THE EFFECT OF THE MICROBIOME
ON DNA METHYLATION

A Thesis
Submitted to
the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree
MASTER OF SCIENCE

---

by
Lauren Cole
December 2017

Thesis Approvals:

Jean-Pierre Issa, MD, Thesis Advisor, Fels Institute, Department of Medicine
Nora Engel, PhD, Fels Institute, Department of Medical Genetics
Italo Tempera, PhD, Fels Institute, Department of Microbiology and Immunology
Çagla Tükel, PhD, Department of Microbiology and Immunology
Carmen Sapienza, PhD, Fels Insitute, Department of Pathology

ABSTRACT

Title: The Effects of the Microbiota and Inflammation on DNA methylation

Candidate's name: Lauren Cole

Degree: Master

TEMPLE UNIVERSITY, 2017

DNA methylation is an epigenetic mark with profound impact on gene expression and regulation. It is known to be altered both in cancer and throughout aging. These aging and cancer related changes are characterized by hyper-methylation of normally unmethylated CpG islands, and global DNA hypomethylation. Microbiota-free mice are known to live longer than their normal counterparts, and microbial dysbiosis is known to be a hallmark of colorectal cancer. In order to determine the microbiota's ability to impact DNA methylation patterns, that in turn may influence aging phenotypes and cancer development, the DNA methylomes of germ-free (GF; no microbiome) and specific pathogen free (SPF; controlled microbiome) mice were analyzed with a DREAM assay. This was done in wild-type mice and in IL-10 KO mice, with or without the addition of the colitis-associated cancer inducing compound, azoxymethane (AOM). We found that individually, inflammation and the microbiota induce moderate changes in the methylation profiles of the large intestine. However, their effects on DNA methylation seem to synergize; in the presence of inflammation, SPF mice have highly different methylation profiles than GF mice. In addition, inflammation causes large methylation changes in specific-pathogen-free mice, but only moderate changes in germ-free mice. The inflammation and microbiota induced changes were characterized by hyper-methylation of sites with low methylation (CpG islands), and hypomethylation of sites with high methylation; these patterns resembled the DNA methylation drift seen during aging and in transformed cells. All sites subject to age-related methylation drift

were vulnerable to the microbiome, but the converse was not true. A subset of sites was vulnerable to only the microbiome/inflammation, indicating multiple mechanisms of action. Overall, this research indicates the microbiota plays a key role in determining host DNA methylation state.

ACKNOWLEDGEMENTS

I would like to firstly acknowledge my thesis adviser Dr. Jean-Pierre Issa. Without his support, guidance, and patience, the completion of this thesis would not have been possible. I would also like to thank my thesis committee members Çagla Tükel**,** and Italo Tempera for their feedback and encouragement. I would like to give special thanks to Dr. Nora Engel, who, in addition to sitting on my committee, provided extra support and encouragement during my journey. Additional thanks go to Dr. Carmen Sapienza for being my external reviewer.

I would like to specially thank Dr. Barbara Hoffman for providing support and helping me grow as a writer.

I would like to specially thank Dr. Jaroslav Jelinek for everything he taught me.

Additional thanks to Fels Insitute faculty, students and staff.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

2HG	2-hydroxyglutarate
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
AID	Activation-induced cytidine deaminase
AML	Acute myeloid leukemia
APOBEC	Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like
AOM	Azoxymethane
APC	Adenomatous polyposis coli
BER	Base excision repair
BRAF	proto-oncogene B-Raf
CDK	cyclin-dependent kinase
CGI	CpG island
CIMP	CpG island methylator phenotype
CRC	Colorectal cancer
DNA	Deoxyribonucleic acid
DREAM	Digitial Restriction Enzyme Analysis of Methylation

DNMT	DNA methyltransferase
EBV	Epstein-Barr virus
FadA	Fusobacterium adhesion A
FOXP3	Forkhead box P3
GCPR	G-protein coupled receptors
GF	Germ-free
FH	Fumarate hydrogenase
HDAC	Histone deacetylase
HGC	High gene count
HLRCC	Hereditary leiomyomatosis and renal cell cancer
IBD	Inflammatory bowel disease
IDH	Isocitrate dehydrogenase
IL10	Interleukin 10
iNTK	invariant Natural killer T cell
KO	Knock out
LGC	Low gene count
LPS	Lipopolysaccharide A

MCT-1	Monocarboxylate transporter 1
MGMT	O-6 methylguanine-DNA methyltransferase
MLH1	Mut-L homolog 1
MMR	Mismatch repair
MSI	microsatellite instability
ncRNA	non-coding ribonucleic acid
NFκB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NLR	Nod-like receptor
PAMP	Pathogen associated molecular pattern
PcG	Polycomb repressor group
PCC	Pheochromocytoma
PGL	Paraganglioma
PRR	Pattern recognition receptor
PSA	Polysaccharide A
RNA	Ribonucleic acid
ROS	Reactive oxygen species
SAM	S-adenylmethionine

SCFA	Short-chain fatty acid
SDH	Succinate dehydrogenase
SFRP	secreted frizzled related protein
SPF	Specific-pathogen free
TET	Ten-eleven translocase
TF	Transcription factor
TLR	Toll-like receptor
TNF	Tumor necrosis factor
TSS	Transcription start site
WT	Wild type

CHAPTER 1

INTRODUCTION

1.1: Epigenetics and DNA Methylation

Epigenetics can be defined as the covalent modifications to DNA and histones that influence changes in gene expression, but are not changes in DNA sequence[1,2]. Three types of enzymes govern epigenetic modifications: writers, readers and erasers. Writers are enzymes that add modifications to DNA or histones; these include, but are not limited to acetyl, methyl, and phosphate groups. These moieties are bound to by reader enzymes, which can alter chromatin conformation, recruit or block transcription activators or repressors, and otherwise enact changes leading to altered gene expression. Finally, eraser enzymes remove epigenetic marks from DNA and histones, undoing the work of the writers. While epigenetic changes are stable, this writer/reader/eraser system allows them to be reversible[1].

DNA methylation is the most thoroughly studied epigenetic mark with a profound impact on gene expression. It is crucial for regulation of embryonic/organismal development, maintenance of cellular identity, establishment of genomic imprinting, and X-inactivation in female cells[3]. This mark consists of the covalent addition of a methyl group to a DNA base. In mammals, this methyl group is added to cytosine residue; this usually occurs only when the cytosine is followed by a guanine[2–4], though cytosine methylation outside CpG dinucletoides has been known to occur[5]. Methyl groups are added to the DNA by the DNA methyltransferase (DNMT) family of enzymes, which remove a methyl group from the methyl-donating complex S-adenylmethionine (SAM)[6], and transfer it to cytosine. DNMT3a and DNMT3b are responsible for establishment of DNA methylation, and methylation is maintained through the cell cycle and mitosis by DNMT1[3,5,6]. A DNMT KO in mice is embryonic lethal, demonstrating the importance of DNA methylation in development[5]. The DNA methylation mark can be removed in a multitude of ways, both active and passive. If the maintenance DNA methyltransferase, DNMT1, is inhibited or repressed during the cell cycle, methylation marks are not maintained through DNA replication, and DNA is passively demethylated. The ten-eleven translocase (TET) family of enzymes act as DNA demethylases by oxidizing the methyl group of 5-methylcytosine (5mC), converting 5mC to 5-hydroxymethylcytosine (5hmC). From there, 5hmC is converted into unmethylated cytosine by other enzymes[3]. In addition, the amine group of 5mC can be acted upon by the AID/APOBEC complex, converting the cytosine into uracil. The base excision repair pathway (BER) then excises and replaces the altered nucleotides[6].

