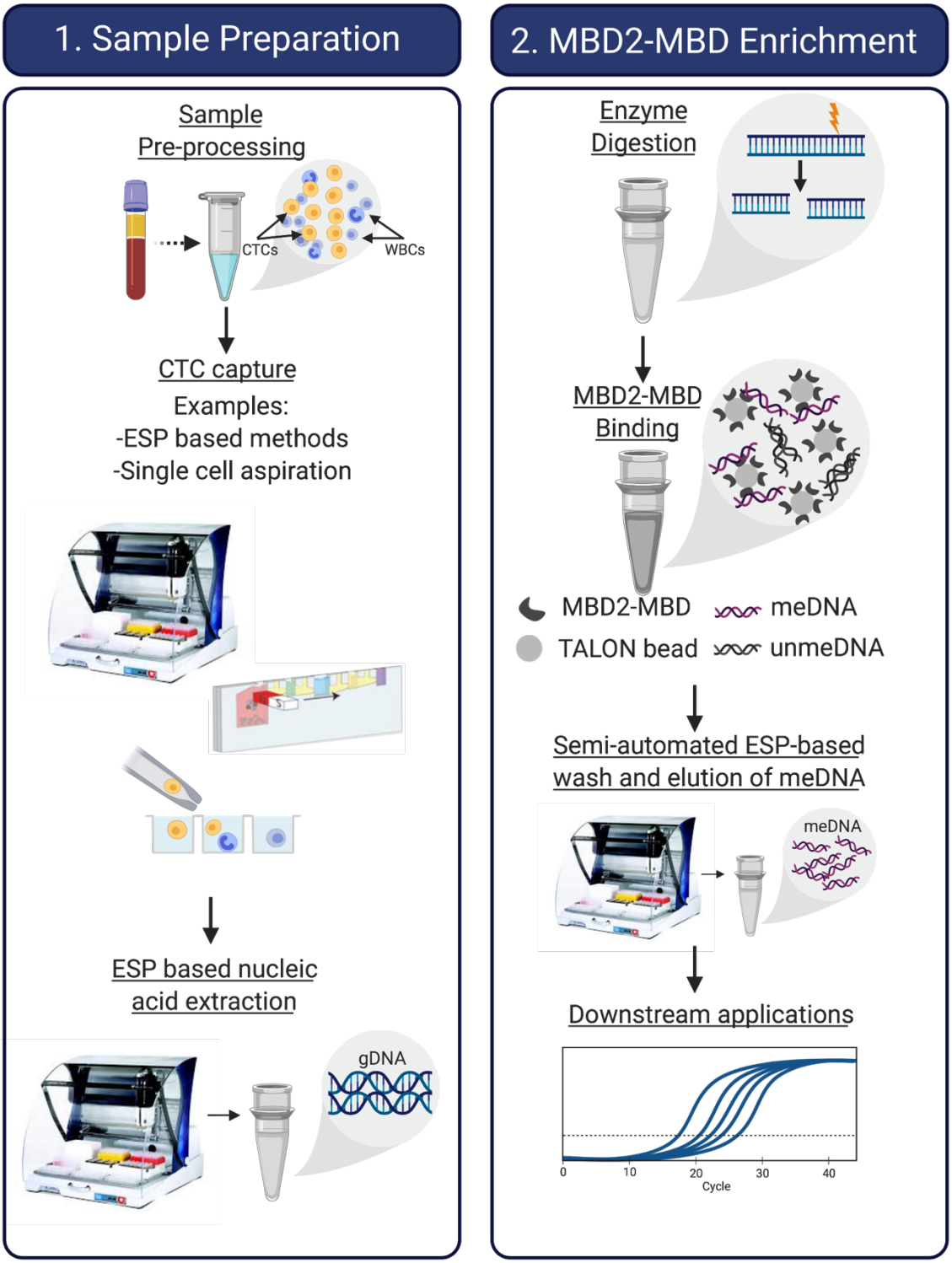


Figure 3.1. SEEMLIS workflow. Samples are pre-processed to enrich for circulating tumor cells. CTCs are captured by methods including ESP-based enrichment and single cell aspiration. Genomic DNA is extracted using ESP-based methods. Genomic DNA is digested with restriction enzymes and enriched using MBD2-MBD bound magnetic beads. Methylated DNA is washed and eluted on a semi-automated ESP-based system and used in downstream applications such as qPCR.

Figure 3.2. Primer locations and genomic context for *GSTP1*, *RASSF1*, *APC*, and *RARB*.

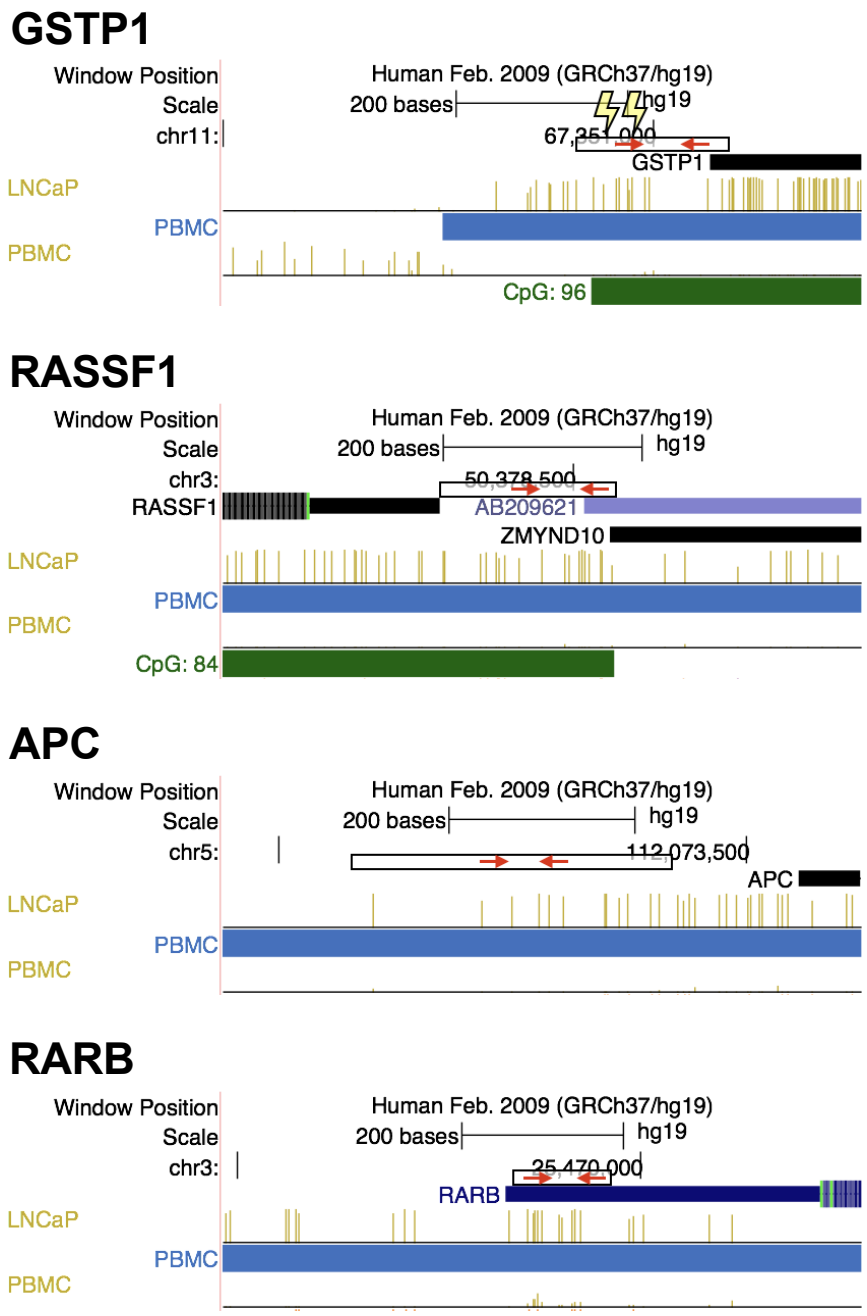


Figure 3.2. Primer locations and genomic context for *GSTP1*, *RASSF1*, *APC*, and *RARB*. Forward and reverse primer locations are shown for each gene (red arrows). The boxes encompassing the red arrows represent the location on either side of the amplicon where the closest restriction enzyme cut site for any of the enzymes used in the assay will cut. Lightning bolts are locations of methylation-sensitive restriction enzyme cut sites in *GSTP1*. Genomic context for each gene was generated using the UCSC genome browser using the GRCh37/hg19 genome. The locations of the CpG island for each gene if present is indicated by the green bar with the total number of CpG dinucleotides in the whole CpG island (not only what is present in image) indicated. LNCaP and WBC (PBMC) methylation data performed by whole genome bisulfite sequencing from studies deposited into the UCSC genome browser is shown by vertical gold bars (176-178).

Figure 3.3. Range of detection of *GSTP1* promoter in MBD2-MBD enriched DNA.

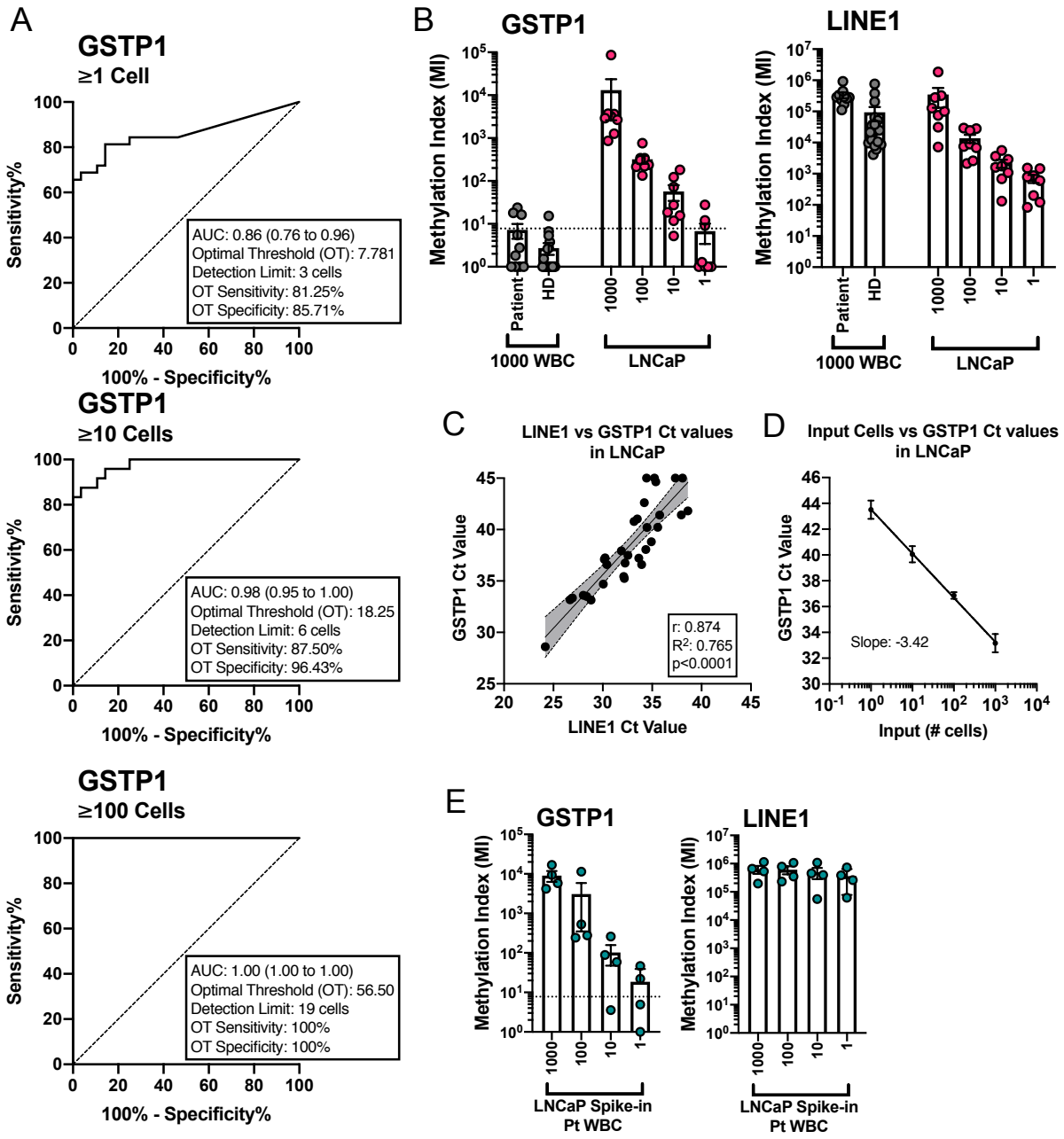


Figure 3.3. Range of detection of *GSTP1* promoter in MBD2-MBD enriched DNA. Methylated DNA was enriched by MBD2-MBD capture from DNA extracted from serially diluted LNCaP cells (n=8 per dilution) and 1000 patient-derived (n=10) or healthy donor (HD) (n=18) white blood cells (WBCs). Quantitative RT-PCR for *GSTP1* and *LINE1* was performed using enriched methylated DNA. A) ROC curves were generated for all LNCaP samples compared to all WBC samples (greater than or equal to 1 cell), only WBC samples greater than or equal to 100 cells, or only WBC samples greater than or equal to 10 cells. Area under the curve (AUC) with 95% confidence interval is indicated. Optimal threshold (OT) values determined by Youden's J statistic are listed with their associated sensitivity and specificity values. Detection limit was calculating using the slope of the best fit line of *GSTP1* Ct values plotted against cell input. B) Relative methylation was calculated by delta Ct relative to a max cycle value (MCV) of 45 with all undetected samples set to a Ct value of 45 for analysis. Optimal threshold as determined by ROC curve is shown as a dotted line. Each dot represents an individual sample taken from a pool of cells diluted to the indicated concentration. C) *LINE1* and *GSTP1* Ct values were plotted against each other for all LNCaP samples. A simple linear regression was performed to determine the best fit line and 95% confidence interval for that line (shaded region). R and R^2 values are listed for the correlation. D) *GSTP1* Ct values were averaged for each cell input and plotted against the cell input values. A semi-log non-linear fit was performed to determine the best fit line and the slope of the best of fit line (-3.42). Each ten-fold dilution of input should result in a gain of 3.32 Ct values, giving a slope of -3.32 for a perfect assay (100% efficiency). E) Relative methylation for serially diluted LNCaP

cells spiked into 1000 WBCs from a patient (n=8) is shown. Performed as described above for (B). All error bars represent standard error of the mean (SEM).

Figure 3.4. Detection of *GSTP1* promoter in additional dilutions of WBCs and limit of *LINE1* detection.

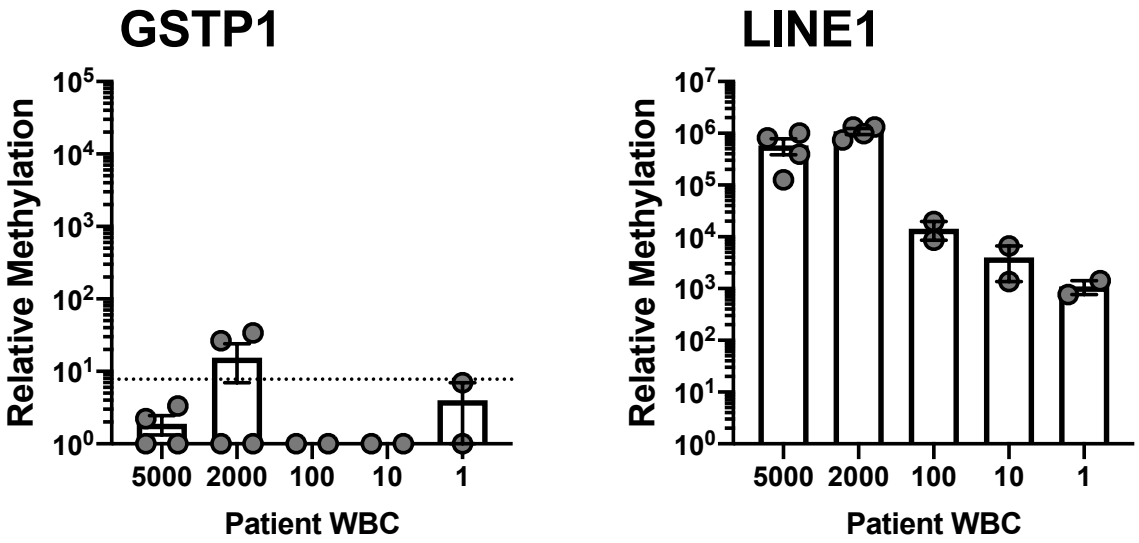


Figure 3.4. Detection of *GSTP1* promoter in additional dilutions of WBCs and lmit of *LINE1* detection. Additional dilutions of WBCs of 5000, 2000, 100, 10, and 1 cell(s) were created to determine the background *GSTP1* level and upper range of detection for the assay based on *LINE1* detection. *GSTP1* detection was not increased in larger dilutions compared to 1000 cells and was under the OT (dotted line) for all dilutions lower than 1000. *LINE1* values were dilution dependent for the 100, 10, and 1 cell(s) dilutions. The upper limit of range of detection is approximately 2000 cells.

Figure 3.5. Detection of *GSTP1* promoter in MBD2-MBD enriched DNA from prostate cancer CTCs.

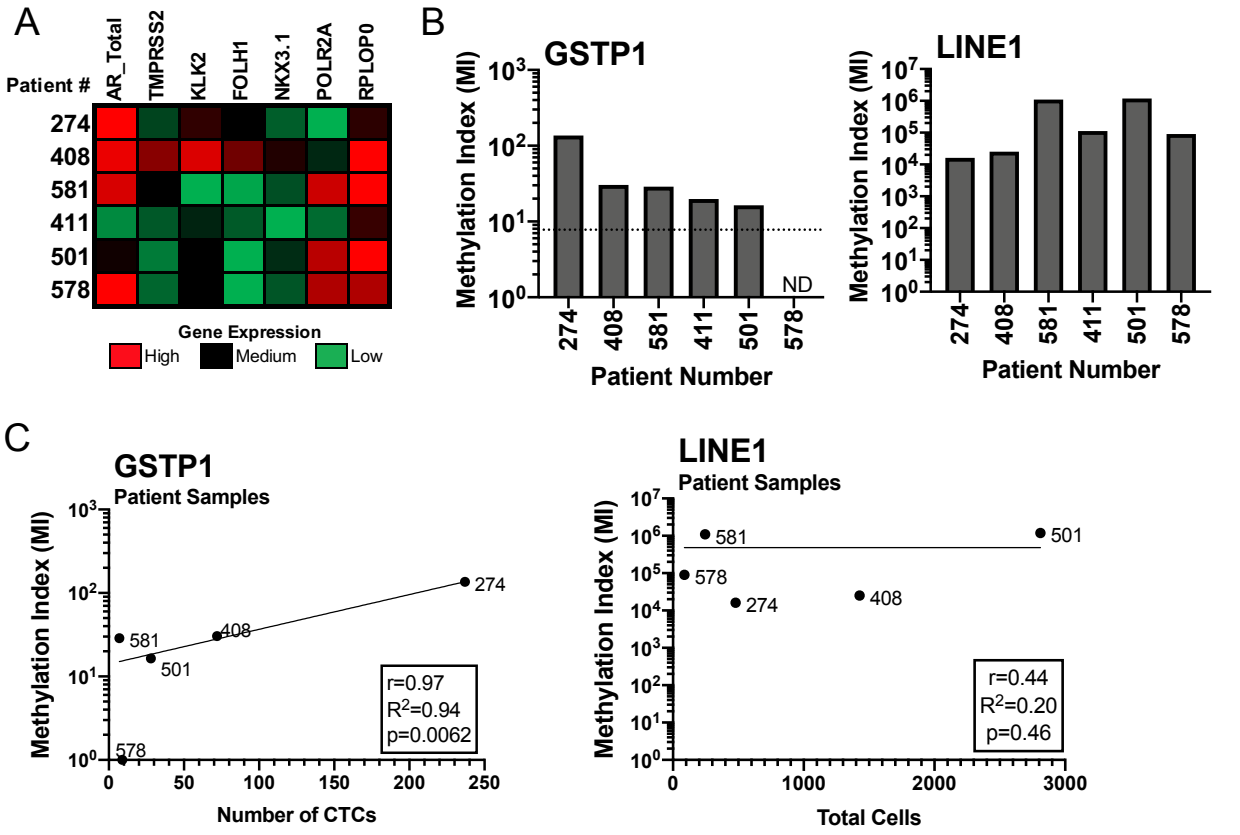


Figure 3.5: Detection of *GSTP1* promoter in MBD2-MBD enriched DNA from prostate cancer CTCs. Circulating tumor cells (CTCs) were enriched by positive selection for EpCAM using ESP. RNA and DNA were extracted from the EpCAM selected population following live cell, on chip imaging of selected cells. A) Gene expression was determined by qPCR for the indicated genes. Raw Ct values were used to create the heat map. Heat map intensity is determined separately for each gene and comparisons can be made within each column, but not across rows. B) DNA extracted from the enriched population was digested with AluI and HhaI restriction enzymes prior to enrichment of methylated DNA by MBD2-MBD. qPCR was performed for *GSTP1* and *LINE1* using the enriched methylated DNA. Relative methylation was calculated by delta Ct relative to a max cycle limit of 45 with all undetected samples (ND) set to a Ct value of 45 for analysis. Optimal threshold for *GSTP1* is shown as a dotted line. C) Relative methylation for *GSTP1* and *LINE1* is plotted against the number of CTCs (*GSTP1*) or total number of cells (*LINE1*). A semi-log non-linear fit was performed to determine the best fit line. *GSTP1* is positively correlated to CTC number with an R^2 value of 0.94 and p value of 0.0062.

Figure 3.6. Effect of non-methylation sensitive enzyme digestion on *GSTP1* enrichment.

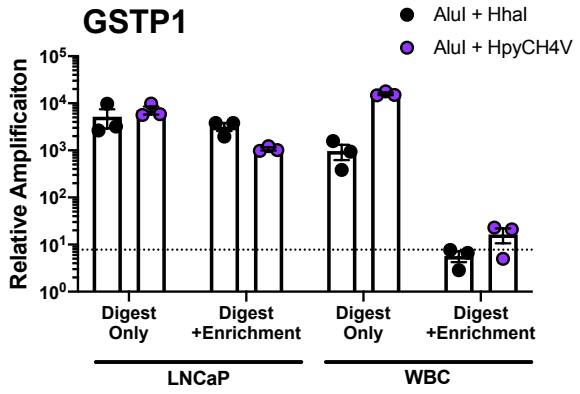


Figure 3.6. Effect of non-methylation-sensitive enzyme digestion on *GSTP1* enrichment. 5ng of LNCaP or WBC DNA was digested with AluI and either HhaI or HpyCH4V. Half of the samples were then enriched for methylated DNA. qPCR for *GSTP1* was performed on the digested and enriched DNA. Relative amplification of *GSTP1* is shown for each condition. Optimal threshold for *GSTP1* is shown as a dotted line.

Figure 3.7. Analysis of pre-amplification for *GSTP1*, *RASSF1*, *APC*, and *RARB*.

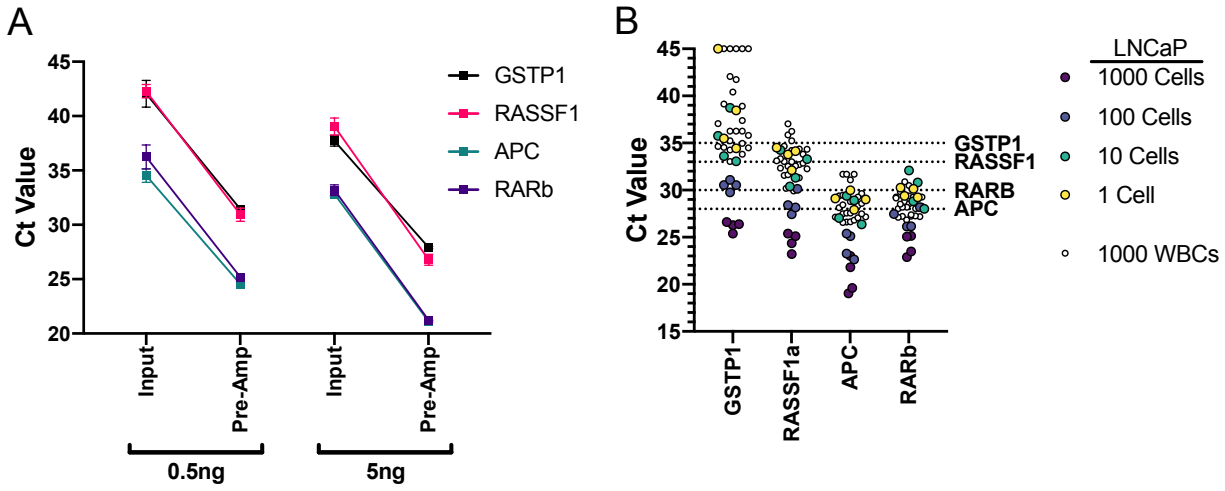


Figure 3.7. Analysis of pre-amplification for *GSTP1*, *RASSF1*, *APC*, and *RARB*. A) 5ng and 0.5ng of LNCaP DNA was digested with AluI and HhaI. Digested DNA was then pre-amplified with primers to the indicated genes. qPCR was performed on pre-amplified (Pre-Amp) and non-pre-amplified (Input) DNA. Ct values are shown for each condition with standard deviation of triplicate qPCR wells shown. B) Ct values for SEEMLIS enriched methylated DNA from serial dilutions of LNCaP DNA and 1000 WBCs were plotted for each gene to determine max cycle value (MCV) cut off Ct value. Dotted lines indicate the determined cut off for each gene as labeled.

Figure 3.8. Detection of multiple genes from MBD2-MBD enriched DNA.