An estimated 80% of the 30 million or so CpG sites located in the human genome are methylated[7]. In general, there is no clearly defined relationship between methylation of CpG sites and gene expression; genomic context is important for determining the effect CpG methylation will have. Gene methylation near a transcription start site (TSS) may directly impact transcription factor (TF) binding, or may alter chromatin conformation in way that determines the TF binding[5], and DNA methylation of promoter regions is generally thought to be repressive. Methylation within a gene body can influence alternative splicing, and there is evidence that gene body methylation is associated with transcriptional activation[6]. Intergenic DNA methylation impacts the ability of enhancers to interact with genes, and can influence the expression of non-protein coding transcription products[2].

While much of the relationship between DNA methylation and gene expression remains inconsistent, CpG island (CGIs) methylation is the exception. CGIs are usually defined as stretches of DNA 200 base pairs or greater in length that consist of 50% or greater CG content, and a ratio of at least 0.6 observed to expected CG dinucleotides[4]. CGIs are often located just upstream of a gene's transcription start site (TSS), in the gene's promoter. Approximately 70% of genes contain a CGI in their promoter regioms[6] and CGIs encompass approximately 0.7% of DNA. Additionally CGIs remain stably unmethylated[4,5], and the genes they govern are actively transcribed. If methylation does occur, it leads to gene silencing.

Studies have shown that CGIs could remain unmethylated due to recruitment of TET enzymes, which are able to bind to these regions via CXXC DNA-binding domains and remove DNA methylation[3,5]; TET1 is known to associate with CGIs[6]. It has also been shown that CGI regions contain fewer nucleosomes, allowing for a relaxed chromatin state. The nucleosomes that are associated with CGI have histone marks that are associated with a permissive chromatin state and gene expression. In addition, CGIs located in promoters that govern actively transcribed genes may be protected from DNA methylation by the presence of RNA polymerase II[5]. TF binding sites frequently are GC-rich, enabling them to associate with CGIs. Mutations in transcription factor binding sites within CGIs have been shown to lead to increased CGI methylation, indicating that transcription factors may be protective against CGI hypermethylation[5].

The complete mechanism by which CGI promoter methylation leads to gene repression has yet to be elucidated, but a great deal is known. Methylation of CGIs has been shown to inhibit transcription factor (TF) binding, but some TFs are not sensitive to DNA methylation[5]. Additionally, histones within CGIs often are modified with repressive marks[5]. It has also been shown that methylated CpG sites are bound by methyl-binding-domain (MBD) proteins, which play a role in gene repression by recruiting histone deaceylases (HDACs)[5]. One example of this is imprinted genes, which are kept repressed by hypermethylation of the CGIs in their promoters[5]. In addition, CGI methylation is used to control the vast changes in gene expression required during development[6] . Overall, CGIs are regions that are rich in CG sites with the potential to be methylated, but are largely methylation-free. When CpG methylation occurs, the regions enter a repressed, heterochromatin state, and as CGIs are methylated, gene expression is repressed.

Within a cell, there is a great deal of epigenetic crosstalk. Histone modifications, non-coding RNAs and DNA methylation have a dynamic relationship, with each one influencing the other. It has been demonstrated that DNMTs can bind to certain repressive marks on histone tails[6], indicating that DNA methylation may have some dependence on histone modifications. Additionally, it was found that in cancers with aberrant CGI methylation, there was an overrepresentation of mutations in chromatin regulators, further indicating that chromatin conformation plays a role in the determination of DNA methylation state[8]. In addition, DNMTs are known to cooperate with other enzymes that are thought to ultimately lead to gene repression[6]. They interact directly with histone modifying enzymes, such as histone deacetylases and histone methyltransferases, and non-coding RNAs. It has been hypothesized that ncRNA binding to DNMTs allows the ncRNAs to guide DNMTs to genes. Additionally, the non-coding RNA mir-209 has been shown to bind to and repress the *DNMT3A* and *DNMT3B* genes. Conversely, in a recent study, DNA methylation was shown to regulate the expression of 10% of detectable miRNAs[6].

Overall, DNA methylation is an epigenetic regulatory mark crucial for embryonic development and maintenance of cellular identity. It is influenced by host genetics, genomic context, and the cellular environment, responding to other epigenetic marks and mutations. Disruption in normal DNA methylation patterns has been associated with disease, including inflammation and cancer, and aging[1,2].

1.2: DNA Methylation Drift with Age

It has been well established that DNA methylation patterns change throughout the life of an organism, but the causes and mechanisms of aging-related changes in DNA methylation remain unproven. Aging-related methylation drift occurs in many species, with some conservation shown between humans and mice[7]. In general, aging related methylation changes are characterized by global loss of DNA methylation, in contrast to CGI promoters which become hypermethylated. Rates of age-related methylation drift vary widely between individuals and the extent of methylation change is highly tissue specific[2,7]. Maegawa *et al* demonstrated that aging-related DNA methylation drift is accompanied by changes in gene expression[7]; this is a possible mechanism behind aging-related changes in tissue function and organ failure in old age. Additionally, as the risk for cancer increases with age, and aging-related epigenetic changes are known to down-regulate expression of tumor-suppressor genes, it has been proposed that aging-related epigenetic changes are responsible for the initiation of tumorigenesis[2].

Tissues with higher rates of proliferation, such as the large intestinal epithelium, experience higher rates of aging-related methylation change. It is possible that replication errors in the maintenance of DNA methylation is responsible for methylation drift in aging, and a cell's DNA methylation levels may serve as a mitotic clock, similar to telomere shortening[2]. However, it has been shown that the rate DNA methylation drift is not uniform within a cell, and age-related methylation drift cannot serve as a perfect mitotic clock; it has been postulated that genomic context affects fidelity of DNA methylation maintenance by DNMT1, making error rate variable across the genome[2]. However, a study of monozygotic twins showed differing rates of aging related methylation drift[9], indicating that more than just the genome influences the rate of aging-related methylation drift. Additionally, genetically identical inbred mice show variability in methylation drift as they age, indicating that this phenomenon is conserved from mice to humans[2].

Aging related methylation drift is likely detrimental to stem cell function, causing a depletion of the stem cell population and aged-tissue phenotypes. In addition, there is evidence of variation in drift between stem cells of the same populations, with some cells experiencing extreme shifts in methylation state of a given gene, and some cells experiencing relatively minor shifts[2]. If extreme shifts lead to silencing of tumor suppressor genes, it is likely that natural selection will favor those cells, leading to a depletion of 'normal' stem cells, and an expansion of stem cells with significant aberrant methylation. In this way, aging related methylation drift can promote tumorigenesis, interfere with stem cell differentiation, and negatively impact tissue function.