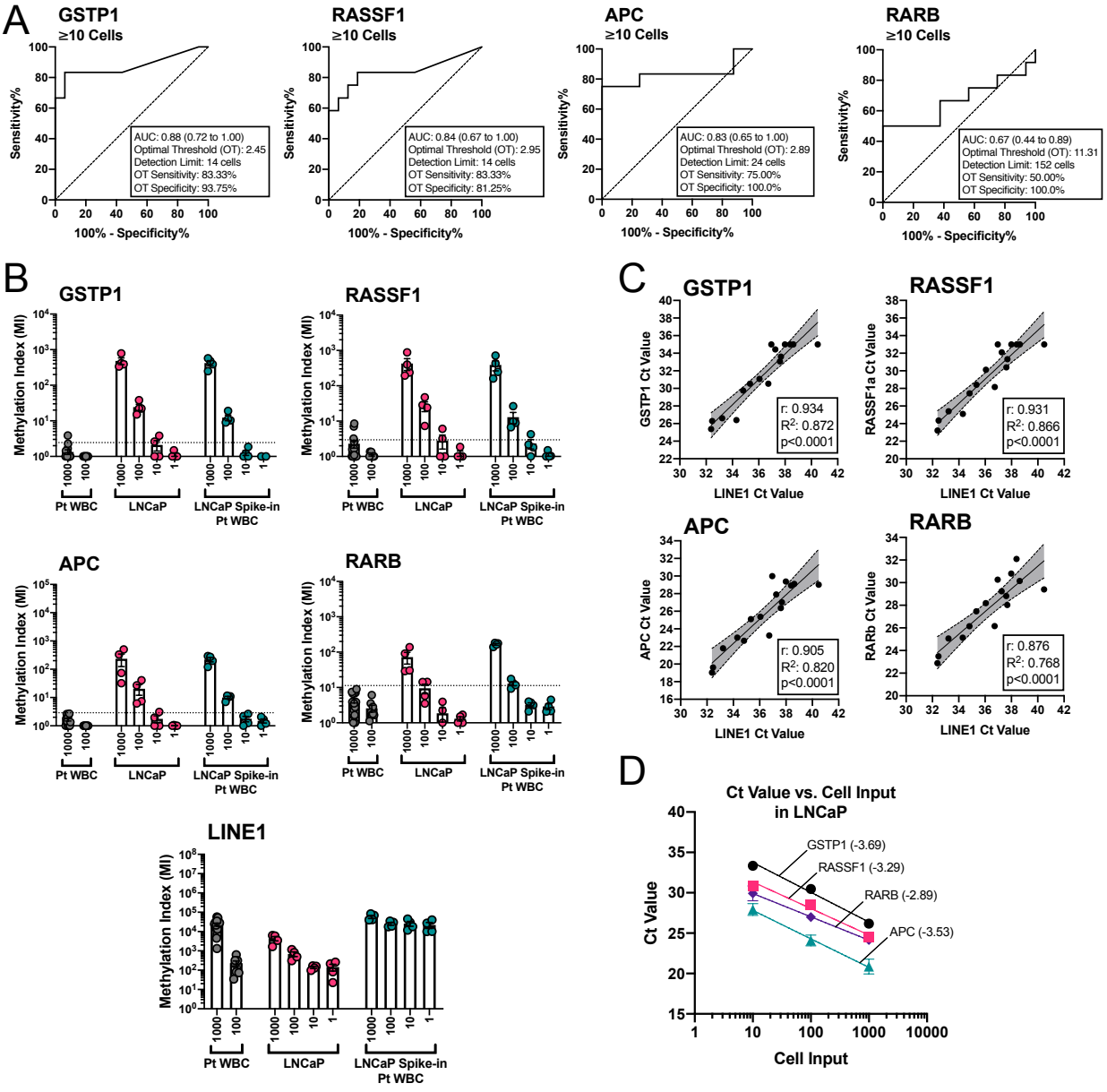


Figure 3.8. Detection of multiple genes from MBD2-MBD enriched DNA. Methylated DNA was enriched by MBD2-MBD capture from DNA extracted from serially diluted LNCaP cells (n=4 per dilution), 1000 (n=16) and 100 (n=8) patient-derived WBCs, and serially diluted LNCaP cells spiked into 1000 patient-derived WBCs (n=4 per dilution). Enriched methylated DNA was pre-amplified with probes to the indicated genes (excluding *LINE1*). Pre-amplified DNA was diluted 1:5 and qPCR was performed with the same probes, including *LINE1*. A) For each gene, ROC curves for WBC samples of 1000 cells and LNCaP samples of 1000, 100, or 10 cells were created. Area under the curve (AUC) with 95% confidence interval is indicated. Optimal threshold (OT) values determined by Youden's J statistic are listed with their associated sensitivity and specificity values. Detection limit was calculating using the slope of the best fit line of Ct values plotted against cell input (D). B) Relative methylation was calculated by delta Ct relative to a max cycles value (MCV) of 35 (*GSTP1*), 33 (*RASSF1*, *APC*, *RARB*), or 45 (*LINE1*) with all undetected samples set to the corresponding MCV for analysis. Optimal threshold as determined by ROC curve is shown as a dotted line. Each dot represents an individual sample taken from a pool of cells at the indicated concentration. C) Ct values for *LINE1* vs. Ct values for each gene were plotted against each other for LNCaP samples of 1000, 100, 10, and 1 cell(s). A simple linear regression was performed to determine the best fit line and 95% confidence interval for that line (shaded region). R and R^2 values are listed for the correlation. Each gene was significantly correlated to *LINE1* values with p values <0.0001. D) Ct values were averaged for each cell input and plotted against the cell input values. A semi-log non-linear fit was performed to determine the best fit line and

the slope of the best of fit line, which are indicated in parentheses for each gene. All error bars represent standard error of the mean (SEM).

Figure 3.9. Detection of multiple genes in MBD2-enriched DNA from cells purified by single cell aspiration.

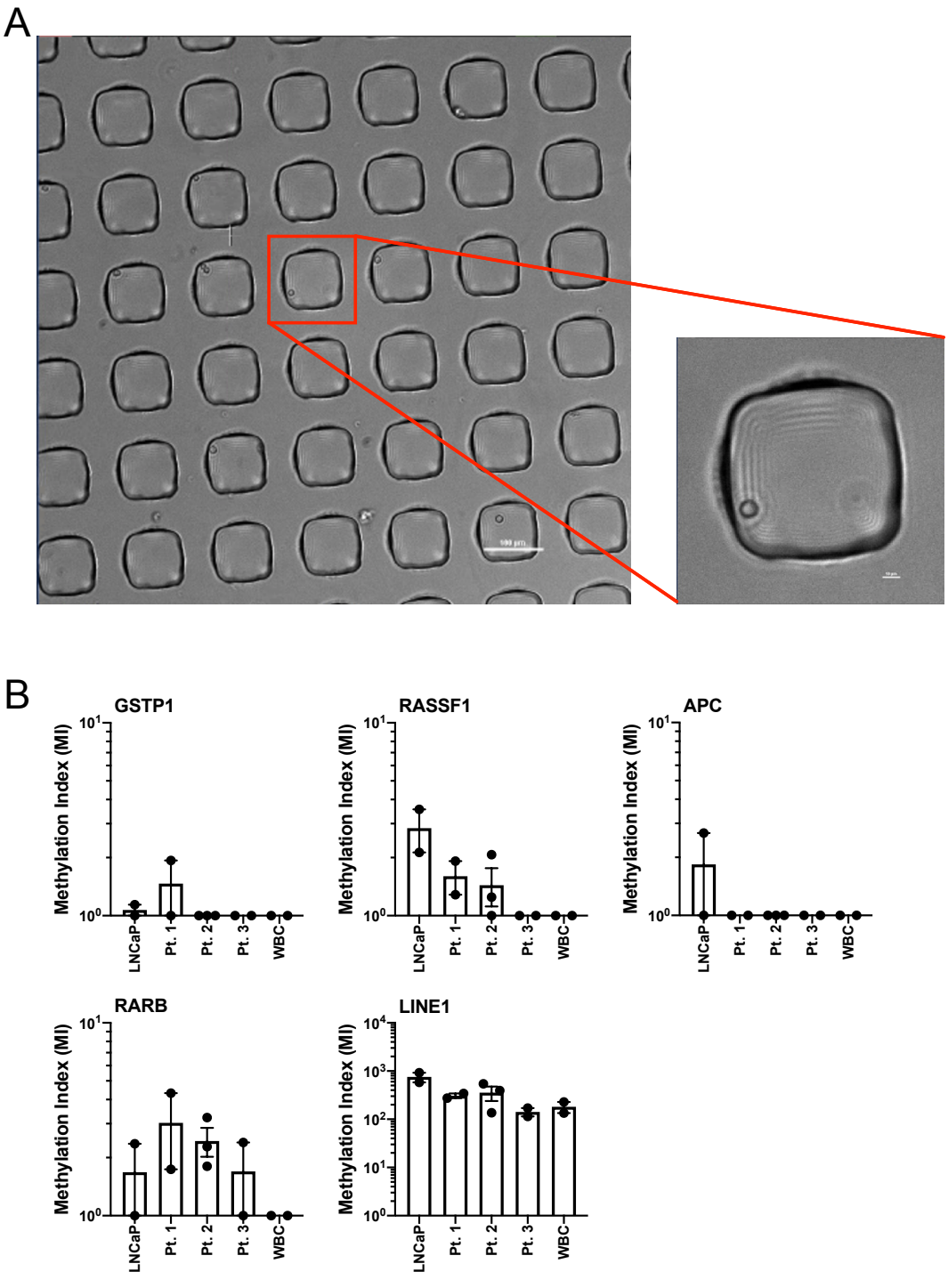


Figure 3.9. Detection of multiple genes in MBD2-enriched DNA from cells purified by single cell aspiration. A) Example image of WBCs seeded into microarray for single cell aspiration. Zoomed image is showing a single microwell within the microarray containing a single WBC. B) Methylation index for *GSTP1*, *RASSF1*, *APC*, *RARB*, and *LINE1* in groups of cells collected by single cell aspiration. Each dot represents a group of 10-15 cells that were aspirated, pooled, and processed with SEEMLIS.

Chapter 4:

HLA-I as a potential biomarker in circulating tumor cells

This chapter is adapted from the following manuscript in preparation: Rodems TS, Heninger E, Stahlfeld CN, Gilsdorf C, Carlson K, Kircher MR, Beebe DJ, McNeel DG, Haffner MC, Lang JM. Targetable epigenetic alterations regulate class I HLA loss in prostate cancer.

Contributions: Figure 4.1A was performed in collaboration with E.H. and K.C. Figure 4.4 was performed in collaboration with C.N.S. and C.G. All other experiments and analyses were performed by T.S.R.

ABSTRACT

Personalized medicine has long been a goal of the oncology community. Using biomarkers to identify patients who would or would not benefit from therapies ensures that patients are not subjected to painful, expensive, and time-consuming treatments that are not effective. In prostate cancer, the hunt for biomarkers has been met with limited success. Part of the lack of robust biomarkers in prostate cancer is the lack of common genomic alterations present in prostate cancer. The most common genomic alterations are only present in half the patient population. On the other hand, epigenetic alterations have been found to be much more common in prostate cancer. For example, DNA methylation in the promoter of GSPT1 and *RASSF1* is present in more than 90% of human prostate cancer. Our studies in this thesis suggest that methylated HLA-I may have utility as a biomarker to identify patients who may benefit from epigenetic therapies either alone or in combination with immunotherapies that rely on HLA-I expression. We have developed a method to assess DNA methylation signatures in low-input samples, including circulating tumor cells (CTCs), called SEEMLIS and descried in Chapter 3 of this thesis. Here, we use SEEMLIS to detect HLA-I methylation in CTCs from patients with prostate cancer. We found that HLA-I was able to be detected in CTCs with low HLA-I expression. This provides a foundation for future biomarker development and studies on HLA-I DNA methylation I circulation during cancer progression.

INTRODUCTION

The World Health Organization defines a biomarker as “any substance, structure or process that can be measured in the body or its products and influence or predict the incidence of outcome or disease” (186). The National Institutes of Health has published official definitions for multiple different types of biomarkers including diagnostic, predicative, prognostic, safety, susceptibility, and pharmacodynamic (187). Biomarkers are varied in nature, ranging from phenotypic to genomic to measures of whole-body system processes. The development of more and better biomarkers will help shape the future of personalized medicine.

Biomarkers have been of particular use in cancer, especially because of the heterogenous nature of human cancer biology and treatment response. Existing and well-known biomarkers include the BRCA genes as susceptibility biomarkers of breast cancer, HER2 as a predictive biomarker for response to HER2-targeted therapies, and PSA as a prognostic biomarker for prostate cancer progression (187). While some biomarkers have proven to be extremely useful in the clinic, other biomarkers have been met with worrisome issues often relating to heterogeneity in patient disease. Therefore, novel biomarkers in cancer are continually being sought after.

One of the most well-known biomarkers, PSA, is a liquid biomarker in prostate cancer. PSA has been reported to have both diagnostic and prognostic uses, however, PSA has been criticized for its lack of sensitivity and specificity, especially for diagnosis (188). Genomic alterations such as AR amplification and TMPRSS2-ERG fusion events have also been proposed as biomarkers in prostate cancer (105,106,189,190). However, the search for other genomic biomarkers has been limited, partly due to the infrequent

nature of genomic alterations in prostate cancer (62). Therefore, there is a need for better biomarkers in the clinic. Epigenetic biomarkers have been proposed as alternatives. Certain methylation events, including *GSTP1*, *RASSF1*, *APC*, and *RARB*, can be identified in more than 80% of patients with prostate cancer and combinations of common epigenetic markers are being investigated for their diagnostic and prognostic potential (65,75-77,157,158,172,191). Identification of methylation signatures in HLA-I in Chapter 2 lead us to hypothesize that methylated HLA-I may have utility as a biomarker in prostate cancer.

The ability to measure HLA-I methylation in tumors is the first step in determining the efficacy of epigenetically silenced HLA-I as a biomarker in patients with prostate cancer. Evaluating primary tumor biopsies is typically limited to single snapshot of early disease biology due to patients receiving prostatectomy as typical first line therapy. Metastatic biopsies require painful and invasive procedures and are generally not repeatable. Liquid biopsies, including blood, urine, and lymph, represent a minimally invasive alternative to traditional solid biopsy sources of tumor material (110). Liquid biopsies can also be performed repeatedly and longitudinally, which allows for dynamic evaluation of biomarkers and treatment response. Multiple types of tumor material can be obtained from liquid biopsies including circulating tumor DNA (ctDNA), exosomes, and circulating tumor cells (CTCs). As intact cells, CTCs represent the most comprehensive source of tumor material and offer the ability to analyze protein, RNA, and DNA from the same sample.

Our lab has a microfluidic platform to enumerate CTCs, analyze protein expression, and extract nucleic acids, however prior to the work in Chapter 3 of this thesis,

we were not able to perform methylated DNA analysis (139,170,171). Traditional methods of methylation analysis are not suitable for targeted evaluation of methylation signatures from rare cell populations such as CTCs. Development of the SEEMLIS assay described in Chapter 3 allows us to perform multiplexed, targeted analysis of HLA-I in CTCs. In this study, we performed the first analysis of HLA-I protein expression in CTCs from patients with prostate cancer and identified methylated HLA-I in CTCs with low HLA-I expression, supporting the potential for methylated HLA-I as a future clinical biomarker, and identifying HLA-I low or negative CTCs as a potentially druggable population.

RESULTS

Evaluation of HLA-I protein expression in CTCs

The data in this thesis as well as data from other groups have shown HLA-I protein and gene expression is lost or downregulated in a subset of primary prostate cancers and the majority of metastatic prostate cancers. Therefore, we hypothesized that HLA-I loss would also be identifiable in CTCs from patients with prostate cancer. Prior to this work, HLA-I expression in CTCs had not been studied. We gathered two cohorts of 8 patients to study HLA-I expression at the cell surface and intracellularly. Samples were collected, stained, and imaged using VERSA technology. CTCs were identified as being Hoechst positive, cytokeratin positive, and negative for multiple white blood cell proteins. White blood cells (WBCs) were identified as being Hoechst positive, cytokeratin negative, and positive for WBC proteins. Mean fluorescent intensity (MFI) of HLA-I was generated per cell and compared to the median HLA-I in the WBC population from the entire cohort, which served as a population control for normal HLA-I expression (Figure 4.1A). A representative image of a CTC and WBC is shown in Figure 4.1B. We found that HLA-I expression was heterogenous both between patients and between individual CTCs from the same patient. In the intracellular expression group, 7 out of 8 patients had at least one CTC with expression below the median WBC level of expression. Half of the patients had intracellular HLA-I expression below the median WBC level in all CTCs. For the surface expression cohort, all 8 samples had at least two CTCs with surface expression of HLA-I below the level of median WBC expression. Only one patient had the majority of CTCs fall above the WBC median. This analysis suggests that surface expression of HLA-I is impaired in CTCs from patients with prostate cancer, but the mechanisms contributing

to loss of surface expression may be varied. Therefore, identifying patients that would benefit from epigenetic therapies may require deeper analysis than evaluating HLA-I protein expression alone. While low intracellular staining of HLA-I may imply transcriptional downregulation, there are multiple molecular mechanisms which may yield the same phenotype. As such, loss of protein expression alone may not be indicative of epigenetic silencing of HLA-I.

Validation of detection of methylated HLA-I in prostate cancer cell lines by SEEMLIS

As described in Chapter 3 of this thesis, we have developed a method for methylated DNA analysis in CTCs, called SEEMLIS. We sought to use this method to measure HLA-I methylation in CTCs to evaluate this signature as a potential biomarker. We first validated the enzymes and primers in cell line models and patient-derived WBCs. As discussed in Chapter 3, using methylation sensitive enzymes to digest DNA for SEEMLIS decreases background signal. However, we were not able to identify a methylation sensitive enzyme with a cut site that was commonly methylated in all of our target DNA regions. For example, evaluation of the cut site for the methylation sensitive enzyme used for *GSTP1* single gene analysis, HhaI, is unmethylated in *HLA-A* in LNCaP cells despite heavy preceding methylation levels (Figure 4.2A). This results in loss of *HLA-A* signal from LNCaPs after digestion with HhaI (Figure 4.2B). We chose to continue with the non-methylation sensitive enzyme, HpyCH4V, to avoid interference from differences in methylated CpG signatures at enzymatic cut sites that may exist between genes and between patients.

We next tested the ability of the assay to detect *HLA-A*, *HLA-B*, and *HLA-C* methylation from various starting DNA and cell inputs and in heterogenous samples. For the following validation experiments, we performed SEEMLIS for multiple gene analysis, which includes a pre-amplification step prior to qPCR. To confirm we could detect differences in methylation level using the enzymes and primers that were chosen, we generated serial dilutions of LNCaP and LAPC4 DNA from 5ng down to 0.005ng as well as dilutions of WBC DNA at 5ng and 0.5ng per run. DNA was digested with AluI and HpyCH4V followed by the methylated DNA enrichment, wash, and elution steps of SEEMLIS. Enriched DNA was subjected to qCPR using primers targeting *HLA-A*, *HLA-B*, *HLA-C*, and *LINE1* (Figure 4.3A). We determined a max cycle threshold (MCV) of 33 by plotting all the LNCaP and LAPC4 raw Ct values and defining a cut-off cycle value near the point where replicate captures were no longer reliably detected (Figure 4.3B). This MCV was used to calculate all relative methylation values in this Chapter. Relative methylation for the LNCaP, LAPC4, and WBC samples is shown in Figure 4.3C. We were able to detect methylated HLA-I from both prostate cancer cell lines down to the 0.05ng level in both duplicate runs. Detection at the 0.005ng level was more successful in the LAPC4 cell line, which is supported by our data in Chapter 2 showing HLA-I is more methylated in LAPC4 cells compared to LNCaP cells. HLA-I methylation was detected in the WBC samples, but at much lower levels than the corresponding cancer cell line dilutions. *LINE1* was detected in an input dependent manner in all cell lines and WBC samples.

Next we tested the whole SEEMLIS platform beginning with the semi-automated DNA capture step. We generated serial dilutions of 1000, 100, 10, and 1 LNCaP cells and

1000 and 100 patient-derived WBCs as control samples. We also generated spike-in samples to represent a patient sample containing CTCs with background WBCs by spiking 1000, 100, 10, or 1 LNCaP cell(s) into 1000 WBCs. Methylated DNA was enriched in each sample by SEEMLIS followed by qPCR using primers targeting *HLA-A*, *HLA-B*, *HLA-C*, and *LINE1*. Relative methylation was calculated for each sample (Figure 4.3D). We were able to detect methylated HLA-I in LNCaP cells at similar levels to the DNA inputs for the corresponding dilutions (i.e. 5ng is on the order of 1000 cells, 0.5ng is on the order of 100 cells, etc). We also detected methylated HLA-I from spike-in samples at similar levels to samples with LNCaP alone. *LINE1* was detected in an input dependent manner in all cell lines and WBC samples as well as spike-in samples.

We next generated ROC curves for each HLA-I gene using the data from Figure 4.3C. We chose to leave the single cell dilution out of the calculations since we are reducing sensitivity of the assay with the addition of the pre-amplification step and do not anticipate performing this assay for HLA-I detection on heterogenous samples with fewer than 10 CTCs due to the relatively high background signal in WBCs shown in Figure 4.3D. We generated two sets of ROC curves, one using the data from 1000 WBCs and one using the data from 100 WBCs, to demonstrate the limitations of using this assay with high background samples (Figure 4.3E,F). The curves generated from the 1000 WBC data had areas under the curve ranging from 0.59 to 0.77. An area under the curve (AUC) of 0.5 represents a test that is not able to distinguish the difference between the two input sample types. Therefore, the ability of this test to distinguish between LNCaP cells and high levels of WBCs based on HLA-I methylation is relatively weak. HLA-B had the highest AUC and lowest detection limit, which was calculated from the optimal threshold

(OT) determined by Youden's J statistic plugged into the equation of the best fit line and for the LNCaP dilution series. The ROC curves generated using the 100 WBC data show that reducing background signal can significantly improve the sensitivity, specificity, and detection limits of this assay. Both *HLA-A* and *HLA-B* had curves with an AUC of 1, representing a perfect test. HLA-C also had a markedly improved AUC of 0.89. Each gene had a calculated detection limit of under 10 cells. Since these detection limits are based on cell lines that we have determined to be relatively uniformly and heavily methylated, we anticipate the real limits for CTC samples would be much more conservative. This may limit our ability to use samples straight from VERSA isolation. In fact, when we attempted to perform this assay on two VERSA-isolated CTC samples with approximately 600-1000 WBCs present, we were unable to detect methylated HLA-I above the thresholds determined by the cell line studies. *HLA-C* was not detected in either sample, but we did detect *HLA-A* and *HLA-B* methylation at levels that were relative to estimated CTC counts. Due to the high threshold that was necessary to apply, we were unable to attribute the *HLA-A* and *HLA-B* signal to CTCs for certain. However, we hypothesized that with more stringent sample clean-up, we would be able to detect methylated HLA-I from CTCs with a lower background threshold applied.

Evaluation of HLA-I methylation in CTCs purified by single-cell aspiration

To eliminate the background signal from WBCs, we utilized a semi-automated single cell aspiration system to individually aspirate CTCs (180). With this technology, we were also able to separate out subgroups of CTCs with high or low HLA-I protein expression to further analyze the association of HLA-I methylation signatures with HLA-I expression in

CTCs. CTCs were first isolated using VERSA and stained with Hoechst and antibodies to HLA-I, EpCAM, and markers of WBCs. Stained cells were seeded onto a microwell array for aspiration (Figure 4.4A). CTCs were then aspirated based on relative levels of HLA-I expression determined by fluorescence microscopy for each patient sample. A representable example image of a cell in each of these populations is presented in Figure 4.4B. Groups of approximately 10-15 HLA-I positive or HLA-I negative CTCs were selected for single cell aspiration. We confirmed HLA-I expression differences in the identified populations by analyzing the mean fluorescent intensity (MFI) of HLA-I expression in each group (Figure 4.4C). HLA-I expression in the HLA-I negative CTC populations was significantly lower than both the HLA-I positive CTC populations and the matched WBC populations. Of note, HLA-I expression in the HLA-I positive CTC groups was lower than matched WBCs. Though the differences were not statistically significant, this lower level of HLA-I expression may be functionally significant and epigenetically regulated. We also collected a control group of approximately 10 WBCs from an additional patient as well as a positive control group of 10 LNCaP cells for comparison.

Methylated DNA was enriched by MBD2 protein-based precipitation from enzymatically digested DNA from the collected groups of CTCs and controls and subjected to qPCR with primers targeting *LINE1* and the HLA-I genes. We were able to successfully detect HLA-I methylation from 10 aspirated LNCaP cells, while no methylation was detected in the WBC group (Figure 4.4D). *LINE1* methylation was detected at comparable levels in all patient samples and controls (Figure 4.4D,E). We were able to detect HLA-I methylation in all three patient samples, though the pattern of detection and association with protein expression varied (Figure 4.4E). Two out of three

patients, patient 568 and 490, had methylation signatures in at least two HLA-I genes that were higher in HLA-I negative CTCs compared to HLA-I positive CTCs. Patient 568 had *HLA-A* and *HLA-B* methylation in the HLA-I negative population at levels similar to LNCaP cells. *HLA-A* methylation was detected in the HLA-I positive population from patient 568, though at a lower level than the HLA-I negative group. We were able to obtain two groups of HLA-I negative CTCs from patient 490, both of which had methylation in *HLA-A* and *HLA-B*. One of the groups also had a small detected amount of *HLA-C* methylation. However, similar to patient 568, *HLA-A* was also detected in the HLA-I positive group. Patient 487 had detectable methylation in both groups of CTCs, with higher levels detected in the HLA-I positive population. As mentioned above and shown in Figure 4.4C, the HLA-I expression level in the HLA-I positive CTC group was lower than the matched WBC expression, suggesting that lower, non-negative HLA-I expression may still be epigenetically regulated in some patients. Importantly, these patients may still benefit from epigenetic therapy to express HLA-I to levels closer to that of WBCs. Overall, this preliminary experiment represents the first analysis of HLA-I methylation in CTCs and demonstrates our ability to detect HLA-I methylation in prostate cancer CTCs with low HLA-I expression levels, providing a foundation for future biomarker studies.

DISCUSSION

Circulating tumor cells are believed to have metastatic potential and provide a relatively non-invasive snapshot of tumor biology, which can be repeatedly sampled by blood draw. The utility of CTCs in the clinic extends to the biological signatures of CTCs that can be used as therapeutic biomarkers. So far, biomarker studies in CTCs have been largely restricted to protein makers and enumeration, but recent advances in technology have opened the door for epigenetic-based biomarker development. This Chapter demonstrates the potential for methylated HLA-I as a biomarker for identifying patients with epigenetically downregulated HLA-I. Identifying these patients may guide future clinician decision about which patients may benefit from certain classes of epigenetic and immunotherapies.

This study represents the first analysis of HLA-I protein expression and DNA methylation in prostate cancer CTCs. HLA-I protein expression has been assessed in primary and metastatic prostate cancer lesions, but no studies have been published on expression of HLA-I in CTCs from patients with prostate cancer. Here we show that both intracellular and extracellular HLA-I protein expression is highly heterogenous in prostate cancer CTCs. However, we were able to identify CTCs with HLA-I expression below the WBC median expression level in all but one patient and several patients had CTCs that were all below the WBC medial expression. The overall lower expression of HLA-I in CTCs in the patients in these two cohorts suggests loss of HLA-I may be important for cell survival in circulation and possibly to their ability to successfully disseminate. Future investigations into how HLA-I loss contributes to the metastatic potential of CTCs and promotes survival in circulation will help support this hypothesis.

Prior to this work, a limited number studies had performed analyses of methylated DNA in CTCs from various tumor types, including two studies in prostate cancer (160,161,192-198). However, all of these studies utilized methods that require bisulfite conversion of DNA, which can damage the input material. Our method, SEEMLIS, is unique in that it does not rely on bisulfite conversion, thus preserving the quality and integrity of the input DNA. Our use of restriction enzymes to cut DNA, rather than relying on random shearing, also preserves our target sequences. These key differences allow for successful enrichment and detection of methylated DNA as evidenced by amplification of *LINE1*, even from very low starting input. For comparison, Pixberg et al. had successful amplification of methylated DNA in only 30-40% of their CTC samples. In contrast, we were able to detect *LINE1* methylation in all 5 CTC samples we tested at levels that were comparable to estimated total cell input. A high assay success rate, such as that achieved here with SEEMLIS, is critically important to for future endeavors to perform analysis of HLA-I at scale in the clinic.

Our initial attempt to detect methylated HLA-I in CTC samples containing large amounts of background WBCs was not successful due to the relatively high cutoff we had to assign to account for the high signal from WBCs determined in our validation experiments. While it is possible that this signal may be non-specific, it is also possible that we are detecting real epigenetic regulation of gene expression in a small subset of the total WBC population. A recent study evaluated HLA-I gene expression in different immune cell types and in hematopoietic cells are various stages of differentiation. This study found significant variability across these populations, where certain cell types

including certain lineages of hematopoietic progenitor cells had much lower HLA-I gene expression compared to other immune cell types. It is possible that epigenetic mechanisms may regulate these differences in expression, from a development or cell differentiation perspective, which may be detected by our assay. Future studies will assess HLA-I methylation in various WBC lineages and developmental stages to determine if this may be impacting our sensitivity. Using this information, we can include additional sample clean-up steps to remove the specific cell types that diminish our ability to detect HLA-I methylation in CTCs.

To overcome the challenges with impure CTC samples, we employed single cell aspiration to generate pure CTC populations. We collected groups of CTCs stratified by HLA-I protein staining and evaluate HLA-I methylation signatures in the two subgroups. While we were able to successfully detect HLA-I methylation in CTCs using this purification method, the association of HLA-I methylation with protein expression was not as clear. We found that *HLA-B* and/or *HLA-C* were associated with HLA-I negative CTCs in two out of three patients, while methylated *HLA-A* was detected in both HLA-I negative and positive CTCs. One possibility is that some of the signal we are detecting may be unrelated to HLA-I loss. The DNA fragment generated by our chosen restriction enzymes contains the entire promoter and CpG island. Our study on HLA-I methylation signatures in prostate cancer cell lines in Chapter 2 revealed that only certain areas of differential methylation in cancer cell lines were correlated to loss of gene and protein expression. Therefore, we may be detecting methylation signatures that are uniquely present in CTCs, but do not have and bearing on HLA-I downregulation. In contrast, our analysis of HLA-I methylation in TCGA samples in Chapter 2 suggested a more uniform effect of differential

methylation on loss of gene expression, especially for HLA-A. Additional investigation into the contribution of specific areas of DNA methylation to HLA-I expression, followed by re-design of enzyme combinations to generate more specific fragments, will alleviate these challenges for future studies.

Another reason we may detect HLA-I methylation in both CTC subgroups is the large amount of intra- and interpatient heterogeneity in HLA-I expression in CTCs and WBCs. This makes it challenging to determine a cut off or range for “normal” and functional HLA-I expression level. The level of HLA-I expression in the HLA-I positive CTC group was overall lower than the WBC group, which may indicate that the HLA-I positive CTCs have impaired and non-functional HLA-I expression that is epigenetically regulated. Determining a cutoff for functional downregulation of HLA-I may make the association of HLA-I protein expression and methylation clearer. Our future studies will include analysis of HLA-I expression levels on HLA-I function to establish the functional expression range of HLA-I in our assay. This will allow us to determine true positive and negative thresholds for HLA-I expression in CTCs.