1.3 Overview of the Gut Microbiome

The human gut microbiome is composed of approximately 10^14 bacteria, far outnumbering human cells in the body[10]. The gut microbiota is essential for host immune development and performing a range of digestive and metabolic reactions by breaking down substrates that host enzymes cannot[11]. In humans, the gut microbiota is dominated by two phyla: *Firmicutes* and *Bacteroidetes*, though other phyla maintain a significant presence. However, it has been difficult to establish the makeup of a 'core' human gut microbiome, as the human microbiome project found that gut microbiota composition varies greatly between individuals[12]. Despite the great variation in microbiota composition among individuals, there appears to be a bimodal distribution of the microbiome in the general population[13,14]. High gene count (HGC) individuals display high microbial diversity, and low gene count (LGC) individuals have a limited microbiome. LGC microbiomes were associated with obesity and inflammation of the gastrointestinal tract[13], known risk factors for colorectal cancer. Additionally, a 2011 study classified the human gut microbiome into three enterotypes, based on variations within levels of three genera: *Bacteroides, Prevotella,* and *Ruminococcus*[15].

The gut microbiota is thought to be established at birth, when bacteria from the mother can colonize the offspring. In infancy, gut microbiota composition is unstable[15], but as an individual develops gut microbiota composition remains relatively stable through life[14]. However, the relationship between the gut microbiota and human host is dynamic, with each one influencing the other. Host genetics, diet and medications are known to effect microbiota composition[16]. Studies comparing identical twins have shown variation, further indicating that the microbiome is highly dependent on environmental factors. There is considerable environmental variation within the gut; pH, oxygen exposure, host nutrient/bile/enzyme secretion, host mucosal surface, and immune interactions are all partially dependent on location along the digestive tract[14], and these factors influence the makeup of the gut microbiome. In return, the microbiota influences host physiology. These interactions include but are not limited to activation and suppression of the immune system and secretion of metabolites that interact with the host in a multitude of ways, including shifting of the epigenome.

1.4: Microbiota-Host Interactions

The gut microbiota and human hosts have a dynamic relationship with the capacity to interact in a multitude of ways. The microbiota has been shown to be involved in epigenetic changes in the host, both beneficial and pathogenic, and in addition, the microbiota is essential for immune development, and is able to up-regulate and down-regulate inflammatory responses[14,17]. Finally, microbes produce thousands of metabolites that interact with host cells[14].

The host immune system must maintain a balance between tolerance towards symbiotic gut microbiota and activation against potential pathogens. Beginning at birth, when the gut microbiota first colonizes a host, the host immune system and the gut microbiota develop together. There is evidence to suggest that the existence of a microbiome is essential to host immune development, as germ-free (GF, contain no microbiome and are housed in a sterile environment) mice do not develop full-fledged immune systems[18]. This "underdeveloped" GF mouse immune system is characterized by underproduction of antibodies, antimicrobial peptides, and T cells, and increased vulnerability to infection[19]. It is interesting to note that immune function is repaired if a microbiome is introduced to GF mice later in life[18]. However, in models of inflammatory bowel disease (IBD) and asthma, it was found that adult GF mice showed more severe inflammation than their specific-pathogen-free (SPF, contain a microbiome) counterparts. This effect was rescued by exposing GF mice to microbiota in the neonatal stage, but not in adulthood, indicating that earlier exposure to a gut microbiota is more beneficial[20].

A major mechanism of host-microbiota interaction is detection of bacterial components via pattern-recognition receptors (PRRs). PRRs are host receptors that bind to conserved molecular features of pathogens known as pathogen associated molecular patterns (PAMPs). Toll-like receptors (TLRs) are a subset of transmembrane PRRs that are expressed on a variety of cell types and frequently interact with microbial components[18]. TLR2 recognizes polysaccharide A (PSA), a component of the common bacterium *Bacteroides fragilis.* TLR4 binds to lipopolysaccharides (LPS), a component of the bacterial cell wall. TLR5 is activated by bacterial flagellin. In addition to TLRs, the cytosolic Nod-like receptors (NLRs) also play a role in mediating the interactions between the microbiota and the immune system. Activation of these pathways by PAMPs is known to upregulate NFκB signaling, which can lead to chronic inflammation and is known to be involved in tumor progression[21].

In addition to stimulating PRRs, bacteria produce a wide range of metabolites that interact with the host in a multitude of ways. A well-studied class of metabolites is short-chain fatty acids (SCFAs), which also have several mechanisms of interacting with the host. The most common SCFAs are butyrate, propionate, and acetate. These compounds are generated by the gut microbiome via fermentation of dietary fiber; consequently, the abundance of these metabolites is dependent on microbiota composition and host diet[18]. Additionally, the relative abundance of these different SCFAs vary based on microbiota composition, availability of substrates, and environmental conditions such as intestinal pH[14]. SCFAs can activate certain G-protein coupled receptors (GCPRs), which can induce protection against inflammation[14]. These compounds also have systemic effects on different tissues. Butyrate is metabolized for energy by colonocytes, whereas propionate is utilized by the liver, and acetate can circulate through the blood[14].

One SCFA of particular interest is butyrate; in addition to being the primary source of energy for colonocytes, it is an important epigenetic regulator[14]. Butyrate can enter the nuclei of host cells and inhibit histone deacetylases (HDACs). However, because colonocytes metabolize butyrate in the cytoplasm, it does not enter the nucleus to exert epigenetic effects in colonocytes as it may in immune cells[22]. Inhibition of HDACs leads to histone hyperacetylation, a phenomenon associated with gene activation, cell differentiation, apoptosis and anti-inflammatory effects[23]. This histone deacetylase inhibition has been shown to down regulate pro-inflammatory cytokines, and open chromatin at the *FOXP3* locus, causing differentiation of regulatory T-cells[24]. Additionally, butyrate is recognized by the GPCR GPR43 and is the only SCFA recognized by GPR109A. Activation of these GPCRs promote differentiation of regulatory T cells and IL10 producing T cells[14]. GPR43 and GPF109a are also known to act as tumor suppressors. Lastly, butyrate has been shown to increase mucosal barrier function[25].

Plant-based diets lead microbes to produce other protective metabolites. They include antioxidants that neutralize reactive oxygen species (ROS), which, when left unchecked, cause mutations that can lead to disease[14]. In addition, phenolic metabolites derived from host diet have anti-inflammatory and antimicrobial effects, which can alter microbiota composition[14].

The microbiota has been cited as a cause of broad epigenetic alterations within a host. Recently, Denu *et al* showed in mice that the presence of a microbiome significantly altered the epigenome over germ-free controls[26]. These changes were characterized by changes in histone methylation states and histone hyper-acetylation. The mass histone reprogramming was diet-dependent, requiring fiber, which is fermented by the gut microbiome into SCFAs. Treatment of these mice with SCFAs (butyrate, propionate and acetate) partially recaptured the epigenetic effects of a microbiome, indicating the importance of bacteria-produced metabolites, particularly the fermentation fiber into SCFAs, in modulating host epigenetic state. Additionally, these epigenetic changes were seen in multiple host tissues, including the colon, liver, and adipose tissue, demonstrating the systemic effects of the microbiome. In addition to affecting histone modifications, the microbiome has been implicated in DNA methylation changes. A 2014 study of 8 pregnant women, which examined fecal microbial composition, divided them into two groups: four women had microbiomes enriched for *Firmicutes* and the other four women had microbiomes enriched for *Bacteroidetes*[27]. It was found that CGI promoter DNA methylation profiles differed based on which phyla dominated the gut microbiome. However, the mechanism behind bacterial-related differences in host methylation profiles is unclear. Additionally, this study was small and needs to be replicated.

While the gut microbiota is essential for proper immune function and nutrient absorption, it can also induce effects that are detrimental to the host. Though the biggest risk factor for development of inflammatory bowel disease (IBD) is family history, known risk loci are unable to explain all of the variance in the development of these inflammatory diseases[28]. It has been hypothesized that much of this unaccounted-for risk may be attributed to the microbiome. Chronic intestinal inflammation is associated with microbial dysbiosis, characterized by a decrease in microbial diversity, a decrease in the abundance of both SCFAs and SCFA-producing *Firmicutes* species, and an increase in pathogenic species of bacteria[29]. Of note is the increase in the adherent-invasive members of the genus *Fusobacterium* and the family *Enterobacteriacae,* which includes the species *E. coli.* In inflammation, the colonic mucosa degrades, leaving the intestinal epithelium vulnerable to invasion by microbes; in addition, *Fusobacterium* are able to contribute to the degradation of the mucosa[30]. *Fusobacterium* has been linked to inflammation via its production of the adhesin FadA, which binds to E-cadherin on host cells, degrading mucosa and leading to downstream activation of NFκB[31]; this can facilitate the invasion of the host cell by the bacterium. *E. coli* is another species known to be increased in abundance in IBD. Certain strains of the adherent-invasive species *E. coli* have been shown to cause colorectal cancer, via production of the compound colibactin, which causes DNA damage[32].

It is important to recognize that the effects of the gut microbiota are not constrained within the digestive tract; the gut microbiome has been linked to juvenile diabetes, allergies and autoimmune disorders, psychiatric disorders, and cancer[33] and obesity[13]. Microbes can have systemic effects on host physiology via the immune system, and intestinal absorption of microbes and microbial products. If the colonic mucosa is degraded during inflammation, the intestinal epithelium becomes permeable, allowing microbes to enter the bloodstream and circulate[33]. In addition, the microbe has significant metabolic effects. Studies of obesity in humans and mice have found that obese and lean individuals have distinct microbiomes[14]. Cohousing experiments in mice have revealed that microbiota transfer between obese and lean mice result in obese mice taking on the 'lean' phenotype, indicating that the microbiota is linked intimately with metabolism[34]. Interestingly, GF mice are protected from diet-induced obesity, indicating that the microbial products play an essential role in the development of the obese phenotype[35].

Denu *et al.* showed that microbial metabolites induced diet-dependent broad changes in histone modifications. Host diet influenced microbial metabolite production, which changed the host epigenome in multiple tissues, further demonstrating the systemic effects of the microbiota on the host, and a two-way relationship between the two[26]. Olszak *et al* found that GF mice showed an elevated amount of invariant natural killer T cells (iNTKs) over conventional mice, and these iNTKs contributed to inflammatory disorders in both the colon and the lung[20]. Overall, the effects of the gut microbiota are not limited to the large intestine, but influence multiple facets of host physiology.

While the microbiota certainly has significant effects on the host, the host also influences bacterial abundance and microbial composition. An important modifier of microbial composition is inflammatory state of the gut. One 2012 study by Arthur *et al* examined the impact of inflammation on the microbiome by comparing the microbiomes of WT and IL10KO mice[32]. IL10 is an anti-inflammatory cytokine, and IL10KO mice experience chronic inflammation. Arthur *et al* found that this IL10KO induced chronic inflammation caused an overall loss in microbial diversity, characterized by decrease in amount of detectable species within the gut. Additionally, principal component analysis showed that the microbiomes of IL10KO mice clustered separately from those of WT mice, and IL10KO mice had increased representation of the pathogenic *Enterobacteriacae* family[32]. In addition to inflammation, diet is known to modulate host gut microbiota[14,28], with long-term dietary patterns being a major determinant of microbial composition. Members of the phylum *Firmicutes* are responsive to changes in carbohydrate intake, whereas *Bacteroides* increased and *Firmicutes* decreased in abundance in response to an animal-based diet[14]. In addition, diet can impact overall microbial diversity, the loss of which is linked to disease[13,14]. Finally, host genetics are also known to play a role in determining microbial composition. A 2013 genome-wide analysis of 1,514 individuals discovered several genetic loci that associate with microbial taxonomy and gene ontology[16].

Overall, the microbiota has a significant effect on host physiology and vice versa. The microbiota metabolizes compounds the host cannot, and produces nutrients the host can utilize. It regulates host epigenetic states, both via metabolites and impacting immune function. Certain bacterial species are capable of inducing inflammation; others reduce it. In return, the microbiota responds to changes made in the host environment, with microbial composition shifting in response to dietary changes, inflammation, and medications. Disruptions in host-microbiota homeostasis certainly have the potential to be pathogenic; studying the mechanisms behind the host-microbiota relationship is crucial to combating these diseases. This research project aims to look at the link between the microbiota and aberrant methylation.

1.5: Aberrant DNA Methylation in disease

Chronic inflammation is strongly associated with aberrant DNA methylation[2]. A 2008 study by Hahn *et al* measured gene methylation in a mouse model of inflammatory bowel disease (IBD). They found that at the age of 8 months, there were 249 genes with increased CGI promoter hypermethylation relative to control mice[36]. Hahn *et al* found that genes that change in aging, as well as genes that do not were affected by inflammation. In addition, the presence of chronic inflammation in the colon has been shown to specifically accelerate aging-related methylation drift[37], indicating that there may be a relationship between inflammation-induced and age-related methylation drift. Another study in gastric epithelial cells found that infection with the pathogenic bacteria *H. pylori* led to CGI hypermethylation in gastric epithelial cells. Evaluation of the mechanism revealed that *H. pylori* didn't cause aberrant methylation directly, but triggered inflammatory processes that altered the DNA methylation state[38]. However, despite the considerable evidence linking inflammation with aberrant methylation, the mechanism by which inflammation impacts DNA methylation is unknown. It has been proposed that because inflammation increases cell turnover rate, the inflammation induced-methylation drift is a result of accelerated aging-related methylation drift[2]. While there is overlap between the genes sites that are affected by inflammation and aging[36], this explanation does not account for the CpG sites and genes affected by only inflammation, or only by aging.

Chronic inflammation is also known to be a major risk factor for the development of cancer. Additionally, aberrant methylation is a common feature of many cancers[39], including breast[3], leukemia[4], liver[40], lung[41], gastric[42] and colorectal[43]. Much like aging, these aberrant patterns are characterized by global hypomethylation, and hypermethylation of CGIs and may lead to the initiation of tumorigenesis in a multitude of ways. Hahn *et al* found that inflammation-associated methylation was more likely to correspond with aberrant methylation in tumors than age-relating methylation, though there was some overlap between aging-related and inflammation-related methylation. This indicates that inflammation-induced methylation changes are pathogenic[36].