Despite these challenges, we have shown that HLA-I methylation can be detected in purified groups of CTCs from patients with prostate cancer. This initial analysis of HLA-I DNA methylation in CTCs is promising for the future of epigenetic biomarker development in rare cells.

MATERIALS AND METHODS

CTC Capture, Imaging, and Analysis

Blood samples were collected from patients after receiving written informed consent under a protocol approved by the IRB at the University of Wisconsin-Madison. CTCs were processed and stained as previously described in Sperger et. al (139). Briefly, PBMCs were isolated from whole blood on Ficoll-Paque PLUS (GE Healthcare, Cat# 45-001-750) gradient and fixed with Cytofix Fixation Buffer (BD Biosciences, Cat# 554655). Fixed cells were incubated with paramagnetic particles (PMPs) (Dynabeads® FlowComp™ Flexi kit, Life Technologies, Cat# 11061D) coated with biotinylated anti-EpCAM antibody (R&D Systems, Cat# AF960, RRID: AB_355745). The Versatile Exclusion-based Rare Sample Analysis (VERSA) platform was used for enrichment and staining of CTCs, as described previously (139,169,199). PMP bound cells were isolated in the VERSA and stained with Hoechst 33342 (Thermo Fisher, Cat# PI62249) and antibodies to the proteins indicated in the corresponding figures, which are summarized in Table 4.1. Pan-cytokeratin (CK) was conjugated to Alexa Fluor™ 790 using the Alexa Fluor™ 790 Antibody Labeling Kit (Life Technologies, Cat# A20189) according to manufacturer's instructions. All other antibodies were purchased pre-conjugated to the fluorophores listed in Table 4.1. CD45, CD34, and CD11b were used on the same channel to serve as a white blood cell (WBC) "exclusion channel". CD14 and CD27 were included in addition to CD45, CD34, and CD11b in the WBC exclusion channel for the experiment measuring extracellular HLA-I in CTCs only. Cells were stained for extracellular markers at 4^{o}C for 30 minutes. Cells were permeabilized and stained for intracellular antibodies with BD Perm/Wash at 4^{o}C overnight (BD Biosciences, Cat# BDB554723). Cells were imaged in the VERSA at 10x

magnification using NIS Elements AR Microscope Imaging Software (NIS-Elements, RRID:SCR_014329) and analyzed using NIS Elements analysis software. CTCs were defined as positive for Hoechst and CK and negative for CD45/34/11b or CD45/34/11b/14/27. All other cells were considered part of the WBC population.

Single Cell Aspiration of CTCs for Methylation Analysis

CTCs enriched using the VERSA platform as described above were stained in the VERSA with Hoechst 33342 (Thermo Fisher, Cat# PI62249) and anti-bodies to HLA-I, EpCAM, CD27, CD45, CD34, and CD11b. Fluorophores, catalog numbers, and other antibody information is summarized in Table 4.1. Cells were then further enriched using a single cell aspiration platform, SASCA, as described by Tokar et al (180). Briefly, cells were seeded into polydimethylsiloxane (PDMS) microwells mounted on a glass microscope slide. The microwell array was imaged on a Nikon Ti-E Eclipse inverted fluorescent microscope, and target cells were identified by phenotypic staining analysis. CTCs were identified as EpCAM positive, exclusion (CD45/CD34/CD11b/CD27) negative cells, whereas WBCs were classified as EpCAM negative, exclusion positive cells. CTCs were further subdivided into groups based on HLA-I positivity compared to WBCs in the same sample. Target cells were aspirated from microwells and dispensed directly into a droplet of PBS in the EXTRACTMAN extraction plate (Gilson, Cat# 22100008) for DNA extraction. Images of the microarray were analyzed using NIS Elements AR Microscope Imaging Software (NIS-Elements, RRID:SCR_014329) to obtain HLA-I mean fluorescent intensity (MFI) values.

DNA Extraction from CTCs

Semi-automated DNA extraction was performed on a Gilson PIPETMAX liquid handling robot enabled for exclusion-based sample preparation (ESP), termed EXTRACTMAX (184). LiDS buffer (90mM Tris-HCL, 500mM lithium chloride, 1% Igepal CA-630, 10mM EDTA, 1mM dithiothreitol) and MagneSil Paramagnetic Particles (PMPs) (Promega, Cat# MD1441) resuspended in GTC buffer (10 mM Tris-HCl, 6 M guanidinium thiocyanate, 0.1 % Igepal CA-630, pH 7.5) are added to the extraction microplate (Gilson, Cat# 22100008) by the robot. Cells were added to the microplate well containing LiDS, GTC, and MagneSil beads and mixed by the robot. Cells were allowed to lyse and DNA was allowed to bind to MagneSil PMPs for 5 minutes. The robot then transferred the MagneSil PMPs with bound DNA by exclusion liquid repellency (ELR) through one PBST (PBS containing 0.1% Tween-20) wash, one PBS wash, and into 15µL of nuclease-free water (Promega, Cat# P1197) for elution. Beads were manually resuspended in the elution well and allowed to elute for 2 minutes. The MagneSil PMPs are magnetically transferred out of the elution well, leaving eluted DNA in water.

MBD2-MBD Enrichment of Methylated DNA from CTCs

The SEEMLIS method described in Chapter 3 was used for all methylation studies. Briefly, 25 µL of TALON magnetic beads (Takara, Cat# 635637) were resuspended in 100 µL MBD2-MBD Coupling Buffer (1x BB, 1x Halt protease inhibitor cocktail (Thermo Scientific, Cat# PI87786), 500 ng Unmethylated Lambda DNA (Promega, Cat# D1521), 5 µL tagged MBD2-MBD (EpiXplore Kit, Takara, Cat# 631962) and placed on shaker at RT for 1 hour to bind MBD2-MBD to the TALON beads. MBD2-MBD bound beads were

added to DNA digested using 1 µL of each restriction enzyme (AluI at 10 units/µL (NEB, Cat# R0137L) and HpyCH4V at 5 units/µL (NEB, Cat# R0620L)) and placed on a shaker at RT for 3 hours to bind methylated DNA to MBD2-MBD conjugated TALON beads. Bead-bound methylated DNA was washed and eluted on the Gilson EXTRACTMAX system and placed into PCR prepared with pre-amplification master mix.

Pre-amplification and qPCR of MBD2-MBD Enriched DNA from CTCs.

Quantitative PCR was performed using custom TaqMan hydrolysis probes (Applied Biosystems) and iTaq Universal Probes Supermix (Bio-Rad, Cat# 1725153). Primer and probe sequences are listed in Table 4.2. Cycling conditions: 5 minutes at 95 ºC for initial denaturation and enzyme activation followed by 45 amplification cycles of 5 seconds at 95 ºC and 30 seconds at 60 ºC. Pre-amplification was performed using the custom hydrolysis probes and TaqMan PreAmp Master Mix (Applied Biosystems, Cat# 4488593) when indicated according to manufacturer specifications. Cycling conditions: 10 minutes at 95 ºC for enzyme activation followed by 14 cycles of 95 ºC for 15 seconds and 60 ºC for 4 minutes. Pre-amplified samples were diluted 1:5 with TE buffer. Ct values were transformed into relative methylation values by the following equation:

$$\text{Methylation Index} = 2^{-(Ct-MCV)}$$

Where MCV is the max cycle value, which is the Ct cut-off determined by plotting data points from serial dilutions and determining where replicate values are no longer reliably detected. If serial dilutions are always reliably detected, the MCV is set at the highest cycle ran for qPCR. For HLA-I, MCV is equal to 33. For *LINE1*, MCV is equal to 45.

Statistical Analysis

All statistical analyses were performed in Prism 8 (GraphPad Prism, RRID: SCR_002798). For CTC MFI experiments, comparisons were made by Kruskal-Wallis test using the Dunn's method for correction of multiple comparisons. Receiver operator characteristic (ROC) curves were generated for each gene by plotting sensitivity vs. 100-specificity for the raw Ct values of LNCaP (true positive) and WBC (true negative). Optimal Threshold (OT) values were determined using Youden's J Statistic which is defined as the maximum value achieved from subtracting 100 from the sum of the sensitivity and 100-specificity values (in percentages). The associated Ct value was then converted into a methylation index (MI) as described above. Area under the curve (AUC) with 95% confidence intervals were found and reported, which indicate the probability that a randomly selected true positive sample will have a greater MI value than a randomly selected true negative sample. Semi-log non-linear fit analyses were performed in prism to calculate detection limits. All error bars represent SEM.

ACKNOWLEDGEMENTS

We would like to thank the patients who donated samples for this study.

Table 4.1. Antibodies used for HLA-I expression studies.

Experiment	Target	Fluorophore	Clone	Manufacturer	Catalog Number	RRID
Intracellular HLA I staining in CTCs	HLA-I	PE	W6/32	BioLegend	311406	AB_314875
	CD45	647	HI30	BioLegend	304018	AB_389336
	CD34	647	581	BioLegend	343508	AB_1877133
	CD11b	647	M1/70	BioLegend	101218	AB_389327
	CK	790	C-11	BioLegend	628602	AB_439775
Extracellular HLA I staining in CTCs	HLA-I	647	W6/32	BioLegend	311414	AB_493135
	CD45	PE	HI30	BioLegend	304008	AB_314396
	CD34	PE	581	BioLegend	343506	AB_1731862
	CD11b	PE	M1/70	BioLegend	101208	AB_312791
	CD14	PE	HCD14	BioLegend	325606	AB_830679
	CD27	PE	O323	BioLegend	302842	AB_2564146
	CK	790	C-11	BioLegend	628602	AB_439775
HLA-I staining in CTCs for single cell aspiration	HLA-I	FITC	W6/32	BioLegend	311404	AB_314873
	CD45	647	HI30	BioLegend	304018	AB_389336
	CD34	647	581	BioLegend	343508	AB_1877133
	CD11b	647	ICRF44	BioLegend	301319	AB_493020
	CD27	647	O323	BioLegend	302812	AB_493082
	EpCAM	PE	VU-1D9	Abcam	ab112068	AB_10861805

Table 4.2. Primers used for HLA-I methylation studies.

	Forward Primer (5'-3')	Reverse Primer (5'-3')	Probe (5'-3')
HLA-A	GTGGACGACACGCAGTTC	GCCCCTCCTGCTCTATCCA	TCGACAGCGACGCCGCG
HLA-B	CAGTTCGTGAGGTTCGACA	CTCTCGGTAAGTCTGTGTGTT	CGGAGTATTGGGACCGGAACACA
HLA-C	GCTTCATCTCAGTGGGCTAC	CGCTTGTACTTCTGTGTCT	TTCGTGCGGTTCGACAGCGA
LINE1	CGCAGGCCAGTGTGTGT	TCCCAGGTGAGGCAATGC	CCGTGCGCAAGCCGA

Figure 4.1. HLA-I expression in circulating tumor cells.

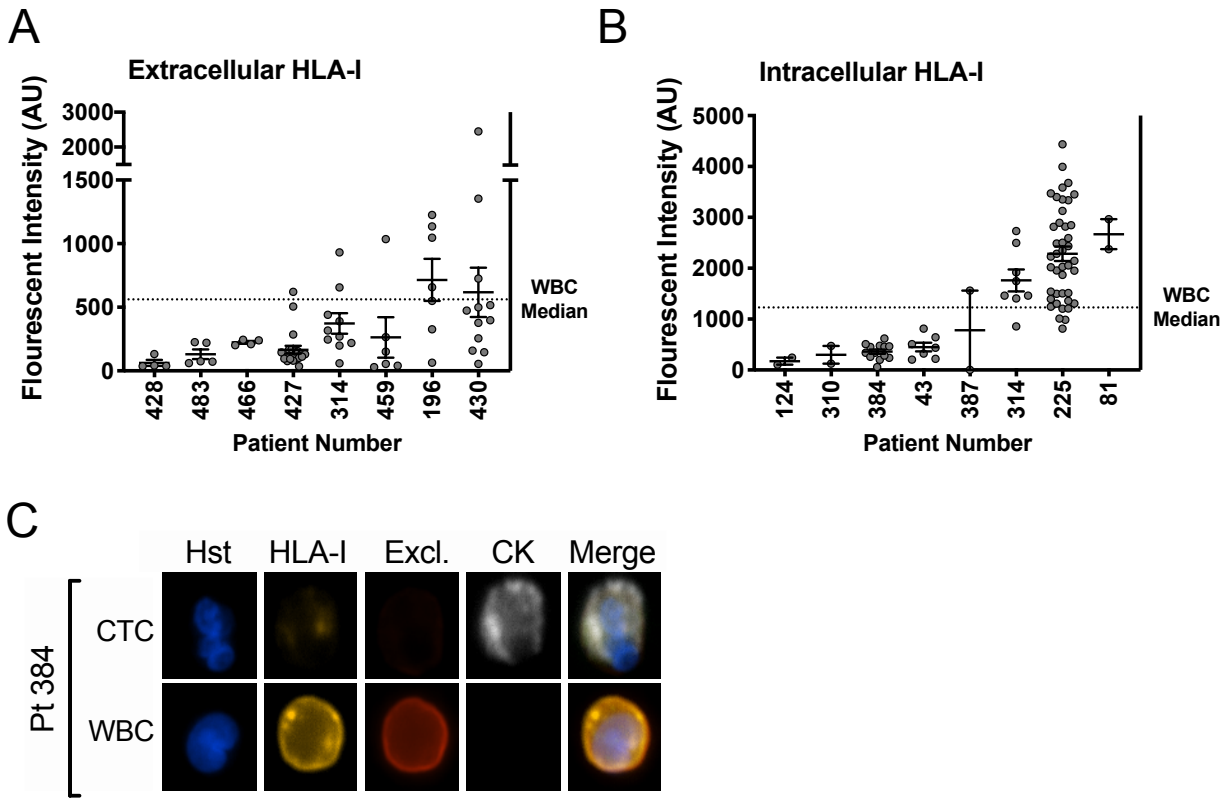


Figure 4.1. HLA-I expression in circulating tumor cells. A) Intra- and extracellular HLA-I protein expression in circulating tumor cells from 8 patients with prostate cancer. Cells in the intracellular expression group were permeabilized prior to HLA-I staining. Dotted line represents median HLA-I expression in all patient matched WBCs. B) Representative images of a CTC with low HLA expression and a WBC near the median expression from patient 384.

Figure 4.2. Evaluating the use of alternative enzymes for methylation analysis by SEEMLIS to detect HLA-I methylation.

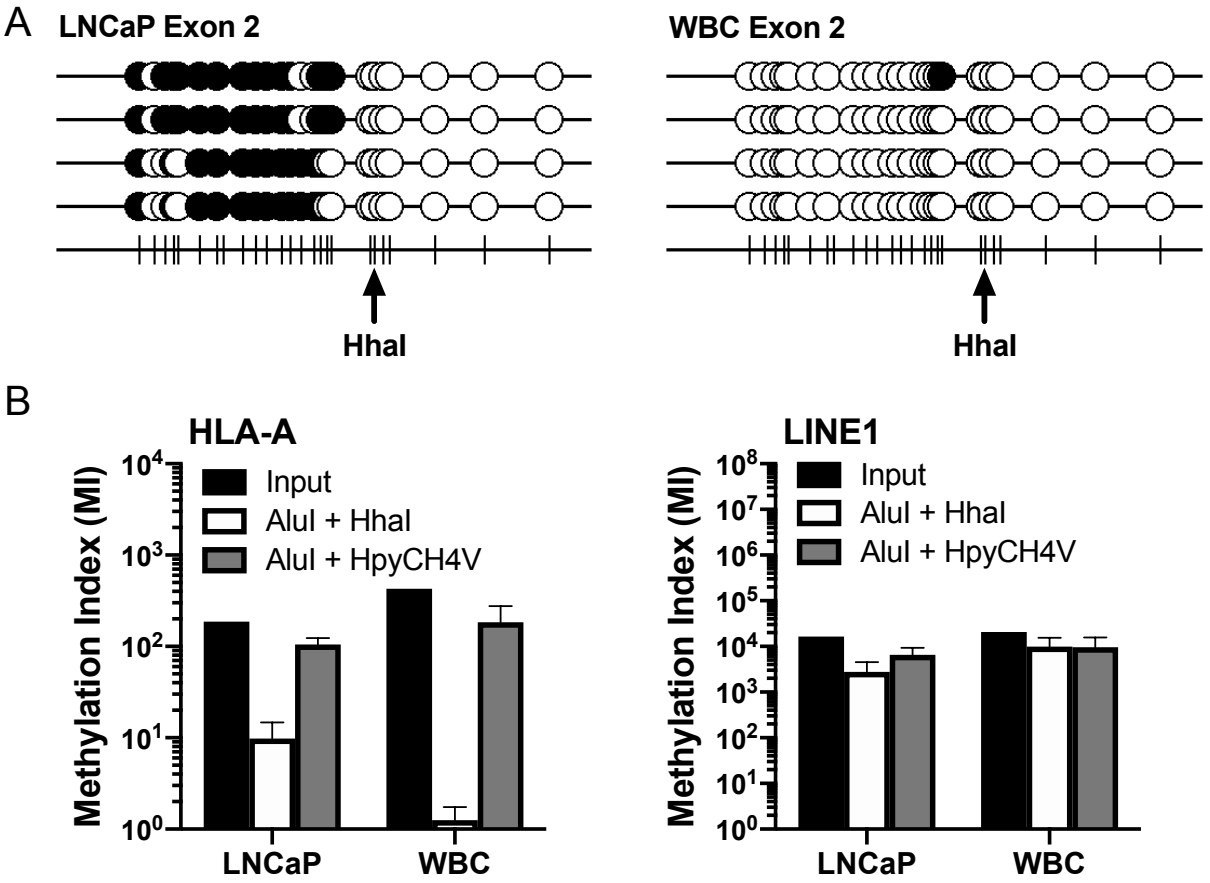


Figure 4.2. Evaluating the use of alternative enzymes for methylation analysis by SEEMLIS to detect HLA-I methylation. A) Bisulfite sequencing analysis of *HLA-A* exon 2 in LNCaP and WBC DNA with the cut site for methylation sensitive restriction enzyme, HhaI, indicated. Black circles represent methylated CpGs. White circles represent unmethylated CpGs. B) Methylation index for *HLA-A* and *LINE1* in two different restriction enzyme digest combinations in LNCaP and WBCs.

Figure 4.3. Initial validation of SEEMLIS for detection of methylated HLA-I.

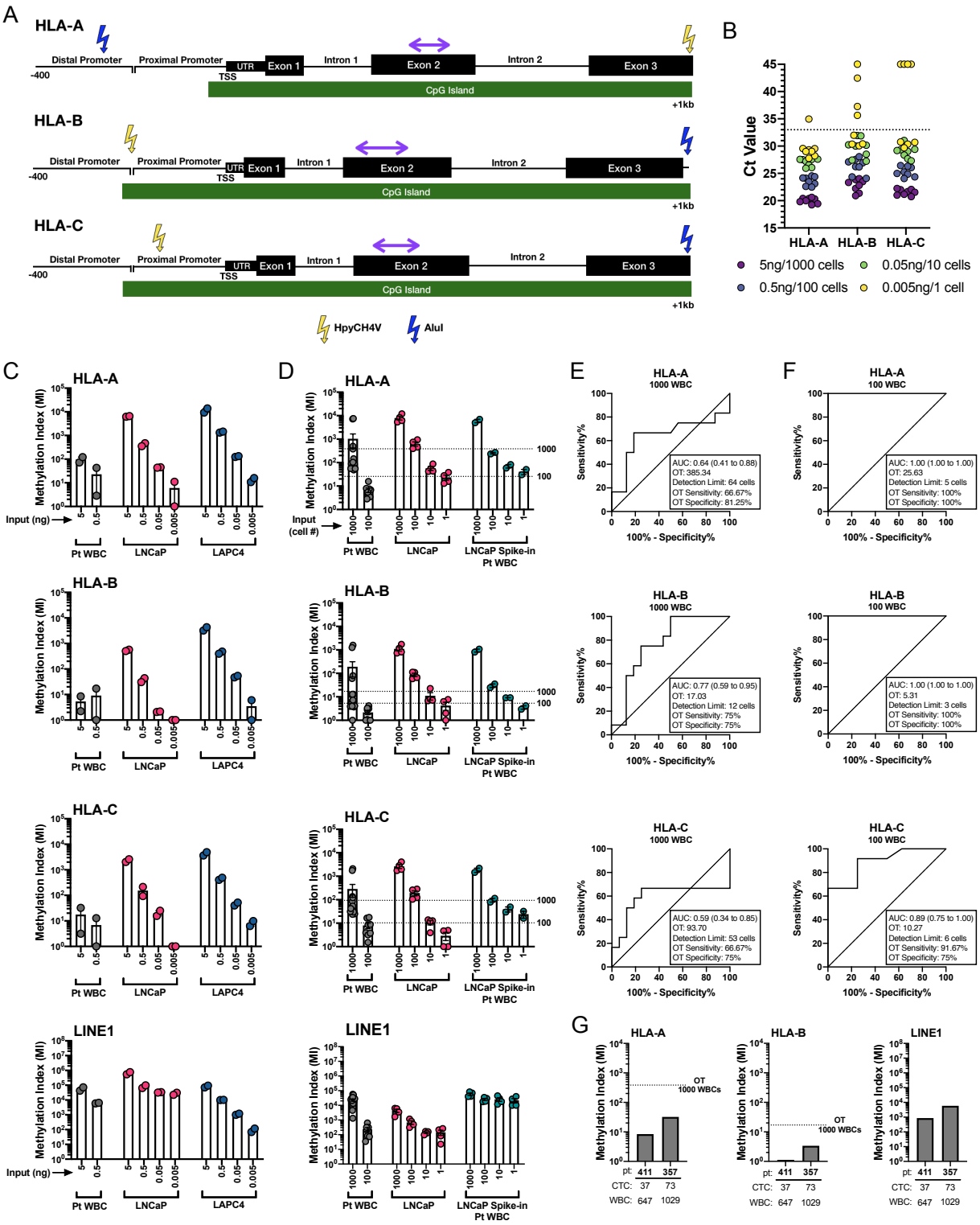


Figure 4.3. Initial validation of SEEMLIS for detection of methylated HLA-I. A) Primer locations for assessing CTC methylation. The cut sites of the restriction enzymes used to digest the DNA are indicated. The fragments used for MBD2-MBD enrichment extend from the first enzyme cut site to the second, with the primer location in the center. B) Raw Ct values from LNCaP and LAPC4 serial dilutions plotted to determine cut off for max cycle value (MCV). Dotted line is at chosen MCV of 33. C) Methylation indexes for MBD2-MBD enriched serially diluted LNCaP and WBC DNA at the indicated ng amounts (amounts indicated are per assay run) D) Methylation indexes for SEEMLIS enriched serially diluted cell amounts (amounts indicated are per assay run). For spike in samples in, serially diluted LNCaP cells spiked into 1000 patient-derived WBCs. For each gene, ROC curves for WBC samples of 1000 cells (E) or 100 cells (F) and LNCaP samples of 1000, 100, or 10 cells were created. Area under the curve (AUC) with 95% confidence interval is indicated. Optimal threshold (OT) values determined by Youden's J statistic are listed with their associated sensitivity and specificity values. Detection limit was calculating using the slope of the best fit line of Ct values plotted against cell input. G) Methylation index of the HLA-I genes and *LINE1* in two CTC samples. Estimated numbers of CTCs and WBCs in each sample are indicated below graph. For (C) and (D), each dot represents an individual sample taken from a pool of cells at the indicated concentration. Error bars represent SEM. For (C), (D), and (E), optimal thresholds as determined by ROC curve are shown as dotted lines.

Figure 4.4. Detection of methylated HLA-I in HLA-I positive and negative CTCs purified by single cell aspiration.

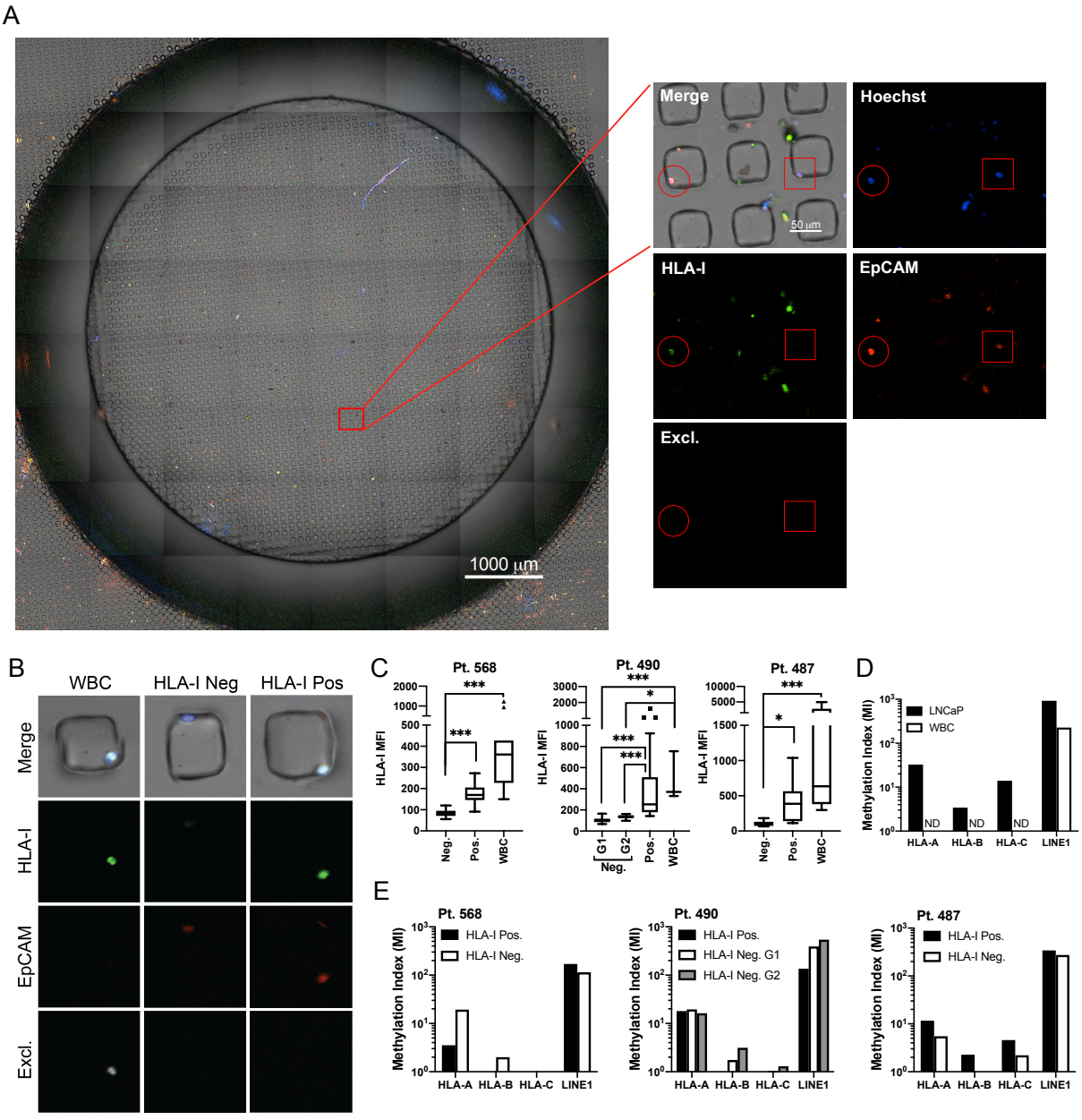


Figure 4.4. Detection of methylated HLA-I in HLA-I positive and negative CTCs purified by single cell aspiration. A) Image of a seeded microwell (pt. 568) with zoomed in images showing one HLA-I positive CTC (circle) and one HLA-I negative CTC (square). B) Representative 10x images of HLA-I negative and HLA-I positive CTCs and a WBC in the single cell aspirator microwells (pt. 490). C) HLA-I mean fluorescent intensity (MFI) of HLA-I in each CTC population and matched WBCs. Approximately 10-15 CTCs were chosen to be aspirated from each group with the exception of the WBC populations, which were not used for analysis. D) HLA-I methylation in 10 WBCs and 10 LNCaP cells purified by single-cell aspiration. E) HLA-I methylation in HLA-I positive and negative groups of approximately 10-15 CTCs. $*$ $p<0.05$, $**$ $p<0.01$, $***$ $p<0.001$.