In some cancers, the causes of aberrant methylation are known. Isocitrate dehydrogenase mutations (IDH1/2) are known to cause hypermethylation in subsets of acute myeloid leukemia and gliomas[1]. These Krebs cycle IDH enzymes acquire a gain-of-function mutation causing them to produce the metabolite 2-hydroxyglutarate (2HG) instead of α-ketoglutarate[23]. 2HG can bind and inhibit the TET enzymes responsible for removing methylation marks from DNA and histones[44]. These enzymes no longer function, and consequently, DNA hypermethylation accumulates[1]. However, IDH mutations do not explain the global hypomethylation phenomenon often seen in cancers. Often, the epigenetic modifiers themselves are mutated: TET and DNMT mutations are common in myeloid malignancies[1]. However, many cancers with aberrant methylation do not experience these epigenetic enzyme mutations.

This thesis aims to investigate the role of the gut microbiome in aberrant DNA methylation; consequently, colorectal cancer (CRC) is of specific interest. Colorectal cancer is the third most common cancer and third leading cause of cancer death worldwide[45]. Most colorectal cancers are not hereditary, and develop because of environmental conditions. Additionally, a common feature of colorectal cancers is over-activation of the Wnt/β-catenin pathway[46]. In terms of DNA methylation, colorectal cancer can be divided into three subsets: CIMP1, CIMP2 and CIMP0. The concept of the CpG-Island Methylator Phenotype (CIMP) was first proposed in 1999 by Toyota *et al*, and evidence in support of this concept has accumulated since[43]. CIMP refers to levels of aberrant DNA methylation, with CIMP1 tumors showing high levels of aberrant methylation, CIMP2 tumors showing moderate aberrant methylation, and CIMP0 tumors showing little to no aberrant methylation[46]. It is interesting to note that CIMP1 and CIMP2 cancers have different molecular features[46]. CIMP1 CRCs are associated with mutations in *BRAF*, hypermutation, and microsatellite instability whereas CIMP2 cancers are associated with *KRAS* mutations. In general, colorectal tumors with aberrant methylation do not harbor *TP53* mutations, whereas CIMP0 tumors often do[47]. Approximately 20% of CRC tumors are CIMP-positive, and they tend to arise from serrated polyps instead of adenomas. Additionally, CIMP-positive tumors tend to be located in the proximal colon, and are more common in women[46]. This suggests CIMP-positive tumors possess key molecular differences from non-CIMP CRC tumors, and arise via a different mechanism(s).

Among the many genes whose promoters are found to be hypermethylated in CIMP-positive tumors are the secreted-frizzled related protein (*SFRP)* genes, *CDKN2A/p16, MGMT*, *CDKN2A (p14ARF)* and *MLH1*[46]. The *SFRP* genes are known negative regulators of the Wnt/β-catenin pathway, the activation of which is associated with a stem-like phenotype and hyper-proliferation[48]. *P16* and *p14* are tumor suppressor genes that down-regulate the cell cycle. *MLH1* and *MGMT* are DNA repair genes: *MLH1* is involved in the mismatch repair pathway (MMR) and *MGMT* repairs alkylation damage products on DNA; dysregulation of these DNA repair pathways results in the hypermutation and microsatellite instability (MSI) seen in CIMP1 tumors[46]. It is easy to visualize how the down-regulation of any of these genes by hypermethylation results in tumorigenesis. The gene expression changes brought about by CGI hypermethylation are thought to convey a selective advantage to the afflicted cells, allowing them to out-compete cells with normal methylation profiles. This clonal growth advantage drives tumorigenesis[46]. Global hypomethylation can cause erroneous activation of transposons, endogenous retroviruses and non-coding RNAs, genomic instability, and loss of imprinting and is associated with microsatellite instability (MSI)[43,46]. Hypermethylation of CGIs causes gene silencing, and is known to affect several tumor suppressor genes. In multiple cancers, hypermethylation has been shown to silence genes involved in DNA repair, including *MGMT* and *MLH1*[43].

1.6: Overview of Proposed Mechanisms of Aberrant Methylation in Colorectal Cancer

The cause of CIMP in CRC is not known, though multiple mechanisms for aberrant methylation have been hypothesized. Additionally, the presence of both CIMP1 and CIMP2, which have different aberrant methylation patterns, and associate with different mutations[46], suggest that there are multiple mechanisms by which aberrant methylation in colorectal cancer may arise. Interestingly, aberrant methylation patterns in colorectal tumors differ from those observed in other tumor types. Aberrant methylation patterns are associated with cell immortalization and are thought to occur early in the malignant transformation process[46]; this is corroborated by the fact that aberrant methylation can be detected in early stage adenomas[49].

There is a large overlap in genes that are methylated with age and genes are that methylated in colorectal cancer. The *APC* gene (down-regulator of the Wnt/β-catenin pathway, often implicated in CRC), is known to be methylated in aging[50]. In addition, age is a significant risk factor in the development of colon cancer[51]. Conversely, it has been proposed that aberrant methylation in cancer reflects the cell's undifferentiated "stem-like" phenotype, rather than an "aged-cell" phenotype[46].

It has also been proposed that aberrant methylation results from a misdirection of DNMTs. Halogenation of DNA may create DNA damage products that mimic 5mC, tricking DNMT1 into adding a methylation mark during DNA replication[36]. Genes with Polycomb repressor group (PcG) marks on histones are known to be erroneously methylated, and PcG proteins are known to direct DNMTs to genes; it has been proposed that PcG errors are in part responsible for CGI hypermethylation[36]. However, neither of these mechanisms can explain the hypomethylation that accompanies CGI hypermethylation. Additionally, DNMT over-expression has been noted in some colorectal cancers, which can induce aberrant methylation[46]. Finally, it has also been proposed that a malfunction in barrier elements (insulators) can cause methylation to spread into different genomic “neighborhoods”[46]. Again, this explanation cannot account for DNA hypomethylation. In addition, there are examples of locus-specific aberrant methylation which the “methylation spreading” hypothesis cannot account for[46]. Overall, it is not clear if, or to what degree, each of these possible mechanisms plays a role in aberrantly methylated colorectal cancers.

Because the DNA methylation changes typically seen are bidirectional (global hypomethylation and CGI promoter hypermethylation), it is unlikely that there is just one mechanism responsible for these erroneous methylation changes. A likely mechanism for aberrant methylation in the large intestine is via the gut microbiome. Some species of bacteria have been proven to be causal both in colorectal cancer, and the microbiome has is linked to epigenetic change.

1.7: Linking the Gut Microbiome to Epigenetics and Cancer

There is considerable evidence that the gut microbiota can play a causal role in the development of cancer. It is known that the gut microbiome can both hinder, promote and respond to the development of disease through various mechanisms, including immune modulation, induction of or protection against inflammation, and via the production of microbial metabolites that interact with the host colonic epithelium[14]. It has been demonstrated that the microbiota can cause changes in host epigenetic states, and a few studies (discussed below) have established links to epigenetic changes in disease to the microbiome.

Colorectal cancer is strongly associated with microbial dysbiosis. A study comparing the fecal microbiota of 19 colorectal cancer patients and 20 healthy control subjects found an enrichment of several bacterial species, including several species of *Fusobacterium*. Additionally, there was an underrepresentation of butyrate-producing bacterial species[52]. In a mouse model of colorectal cancer, it was found that pro-biotic supplementation with *B. fibrisolvens,* a butyrate-producing bacterial species, accompanied by a high-fiber diet, led to increased production of butyrate and significant reduction in tumor burden[22]. This demonstrates that bacterial production of butyrate plays a preventative role in the development of CRC, and indicates that a lack of such butyrate-producing species may be carcinogenic.

It was demonstrated that *H. pylori*, a well-known pathogenic bacterium that causes inflammation, also induces changes in DNA methylation at CGIs in gastric epithelial cells[38]. It was also demonstrated that the aberrant methylation caused by *H. pylori* infection increased the risk for gastric cancer[53]. Also in gastric cancer, CIMP has been linked to Epstein-Barr virus (EBV) presence[54]. In addition, *E. coli* down-regulates MCT-1, a receptor responsible for uptake of anti-inflammatory, epigenetic regulator butyrate[55]. *E. coli* has also been shown to cause transformation via production of colibactin, which causes DNA damage in host cells[32]. As previously stated, *Fusobacteria* is thought to promote tumorigenesis through its production of FadA[56]. *Fusobacteria* also has an epigenetic link; it is enriched in in CRCs with CIMP, but not in cancers with a normal methylation profile[49].

1.8: Hypothesis and Goals

DNA methylation profiles shift aberrantly in aging, cancer, and inflammation, the latter two of which are accompanied by shifts in the microbiome. In addition, GF mice experience longer lifespans than their microbiota-containing counterparts.

Based on the literature above, we hypothesized that:

1. a microbiota would induce broad changes in DNA methylation in mice
2. These changes would be similar to those seen in aging and inflammation

Therefore, the goals of this project are:

1. To evaluate and compare the DNA methylation profiles of mice with and without a microbiota
2. Test the role of inflammation in this process by including both GF and SPF IL10KO mice in our analysis
3. Compare the effects of inflammation and the microbiome to the effects of aging, based on previous aging data

CHAPTER 2

METHODS

2.1: Mouse Models

Mouse tissues used in microbiome and IL10KO DREAM analyses were obtained from previous studies conducted by Christian Jobin, and husbandry techniques described in Arthur *et al*[32]. Germ-free (GF) mice were raised in sterile conditions (germ-free isolators) and kept free of microbiota, including mother-pup transmission of microbes. Specific-pathogen-free (SPF) mice were kept in controlled conditions and free of pathogenic microbes, including fungi and viruses. IL10KO mice experience inflammation, and AOM is a carcinogenic inflammatory agent. 42 mice were used in our analysis; the table below shows the breakdown of mice and treatment group.

Mouse Model	Number of Mice
WT/SPF	6
WT/GF	6
IL10KO/GF	6
IL10KO/SPF	6
WT/SPF + AOM	6
IL10KO/GF + AOM	6
IL10KO/SPF + AOM	6

2.2: Digital Restriction Enzyme Analysis of Methylation (DREAM)

Digital Restriction Enzyme Analysis of Methylation (DREAM) was performed on DNA extracted from whole mouse intestinal tissue, as described in Jelinek *et al*[4]. It is a quantitative, deep-sequencing based method of measuring DNA methylation. It is based on sequential digestion by two restriction enzymes, which cut the same consensus sequence: CCCGGG. DREAM detects 175,000 CpG sites located within the CCCGGG sequence. The DNA is first treated with *SmaI*, the activity of which is blocked by DNA methylation. *SmaI* leaves the CCC/GGG blunt ends after cutting, and target sites with a methylated CpG are not cut. The DNA is then digested with *XmaI*, which cuts at C/CCGGG and is not blocked by methylation. It leaves a CCGGG 5' overhang. Sequencing adaptors are added to the digested DNA fragments, and libraries are prepared for deep sequencing. Libraries used in this study were sequenced on Illumina HiSeq 2000. To ensure precision, we only analyzed CpG sites that had greater than 100 reads in at least 4/6 mice in each group.

DREAM detects approximately 170,000 CpG sites located in the mouse genome. It is unable to detect CpG methylation in CpG sites located outside of the CCCGGG sequence targeted by the restriction enzymes used.

2.3 Statistical Analysis

Statistical analyses were performed, and figures were generated using R. Volcano plots were generated by plotting the average difference in methylation for a given CpG site against the negative log10 of the p-value. Average methylation change was calculated by subtracting average methylation in GF mice from average methylation in SPF mice, or

WT mice from KO mice, etc. A student t-test was used to calculate a p-value. All double Venn diagrams were generated using the VennDiagram package in R; p-values were calculated using a hypergeometric test for enrichment. Triple Venn diagrams were created using the BioInfoRX Venn diagram tool, and p-values for each overlap were calculated using the SuperExactTest package in R. All bar graphs were generated using the ggplots2 package in R. Pathway analyses were performed using the Modular and Singular Enrichment Analysis tool provided by GeneCodis, and results were visualized using R. Scatterplots were also generated in R.

CHAPTER 3

RESULTS

3.1: The microbiome influences DNA methylation

We used the quantitative, deep-sequence based method digital restriction enzyme analysis of methylation (DREAM) to examine how the presence of a microbiome affected global DNA methylation. To do this, DREAM data for wild-type germ-free (WT/GF, no gut microbiome) and wild-type specific-pathogen-free (WT/SPF, containing controlled gut microbiome free of pathogens) mice was analyzed and a volcano plot depicting the methylation differences between them was generated (Figure 1a). For most experiments, each group of mice consists of 6 animals. We considered sites to be "changed" by the microbiome if there was a statistically significant increase or decrease