Chapter 5:

Discussion and Future Directions

AIM OF THESIS OVERVIEW

Prostate cancer is the second most commonly diagnosed cancer in men and accounts for just over 10% of all new cancer cases in the United States (3). While men who are diagnosed with local and regional disease have extremely good prognoses, men diagnosed with metastatic prostate cancer have a 5-year survival rate of only 30% (3). The overarching goal of many prostate cancer research groups has been to understand the molecular mechanisms of metastasis in an effort to improve long term survival rates for men with metastatic prostate cancer. An area of increasing interest in prostate cancer, and the larger oncology research community, is the contribution of the immune system to cancer progression and metastasis. Immune evasion, or the ability of cancer cells to hide from the immune system, has been attributed to the promotion of cancer progression and metastasis and also plays a role in resistance to therapy, especially immunotherapies (34,35,200). Understanding how cancer cells evade the immune system is critical to improve patient outcomes and develop better therapies.

One mechanism of immune evasion is downregulation of immune regulatory molecules by cancer cells, including the class I major histocompatibility complex (MHC-I). MHC-I is essential for tumor cell recognition by T-cells through the T-cell receptor (TCR). Loss of MHC-I at the cell surface reduces tumor immunogenicity and renders certain immunotherapies ineffective. Downregulation or loss of expression of class I human leukocyte antigens (HLA-I), one component of MHC-I molecules, represents a major contributing factor to loss of MHC-I at the surface of tumor cells. Downregulation of HLA-I has been reported in many types of cancer, including prostate cancer (36,37). However, prior to this thesis, the epigenetic mechanisms regulating HLA-I downregulation

in prostate cancer had not been widely explored. Additionally, no research had been done to describe HLA-I expression in circulating tumor cells, which are a valuable tool for studying solid tumor dissemination and are thought to have metastatic potential. The goal of this thesis work was to characterize and investigate the epigenetic mechanisms involved in HLA-I downregulation in prostate cancer and develop a method to describe HLA-I methylation in circulating tumor cell (CTC) samples to aid future biomarker development. Elucidating mechanisms of HLA-I regulation in primary, metastatic, and circulating prostate tumor cells can increase understanding of immune evasion mechanisms during prostate cancer progression and generate new molecular targets for therapeutic intervention. In addition, the ability to measure and monitor HLA-I methylation in CTCs can aid physicians in decision-making for treatment regimens, allowing for more personalized therapy options.

DISCUSSION

Epigenetic regulation of HLA-I in prostate cancer

HLA-I loss has been well defined at the protein level in multiple cancer types, including prostate cancer (36,37). However, prior to this thesis work, limited research had been done to explore the molecular mechanisms responsible for HLA-I loss. A previous study had identified loss of heterozygosity (LOH) as responsible for HLA-I loss in breast and non-small cell lung cancer (132,133). Previous studies in colon cancer and melanoma had also identified genomic alterations that contributed to HLA-I loss (201). Yet another study identified mutations that caused loss of HLA-I expression in cervical cancer (202,203). In contrast, we found a striking lack of genomic alterations and mutations in the HLA-I genes in patients with primary or metastatic prostate cancer. More importantly, we found no correlation of genomic alterations to HLA-I gene expression. One caveat to this analysis is that the HLA-I genes are significantly polymorphic, which can lead to certain alterations being masked when analyzing sequencing data. One study addressed this issue in lung cancer by aligning sequencing reads to patient matched germline DNA, rather than a reference genome, to account for variations in each individual's HLA haplotype (133). This study found the percentage of NSCLC tumors with LOH was approximately 5 to 7 times higher than what is reported in TCGA data for NSCLC. However, even if we apply this increase to the number of deep deletion events in the data in Figure 2.1A, there would still be less than 10% of samples harboring deep deletions in any HLA-I gene. This is, of course, speculative and further investigations into the true percentage of prostate tumors with LOH in the HLA-I genes may be warranted.

Nevertheless, it seems likely that genomic alterations in HLA-I genes are not a significant contributor to HLA-I loss in prostate cancer.

We next measured HLA-I gene expression in prostate cancer to determine if HLA-I loss occurs at the transcriptional level. To do this, we analyzed HLA-I gene expression data from the TCGA-PRAD (PRAD) data set and a study by Taylor et al (134). Comparison of HLA-I gene expression in primary tumor samples from both studies relative to gene expression in normal prostate tissue identified a subgroup of primary tumor samples with diminished HLA-I gene expression. Interestingly, in the Taylor data set, the percentage of samples with low HLA-I expression was much higher in metastatic samples, suggesting HLA-I gene expression may be lost during the metastatic process or that tumor cells with low HLA-I expression may have a survival advantage. Patients from the Taylor data set with low HLA-I gene expression had significantly decreased disease-free survival time based on biochemical recurrence (two occurrences of PSA $\geq$ 0.2ng/mL) after radical prostatectomy. Together, these analyses identified that there is a subgroup of patients with prostate cancer with low HLA-I gene expression, which is associated with negative patient outcomes. This is important information for oncologists due to the prevalence of immunotherapies in use in the clinic. Not only would these patients not benefit from immunotherapies that rely on MHC-I expression on tumor cells, including certain cancer vaccines and PD-1/PD-L1 targeted therapies, but they also likely represent a patient population that may progress faster without effective treatment due to increased immune evasion capabilities of the tumors (40,42).

It is important to consider the different methods of data collection and analysis in the PRAD and Taylor data sets. The PRAD gene expression data was generated by RNA-

seq while the Taylor data set was generated by cDNA microarray. These two methods differ in that RNA-seq measures expression across the entire transcriptome, while cDNA microarrays only measure expression by pre-defined probes included in the array. In this way, RNA-seq is much more comprehensive. A recent study looked at concordance between gene expression data generated by Affymetrix microarray and RNA-seq. This study found that RNA-seq had increased dynamic range compared to the Affymetrix microarray and determined a concordance of 40-60% between the platforms for differentially expressed genes (204). Another caveat to consider is sample purity. In both the PRAD and Taylor studies, the samples collected for analyses were not always 100% pure, meaning non-tumor cells were included in the sample preparation and subsequent analyses. Non-tumor cell types that would be included in prostate cancer samples include immune cells and prostate basal cells, which are likely to have high expression of HLA-I genes. The Taylor et al. and TCGA tumor samples were required to be at least 70% and 60% tumor material, respectively (205). One study analyzed tumor purity in multiple TCGA studies and found that the PRAD data set had a median purity level of 90%, with a range from ~40% to 100% (206). This study also found that enrichment of genes related to immune function was significantly altered when accounting for tumor purity in multiple TGCA data sets, however the HLA-I genes were not among those that were altered. Despite these considerations, these data sets provide valuable information for the biological relevance of HLA-I gene expression in patients. We hypothesized that understanding the mechanisms that contribute to loss of HLA-I gene expression may help identify effective therapies for patients in the HLA-I low subgroup.

Loss of gene expression can occur as the result of many different mechanisms. Having ruled out genomic alterations as a significant source of HLA-I regulation in prostate cancer, we hypothesized that transcriptional regulation by epigenetic mechanisms could be responsible. Previous studies in esophageal cancer had reported that DNA methylation in the promoter regions of the HLA-I genes was a major contributing factor to loss of HLA-I gene expression (128). Additionally, studies from our lab had found that treating cell lines with epigenetic modifying agents could induce expression of cancer testis antigens (CTAs), which are peptides presented to T-cells by MHC-I molecules and are important for cancer vaccine therapies and general immune surveillance (101). At the global level, epigenetic changes are common in prostate cancer. There is a global loss of methylation across the genome leading to increased genomic instability and expression of oncogenes (67,69-71). In contrast, DNA methylation at specific gene promoters has been found to be significantly increased during prostate cancer progression (65). In addition, the expression of epigenetic readers and writers is altered in prostate cancer. DNA methyltransferase (DNMT) and histone deacetylase (HDAC) protein expression has been reported to be increased in prostate cancer tissue (95). While protein expression was not available, DNMT and HDAC gene expression was increased in the PRAD and Taylor data sets. We found DNMT and HDAC protein expression was increased in prostate cancer cell lines compared to a benign prostate epithelial cell line, however gene expression was not increased. Importantly, expression of members of the DNMT and HDAC families was correlated to HLA-I expression in the PRAD and Taylor data sets and our cell line models. Therefore, we hypothesized that HLA-I genes may harbor epigenetic signatures which contribute to loss of HLA-I expression by blocking HLA-I transcription.

Our analysis of the epigenetic landscape of the HLA-I genes in prostate cancer cell lines revealed differential DNA methylation, H3K27 acetylation, and H3K27 tri-methylation levels across the HLA-I genes when compared to a benign prostate epithelial cell line. These epigenetic changes were highly correlated to HLA-I gene and protein expression in the cell lines. Treatment with DNMT and HDAC inhibitors reversed DNA methylation and histone acetylation signatures, respectively. DNMT and HDAC inhibition also induced HLA-I expression in cancer cell lines and prostate cancer biopsies cultured *ex vivo*. Increased binding of RNA polymerase II at HLA-I promoters after treatment with DNMT and HDAC inhibitors supports our hypothesis that repressive epigenetic signatures are regulating HLA-I transcription. We confirmed functionality of the induced HLA-I protein by co-culturing PSMA positive LNCaP cells pre-treated with DNMT and/or HDAC inhibitors with T-cells from PSMA-vaccinated mice. Activation of T-cells was increased when co-cultured with treated cells compared to DMSO controls and was dependent on HLA-I accessibility. These results show that HLA-I is epigenetically regulated in prostate cancer and inhibition of DNMT and HDAC proteins leads to functional re-expression of HLA-I at the cell surface.

While this study focused on epigenetic regulation of HLA-I genes, there are many other genes involved in maintaining MHC-I expression at the cell surface. Multiple genes involved in antigen processing, including genes encoding the subunits of the immunoproteasome and genes involved in MCH-I trafficking, are downregulated in the prostate cancer cell lines used in this study. Preliminary studies suggest that DNMT and HDAC inhibition may induce expression of the antigen processing machinery (APM) genes (Appendix A). Epigenetic silencing in genes involved in antigen presentation may

drive the expression changes in APM proteins observed in prostate cancer, which could contribute to loss of MHC-I expression at the cell surface. It is a goal of future studies to explore the interconnected pathways that contribute to functional MHC-I expression. Analysis of genome-wide methylation and histone modifications in cell lines and patient samples would give insight into the pattern of epigenetic silencing that contributes to loss of MHC-I expression overall.

Overall, these results indicate a key role for epigenetic regulation of HLA-I expression in prostate cancer and that HLA-I loss is reversible with epigenetic therapy (Figure 5.1). Future studies will need to focus on the utility of this type of treatment in patients. Identifying patients with epigenetically silenced HLA-I is necessary to perform these studies and would allow for more personalized therapy decisions in the future, specifically in regards to epigenetic and immunotherapy.

Development of SEEMLIS

Sensitively and specifically measuring epigenetic signatures in patients is essential to the development of epigenetic-based biomarkers. Tumor tissue is difficult to obtain from patients with prostate cancer. Biopsies from the primary tumor offer the most readily available source of tumor DNA, but is often only an option in early stage prostate cancer due to prostatectomy. In later stage disease, metastatic lesions can also serve as a source of tumor DNA, however prostate cancer most often metastasizes to the bone, which requires a painful and invasive procedure to biopsy. Liquid biopsies, such as blood draws, have been proposed as a minimally invasive alternative to traditional tissue biopsies with the added benefit of being able to be performed repeatedly. Our lab has a

semi-automated platform for isolation of CTCs from whole blood by exclusion-based sample preparation (ESP), which enables CTC enumeration, protein staining, and nucleic acid extraction (139,171). In order to detect epigenetic signatures in CTCs, we sought to expand the capabilities of this platform to enable methylated DNA enrichment.

A method called COMPARE-MS had been developed by a collaborator of our lab to enrich for methylated DNA from heterogenous samples including tumor biopsies using a combination of methylation sensitive restriction enzyme digestion and enrichment with a peptide derived from the methyl DNA binding domain of methyl-CpG-binding domain protein 2 (MBD2-MBD) (168). We took advantage of the sensitivity and specify afforded by COMPARE-MS with the capabilities of our semi-automated CTC isolation system to develop SEEMLIS, semi-automated ESP-based enrichment of methylated DNA from low-input samples. Using this assay, we were able to successfully detect *GSTP1* methylation from CTC samples estimated at ≥1% purity. To detect methylation in multiple genes to enable multiplexed analysis of DNA methylation in the same cells, we generated pure CTC populations by single cell aspiration and added a pre-amplification step. In summary, SEEMLIS was able to sensitively and specifically detect methylation signatures from low-input CTC samples, suggesting this assay could be used to develop epigenetic biomarkers, such as methylated HLA-I, in prostate cancer.

SEEMLIS was developed with analysis of prostate cancer CTCs, and specifically for analysis of HLA-I methylation, in mind, however there are far reaching applications for this assay. SEEMLIS is inherently flexible in both the sample types that can be used and the endpoints that can be measured. While we developed this assay using CTCs as our input, any source of DNA could but used provided methylation signatures remain intact.

For example, we have used this assay to measure methylation in biopsies, spheroid cultures, and cfDNA from plasma. Other sources of DNA that may be applicable include urine, seminal fluid, cerebrospinal fluid, pleural effusion, and ascites. Each of these sources of DNA may provide unique insights into the course of disease and disease biology. Additionally, this assay is not limited to studying methylation in cancer. Important methylation changes occur in other disease states as well as non-disease related cellular development and differentiation and aging. The ability to perform methylation analysis with such low-inputs also allows for multiple analyses to be performed on one sample, increasing the amount of information that can be obtained from one patient sample. In addition to flexibility in sample type, the types of analyses that can be performed on the enriched DNA can also be interchanged. Enriched DNA can be used in any downstream assay that is compatible with measuring fragmented DNA. Methods of analyzing enriched DNA include gel-based analyses to look at global methylation levels, traditional and multiplexed qPCR, and sequencing. Overall, the flexibility of SEEMLIS makes this a powerful tool to study disease biology using rare and low-input samples.

HLA-I methylation as a biomarker in circulating tumor cells

Prior to the work described in this thesis, loss of HLA-I protein expression had been observed in primary and metastatic prostate tumor lesions. However, the status of HLA-I expression in tumor cells in circulation was unknown. Here we performed the first analysis of HLA-I protein expression in circulating tumor cells (CTCs). We found that while CTCs had heterogenous expression of both total and surface HLA-I, we were able to detect populations of CTCs with strikingly low HLA-I expression when compared to the white

blood cell (WBC) population. Cells in circulation with low HLA-I expression are likely not able to be located and killed by circulating T-cells, contributing to persistence and survival of this cell population. In this way, HLA-I loss may be contributing to the development of metastases since CTCs are thought to have metastatic potential. Treatments that reverse HLA-I loss may help boost the ability of the immune system to locate tumor cells before they are able to seed at a metastatic site.

Prior to this study, HLA-I methylation had not been assessed in prostate cancer CTCs. The data in this thesis show that epigenetic loss of HLA-I can be reversed to functionally induce HLA-I expression. However, we also found that epigenetic signatures and responses to different classes of epigenetic modifying agents vary. Therefore, it is important to know whether HLA-I loss in each patient is due to epigenetic regulation and if so, what kind of regulation. Understanding the regulatory mechanisms involved in HLA-I loss may allow physicians to personalize therapies to match the epigenetic signature in HLA-I from each patient. We first sought to use the SEEMLIS assay to detect methylated HLA-I as a biomarker for patients with HLA-I regulated by methylation. These patients may be candidates for future clinical trials with DNMT inhibitors to induce expression of HLA-I. To assess HLA-I methylation in CTCs, we generated pure CTC populations by single cell aspiration stratified by HLA-I protein and measured methylation in *HLA-A*, *HLA-B*, and *HLA-C*. To our knowledge, this represents the first targeted analysis of HLA-I methylation in CTC samples. We found that methylation in the HLA-I genes varied by patient. We were able to identify three patients with evidence of methylation in one or more HLA-I genes in CTC populations with low HLA-I expression. HLA-I methylation was also detected in some of the HLA-I positive CTC subset. HLA-I expression in the HLA-I

positive CTC subsets was still lower than the mean expression in matched PBMCs, suggesting diminished, non-negative HLA-I expression may also be regulated by epigenetic mechanisms. Importantly, these patients may still respond to epigenetic therapy. Alternatively, we hypothesize that further optimization of the assay specifically to target only areas of methylation that are necessary or required for HLA-I inactivation may reveal clearer differences in the HLA-I positive and negative groups. Investigation into the correlation of HLA-I methylation in CTCs and clinical outcomes is also warranted and a goal of our future studies. In summary, we were able to use SEEMLIS to perform the first analysis of HLA-I methylation in prostate cancer CTCs and discovered that HLA-I methylation can be detected in CTCs with low HLA-I protein expression, proving potential for development of a clinical biomarker.

FUTURE DIRECTIONS

Evaluate the effect of EZH2 inhibition on HLA-I induction and epigenetic signatures

The data in this thesis show that epigenetic mechanisms, including DNA methylation within the CpG islands and a loss of normal levels of histone H3 lysine 27 acetylation, regulate HLA-I expression in prostate cancer. We demonstrated the inducibility of HLA-I expression by two DNMT inhibitors and an HDAC inhibitor. However, these drugs represent only a small subset of the epigenetic modifying agents that are currently being studied in pre-clinical models and clinical trials. Therefore, we are interested in evaluating the effect of other classes of epigenetic modifying agents for their ability to re-express epigenetically silenced HLA-I in prostate cancer.

One target that has been increasingly studied for its role in epigenetic regulation in many cancer types is EZH2. EZH2 is a methyltransferase responsible for writing lysine 27 methylation on histone H3 and is a part of the polycomb repressive complex 2 (PRC2). There are currently dozens of active trials for drugs targeting EZH2, including multiple for treatment of prostate cancer (207). Tazemetostat, a first-in-class EZH2 inhibitor, was approved this year for treatment of epithelioid sarcoma (141,142). This is the first epigenetic therapy approved by the FDA for use in a solid tumor. One study found that EZH2 was a critical regulator of MHC-I loss in neuroblastoma, Merkel cell carcinoma, and small cell lung cancer (208). We discovered an increase in H3K27me3 in the HLA-I genes in prostate cancer cell lines compared to non-cancerous cells, which was highly negatively correlated to HLA-I gene and protein expression. We hypothesize that EZH2 inhibitor treatment of the prostate cancer cell lines identified in this thesis as having increased levels of H3K27me3 in the HLA-I genes would result in decreased levels of

H3K27me3 and increased levels of H3K27ac with corresponding increases in HLA-I gene and protein expression. Preliminary studies from our lab have shown promising results in the ability of EZH2 inhibition to induce HLA-I expression similarly to DNMT and HDAC inhibition. Future studies will focus on treatments with EZH2 inhibitors alone and in combination with DNMT and/or HDAC inhibitors and their effect on improving immunogenicity of prostate tumors.

Investigate how HLA-I DNA methylation signatures are written and maintained in prostate cancer cells by DNMTs

One of the major questions that has arisen after we characterized the DNA methylation and H3K27 signatures in prostate cancer cell lines was how these signatures are being written and maintained. We determined that DNMT and HDAC proteins are involved in HLA-I regulation from the data presented in this thesis on DNMT and HDAC inhibition in cell lines. However, because these inhibitors are not targeted to a specific DNMT or HDAC family member, we do not know which members are responsible for writing and erasing the corresponding epigenetic marks. We obtained preliminary data using siRNA to knock down DNMT3a and DNMT3b alone and in combination followed by qPCR and MBD2-MBD enrichment to measure HLA-I induction and changes in DNA methylation (Appendix A). Interestingly, HLA-I was not significantly induced in LNCaP and LAPC4 cells despite decreases in DNA methylation. It is possible that repressive histone signatures are still preserved at the HLA-I genes in these cells, causing the genes to remain silenced. The opposite effect was seen in PC3 cells where siDNMT treatment led to significant increases in HLA-I gene expression, but DNA methylation was increased

rather than decreased. These results suggest complicated and varying biological mechanisms regulate HLA-I epigenetic changes. We need to further investigate this mechanism and perform necessary control experiments to understand these differences.

Investigate the contribution of HLA-I promoter and exon 2 DNA methylation to loss of HLA-I expression.

The CpG islands of the HLA-I genes are very large at around 1kb each. We found that methylation occurs across the entirety of the CpG islands and is not confined to the promoter region. Traditionally, promoter methylation and methylation within exon 1 has been associated with repressed genes while intragenic methylation has been associated with actively transcribed genes, preventing alternative splicing, and preventing alternative transcription start sites from being used (209-211). This paradigm has started to shift as methylation and its control of gene expression are beginning to be better understood. For example, we found that methylation within exon 2 of the HLA-I genes was highly negatively correlated with HLA-I expression. Interestingly, our study found that promoter methylation in HLA-I was not always negatively correlated with gene expression and generally occurred at lower levels than exonic methylation. We are interested in determining if promoter and exonic methylation in HLA-I is necessary and/or sufficient for HLA-I silencing. Knowing which areas of methylation are required for HLA-I repression would help us to develop a better biomarker and better understand epigenetic regulation of HLA-I.

To study the specific areas of methylation that may control HLA-I expression, we want to employ a CRISPR-based system for targeted DNA de-methylation as described

in a paper by Xu et al. (212). This method utilizes a catalytically dead Cas9 protein fused to the catalytic domain of TET1 (dCas9-TET1-CD) to specifically de-methylate regions of DNA. dCas9-TET1-CD is guided to the region of interest by a small guide RNA (sgRNA) and TET1-CD will catalyze the de-methylation of methyl groups in the region. In this way, we can specifically de-methylate either the promoter or exon 2 regions of HLA-I genes and evaluate gene expression as a result of de-methylation. Additionally, this experiment can determine whether removing methylation in the HLA-I genes alone is enough to re-express the genes, since drug treatments and siRNA experiments will affect methylation genome wide.

If this experiment is successful, there are other dCas9 fusion constructs that might be of interest. For example, dCas9 fused to histone acetyltransferase (HAT) catalytic domains could reveal whether increased histone acetylation at HLA-I alone is enough to induce gene expression of HLA-I. Another option is to transfect dCas9 fused to HDAC or DNMT proteins into cells that express normal levels of HLA-I to introduce repressive epigenetic signatures at specific regions of the gene. Adding repressive modifications is an alternative way to demonstrate what epigenetic modifications are required for gene silencing as well as the locations that are most effective.

Preliminary experiments using dCas9-TET1-CD with sgRNAs targeting the promoter and exon 2 regions of *HLA-A* have shown very little impact on gene expression (Appendix B). However, it is possible we need to alter technical aspects of the experiment such as cell lines used or titration of reagents and plasmid for transfection in order to see the effects of methylation loss on HLA-I. Additional control experiments also need to be performed to confirm dCas9 is reaching the target location in the cell. Nevertheless, this

study would provide valuable insight into the mechanism behind epigenetic regulation of HLA-I as well as help optimize HLA-I methylation as a biomarker.

Investigate epigenetic regulation of MHC-I assembly and antigen processing machinery

Loss of expression of HLA-I is only one aspect of MHC-I downregulation in prostate cancer cells. Dysregulation or loss of expression of any of the numerous proteins involved in antigen processing, assembly of MHC-I molecules, or transport of MHC-I molecules to the cell surface can result in decreased MHC-I expression. Many of these proteins, including TAP2 and B2M, have been reported to be downregulated in prostate cancer (36). The inhibitors that were used in this thesis will have cell-wide and genome-wide effects. While we have focused on the response of HLA-I expression and epigenetics to these inhibitors, it stands to reason that re-expression of HLA-I is enhanced by other genes affected by the inhibitors. The dCas9-TET1-CD study described above will shed some light on the contribution of other induced genes to HLA-I induction by DNMT inhibitors since HLA-I should be the only de-methylated gene in these experiments. However, it would also be prudent to evaluate induction of a panel of these genes in response to DNMT and HDAC inhibitors followed by epigenetic analysis of genes that are induced by these inhibitors. Preliminary data from our lab suggests some of the APM genes are indeed induced by DNMT and HDAC inhibition, however we have not examined the epigenetic signatures present in these genes (Appendix A). The epigenetic signatures in these genes will be further investigated in future studies.

Further develop the SEEMLIS assay to be completely automated

One of the greatest advantages of SEEMLIS over traditional tube-based approaches to methylated DNA enrichment is the inclusion of automated steps. Automation is a powerful tool for biological assays due to the reduced hands on time for the researcher and reduction in human error. The increased accuracy and precision afforded when using automation is crucial when working with low-input samples such as DNA from single cells or rare cell populations including CTCs. Therefore, increasing the level of automation in the SEEMLIS assay would potentially increase the assay sensitivity even further.

The Gilson EXTRACTMAX platform on which SEEMLIS was designed has inherent flexibility due to multiple deck spaces for various, interchangeable uses including assay plates, tip boxes, waste boxes, and tube racks of various sizes. Additionally, our lab has been successful in working with engineers to 3D print custom devices to address non-traditional assay needs. Therefore, it is within the capabilities of our lab to have the SEEMLIS assay be virtually completely automated from CTC capture through methylated DNA enrichment.

Currently, we are able to capture CTCs and extract DNA in an entirely automated manner. At this stage, the DNA is removed from the assay plate by manually pipetting the elution volume out of the elution well and into PCR strip tubes in which DNA digestion and MBD2-MBD binding is performed. The MBD2-MBD-bound methylated DNA and beads is then washed and eluted on the automated platform. In order to move the DNA digestion and MBD2-MBD binding steps onto the robot, new elements need to be added to the robot capabilities. For restriction enzyme digestion, samples will need to be heated to 37^{o}C for 15 minutes so a heating element would need to be added to the robot deck.

We would also need to confirm minimal loss of sample volume due to evaporation will take place during this step. If evaporation proves to be problematic, other steps would need to be taken to ensure volume retention, such as the addition of a plate cover or increased digestion volume. For MBD2-MBD enrichment, a method for continuous sample mixing would need to be included in the automated protocol. Continuous mixing may be accomplished by magnetic-based mixing where we employ the built-in magnetic system to move the beads up and down through the solution. Another method may be to have the robot periodically mix the samples by pipetting up and down. It may also be possible to have efficient binding if the volume of the binding reaction is low enough. We have successfully enriched methylated DNA in volumes lower than 5uL without mixing using a patent-pending method for volume free reagent addition, suggesting that lowering our reaction volume may eliminate or lessen the need for on-robot mixing. Implementing and optimizing these additions may increase sensitivity and ease of use of the assay and are worth pursuing, especially for clinical applications.

Further develop HLA-I as a biomarker

Chapter 4 of this thesis delves into the possibility that DNA methylation in HLA-I could be used as a biomarker to identify patients with epigenetically silenced HLA-I. Being able to identify these patients may be useful to determine which patients would not benefit from certain types of immunotherapy or which patients may benefit from epigenetic therapy alone or in combination with immunotherapy. We were able to successfully detect HLA-I methylation in CTCs from patients with prostate cancer with low HLA-I expression. While

these results are promising, changes can be made to improve sensitivity and specificity of detection.

Two major areas of improvement are choice of restriction enzymes and primer design. The restriction enzymes used in the study in Chapter 4 were chosen for their compatibility with a larger panel of genes and could be further optimized specifically for the HLA-I genes. The DNA fragments generated by these enzymes are approximately 1kb long, extending from the promoter to exon 3 and encompassing the majority of the CpG island of each gene. The advantage of a larger fragment is that we are able to capture and measure methylation anywhere within the CpG island and proximal promoter regions of the genes, which was ideal for early stage studies since we did not have a clear idea of which areas of methylation were common in the patient population. However, having such a large DNA fragment also has disadvantages. One disadvantage is that we may be detecting methylation that is not important for gene silencing. The results of future studies exploring targeted de-methylation or methylation of the HLA-I genes as described above may inform us of the ideal locations in each gene to target for biomarker development. Enzymes could then be chosen that frame the area of interest better. Another disadvantage is that the HLA-I genes are methylated in normal cells beginning near the end of exon 3. This may explain why HLA-I methylation is detected at all in WBC samples. We have begun to test a new restriction enzyme, BstY1, that would cut the DNA closer to exon 2 (Figure 5.2A). Adding BstY1 would potentially improve specificity by removing areas from the *HLA-A* and *HLA-C* target DNA fragments near exon 3 where normal cells are methylated. Preliminary data shows that adding BstYI reduced the methylated *HLA-A* signal from WBCs to undetectable levels, while preserving the signal

from LNCaP cells (Figure 5.2B). *LINE1* signal was comparable across all digestion and capture conditions (Figure 5.2C). Increasing the sensitivity of the assay in this way may allow us to detect HLA-I from CTC samples without the need to employ the single cell aspirator, saving time and effort.

To improve primer design, we need to take into consideration the polymorphic nature of HLA-I. The primers used in this study will be optimized only for HLA-I haplotypes that share the same polymorphisms. While we know primers can handle some degree of mismatch, we would not know the level of mismatch or location of mismatches without knowing the HLA-I type of the patient. Therefore, it may be necessary to type each patient using patient matched WBCs and have a bank of HLA-I primers for SEEMLIS analysis to choose from based on the individual HLA-I signature from each patient. Another solution would be to rely on sequencing each sample rather than performing qPCR. We could then also sequence the genome from patient matched WBCs to account for issues aligning polymorphic regions to a reference genome. In either case, accounting for the polymorphic nature of HLA-I will allow us to develop a better biomarker.

Perform ChIP and/or ATAC-seq on CTCs

Prior to this thesis work, HLA-I expression and epigenetic signatures in CTCs had not been characterized. As described in Chapter 4 of this thesis, we identified populations of CTCs from patients with prostate cancer that had reduced HLA expression compared to white blood cells. We were also able to detect methylation of HLA-I genes in CTCs from patients with prostate cancer. While methylated HLA-I as a biomarker has promise, we

are also interested in pursuing alternative epigenetic biomarkers as well as investigating the epigenetic landscape in CTCs at a deeper level.

In Chapter 2 of this thesis we found that HLA-I gene and protein expression was more tightly correlated to histone modifications rather than methylation signatures. This suggests that a histone modification may be a more predictive biomarker for epigenetic silencing of HLA-I than DNA methylation. However, analysis of histone modifications has traditionally required large amounts of starting material. Traditional ChIP assays typically require cell inputs in the range of millions of cells. More recently, there has been a push to generate protocols for low-input ChIP. However, even low-input ChIP protocols still require cell inputs of around 1000 cells, which is 10-100 times more than a typical CTC yield (213). While single cell methods for ChIP-seq exist, these methods require sequencing of 1000s of single cells to generate an overall picture of the cell population and are unable to accomplish targeted analysis of specific genes in single cells due to low coverage per individual cell (117,214). Therefore, in order to evaluate histone modifications in the HLA-I genes as biomarkers, a new method for targeted ChIP analysis would need to be developed.

As described above, one of the advantages of the ESP-enabled Gilson EXTRACTMAX automated system is its flexibility. Additionally, many of the steps of MBD2 peptide-based enrichment of methylated DNA used in the SEEMLIS assay have parallels to ChIP. Both methods require shearing of DNA, binding of DNA to beads linked to a capture molecule, and several wash steps. As such, the SEEMLIS method can be used as a backbone to develop a ChIP assay for CTCs. One important difference between the two assays that will need to be considered is the use of an antibody for precipitation

for ChIP compared to a peptide for SEEMLIS. The differences in binding dynamics and non-specific binding capacity of antibodies and the MBD2-MBD peptide will need to be characterized and accounted for. Another key difference in the two assays is the method of DNA shearing used. SEEMLIS employs restriction enzymes to cut the DNA in predictable ways, which theoretically allows every available piece of target DNA to be analyzed. This is crucial for low-input assays. On the other hand, ChIP methods typically use sonication or MNase digestion to shear the DNA. These methods cut the DNA randomly and do not necessarily preserve every piece of target DNA. However, digesting CTC DNA with restriction enzymes prior to immunoprecipitation may be a viable option to combat this issue. With these modifications applied to the SEEMLIS method on our automated system, it may be possible to develop an assay for ChIP in CTCs.

While targeted ChIP analysis of CTCs would allow for gene specific biomarker development, it is not ideal for discovery-based experiments to learn about CTC biology and heterogeneity. ChIP-seq can be used for these types of studies to identify epigenetic signatures that are associated with various CTC characteristics. We could perform low coverage ChIP-seq on thousands of single CTCs or smaller pools of CTCs using existing methods for low-input ChIP-seq (87,213,215). ChIP-seq would provide insight into the epigenetics of CTCs at the population level. Additionally, if CTC samples are stratified on certain characteristics such as treatment response or expression of specific genes or protein, ChIP-seq may be able to inform on key differences in histone modification patterns in the different groups.

Another method for analyzing chromatin-based epigenetic signatures in CTCs is ATAC-seq. ATAC-seq is a way to measure chromatin accessibility in the genome. Similar

to single cell ChIP-seq, single cell ATAC-seq requires hundreds or thousands of cells sequenced at low coverage to be informative. However, if we are able to perform this analysis on enough CTCs, ATAC-seq could be very informative on the general epigenetic state across the genome in CTCs. One advantage to ATAC-seq over ChIP-seq is not having to choose a single histone modification to work with. While ATAC-seq does not specifically look at histone modifications, in many cases the amount of chromatin accessibility at a given gene promoter measured by ATAC-seq can give a general idea of what chromatin modifications may be present at that gene. For example, if the ATAC-seq signal is low at a specific gene, we may expect to find a repressive histone signature including H3K27me3 at this gene. Global ATAC-seq analysis could therefore point us in the direction of genes that may be important for prostate cancer progression and metastasis or for treatment response, which may then may be developed as biomarkers using targeted ChIP analysis.

Further develop the SEEMLIS assay to support MBD-seq capabilities

In addition to ChIP-seq and ATAC-seq to evaluate chromatin-based epigenetic signatures, there are methods to evaluate DNA methylation signatures genome-wide. One method, MBD-seq, utilizes the MBD2-MBD peptide to enrich methylated DNA prior to sequencing the enriched DNA. This method identifies enriched areas of the genome, which are considered to have been methylated in the original cell population.

We are currently working to adapt a next generation sequencing method to include a step for MBD2-MBD enrichment by SEEMLIS to study DNA methylation in CTCs. The method we are developing uses an enzyme-based approach to shear the DNA prior to

adaptor ligation to generate libraries. After adaptor ligation, the libraries are run through SEEMLIS beginning with the MBD2-MBD enrichment step. Libraries are then enriched by PCR with unique dual index primers to prepare them for sequencing. As a control, each sample will be split and run in parallel with one enriched using MBD2-MBD and one un-enriched. I have successfully generated enriched and un-enriched libraries from cell line samples and am working to make necessary adjustments based on the quality and quantity of the libraries to generate libraries from CTC samples.

One potential pitfall of this method is the presence of WBCs in the CTC samples. To attempt to combat false positive readings of tumor methylation, we are also including a matched WBC sample from each patient. This inclusion will allow us to know whether a methylation signature seen in the CTC sample is unique to the CTCs or present in the WBC population. If the background WBC signal turns out to mask signal from the CTCs, we can use the single cell aspirator to further purify the CTCs prior to sequencing. Another potential issue is differences in methylated DNA capture efficiency between samples. This could cause biases in sequencing or problems with false positive or negative results during analysis. To alleviate this, we are spiking enzymatically methylated lambda phage DNA into each sample to have an internal control for the efficiency of capture of methylated DNA.

Adding MBD-seq to the SEEMLIS repertoire would take the assay to a higher level of utility since genome-wide analyses on low-input samples and single cells is quickly becoming the gold standard in evaluating tumor biology and heterogeneity. Prostate cancer is generally a heterogenous disease and it stands to reason that differences in epigenetic regulation could contribute to this heterogeneity. Generating a picture of the

CTC DNA methylation landscape in prostate cancer would also lead to important advances in the understanding of prostate cancer metastasis.

Figure 5.1. Schematic of epigenetic and gene expression changes in HLA-I in response to DNMT and HDAC inhibition.

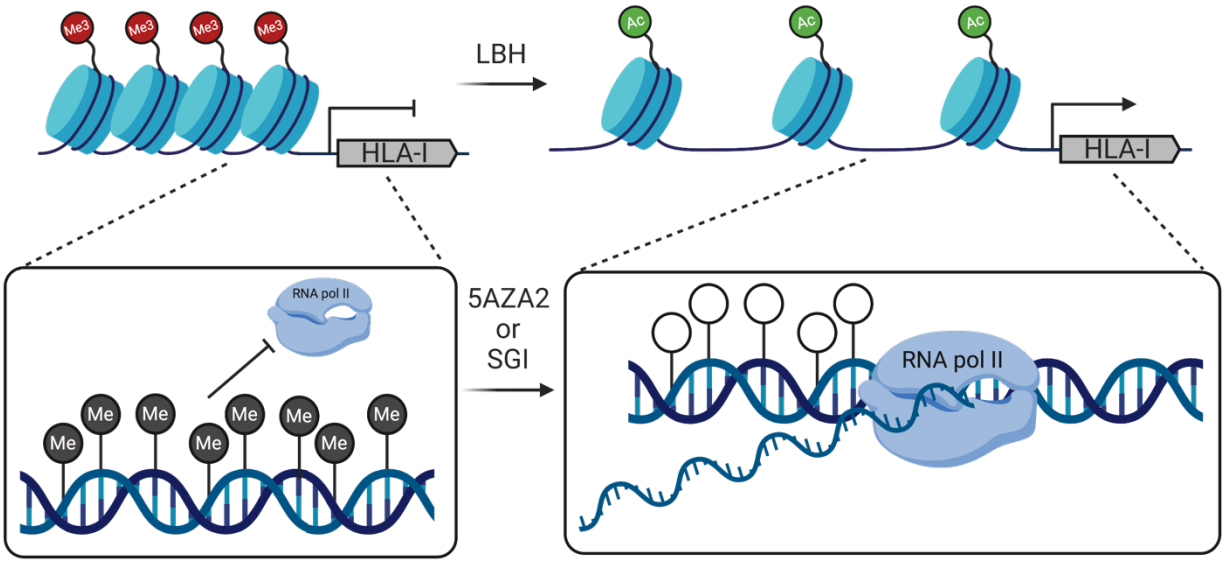


Figure 5.1. Schematic of epigenetic and gene expression changes in HLA-I in response to DNMT and HDAC inhibition. At baseline, HLA-I genes are epigenetically silenced by repressive histone and DNA methylation signatures. LBH treatment inhibits HDACs and leads to increased levels of H3K27 acetylation, which is associated with an open chromatin state. 5AZA2 or SGI treatment inhibits DNMTs and leads to hypomethylation at the HLA-I genes. Both of these epigenetic changes allow RNA polymerase II to bind to the HLA-I promoters and induce gene transcription. Created with BioRender.com.

Figure 5.2. Potential alternative restriction enzyme combination for HLA-I biomarker studies

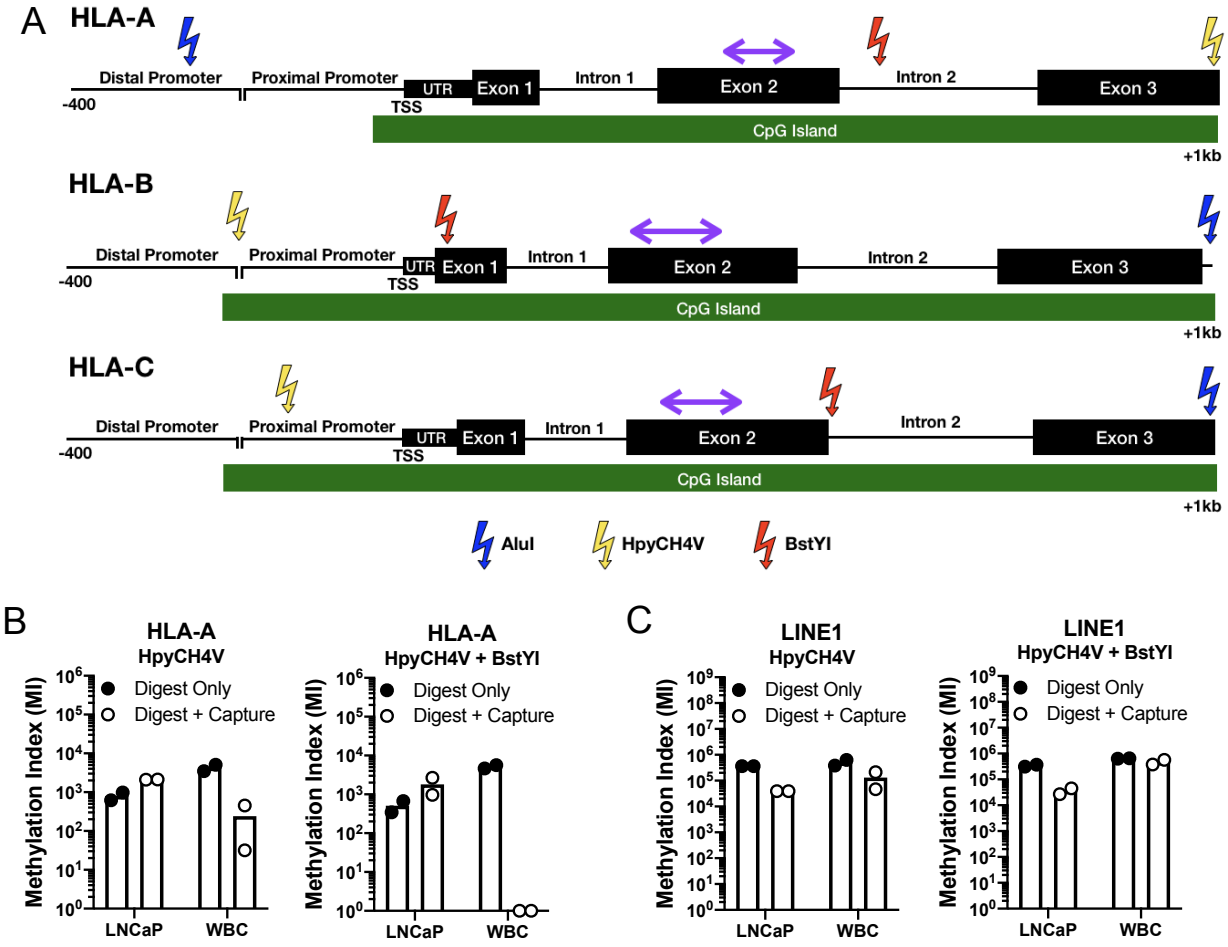


Figure 5.2. Potential alternative restriction enzyme combination for HLA-I. A) Schematic showing cut sites for BstYI (red), HpyCH4V (yellow), and AluI (blue) in the HLA-I genes. Primer location is shown with double sided arrow. B) Methylation index for (B) *HLA-A* and (C) *LINE1* is shown for LNCaP and WBC samples digested with AluI and HpyCH4V with and without the addition of BstYI.

Appendix A:

Additional Cell Line Epigenetic Signature Data and Epigenetic Modifying Drug Treatments

Portions of this appendix are adapted from the following publication in preparation: Rodems TS, Heninger E, Stahlfeld CN, Gilsdorf C, Carlson K, Kircher MR, Beebe DJ, McNeel DG, Haffner MC, Lang JM. Targetable epigenetic alterations regulate class I HLA loss in prostate cancer.

Contributions: Figure A.2A was performed in collaboration with C.G. Figure A.3A was performed by E.H. Figure A.3B was performed by E.H., K.C, and M.R.K. All other experiments and analyses were performed by T.S.R.

This appendix contains data that is supplementary to the studies in Chapter 2. I performed a more extensive analysis of DNA methylation and H3K27 modifications in the HLA-I genes. This was not essential to the story in Chapter 2, but provides extra proof of epigenetic control of HLA-I expression and more support for what I found in Chapter 2. This appendix also contains preliminary data demonstrating the ability of DNMT and HDAC inhibition to induce expression of other genes related to MHC-I expression. Additionally, I have included experiments using siRNA to knockdown the de novo DNMTs in LNCaP, PC3, and LAPC4 cells.

The DNA methylation experiments in this thesis were performed using MBD2-MBD peptide-based precipitation of methylated DNA fragments followed by qPCR. While this method was useful to gain an overall picture of DNA methylation in multiple cell lines and perform targeted analysis of DNA methylation in CTCs, it lacks single base pair resolution. In order to achieve this level of resolution, a method such as bisulfite sequencing must be used. I performed bisulfite sequencing on LNCaP and LAPC4 DNA for *HLA-A*, *HLA-B*, and *HLA-C* (Figure A.1). Technical issues related to this experiment precluded me from getting data for some of the regions. We decided not to pursue the experiment further after that. However, this data does confirm that LNCaP and LAPC4 cells have the heaviest methylation in exon 2 and 3, which was seen in the MBD2-MBD experiments and in TCGA methylation array data from Chapter 2. This data also shows the relatively lower level of methylation overall in LNCaP cells, which is also corroborated by the MBD2-MBD data. Also in line with our data in Chapter 2 is the lower level of methylation in the promoter regions compared to the exonic and intronic regions. Future studies to investigate the contribution of exonic and intronic methylation to gene expression will be

useful to determine whether these areas are drivers or passengers of HLA-I loss. Additionally, we were able to see that WBCs have a small amount of methylation late in Exon 3, confirming the utility of a new enzyme combination for methylation analysis in CTCs to avoid detecting this normal level of methylation in our patient samples.

Chapter 2 of this thesis demonstrates that prostate cancer cell lines with low HLA-I expression have repressive epigenetic signatures present at the HLA-I genes, including DNA and H3K27 methylation. In that Chapter, I showed selected primer sets and cell lines from the more comprehensive analysis that was performed. Figure A.2A shows the locations of additional primers that were used in this analysis. MBD2-MBD enrichment of DNA was performed in two additional cell lines a lymphoblastoid cell line (LCL), another prostate cancer cell line, Du145. We also did this analysis on two control samples: DNA from a DNMT double knockout cell line (DKO HCT116) and enzymatically methylated HCT116 DNA as negative and positive controls for DNA methylation respectively. Figure A.2B shows a heat map with our entire data set for MBD2-MBD enrichment. We see that across all three genes, LAPC4, LNCaP, and PC3 have the highest levels of methylation overall, in accordance with the conclusions in Chapter 2. Du145 has increased levels of methylation compared to the non-cancerous cell lines in certain gene regions as well. LCL cells had overall low levels of methylation, with the exception of regions towards the end of the *HLA-A* and *HLA-B* CpG islands, in line with our previous observations for WBC HLA-I methylation. We also had included an extra set of primers for *HLA-A*, *HLA-B*, and *HLA-C* in our ChIP experiments from Chapter 2. The data for baseline expression with the extra primers is shown in Figure A.2C. This shows the same trends as we found in Chapter 2, where H3K27ac is lost in prostate cancer cell lines and H3K27me3 is more

enriched compared with RWPE1. I also included control genes in these experiments. MYOD1 is a control for H3K27me3 enrichment and lack of H3K27ac, while RPL30 is a control for the inverse signature. I also included total H3 and IgG as controls in my ChIP experiments. I did not see significant changes in total H3 capture between the cell lines in the HLA-I genes (Figure A.2D). IgG signal was low across all experiments (Figure A.2E).

In Chapter 2, we measured the effects of SGI and LBH on HLA-I gene and protein induction and epigenetic signatures. We also did these same experiments using a different DNMT inhibitor, 5-aza-2-deoxycytidine (5AZA2). We saw similar results to our SGI experiments, which makes sense since the active form of SGI is the same molecule as 5AZA2. We saw that the combination of 5AZA2 and LBH was the most successful in inducing HLA-I gene and protein expression (Figure A.3A,B). We also again saw that 22rv1 cells did not respond to DNMT inhibition alone, but did respond to LBH alone. LNCaP cells were less responsive in this experiment, but overall showed similar patterns as the SGI/LBH study. PC3 cells were less responsive to 5AZA2 than SGI, though did respond to combination treatment. I also measured DNA methylation in LNCaP, 22rv1, and LAPC4 cells treated with DMSO and 5AZA2 (Figure A.3C). We saw an overall reduction in DNA methylation after treatment with 5AZA2 compared to DMSO treatment, supporting the results in Chapter 2.

This thesis focused on induction of HLA-I by epigenetic modifying agents. However, there are many other proteins involved in antigen processing and MHC-I expression at the surface. Three of these are beta-2-microglobulin (B2M), the immunoproteasome subunit LMP7, and calreticulin (CALR). As shown in Figure 1.2, B2M

is part of the MHC-I complex with HLA-I, LMP7 is a subunit of the immunoproteasome which digests proteins to generate the peptides that are loaded into complete MHC-I molecules, and CALR is a chaperone of MHC-I complex formation. One of the future directions of this project is to investigate how epigenetic mechanisms affect expression of these proteins. Figure A.4 shows preliminary data measuring induction of gene expression of *B2M*, *PSMB8*, the gene that encodes LMP7, and *CALR* in response to SGI and LBH. For *B2M* and *PSMB8*, a similar trend of induction was seen in LNCaP, 22rv1, and LAPC4 cells, where the combination treatment induced expression the most. Significant induction of *B2M* by LBH was seen in LNCaP and LAPC4 cells as well as by the combination treatment in LAPC4 cells. *PSMB8* was significantly induced by SGI alone in LNCaP cells and by LBH alone in LAPC4 cells. Again, the combination treatment had the highest overall level of induction, though the results were not significant due to variability in response across replicates. *PSMB8* expression was increased by LBH and the combination treatment in 22rv1, but again the results were not significant. Unlike *B2M* and *PSMB8*, *CALR* expression was not reduced in the cancer cell lines compared to RWPE1 and was not significantly induced by either drug or the combination in any of the cell lines, This demonstrates that not all genes involved in MHC-I expression are affected by epigenetic treatment. Overall, these preliminary results suggest that DMNT and HDAC inhibition has a wide impact on MHC-I processing, affecting other genes involved in MHC-I expression at the cell surface in addition to HLA-I. Investigations into the entire MHC-I processing network will reveal important interactions between the genes that are affected by epigenetic regulation changes in prostate cancer.

The DNMT inhibitors used in this thesis affect DNMT1, DNMT3a, and DNMT3b activity. As the maintenance methyltransferase, DNMT1 regulates the copying of methylation signatures from parent to daughter cells during cell replication. DNMT3a and DNMT3b are able to write methylation onto CpGs in response to cellular signals during cellular growth phase. This means that the inhibitors we used will affect the copying of methylation signatures as well as the writing of new methylation signatures. However, we are not able to easily distinguish between the two types of methylation loss. To address this ambiguity, I utilized siRNA to knockdown the de novo methyltransferases alone and in combination with each other in LNCaP, PC3, and LAPC4 cells and looked at gene expression and DNA methylation of HLA-I. These three cell lines were the ones identified in Chapter 2 to have increased methylation in the HLA-I genes compared to RWPE1. I was able to achieve knockdown of DNMT3a and DNMT3b at the gene expression level in all three cell lines (Figure A.5A). However, we have not yet been able to successfully measure knockdown at the protein level. Each cell line had a different response to the knockdowns (Figure A.5A). HLA-I gene expression was not induced in LNCaP cells by knockdown of either methyltransferase. HLA-I gene expression was induced in PC3 cells in response to knockdown of both methyltransferases, with the dual knockdown being most effective. In LAPC4 cells, *HLA-B* was increased in response to DNMT3a and the dual knockdown, but the results were not significant. Interestingly, *HLA-A* gene expression was actually slightly reduced in response to siDNMT3b in LAPC4 cells.

Next I looked at HLA-I DNA methylation in response to knockdown of each methyltransferase. Primer locations for this analysis are indicated in Figure A.5B. A heat map showing methylation levels across cell lines and conditions for each primer is shown

in Figure A.5C. Methylation was lost with knockdown of both methyltransferases and the combination in LNCaP cells, especially in the promoter regions of HLA-I. This would be the expected result if the de novo DNMTs are responsible are activity writing DNA methylation signatures on the HLA-I genes in the cell. In contrast, HLA-I DNA methylation was increased across all three genes in response to de novo DNMT knockdown in PC3 cells. This result was surprising due to the increase in gene expression I observed in response to these conditions. One explanation could be that DNMT1 is compensating for DNMT3a and DNMT3b activity when one or the other is lost, accounting for increased methylation levels. These results also suggest that DNA methylation in PC3 do not significantly regulate HLA-I expression, which would be in line with the minimal, statistically insignificant levels of HLA-I induction we observed in response to DNMT inhibitors in PC3 cells. HLA-I methylation in LAPC4 was reduced in response to knockdown of DNMT3a and DNMT3b alone, but was not decreased in the dual knockdown condition. Again, this may suggest that DNMT1 may be able to compensate for de novo DNMT activity in these cells. HLA-I induction patterns in LAPC4 support this idea since siDNMT3a increased HLA-A and HLA-B expression more than the combination conditions. Interestingly, siDNMT3b did not increase LAPC4 gene expression at all, but did reduce methylation levels, suggesting there are other regulatory mechanisms that are affected by DNMT3b knockdown that influence HLA-I expression. Overall, this experiment generates many new questions related to the regulation of HLA-I in prostate cancer by epigenetic mechanisms, which will be addressed in future studies.

.

Figure A.1. Bisulfite sequencing of the HLA-I genes in LNCaP, LAPC4, and WBC.

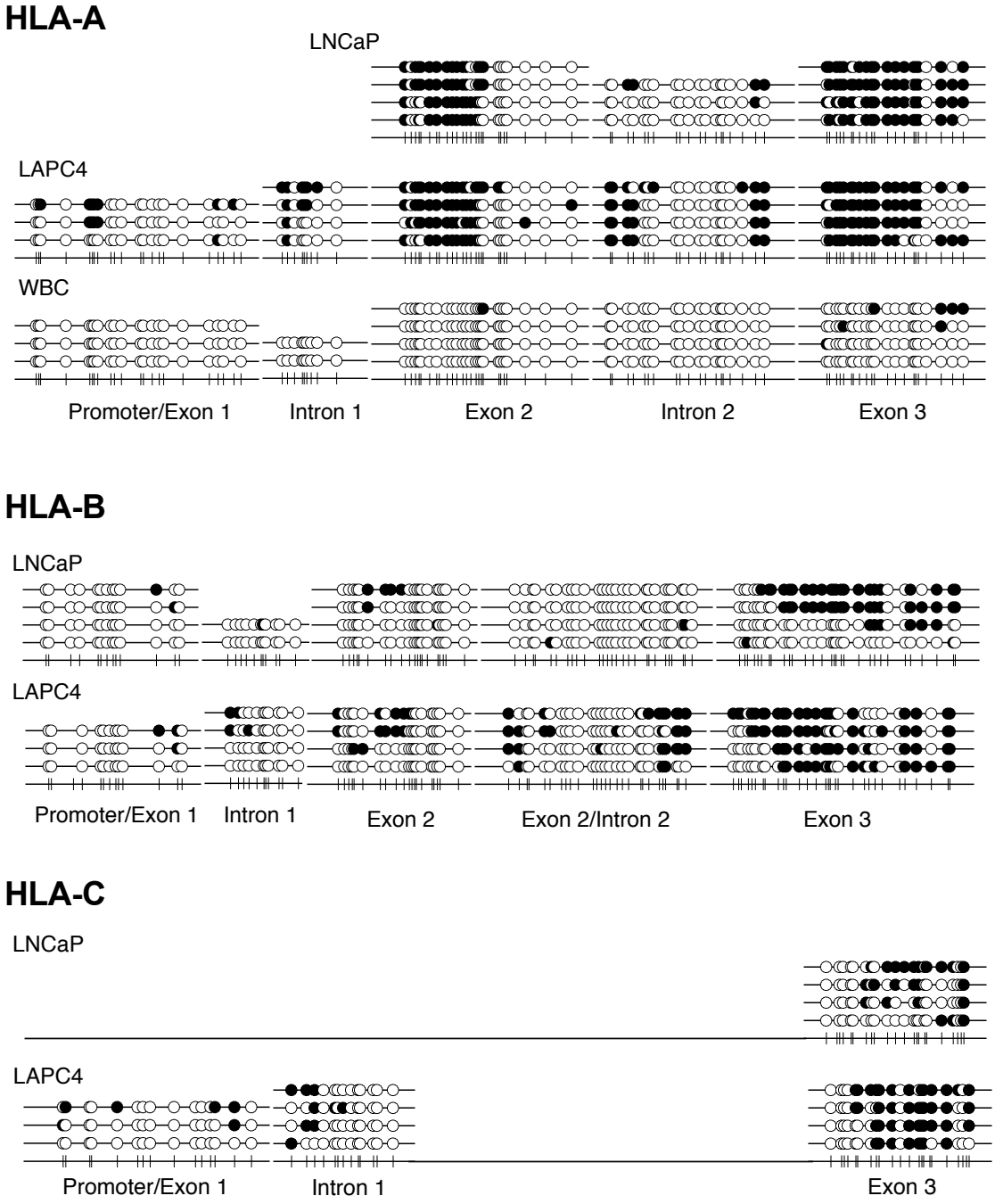


Figure A.1. Bisulfite sequencing of the HLA-I genes in LNCaP, LAPC4, and WBC. Bisulfite sequencing of three regions of *HLA-A*, *HLA-B, and HLA*-C in LNCaP, LAPC4, and patient-derived WBCs (*HLA-A* only). Black circles represent methylated CpGs, white circles represent unmethylated CpGs.