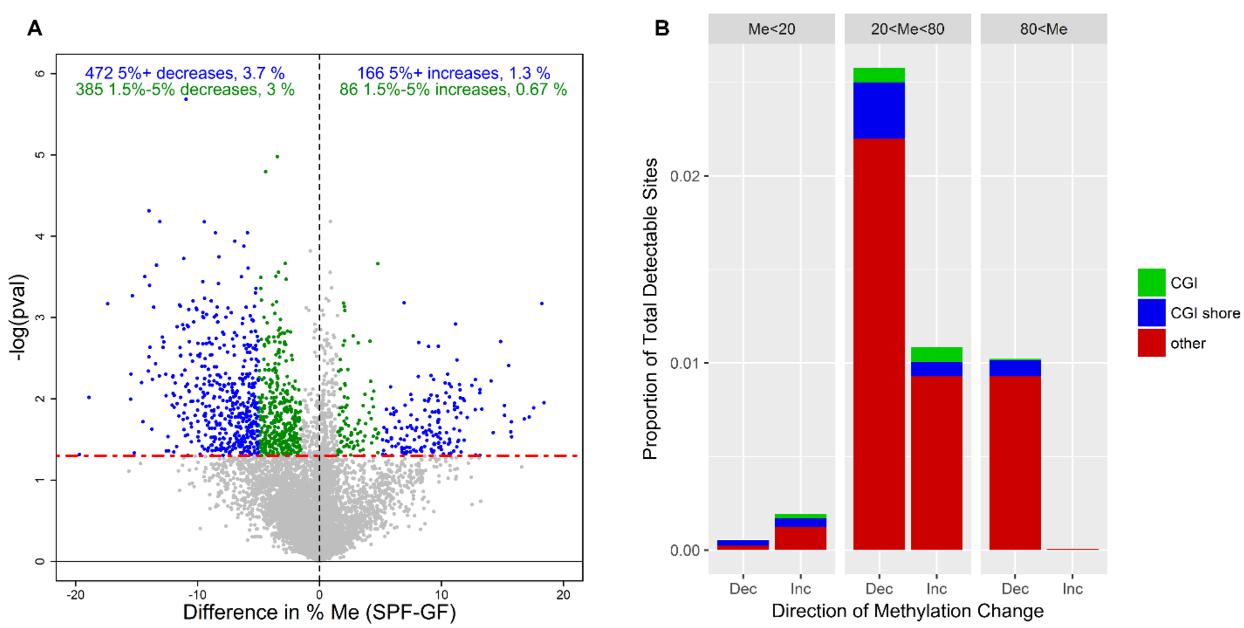


Figure 1: The microbiome influences DNA methylation. A) Volcano plot analysis showing methylation differences between SPF and GF mice. The x-axis shows the difference in average methylation between SPF and GF mice for a given site. The y-axis is the negative log(10) of the p-value, which was determined with a student t-test. All sites above the red line are significant at p=0.05. Green sites change between 1.5% and 5% and blue sites at a magnitude of 5% or greater. B) Bar chart showing the proportion and type of CpG sites that change at least 5% between SPF and GF mice in 1a. Shore indicates the sites are not in CGIs, but within 2000 bp of them.

in average methylation of 5% or more. To ensure precision, we only analyzed CpG sites that had greater than 100 reads in at least 4 out of 6 mice in each group. Overall, of 12,919 detectable sites, 3.7% decreased, and 1.3% of sites increased in methylation.

Sites that are greater than 80% methylated or less than 20% at baseline tend to be more stable than intermediately methylated sites[4]. Therefore, we examined how the microbiome affects different CpG compartments. Volcano plots were generated that looked only at sites with greater than 80% average methylation in GF mice (Figure 2a), sites with less than 20% methylation in GF mice (Figure 2b) and sites with between 20 and 80 percent methylation (Figure 2c). Overall, 0.3% of unmethylated sites were affected by the microbiome, compared to 5.2% of highly methylated sites, and 22.3% of intermediately methylated sites. Among this latter compartment (20-80% methylation), 15.7% decreased compared to 6.6% increase. Figure 1b shows the distribution of sites that were changed at least 5% between GF and SPF mice, stratified by baseline methylation. CGI and CGI shore sites appear to be largely stable. The most vulnerable sites seem to be gene body or intergenic areas with baseline methylation of between 20% and 80%. Thus the presence of a microbiome significantly affects DNA methylation with more pronounced hypomethylation than hypermethylation.

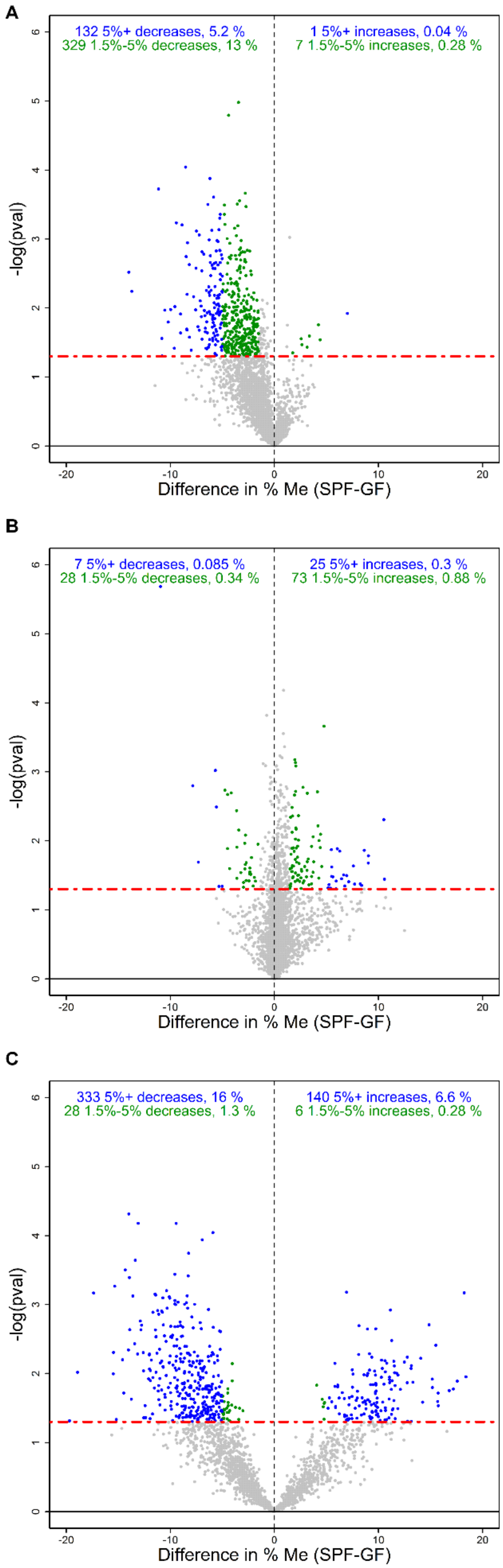


Figure 2: The microbiome influences DNA methylation, stratified by baseline methylation. A) Volcano plot analysis showing methylation differences between WT/GF and WT/SPF mice, only analyzing sites with greater than 80% methylation. B) Volcano plot analysis showing methylation differences between SPF and GF mice, only analyzing sites with less than 20% methylation. C) Volcano plot analysis showing methylation differences between SPF and GF mice, only analyzing sites with between 20% and 80% methylation.

3.2: The presence of inflammation independently affects DNA methylation

IL10 is an anti-inflammatory cytokine and IL10KO mice are used as a model of spontaneous inflammation. To determine the effects of inflammation on DNA methylation in the absence of a microbiome, DREAM data from GF/WT and GF/IL10KO mice were analyzed; Figure 3a shows a volcano plot analysis. Of all detectable sites, 2.4% decreased and 1.75% increased in methylation Figure 3b shows the distribution of affected sites located in CGIs, CGI shores or other sites; Like Figure 1b, Figure 3b shows that relatively few CGI/shore sites are subject to change, and most of the changes that do occur are at variable sites with between 20% and 80% methylation.