Figure A.2. Additional analysis of the epigenetic landscape of HLA-I in prostate cancer cell lines.

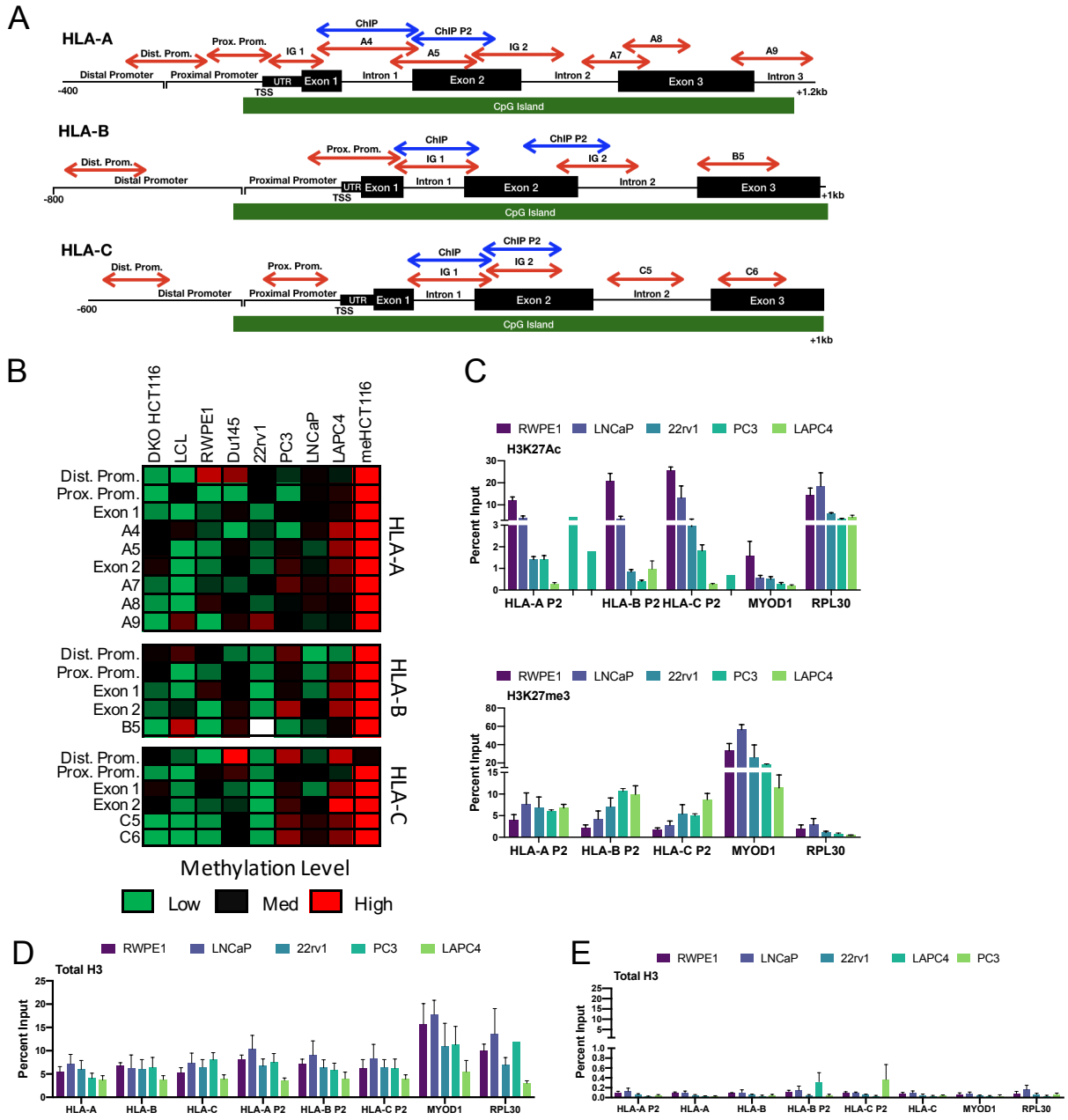


Figure A.2. Additional analysis of the epigenetic landscape of HLA-I in prostate cancer cell lines. A) Primer locations for DNA methylation and histone modification experiments. B) Heat map of DNA methylation in prostate cancer cell lines, RWPE1, a lymphoblastoid cell line (LCL). Enzymatically methylated HCT116 DNA and DNA from HCT116 cells with a double DNMT knockout phenotype are included as positive and negative assay controls, respectively. Data from Chapter 2 is included for reference with the additional cell lines and primers. ChIP with antibodies targeting H3K27ac and H3K27me3 (C) total histone H3 (D) and IgG (E) in prostate cancer cell lines and RWPE1.

Figure A.3. HLA-I induction and epigenetic changes by 5AZA2, SGI, and LBH.

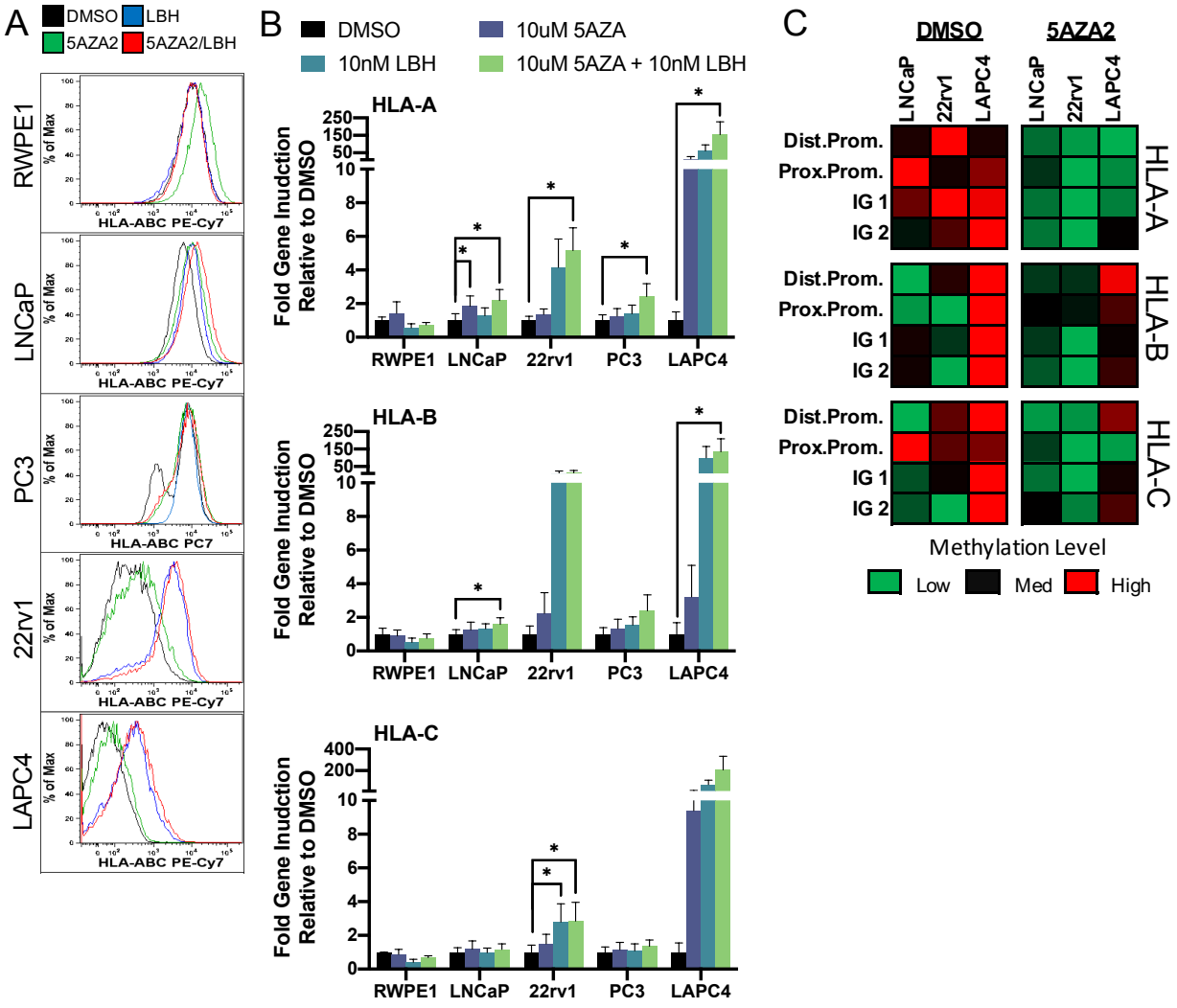


Figure A.3. HLA-I induction and epigenetic changes by 5AZA2, SGI, and LBH. A) Representative histograms of flow cytometry analysis of HLA-I protein induction by 5AZA2, LBH, or combination treatment with DMSO as a control. B) Gene induction by 5AZA2, LBH, or combination treatment relative to DMSO treated cells. C) Heat map showing DNA methylation in DMSO and 5AZA2 treated cells.

Figure A.4. Induction of *B2M* and *LMP7* by DNMT and HDAC inhibition in prostate cancer cell lines.

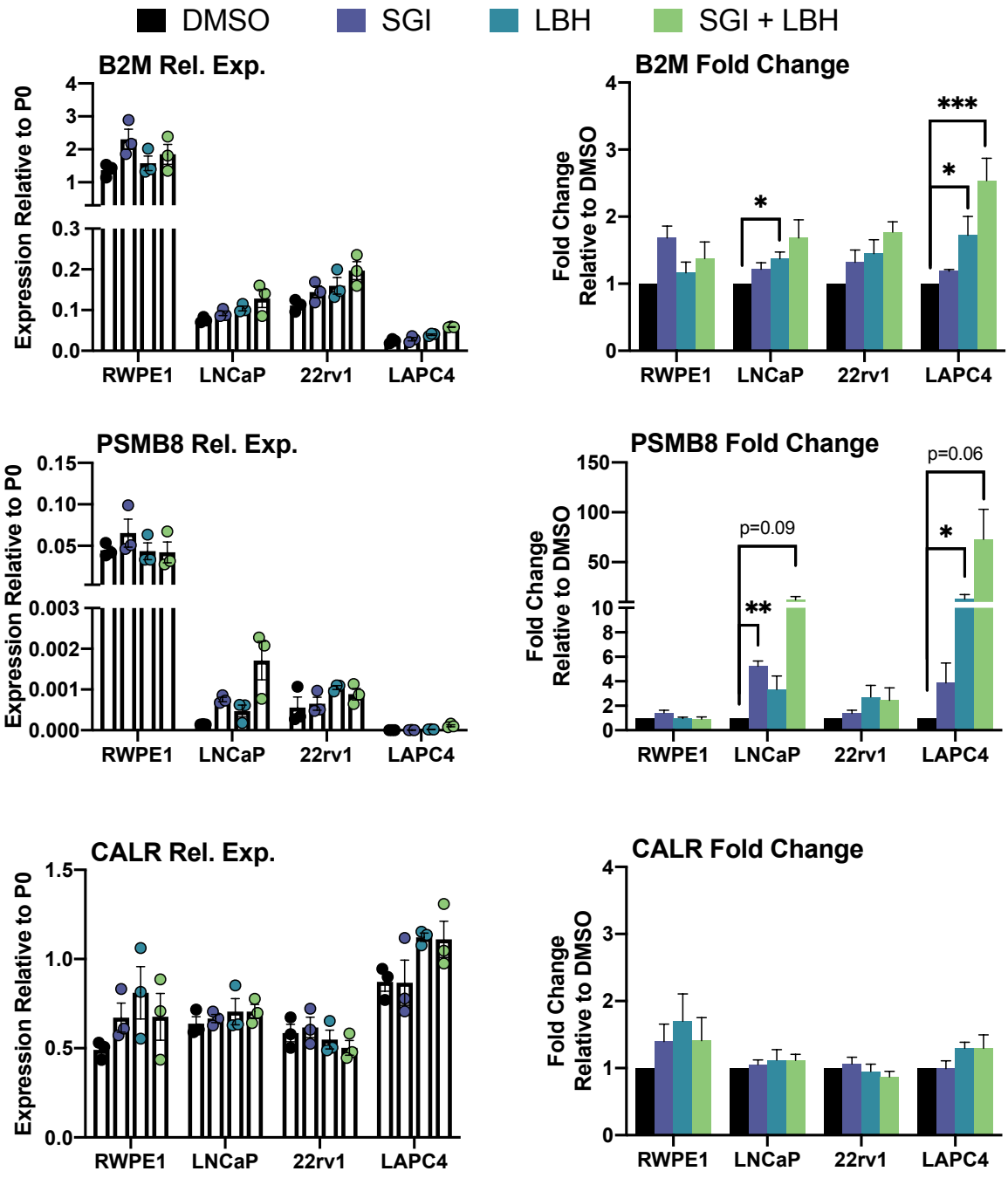


Figure A.4. Induction of *B2M* and *LMP7* by DNMT and HDAC inhibition in prostate cancer cell lines. Relative expression and fold change of *B2M*, *PSBM8*, and *CALR* in RWPE1, LNCaP, 22rv1, and LAPC4 cells treated with DMSO, SGI or LBH alone or in combination. Error bars represent SEM. $*$ $p<0.05$, $**$ $p<0.01$, $***$ $p<0.001$.

Figure A.5. Effect of siDNMT3a and siDNMT3b on HLA-I expression and DNA methylation in prostate cancer cell lines.

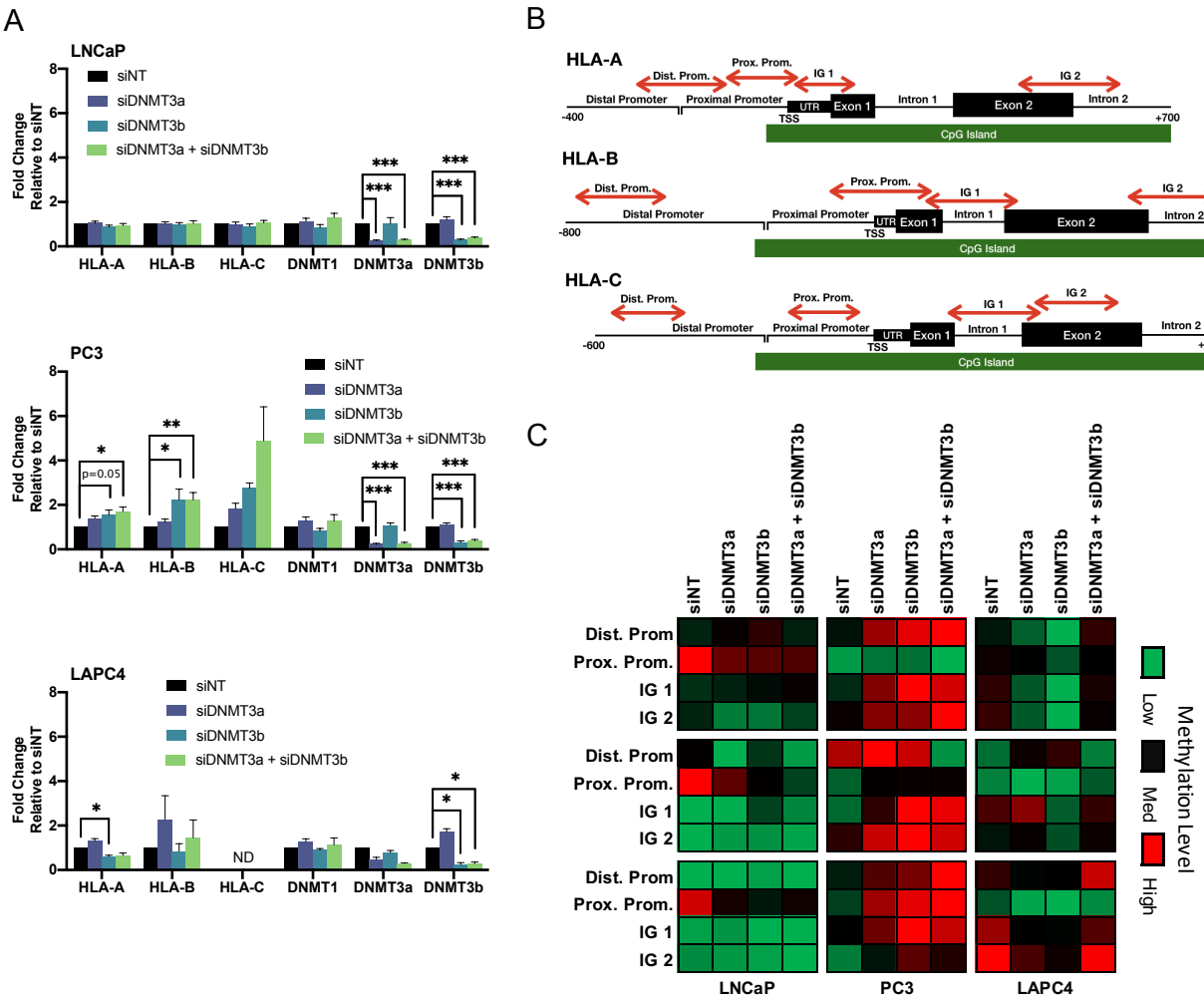


Figure A.5. Effect of siDNMT3a and siDNMT3b on HLA-I expression and DNA methylation in prostate cancer cell lines. A) Fold change in gene expression in cell lines transfected with siDNMT3a, siDNMT3b alone or in combination compared to cells transfected with non-targeting siRNA (siNT). B) Primer locations for DNA methylation analysis. C) Heat map showing DNA methylation in cell lines transfected with siNT, siDNMT3a, or siDNMT3b alone or in combination.

Appendix B:

CRISPR-dCas9 Mediated Targeted De-methylation of *HLA-A* by TET1-CD

Appendix B contains preliminary data related to the work in this thesis regarding the aberrant methylation of *HLA-A* in prostate cancer. The goal of these experiments were to determine which region or regions of the extensive methylation present in the *HLA-A* CpG island and promoter is essential for *HLA-A* silencing. Although the data presented in this thesis demonstrate that removal of methylation within the HLA-I genes by inhibition of the DNMT proteins is associated with an increase in gene expression, we cannot conclude that removal of DNA methylation in the HLA-I genes is sufficient for re-expression of HLA-I due to the global effect DNMT inhibitors have in the cell. Additionally, it would be prudent to biomarker development to know which regions of methylation are most important to gene expression regulation. Therefore, I sought to design an experiment where I could systematically de-methylate specific regions of the *HLA-A*, while not affecting methylation levels elsewhere in the cell.

To do this, I used a CRISPR construct with a catalytically dead Cas9 component (dCas9) which had been fused to the catalytic domain of TET1 (TET1-CD). This was developed by Xu et al. as a method for targeted removal of DNA methylation in the region where dCas9 is directed (212). Demethylation is accomplished by the fused TET1 catalytic domain. TET proteins begin the demethylation process of CpGs endogenously in the cell by catalyzing the conversion of 5mC to 5hmC. Xu et al. was able to use this system to successfully demethylate and re-express *RANKL*, *MAGEB2*, and *MMP2* in cell lines.

I designed two small guide RNAs (sgRNAs) targeted to different areas of the *HLA-A* gene: the promoter and exon 2. The locations of the sgRNAs are shown in Figure B.1A. I cloned the sgRNAs into the CRISPR-TET1-CD backbone and propagated the plasmids

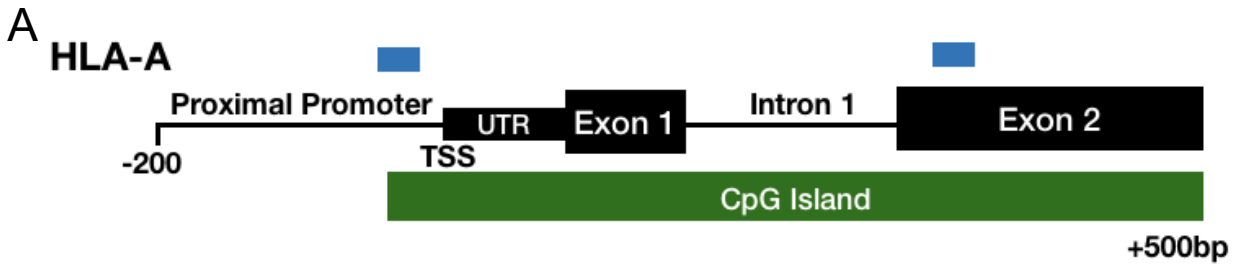
in E.coli. I mini-prepped the plasmid while growing a culture for midi-prep and sequencing the resulting DNA to confirm the insertion of my sgRNA. Annotated chromatograms for the results of sequencing are shown in Figure B.1B.

Next I transfected the CRSPR-TET1-CD plasmids into RWPE1, PC3 and LAPC4 cells. I measured gene expression of *HLA-A* in the cells transfected with the promoter or exon 2 sgRNA plasmids compared to cells transfected with the original plasmid which has a scrambled sequence inserted in the sgRNA plasmid location (-sgRNA). No change in gene expression was seen in cells transfected with exon 2 sgRNA compared to the -sgRNA condition (Figure B.2A). However, there was an approximately 50% decrease in gene expression in LAPC4 cells containing the promoter targeting sgRNA (Figure B.2B). It is possible that the reduction in gene expression could be due to the proximity of the promoter sgRNA to the TSS, where the dCas9 protein may be blocking the binding of RNA polymerase. This is supported by published studies that have used dCas9 targeted to gene TSSs to effectively block transcription (216,217). However, this reduction in gene expression was not seen in the other two cell lines. Experiments need to be done to confirm that the dCas9-TET1-CD protein is present at *HLA-A*, which can be done by ChIP analysis using a Cas9 antibody.

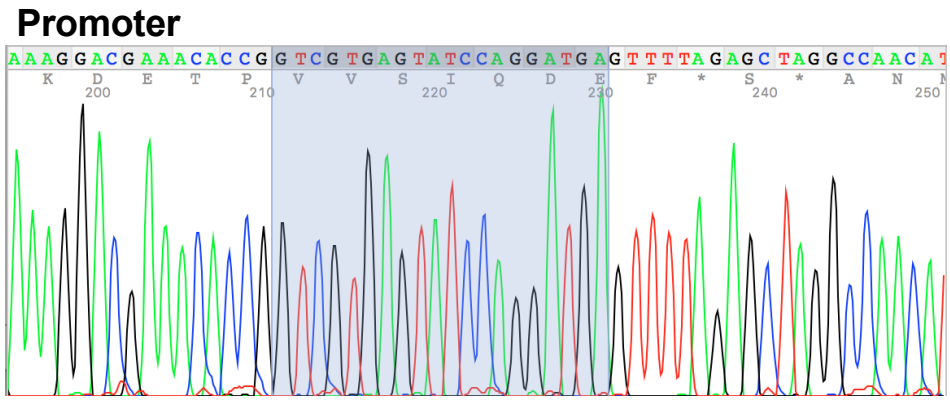
I measured methylation at three locations in *HLA-A* in cells transfected with the exon 2 or scrambled sgRNA plasmid to see if methylation was successfully removed. There was an approximately 50% reduction in methylation at the three regions in the exon 2 sgRNA transfected cells (Figure B.2C). This experiment has only been conducted once and only in one cell line, so additional experiments will need to be completed to draw conclusions from this data. However, preliminarily, it seems that loss of methylation alone

is not enough to induce expression of *HLA-A* in LAPC4 cells. This may be because repressive histone signatures are still present even if DNA methylation is removes. Future studies will also investigate the impact of HDAC inhibition or targeted histone tag modulation using a similar CRISPR-based system in combination with targeted de-methylation of *HLA-A*.

Figure B.1. Locations and Sanger sequencing of sgRNAs



B



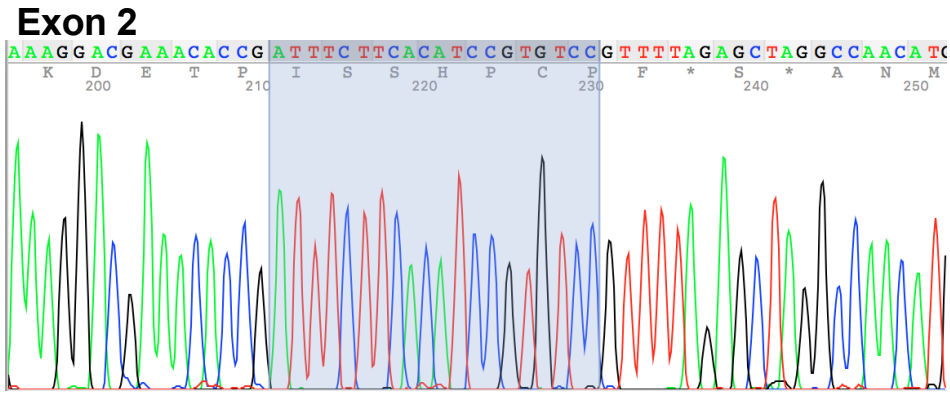


Figure B.1. Locations and Sanger sequencing of sgRNAs. A) sgRNA locations are indicated by blue boxes. B) Chromatograms from Sanger sequencing of mini-preps of clones generated for each sgRNA.

Figure B.2. HLA-A gene expression and methylation in cell lines transfected with dCas9-TET1-CD

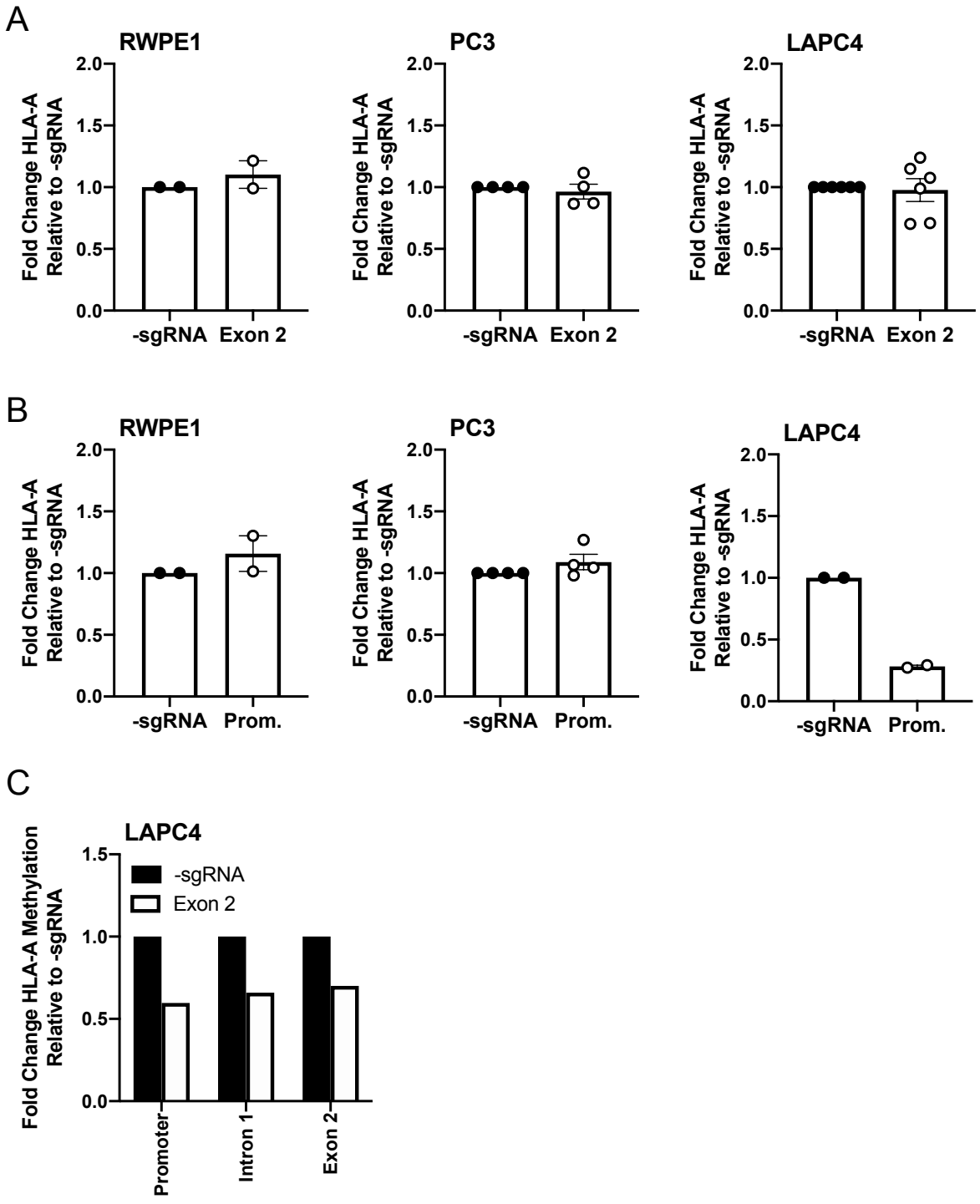


Figure B.2. *HLA-A* gene expression and methylation in cell lines transfected with dCas9-TET1-CD. Fold change of *HLA-A* gene expression in cells transfected with dCas9-TET1-CD containing sgRNA targeting exon 2 (A), promoter (B), or scrambled sequence (-sgRNA). B) *HLA-A* methylation at three gene regions in cells transfected with exon 2 or scrambled sequence (-sgRNA).

Appendix C:

Additional Experiments for the Development of SEEMLIS

This appendix contains extra experiments that were done during the development and validation of SEEMLIS. During the early development stages of this assay, we were using standard DNA oligo primers with SYBR green chemistry for qPCR. We switched to FAM labeled TaqMan probes so that we could perform pre-amplification. TaqMan also offers increased specificity due to the addition of an internal probe sequence for detection. Using the standard DNA oligos and SYBR chemistry, we had similar results to the final version of SEEMLIS in terms of sensitivity for *GSTP1* analysis where methylated *GSTP1* was detected in an input dependent manner from serially diluted LNCaP DNA, but was detected at much lower levels, if at all in WBC DNA (Figure C.1A). Using this version of the assay we measured *GSTP1* methylation in 7 CTC samples and cell free DNA (cfDNA) collected from matched plasma samples (Figure C.1B). Patient 314 did not have matched plasma available. *LINE1* was detected in all samples. *GSTP1* was detected in 5 CTC samples. In four of these samples, *GSTP1* was also detected in the plasma samples. *GSTP1* was detected in the plasma of one patient, but not in the CTC DNA. This experiment suggests that cfDNA could also be used for biomarker analysis using SEEMLIS.

Figure C.2 contains additional data from validation of the multiplexed SEEMLIS assay. In this validation experiment, I serially diluted LNCaP, LAPC4, and WBC DNA and performed SEEMLIS from the MBD2-MBD enrichment step. We obtained similar results to what was seen in Chapter 3 with WBC and LNCaP whole cell dilutions. We also performed the multiplexed version of SEEMLIS on CTC samples that were processed using VERSA and contained high numbers of background WBCs (Figure C.3). We did not detect *GSTP1* or *RASSF1* above the Ct value thresholds determined in Chapter 4.

We able to detect *APC* and *RARB*, but could not definitively say that these signals were from CTCs as opposed to WBCs due to the high level of WBCs in the sample. The dotted lines on the graph represent the thresholds determined in Chapter 3 for these genes based on background level of 1000 WBCs. This data led us to purify CTC samples further by single aspiration for analysis, as seen in Chapter 3.

Figure C.1. Validation of methylated *GSTP1* detection in LNCaP and WBC DNA and detection of methylated *GSTP1* in CTC and plasma from patients with prostate cancer.

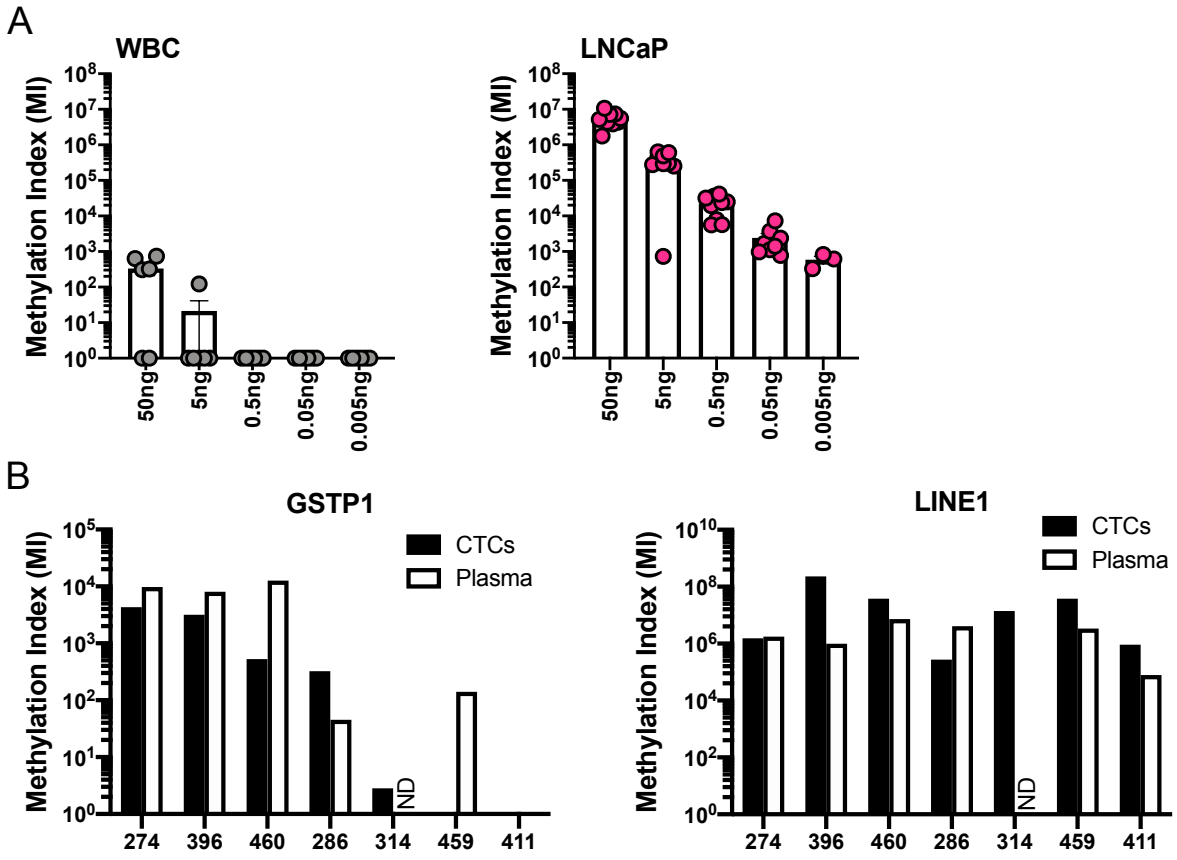


Figure C.1. Validation of methylated *GSTP1* detection in LNCaP and WBC DNA and detection of methylated *GSTP1* in CTC and plasma from patients with prostate cancer. A) Methylation index for GSPT1 in serially diluted WBC and LNCaP DNA enriched by MBD2-MBD. B) Methylation index for GSPT1 and *LINE1* in DNA from CTCs and cfDNA from plasma from patients with prostate cancer.

Figure C.2. Validation of SEEMLIS in serially diluted WBC, LNCaP, and LAPC4 DNA.

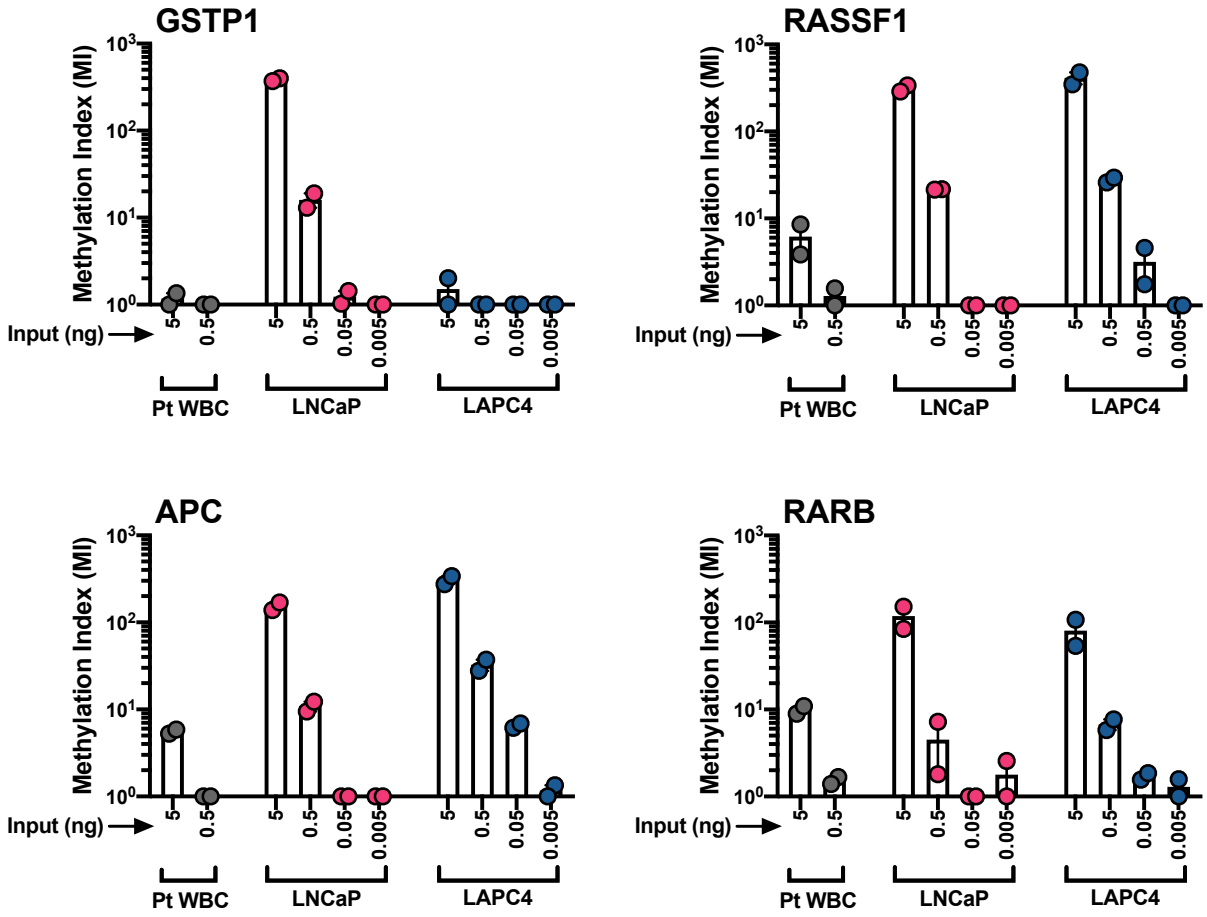


Figure C.2. Validation of SEEMLIS in serially diluted WBC, LNCaP, and LAPC4 DNA. Methylation index for serially diluted DNA samples in the indicated cell types are shown for *GSTP1*, *RASSF1*, *APC*, and *RARB*. *LINE1* data for these samples can be found in Chapter 4.

Figure C.3. Background WBC population interferes with the ability to detect methylation of multiple genes in prostate cancer CTCs.

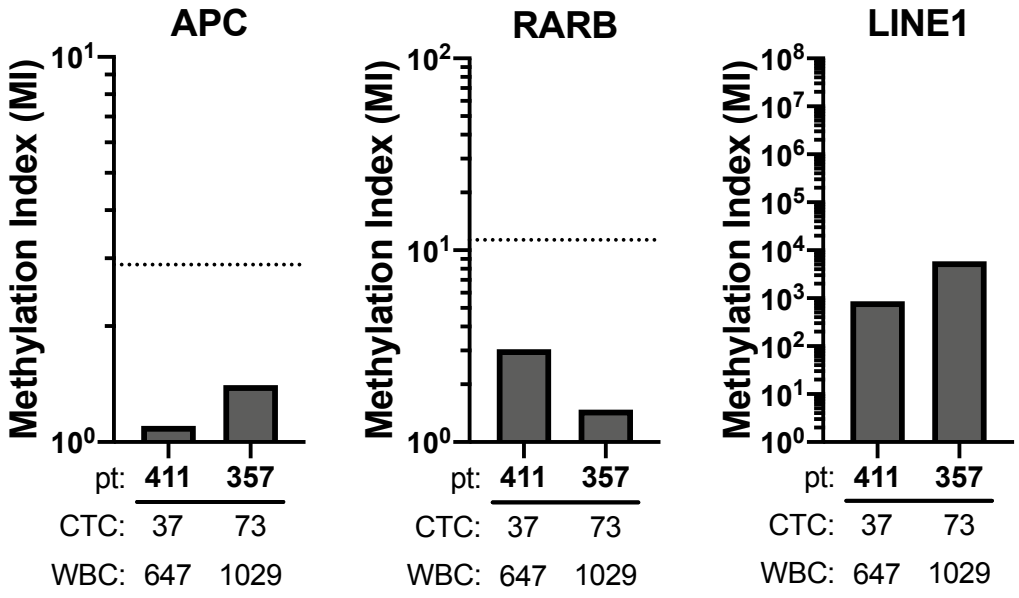


Figure C.3. Background WBC population interferes with the ability to detect methylation of multiple genes in prostate cancer CTCs. Methylation index is indicated for *APC*, *RARB*, and *LINE1* in two patient samples. Dotted lines indicate the cut off based on 1000 WBC background. Estimated numbers of CTCs and WBCs are shown below each graph. *GSTP1* and *RASSF1* were not detected in these samples.

REFERENCES

1. Heinlein CA, Chang C. Androgen receptor in prostate cancer. Endocr Rev **2004**;25:276-308

2. Litwin MS, Tan HJ. The Diagnosis and Treatment of Prostate Cancer: A Review. Jama **2017**;317:2532-42

3. Howlader N, Noone AM, Krapcho M, Miller D, Brest A, Yu M*, et al.* SEER Cancer Statistics Review. based on November 2019 SEER data submission, posted to the SEER web site, April 2020 ed. Bethesda, MD: National Cancer Institute; 1975-2017.

4. Karantanos T, Corn PG, Thompson TC. Prostate cancer progression after androgen deprivation therapy: mechanisms of castrate resistance and novel therapeutic approaches. Oncogene **2013**;32:5501-11

5. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell **2011**;144:646-74

6. Vinay DS, Ryan EP, Pawelec G, Talib WH, Stagg J, Elkord E*, et al.* Immune evasion in cancer: Mechanistic basis and therapeutic strategies. Semin Cancer Biol **2015**;35 Suppl:S185-98

7. Tesi RJ. MDSC; the Most Important Cell You Have Never Heard Of. Trends Pharmacol Sci **2019**;40:4-7

8. Liu T, Han C, Wang S, Fang P, Ma Z, Xu L*, et al.* Cancer-associated fibroblasts: an emerging target of anti-cancer immunotherapy. J Hematol Oncol **2019**;12:86

9. Shitara K, Nishikawa H. Regulatory T cells: a potential target in cancer immunotherapy. Ann N Y Acad Sci **2018**;1417:104-15

10. Petty AJ, Yang Y. Tumor-associated macrophages: implications in cancer immunotherapy. Immunotherapy **2017**;9:289-302

11. Duan S, Guo W, Xu Z, He Y, Liang C, Mo Y*, et al.* Natural killer group 2D receptor and its ligands in cancer immune escape. Mol Cancer **2019**;18:29

12. Schmiedel D, Mandelboim O. NKG2D Ligands–Critical Targets for Cancer Immune Escape and Therapy. Front Immunol **2018**;9

13. Mou X, Zhou Y, Jiang P, Zhou T, Jiang Q, Xu C*, et al.* The Regulatory Effect of UL-16 Binding Protein-3 Expression on the Cytotoxicity of NK Cells in Cancer Patients. Sci Rep **2014**;4

14. Chen L, Flies DB. Molecular mechanisms of T cell co-stimulation and co-inhibition. Nat Rev Immunol **2013**;13:227-42

15. Héninger E, Krueger TE, Lang JM. Augmenting antitumor immune responses with epigenetic modifying agents. Front Immunol **2015**;6:29

16. Blank C, Gajewski TF, Mackensen A. Interaction of PD-L1 on Tumor Cells With PD-1 on Tumor-Specific T Cells as a Mechanism of Immune Evasion: Implications for Tumor Immunotherapy. Cancer immunology, immunotherapy : CII **2005**;54

17. Driessens G, Kline J, Gajewski TF. Costimulatory and coinhibitory receptors in anti-tumor immunity. Immunol Rev **2009**;229:126-44

18. Spranger S. Mechanisms of tumor escape in the context of the T-cell-inflamed and the non-T-cell-inflamed tumor microenvironment. Int Immunol **2016**;28:383-91

19. Shiina T, Hosomichi K, Inoko H, Kulski JK. The HLA genomic loci map: expression, interaction, diversity and disease. J Hum Genet **2009**;54:15-39

20. Rock KL, Reits E, Neefjes J. Present Yourself! By MHC Class I and MHC Class II Molecules. Trends Immunol **2016**;37:724-37

21. Cruz-Tapias P, Castiblanco J, Anaya J-M. Major histocompatibility complex: Antigen processing and presentation. **2013**

22. Karr RW. The HLA Complex. In: Jameson JL, editor. Principles of Molecular Medicine. Totowa, NJ: Humana Press; 1998. p 273-81.

23. Knapp LA. The ABCs of MHC. Evolutionary Anthropology **2005**;14:28-37

24. Hewitt EW. The MHC class I antigen presentation pathway: strategies for viral immune evasion. Immunology **2003**;110:163-9

25. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. T Cells and MHC Proteins. Molecular Biology of the Cell. New York: Garland Science; 2002.

26. Ellgaard L, Frickel EM. Calnexin, calreticulin, and ERp57: teammates in glycoprotein folding. Cell Biochem Biophys **2003**;39:223-47

27. Diedrich G, Bangia N, Pan M, Cresswell P. A role for calnexin in the assembly of the MHC class I loading complex in the endoplasmic reticulum. J Immunol **2001**;166:1703-9

28. Hughes EA, Cresswell P. The thiol oxidoreductase ERp57 is a component of the MHC class I peptide-loading complex. Curr Biol **1998**;8:709-12

29. Ritz U, Seliger B. The transporter associated with antigen processing (TAP): structural integrity, expression, function, and its clinical relevance. Mol Med **2001**;7:149-58

30. Li S, Paulsson KM, Chen S, Sjogren HO, Wang P. Tapasin is required for efficient peptide binding to transporter associated with antigen processing. J Biol Chem **2000**;275:1581-6

31. Houghton AN, Guevara-Patiño JA. Immune recognition of self in immunity against cancer. J Clin Invest **2004**;114:468-71

32. Petersen TR, Dickgreber N, Hermans IF. Tumor Antigen Presentation by Dendritic Cells. Critical reviews in immunology **2010**;30

33. Xia A, Zhang Y, Xu J, Yin T, Lu XJ. T Cell Dysfunction in Cancer Immunity and Immunotherapy. Front Immunol **2019**;10

34. Seliger B. Molecular mechanisms of HLA class I-mediated immune evasion of human tumors and their role in resistance to immunotherapies. HLA **2016**;88:213-20

35. Thakur A, Vaishampayan U, Lum LG. Immunotherapy and immune evasion in prostate cancer. Cancers (Basel) **2013**;5:569-90

36. Carretero FJ, Del Campo AB, Flores-Martin JF, Mendez R, Garcia-Lopez C, Cozar JM*, et al.* Frequent HLA class I alterations in human prostate cancer: molecular mechanisms and clinical relevance. Cancer Immunol Immunother **2016**;65:47-59

37. Hicklin DJ, Marincola FM, Ferrone S. HLA class I antigen downregulation in human cancers: T-cell immunotherapy revives an old story. Mol Med Today **1999**;5:178-86

38. Blades RA, Keating PJ, McWilliam LJ, George NJ, Stern PL. Loss of HLA class I expression in prostate cancer: implications for immunotherapy. Urology **1995**;46:681-7

39. Kitamura H, Torigoe T, Asanuma H, Honma I, Sato N, Tsukamoto T. Down-regulation of HLA class I antigens in prostate cancer tissues and up-regulation by histone deacetylase inhibition. J Urol **2007**;178:692-6

40. Garrido F. HLA Class-I Expression and Cancer Immunotherapy. Adv Exp Med Biol **2019**;1151:79-90

41. June CH. Adoptive T cell therapy for cancer in the clinic. J Clin Invest **2007**;117:1466-76

42. Lopes A, Vandermeulen G, Préat V. Cancer DNA vaccines: current preclinical and clinical developments and future perspectives. J Exp Clin Cancer Res **2019**;38

43. Garrido F, Aptsiauri N, Doorduijn EM, Garcia Lora AM, van Hall T. The urgent need to recover MHC class I in cancers for effective immunotherapy. Curr Opin Immunol **2016**;39:44-51

44. Paulson KG, Voillet V, McAfee MS, Hunter DS, Wagener FD, Perdicchio M*, et al.* Acquired cancer resistance to combination immunotherapy from transcriptional loss of class I HLA. Nat Commun **2018**;9:3868

45. Inbar-Feigenberg M, Choufani S, Butcher DT, Roifman M, Weksberg R. Basic concepts of epigenetics. Fertil Steril **2013**;99:607-15

46. Moore LD, Le T, Fan G. DNA methylation and its basic function. Neuropsychopharmacology **2013**;38:23-38

47. Deaton AM, Bird A. CpG islands and the regulation of transcription. Genes Dev **2011**;25:1010-22

48. Vavouri T, Lehner B. Human genes with CpG island promoters have a distinct transcription-associated chromatin organization. Genome Biol **2012**;13:R110

49. Lyko F. The DNA methyltransferase family: a versatile toolkit for epigenetic regulation. Nat Rev Genet **2018**;19:81-92

50. Chedin F. The DNMT3 family of mammalian de novo DNA methyltransferases. Prog Mol Biol Transl Sci **2011**;101:255-85

51. Pacaud R, Sery Q, Oliver L, Vallette FM, Tost J, Cartron PF. DNMT3L interacts with transcription factors to target DNMT3L/DNMT3B to specific DNA sequences: role of the DNMT3L/DNMT3B/p65-NFkappaB complex in the (de-)methylation of TRAF1. Biochimie **2014**;104:36-49

52. Deplus R, Brenner C, Burgers WA, Putmans P, Kouzarides T, de Launoit Y*, et al.* Dnmt3L is a transcriptional repressor that recruits histone deacetylase. Nucleic Acids Res **2002**;30:3831-8

53. Handy DE, Castro R, Loscalzo J. Epigenetic Modifications: Basic Mechanisms and Role in Cardiovascular Disease. Circulation **2011**;123:2145-56

54. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. Cell Res **2011**;21:381-95

55. Hyun K, Jeon J, Park K, Kim J. Writing, erasing and reading histone lysine methylations. Exp Mol Med **2017**;49:e324

56. Verdone L, Caserta M, Di Mauro E. Role of histone acetylation in the control of gene expression. Biochem Cell Biol **2005**;83:344-53

57. Sterner DE, Berger SL. Acetylation of histones and transcription-related factors. Microbiol Mol Biol Rev **2000**;64:435-59

58. Seto E, Yoshida M. Erasers of histone acetylation: the histone deacetylase enzymes. Cold Spring Harb Perspect Biol **2014**;6:a018713

59. Dimitrova E, Turberfield AH, Klose RJ. Histone demethylases in chromatin biology and beyond. EMBO Rep **2015**;16:1620-39

60. Marmorstein R, Trievel RC. Histone Modifying Enzymes: Structures, Mechanisms, and Specificities. Biochim Biophys Acta **2009**;1789:58-68

61. Rennie PS, Nelson CC. Epigenetic mechanisms for progression of prostate cancer. Cancer Metastasis Rev **1998**;17:401-9

62. Lawrence MS, Stojanov P, Mermel CH, Garraway LA, Golub TR, Meyerson M*, et al.* Discovery and saturation analysis of cancer genes across 21 tumor types. Nature **2014**;505:495-501

63. Barbieri CE, Bangma CH, Bjartell A, Catto JW, Culig Z, Gronberg H*, et al.* The Mutational Landscape of Prostate Cancer. Eur Urol **2013**;64:567-76

64. Jernberg E, Bergh A, Wikstrom P. Clinical relevance of androgen receptor alterations in prostate cancer. Endocr Connect **2017**;6:R146-R61

65. Yegnasubramanian S, Kowalski J, Gonzalgo ML, Zahurak M, Piantadosi S, Walsh PC*, et al.* Hypermethylation of CpG islands in primary and metastatic human prostate cancer. Cancer Res **2004**;64:1975-86

66. Majumdar S, Buckles E, Estrada J, Koochekpour S. Aberrant DNA Methylation and Prostate Cancer. Curr Genomics **2011**;12:486-505

67. Yegnasubramanian S, Haffner MC, Zhang Y, Gurel B, Cornish TC, Wu Z*, et al.* DNA hypomethylation arises later in prostate cancer progression than CpG island hypermethylation and contributes to metastatic tumor heterogeneity. Cancer Res **2008**;68:8954-67

68. Zelic R, Fiano V, Grasso C, Zugna D, Pettersson A, Gillio-Tos A*, et al.* Global DNA hypomethylation in prostate cancer development and progression: a systematic review. Prostate Cancer Prostatic Dis **2015**;18:1-12

69. Bedford MT, van Helden PD. Hypomethylation of DNA in pathological conditions of the human prostate. Cancer Res **1987**;47:5274-6

70. Eden A, Gaudet F, Waghmare A, Jaenisch R. Chromosomal instability and tumors promoted by DNA hypomethylation. Science **2003**;300:455

71. Feinberg AP, Vogelstein B. Hypomethylation of ras oncogenes in primary human cancers. Biochem Biophys Res Commun **1983**;111:47-54

72. Kazanets A, Shorstova T, Hilmi K, Marques M, Witcher M. Epigenetic silencing of tumor suppressor genes: Paradigms, puzzles, and potential. Biochim Biophys Acta **2016**;1865:275-88

73. Jarrard DF, Kinoshita H, Shi Y, Sandefur C, Hoff D, Meisner LF*, et al.* Methylation of the androgen receptor promoter CpG island is associated with loss of androgen receptor expression in prostate cancer cells. Cancer Res **1998**;58:5310-4

74. Kirby MK, Ramaker RC, Roberts BS, Lasseigne BN, Gunther DS, Burwell TC*, et al.* Genome-wide DNA methylation measurements in prostate tissues uncovers novel prostate cancer diagnostic biomarkers and transcription factor binding patterns. BMC Cancer **2017**;17

75. Mahon KL, Qu W, Devaney J, Paul C, Castillo L, Wykes RJ*, et al.* Methylated Glutathione S-transferase 1 (mGSTP1) is a potential plasma free DNA epigenetic marker of prognosis and response to chemotherapy in castrate-resistant prostate cancer. British journal of cancer **2014**;111:1802-9

76. Matuschek C, Bolke E, Lammering G, Gerber PA, Peiper M, Budach W*, et al.* Methylated APC and GSTP1 genes in serum DNA correlate with the presence of circulating blood tumor cells and are associated with a more aggressive and advanced breast cancer disease. Eur J Med Res **2010**;15:277-86

77. Nakayama M, Gonzalgo ML, Yegnasubramanian S, Lin X, De Marzo AM, Nelson WG. GSTP1 CpG island hypermethylation as a molecular biomarker for prostate cancer. J Cell Biochem **2004**;91:540-52

78. Lee WH, Morton RA, Epstein JI, Brooks JD, Campbell PA, Bova GS*, et al.* Cytidine methylation of regulatory sequences near the pi-class glutathione S-transferase gene accompanies human prostatic carcinogenesis. Proc Natl Acad Sci U S A **1994**;91:11733-7

79. Zhang W, Jiao H, Zhang X, Zhao R, Wang F, He W*, et al.* Correlation between the expression of DNMT1, and GSTP1 and APC, and the methylation status of GSTP1 and APC in association with their clinical significance in prostate cancer. Mol Med Rep **2015**;12:141-6

80. Van der Auwera I, Elst HJ, Van Laere SJ, Maes H, Huget P, van Dam P*, et al.* The presence of circulating total DNA and methylated genes is associated with circulating tumour cells in blood from breast cancer patients. Br J Cancer **2009**;100:1277-86

81. Gravina GL, Ranieri G, Muzi P, Marampon F, Mancini A, Di Pasquale B*, et al.* Increased levels of DNA methyltransferases are associated with the tumorigenic capacity of prostate cancer cells. Oncol Rep **2013**;29:1189-95

82. Patra SK, Patra A, Zhao H, Dahiya R. DNA methyltransferase and demethylase in human prostate cancer. Mol Carcinog **2002**;33:163-71

83. Kamdar S, Isserlin R, Kwast TVd, Zlotta AR, Bader GD, Fleshner NE*, et al.* Exploring targets of TET2-mediated methylation reprogramming as potential discriminators of prostate cancer progression. Clinical Epigenetics **2019**;11:1-19

84. Smeets E, Lynch AG, Prekovic S, Van den Broeck T, Moris L, Helsen C*, et al.* The role of TET-mediated DNA hydroxymethylation in prostate cancer. Mol Cell Endocrinol **2018**;462:41-55

85. Spans L, Van den Broeck T, Smeets E, Prekovic S, Thienpont B, Lambrechts D*, et al.* Genomic and epigenomic analysis of high-risk prostate cancer reveals changes in hydroxymethylation and TET1. Oncotarget **2016**;7:24326-38

86. Rycaj K, Tang DG. Molecular determinants of prostate cancer metastasis. Oncotarget **2017**;8:88211-31

87. Dahl JA, Gilfillan GD. How low can you go? Pushing the limits of low-input ChIP-seq. Brief Funct Genomics **2018**;17:89-95

88. Gilfillan GD, Hughes T, Sheng Y, Hjorthaug HS, Straub T, Gervin K*, et al.* Limitations and possibilities of low cell number ChIP-seq. BMC Genomics **2012**;13:645

89. Chen Z, Wang L, Wang Q, Li W. Histone modifications and chromatin organization in prostate cancer. Epigenomics **2010**;2:551-60

90. Braadland PR, Urbanucc iA. Chromatin Reprogramming as an Adaptation Mechanism in Advanced Prostate Cancer. Endocrine-related cancer **2019**;26

91. Heemers HV, Debes JD, Tindall DJ. The Role of the Transcriptional Coactivator p300 in Prostate Cancer Progression. Advances in experimental medicine and biology **2008**;617

92. Jin L, Garcia J, Chan E, de la Cruz C, Segal E, Merchant M*, et al.* Therapeutic Targeting of the CBP/p300 Bromodomain Blocks the Growth of Castration-Resistant Prostate Cancer. Cancer research **2017**;77

93. Lasko LM, Jakob CG, Edalji RP, Qiu W, Montgomery D, Digiammarino EL*, et al.* Discovery of a potent catalytic p300/CBP inhibitor that targets lineage-specific tumors. Nature **2017**;550:128-32

94. Kondo Y, Shen L, Cheng AS, Ahmed S, Boumber Y, Charo C*, et al.* Gene Silencing in Cancer by Histone H3 Lysine 27 Trimethylation Independent of Promoter DNA Methylation. Nature genetics **2008**;40

95. Weichert W, Röske A, Gekeler V, Beckers T, Stephan C, Jung K*, et al.* Histone deacetylases 1, 2 and 3 are highly expressed in prostate cancer and HDAC2 expression is associated with shorter PSA relapse time after radical prostatectomy. Br J Cancer **2008**;98:604-10

96. Melling N, Thomsen E, Tsourlakis MC, Kluth M, Hube-Magg C, Minner S*, et al.* Overexpression of enhancer of zeste homolog 2 (EZH2) characterizes an aggressive subset of prostate cancers and predicts patient prognosis independently from pre- and postoperatively assessed clinicopathological parameters. Carcinogenesis **2015**;36:1333-40

97. Yang YA, Yu J. EZH2, an epigenetic driver of prostate cancer. Protein Cell **2013**;4:331-41

98. Hoffmann MJ, Engers R, Florl AR, Otte AP, Muller M, Schulz WA. Expression changes in EZH2, but not in BMI-1, SIRT1, DNMT1 or DNMT3B are associated with DNA methylation changes in prostate cancer. Cancer Biol Ther **2007**;6:1403-12

99. Ezponda T, Popovic R, Shah MY, Martinez-Garcia E, Zheng Y, Min DJ*, et al.* The Histone Methyltransferase MMSET/WHSC1 Activates TWIST1 to Promote an Epithelial-Mesenchymal Transition and Invasive Properties of Prostate Cancer. Oncogene **2013**;32:2882-90

100. Kim KH, Roberts CWM. Targeting EZH2 in cancer. Nat Med **2016**;22:128-34

101. Heninger E, Krueger TE, Thiede SM, Sperger JM, Byers BL, Kircher MR*, et al.* Inducible expression of cancer-testis antigens in human prostate cancer. Oncotarget **2016**;7:84359-74

102. Kortenhorst MS, Wissing MD, Rodriguez R, Kachhap SK, Jans JJ, Van der Groep P*, et al.* Analysis of the genomic response of human prostate cancer cells to histone deacetylase inhibitors. Epigenetics **2013**;8:907-20

103. Mao W, Ghasemzadeh A, Freeman ZT, Obradovic A, Chaimowitz MG, Nirschl TR*, et al.* Immunogenicity of prostate cancer is augmented by BET bromodomain inhibition. J Immunother Cancer **2019**;7:277

104. Wang G, Zhao D, Spring DJ, DePinho RA. Genetics and biology of prostate cancer. Genes & development **2018**;32:1105-40

105. Fleischmann A, Saramaki OR, Zlobec I, Rotzer D, Genitsch V, Seiler R*, et al.* Prevalence and prognostic significance of TMPRSS2-ERG gene fusion in lymph node positive prostate cancers. Prostate **2014**;74:1647-54

106. Pettersson A, Graff R, Bauer S, Pitt M, Lis R, Stack E*, et al.* The TMPRSS2:ERG Rearrangement, ERG Expression, and Prostate Cancer Outcomes: A Cohort Study and Meta-Analysis. Cancer Epidemiol Biomarkers Prev **2012**;21

107. Armstrong AJ, Halabi S, Luo J, Nanus DM, Giannakakou P, Szmulewitz RZ*, et al.* Prospective Multicenter Validation of Androgen Receptor Splice Variant 7 and Hormone Therapy Resistance in High-Risk Castration-Resistant Prostate Cancer: The PROPHECY Study. J Clin Oncol **2019**;37:1120-9

108. Sharp A, Coleman I, Yuan W, Sprenger C, Dolling D, Rodrigues DN*, et al.* Androgen receptor splice variant-7 expression emerges with castration resistance in prostate cancer. J Clin Invest **2019**;129:192-208

109. Scher HI, Lu D, Schreiber NA, Louw J, Graf RP, Vargas HA*, et al.* Association of AR-V7 on Circulating Tumor Cells as a Treatment-Specific Biomarker With Outcomes and Survival in Castration-Resistant Prostate Cancer. JAMA Oncol **2016**;2:1441-9

110. Lianidou E, Pantel K. Liquid biopsies. Genes Chromosomes Cancer **2019**;58:219-32

111. Zainfeld D, Goldkorn A. Liquid Biopsy in Prostate Cancer: Circulating Tumor Cells and Beyond. Cancer Treat Res **2018**;175:87-104

112. Millner LM, Linder MW, Valdes R. Circulating Tumor Cells: A Review of Present Methods and the Need to Identify Heterogeneous Phenotypes. Ann Clin Lab Sci **2013**;43:295-304

113. Lang JM, Casavant BP, Beebe DJ. Circulating tumor cells: getting more from less. Sci Transl Med **2012**;4:141ps13

114. Darst RP, Pardo CE, Ai L, Brown KD, Kladde MP. Bisulfite Sequencing of DNA. Curr Protoc Mol Biol **2010**;CHAPTER:Unit-7 917

115. Kurdyukov S, Bullock M. DNA Methylation Analysis: Choosing the Right Method. Biology (Basel) **2016**;5

116. Shen L, Waterland RA. Methods of DNA methylation analysis. Curr Opin Clin Nutr Metab Care **2007**;10:576-81

117. Karemaker ID, Vermeulen M. Single-Cell DNA Methylation Profiling: Technologies and Biological Applications. Trends Biotechnol **2018**;36:952-65

118. Kelly SP, Anderson WF, Rosenberg PS, Cook MB. Past, Current, and Future Incidence Rates and Burden of Metastatic Prostate Cancer in the United States. Eur Urol Focus **2018**;4:121-7

119. Boettcher AN, Usman A, Morgans A, VanderWeele DJ, Sosman J, Wu JD. Past, Current, and Future of Immunotherapies for Prostate Cancer. Front Oncol **2019**;9:884

120. Schweizer MT, Drake CG. Immunotherapy for prostate cancer: recent developments and future challenges. Cancer Metastasis Rev **2014**;33:641-55

121. Lee JH, Shklovskaya E, Lim SY, Carlino MS, Menzies AM, Stewart A*, et al.* Transcriptional downregulation of MHC class I and melanoma de- differentiation in resistance to PD-1 inhibition. Nature communications **2020**;11:1897-

122. Fares CM, Van Allen EM, Drake CG, Allison JP, Hu-Lieskovan S. Mechanisms of Resistance to Immune Checkpoint Blockade: Why Does Checkpoint Inhibitor Immunotherapy Not Work for All Patients? American Society of Clinical Oncology Educational Book **2019**:147-64

123. Blum JS, Wearsch PA, Cresswell P. Pathways of antigen processing. Annu Rev Immunol **2013**;31:443-73

124. Ruggero K, Farran-Matas S, Martinez-Tebar A, Aytes A. Epigenetic Regulation in Prostate Cancer Progression. Curr Mol Biol Rep **2018**;4:101-15

125. Zhao SG, Chen WS, Li H, Foye A, Zhang M, Sjöström M*, et al.* The DNA methylation landscape of advanced prostate cancer. **2020**;52:778-89

126. Kim NH, Kim SN, Kim YK. Involvement of HDAC1 in E-cadherin expression in prostate cancer cells; its implication for cell motility and invasion. Biochem Biophys Res Commun **2011**;404:915-21

127. Ye Q, Shen Y, Wang X, Yang J, Miao F, Shen C*, et al.* Hypermethylation of HLA class I gene is associated with HLA class I down-regulation in human gastric cancer. Tissue Antigens **2010**;75:30-9

128. Nie Y, Yang G, Song Y, Zhao X, So C, Liao J*, et al.* DNA hypermethylation is a mechanism for loss of expression of the HLA class I genes in human esophageal squamous cell carcinomas. Carcinogenesis **2001**;22:1615-23

129. Liao Y, Xu K. Epigenetic regulation of prostate cancer: the theories and the clinical implications. Asian J Androl **2019**;21:279-90

130. Villanueva L, Álvarez-Errico D, Esteller M. The Contribution of Epigenetics to Cancer Immunotherapy. Trends Immunol **2020**;41:676-91

131. Armenia J, Wankowicz S, Liu D, Gao J, Kundra R, Reznik E*, et al.* The Long Tail of Oncogenic Drivers in Prostate Cancer. Nature genetics **2018**;50

132. Garrido M, Rodriguez T, Zinchenko S, Maleno I, Ruiz-Cabello F, Concha Á*, et al.* HLA Class I Alterations in Breast Carcinoma Are Associated With a High

Frequency of the Loss of Heterozygosity at Chromosomes 6 and 15. Immunogenetics **2018**;70

133. McGranahan N, Rosenthal R, Hiley C, Rowan A, Watkins T, Wilson G*, et al.* Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. Cell **2017**;171

134. Taylor BS, Schultz N, Hieronymus H, Gopalan A, Xiao Y, Carver BS*, et al.* Integrative genomic profiling of human prostate cancer. Cancer Cell **2010**;18:11-22

135. Seligson DB, Horvath S, Shi T, Yu H, Tze S, Grunstein M*, et al.* Global histone modification patterns predict risk of prostate cancer recurrence. **2005**;435:1262-6

136. Zhu A, Hopkins KM, Friedman RA, Bernstock JD, Broustas CG, Lieberman HB. DNMT1 and DNMT3B regulate tumorigenicity of human prostate cancer cells by controlling RAD9 expression through targeted methylation. Carcinogenesis **2020**

137. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature **2012**;489:57-74

138. Aryee MJ, Liu W, Engelmann JC, Nuhn P, Gurel M, Haffner MC*, et al.* DNA methylation alterations exhibit intraindividual stability and interindividual heterogeneity in prostate cancer metastases. Sci Transl Med **2013**;5:169ra10

139. Sperger JM, Strotman LN, Welsh A, Casavant BP, Chalmers Z, Horn S*, et al.* Integrated analysis of multiple biomarkers from circulating tumor cells enabled by exclusion-based analyte isolation. Clin Cancer Res **2017**;23:746-56

140. Cheng Y, He C, Wang M, Ma X, Mo F, Yang S*, et al.* Targeting epigenetic regulators for cancer therapy: mechanisms and advances in clinical trials. Signal Transduct Target Ther **2019**;4:62

141. Rothbart SB, Baylin SB. Epigenetic Therapy for Epithelioid Sarcoma. Cell **2020**;181:211

142. Italiano A. Targeting epigenetics in sarcomas through EZH2 inhibition. J Hematol Oncol **2020**;13:33

143. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO*, et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal **2013**;6:pl1

144. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA*, et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov **2012**;2:401-4

145. Goldman M, Craft B, Hastie M, Repecka K, McDade F, Kamath A*, et al.* The UCSC Xena platform for public and private cancer genomics data visualization and interpretation. BioRxiv **2019**

146. Corces MR, Granja JM, Shams S, Louie BH, Seoane JA, Zhou W*, et al.* The Chromatin Accessibility Landscape of Primary Human Cancers. Science (New York, NY) **2018**;362

147. Diez-Villanueva A, Mallona I, Peinado MA. Wanderer, an interactive viewer to explore DNA methylation and gene expression data in human cancer. Epigenetics Chromatin **2015**;8:22

148. Pascolo S, Bervas N, Ure JM, Smith AG, Lemonnier FA, Pérarnau B. HLA-A2.1-restricted education and cytolytic activity of CD8(+) T lymphocytes from beta2 microglobulin (beta2m) HLA-A2.1 monochain transgenic H-2Db beta2m double knockout mice. J Exp Med **1997**;185:2043-51

149. Murrell A, Hurd PJ, Wood IC. Epigenetic mechanisms in development and disease. Biochem Soc Trans **2013**;41:697-9

150. Greenberg MVC, Bourc'his D. The diverse roles of DNA methylation in mammalian development and disease. Nat Rev Mol Cell Biol **2019**;20:590-607

151. Razin A, Riggs AD. DNA methylation and gene function. Science (New York, NY) **1980**;210:604-10

152. Ramchandani S, Bhattacharya SK, Cervoni N, Szyf M. DNA methylation is a reversible biological signal. Proceedings of the National Academy of Sciences of the United States of America **1999**;96:6107-12

153. Almouzni G, Cedar H. Maintenance of Epigenetic Information. Cold Spring Harbor perspectives in biology **2016**;8

154. Strand SH, Orntoft TF, Sorensen KD. Prognostic DNA methylation markers for prostate cancer. Int J Mol Sci **2014**;15:16544-76

155. Yang M, Park JY. DNA methylation in promoter region as biomarkers in prostate cancer. Methods in molecular biology (Clifton, NJ) **2012**;863:67-109

156. Kobayashi Y, Absher DM, Gulzar ZG, Young SR, McKenney JK, Peehl DM*, et al.* DNA methylation profiling reveals novel biomarkers and important roles for DNA methyltransferases in prostate cancer. Genome Res **2011**;21:1017-27

157. Ahmed H. Promoter Methylation in Prostate Cancer and its Application for the Early Detection of Prostate Cancer Using Serum and Urine Samples. Biomark Cancer **2010**;2010:17-33

158. Zhao F, Olkhov-Mitsel E, van der Kwast T, Sykes J, Zdravic D, Venkateswaran V*, et al.* Urinary DNA Methylation Biomarkers for Noninvasive Prediction of Aggressive Disease in Patients with Prostate Cancer on Active Surveillance. J Urol **2017**;197:335-41

159. Yegnasubramanian S, Wu Z, Haffner MC, Esopi D, Aryee MJ, Badrinath R*, et al.* Chromosome-wide mapping of DNA methylation patterns in normal and malignant prostate cells reveals pervasive methylation of gene-associated and conserved intergenic sequences. BMC Genomics **2011**;12:313

160. Chimonidou M, Tzitzira A, Strati A, Sotiropoulou G, Sfikas C, Malamos N*, et al.* CST6 promoter methylation in circulating cell-free DNA of breast cancer patients. Clin Biochem **2013**;46:235-40

161. Chimonidou M, Strati A, Tzitzira A, Sotiropoulou G, Malamos N, Georgoulias V*, et al.* DNA methylation of tumor suppressor and metastasis suppressor genes in circulating tumor cells. Clinical chemistry **2011**;57:1169-77

162. Kloten V, Becker B, Winner K, Schrauder MG, Fasching PA, Anzeneder T*, et al.* Promoter hypermethylation of the tumor-suppressor genes ITIH5, DKK3, and RASSF1A as novel biomarkers for blood-based breast cancer screening. Breast Cancer Res **2013**;15:R4

163. Ellinger J, Haan K, Heukamp LC, Kahl P, Buttner R, Muller SC*, et al.* CpG island hypermethylation in cell-free serum DNA identifies patients with localized prostate cancer. Prostate **2008**;68:42-9

164. Han D, Lu X, Shih AH, Nie J, You Q, Xu MM*, et al.* A Highly Sensitive and Robust Method for Genome-wide 5hmC Profiling of Rare Cell Populations. Mol Cell **2016**;63:711-9

165. Taiwo O, Wilson GA, Morris T, Seisenberger S, Reik W, Pearce D*, et al.* Methylome analysis using MeDIP-seq with low DNA concentrations. Nat Protoc **2012**;7:617-36

166. Brinkman AB, Simmer F, Ma K, Kaan A, Zhu J, Stunnenberg HG. Whole-genome DNA methylation profiling using MethylCap-seq. Methods (San Diego, Calif) **2010**;52:232-6

167. Sonnet M, Baer C, Rehli M, Weichenhan D, Plass C. Enrichment of methylated DNA by methyl-CpG immunoprecipitation. Methods in molecular biology (Clifton, NJ) **2013**;971:201-12

168. Yegnasubramanian S, Lin X, Haffner MC, DeMarzo AM, Nelson WG. Combination of methylated-DNA precipitation and methylation-sensitive restriction enzymes (COMPARE-MS) for the rapid, sensitive and quantitative detection of DNA methylation. Nucleic Acids Res **2006**;34:e19

169. Casavant BP, Guckenberger DJ, Berry SM, Tokar JT, Lang JM, Beebe DJ. The VerIFAST: an integrated method for cell isolation and extracellular/intracellular staining. Lab Chip **2013**;13:391-6

170. Strotman L, O'Connell R, Casavant BP, Berry SM, Sperger JM, Lang JM*, et al.* Selective nucleic acid removal via exclusion (SNARE): capturing mRNA and DNA from a single sample. Anal Chem **2013**;85:9764-70

171. Pezzi HM, Guckenberger DJ, Schehr JL, Rothbauer J, Stahlfeld C, Singh A*, et al.* Versatile exclusion-based sample preparation platform for integrated rare cell isolation and analyte extraction. Lab Chip **2018**;18:3446-58

172. Maldonado L, Brait M, Loyo M, Sullenberger L, Wang K, Peskoe SB*, et al.* GSTP1 promoter methylation is associated with recurrence in early stage prostate cancer. J Urol **2014**;192:1542-8

173. Yang AS, Estecio MR, Doshi K, Kondo Y, Tajara EH, Issa JP. A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA elements. Nucleic Acids Res **2004**;32:e38

174. Cordaux R, Batzer MA. The impact of retrotransposons on human genome evolution. Nat Rev Genet **2009**;10:691-703

175. Fu LJ, Ding YB, Wu LX, Wen CJ, Qu Q, Zhang X*, et al.* The Effects of Lycopene on the Methylation of the GSTP1 Promoter and Global Methylation in Prostatic Cancer Cell Lines PC3 and LNCaP. Int J Endocrinol **2014**;2014:620165

176. Pidsley R, Zotenko E, Peters TJ, Lawrence MG, Risbridger GP, Molloy P*, et al.* Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. Genome Biol **2016**;17:208

177. Li Y, Zhu J, Tian G, Li N, Li Q, Ye M*, et al.* The DNA methylome of human peripheral blood mononuclear cells. PLoS Biol **2010**;8:e1000533

178. Song Q, Decato B, Hong EE, Zhou M, Fang F, Qu J*, et al.* A reference methylome database and analysis pipeline to facilitate integrative and comparative epigenomics. PloS one **2013**;8:e81148

179. Johnson BP, Vitek RA, Geiger PG, Huang W, Jarrard DF, Lang JM*, et al.* Vital ex vivo tissue labeling and pathology-guided micropunching to characterize cellular heterogeneity in the tissue microenvironment. Biotechniques **2018**;64:13-9

180. Tokar JJ, Stahlfeld CN, Sperger JM, Niles DJ, Beebe DJ, Lang JM*, et al.* Pairing Microwell Arrays with an Affordable, Semiautomated Single-Cell Aspirator for the Interrogation of Circulating Tumor Cell Heterogeneity. SLAS Technol **2020**;25:162-76

181. Guo H, Zhu P, Guo F, Li X, Wu X, Fan X*, et al.* Profiling DNA methylome landscapes of mammalian cells with single-cell reduced-representation bisulfite sequencing. Nat Protoc **2015**;10:645-59

182. Adey A, Shendure J. Ultra-low-input, tagmentation-based whole-genome bisulfite sequencing. Genome Res **2012**;22:1139-43

183. Aberg KA, Chan RF, Shabalin AA, Zhao M, Turecki G, Staunstrup NH*, et al.* A MBD-seq protocol for large-scale methylome-wide studies with (very) low amounts of DNA. Epigenetics **2017**;12:743-50

184. Guckenberger DJ, Pezzi HM, Regier MC, Berry SM, Fawcett K, Barrett K*, et al.* Magnetic System for Automated Manipulation of Paramagnetic Particles. Anal Chem **2016**;88:9902-7

185. Schehr JL, Schultz ZD, Warrick JW, Guckenberger DJ, Pezzi HM, Sperger JM*, et al.* High Specificity in Circulating Tumor Cell Identification Is Required for Accurate Evaluation of Programmed Death-Ligand 1. PLoS One **2016**;11:e0159397

186. World Health Organization. WHO International Programme on Chemical Safety Biomarkers in Risk Assessment: Validity and Validation. Geneva: WHO; 2001.

187. Food and Drug Administration. BEST (Biomarkers, EndpointS, and other Tools) Resource. Bethesda, MD: Co-published by National Institutes of Health (US); 2016.

188. Neuhaus J, Yang B. Liquid Biopsy Potential Biomarkers in Prostate Cancer. Diagnostics (Basel, Switzerland) **2018**;8:68

189. Perner S, Mosquera JM, Demichelis F, Hofer MD, Paris PL, Simko J*, et al.* TMPRSS2-ERG fusion prostate cancer: an early molecular event associated with invasion. Am J Surg Pathol **2007**;31:882-8

190. Rajput AB, Miller MA, Luca AD, Boyd N, Leung S, Hurtado-Coll A*, et al.* Frequency of the TMPRSS2:ERG gene fusion is increased in moderate to poorly differentiated prostate cancers. **2007**

191. Gonzalgo ML, Pavlovich CP, Lee SM, Nelson WG. Prostate cancer detection by GSTP1 methylation analysis of postbiopsy urine specimens. Clin Cancer Res **2003**;9:2673-7

192. Pixberg CF, Raba K, Muller F, Behrens B, Honisch E, Niederacher D*, et al.* Analysis of DNA methylation in single circulating tumor cells. Oncogene **2017**

193. Friedlander TW, Ngo VT, Dong H, Premasekharan G, Weinberg V, Doty S*, et al.* Detection and characterization of invasive circulating tumor cells derived from

men with metastatic castration-resistant prostate cancer. Int J Cancer **2014**;134:2284-93

194. Gkountela S, Castro-Giner F, Szczerba BM, Vetter M, Landin J, Scherrer R*, et al.* Circulating Tumor Cell Clustering Shapes DNA Methylation to Enable Metastasis Seeding. Cell **2019**;176:98-112 e14

195. Mastoraki S, Strati A, Tzanikou E, Chimonidou M, Politaki E, Voutsina A*, et al.* ESR1 Methylation: A Liquid Biopsy-Based Epigenetic Assay for the Follow-up of Patients with Metastatic Breast Cancer Receiving Endocrine Treatment. Clin Cancer Res **2018**;24:1500-10

196. Chimonidou M, Kallergi G, Georgoulias V, Welch DR, Lianidou ES. Breast cancer metastasis suppressor-1 promoter methylation in primary breast tumors and corresponding circulating tumor cells. Mol Cancer Res **2013**;11:1248-57

197. Lyberopoulou A, Galanopoulos M, Aravantinos G, Theodoropoulos GE, Marinos E, Efstathopoulos EP*, et al.* Identification of Methylation Profiles of Cancer-related Genes in Circulating Tumor Cells Population. Anticancer Res **2017**;37:1105-12

198. Chimonidou M, Strati A, Malamos N, Georgoulias V, Lianidou ES. SOX17 promoter methylation in circulating tumor cells and matched cell-free DNA isolated from plasma of patients with breast cancer. Clin Chem **2013**;59:270-9

199. Casavant BP, Strotman LN, Tokar JJ, Thiede SM, Traynor AM, Ferguson JS*, et al.* Paired diagnostic and pharmacodynamic analysis of rare non-small cell lung cancer cells enabled by the VerIFAST platform. Lab Chip **2014**;14:99-105

200. Jones D, Pereira ER, Padera TP. Growth and Immune Evasion of Lymph Node Metastasis. Front Oncol **2018**;8:36

201. Browning M, Petronzelli F, Bicknell D, Krausa P, Rowan A, Tonks S*, et al.* Mechanisms of loss of HLA class I expression on colorectal tumor cells. Tissue Antigens **1996**;47:364-71

202. Koopman LA, Corver WE, van der Slik AR, Giphart MJ, Fleuren GJ. Multiple genetic alterations cause frequent and heterogeneous human histocompatibility leukocyte antigen class I loss in cervical cancer. J Exp Med **2000**;191:961-76

203. Koopman LA, van Der Slik AR, Giphart MJ, Fleuren GJ. Human leukocyte antigen class I gene mutations in cervical cancer. J Natl Cancer Inst **1999**;91:1669-77

204. Rao MS, Van Vleet TR, Ciurlionis R, Buck WR, Mittelstadt SW, Blomme EAG*, et al.* Comparison of RNA-Seq and Microarray Gene Expression Platforms for the Toxicogenomic Evaluation of Liver From Short-Term Rat Toxicity Studies. Frontiers in Genetics **2019**;9

205. Cancer Genome Atlas Research Network. The Molecular Taxonomy of Primary Prostate Cancer. Cell **2015**;163:1011-25

206. Aran D, Sirota M, Butte AJ. Systematic pan-cancer analysis of tumour purity. **2015**;6:8971

207. Duan R, Du W, Guo W. EZH2: a novel target for cancer treatment. Journal of hematology & oncology **2020**;13:104-

208. Burr ML, Sparbier CE, Chan KL, Chan YC, Kersbergen A, Lam EYN*, et al.* An Evolutionarily Conserved Function of Polycomb Silences the MHC Class I Antigen Presentation Pathway and Enables Immune Evasion in Cancer. Cancer Cell **2019**;36:385-401 e8

209. Anastasiadi D, Esteve-Codina A, Piferrer F. Consistent inverse correlation between DNA methylation of the first intron and gene expression across tissues and species. Epigenetics Chromatin **2018**;11:37

210. Ehrlich M, Lacey M. DNA methylation and differentiation: silencing, upregulation and modulation of gene expression. Epigenomics **2013**;5:553-68

211. Maunakea AK, Chepelev I, Cui K, Zhao K. Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. Cell Res **2013**;23:1256-69

212. Xu X, Tao Y, Gao X, Zhang L, Li X, Zou W*, et al.* A CRISPR-based approach for targeted DNA demethylation. Cell Discovery **2016**;2:16009

213. Brind'Amour J, Liu S, Hudson M, Chen C, Karimi MM, Lorincz MC. An ultra-low-input native ChIP-seq protocol for genome-wide profiling of rare cell populations. Nat Commun **2015**;6:6033

214. Rotem A, Ram O, Shoresh N, Sperling RA, Goren A, Weitz DA*, et al.* Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. Nat Biotechnol **2015**;33:1165-72

215. Dahl JA, Collas P. MicroChIP--a rapid micro chromatin immunoprecipitation assay for small cell samples and biopsies. Nucleic Acids Res **2008**;36:e15

216. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP*, et al.* Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell **2013**;152:1173-83

217. Larson MH, Gilbert LA, Wang X, Lim WA, Weissman JS, Qi LS. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. Nature protocols **2013**;8:2180-96