Next we compared the sites that were affected by the microbiome to the sites that were affected in IL10KO mice. We compared sites changed from GF to SPF mice in Figure 1 to the sites change in the IL10KO mice in Figure 2. Out of 11,377 total detectable sites, 409 hypomethylated as a result of the microbiome, 242 sites decreased as a result of IL10KO, and 43 sites were changed under both conditions (Figure 3c, hypergeometric test, p=2.7x10^-14). Similar data were seen for hypermethylation. 140 sites hypermethylated by the microbiome, 264 sites hypermethylated by the IL10KO and 11 sites increased in both (Figure 3d, p=0.001). Thus, while the overlaps are statistically significant, there are relatively few sites that actually overlap between the IL10KO and the microbiome, which may indicate that inflammation and the microbiome act partially on separate sites.

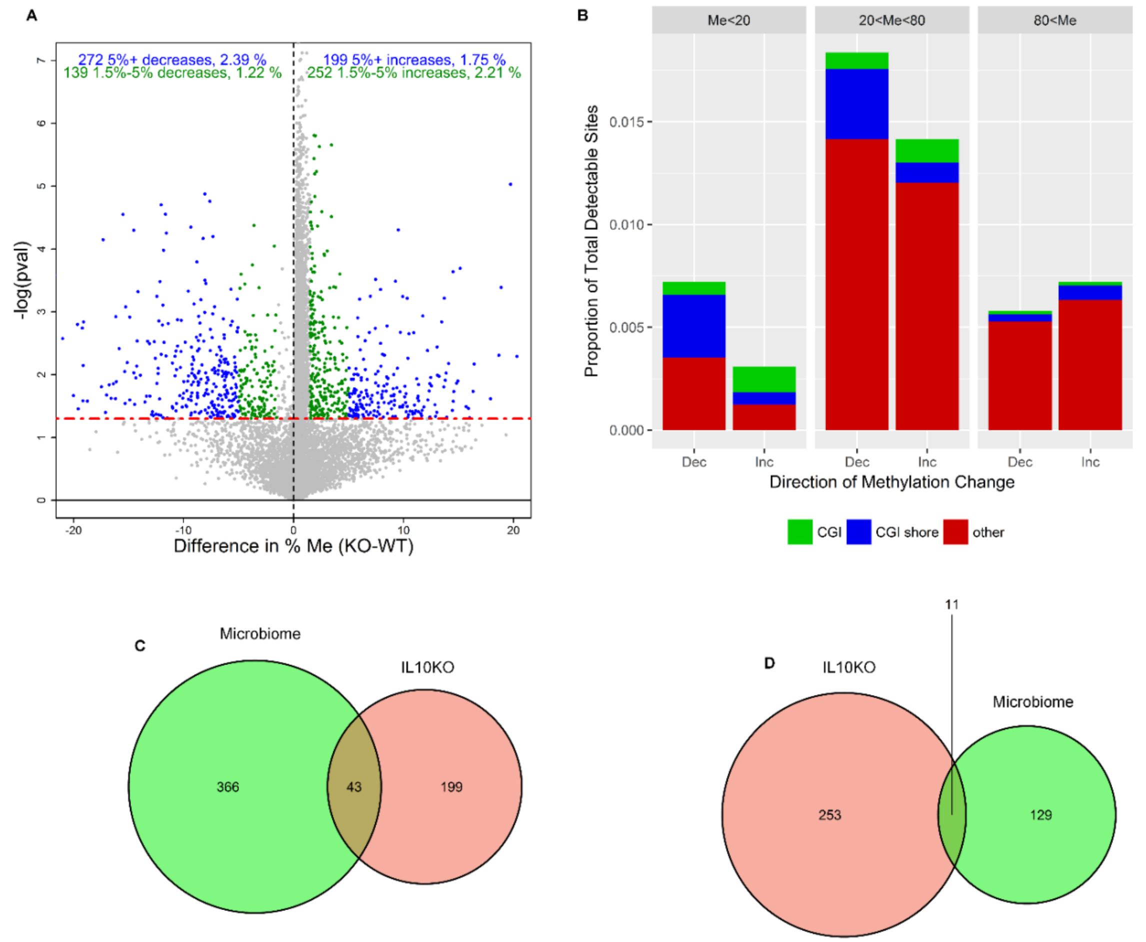


Figure 3 The presence of IL10KO affects DNA methylation. A) Volcano plot analysis showing methylation differences between WT/GF and IL10KO/GF mice. The x-axis shows the difference in average methylation between WT and KO mice for a given site. The y-axis is the negative log(10) of the p-value, which was determined with a student t-test. All sites above the red line change are significant at p=0.05. Green sites change between 1.5% and 5% and blue sites at a magnitude of 5% or greater. B) Bar chart showing the proportion and type of CpG sites that change at least 5% between SPF and GF mice in 2a. Shore indicates the sites are not in CGIs, but within 2000 bp of them. C) Overlap of CpG sites that lose methylation by at least 5% by the microbiome or by inflammation. GF/WT vs SPF/WT and GF/WT vs GF/IL10KO. N=11377, p=2.7x10^-14. D) same as C, looking at sites that gain methylation by at least 5%. N=11377, p=0.001.

3.3: The microbiome and inflammation synergize to have a large effect on DNA methylation.

We have shown that individually, the microbiome and IL10 deficiency each have a moderate effect on DNA methylation. By comparing the DREAM data from GF/WT and SPF/IL10KO mice, we were able to observe the effects of both conditions occurring simultaneously. In a volcano plot analysis, we found that 8.1% of sites increased in average methylation and 9.9% of sites decreased (Figure 4a). In addition to more sites changing average methylation, the combined effects of inflammation and the microbiome appear to cause a specific increase in methylation at CGIs and CGI shores. Figure 4b shows that 2% of all sites are CGI/CGI shore sites that increase in average methylation between WT/GF and SPF/IL10KO mice. This contrasts the individual effects IL10 deficiency or the microbiome, which had little effect on CGIs or CGI shores. Figure 4c gives a broad overview of the effects of the microbiome and IL10KO on DNA methylation individually and together. With a threshold of 5% change in average methylation, both the microbiome and IL10 deficiency change methylation at approximately 5% of sites. When combined, approximately 18% of sites experience average methylation changes, with more sites decreasing in average methylation than increasing. When we examined CGIs exclusively (Figure 4d), 0.4% were changed by SPF, and 0.4% were changed by IL10KO while 4% were changed by both simultaneously.

We examined overlap between sites changed by the microbiome, sites changed by the lack of IL10, and sites changed by the combination of both conditions. To do this triple Venn diagrams were generated, and p-values for each overlap were calculated using the

SuperExactTest package in R. Figure 4e shows the overlap of sites that hypermethylated comparing IL10KO/GF mice, WT/SPF mice, and IL10KO/SPF mice to the WT/GF baseline. It is interesting to note that the sites being acted on by just the microbiome are also acted on by the microbiome-inflammation combination; approximately 75% of the sites that are changed in WT SPF mice are also changed in IL10KO SPF mice. Approximately 50% of the sites that change in KO/GF mice are also changed in KO/SPF mice. All overlaps are statistically significant at p=0.01 (see Figure 4 legend for p-values). The same pattern was observed when looking at sites that decrease at least 5% in average methylation (Figure 4f).

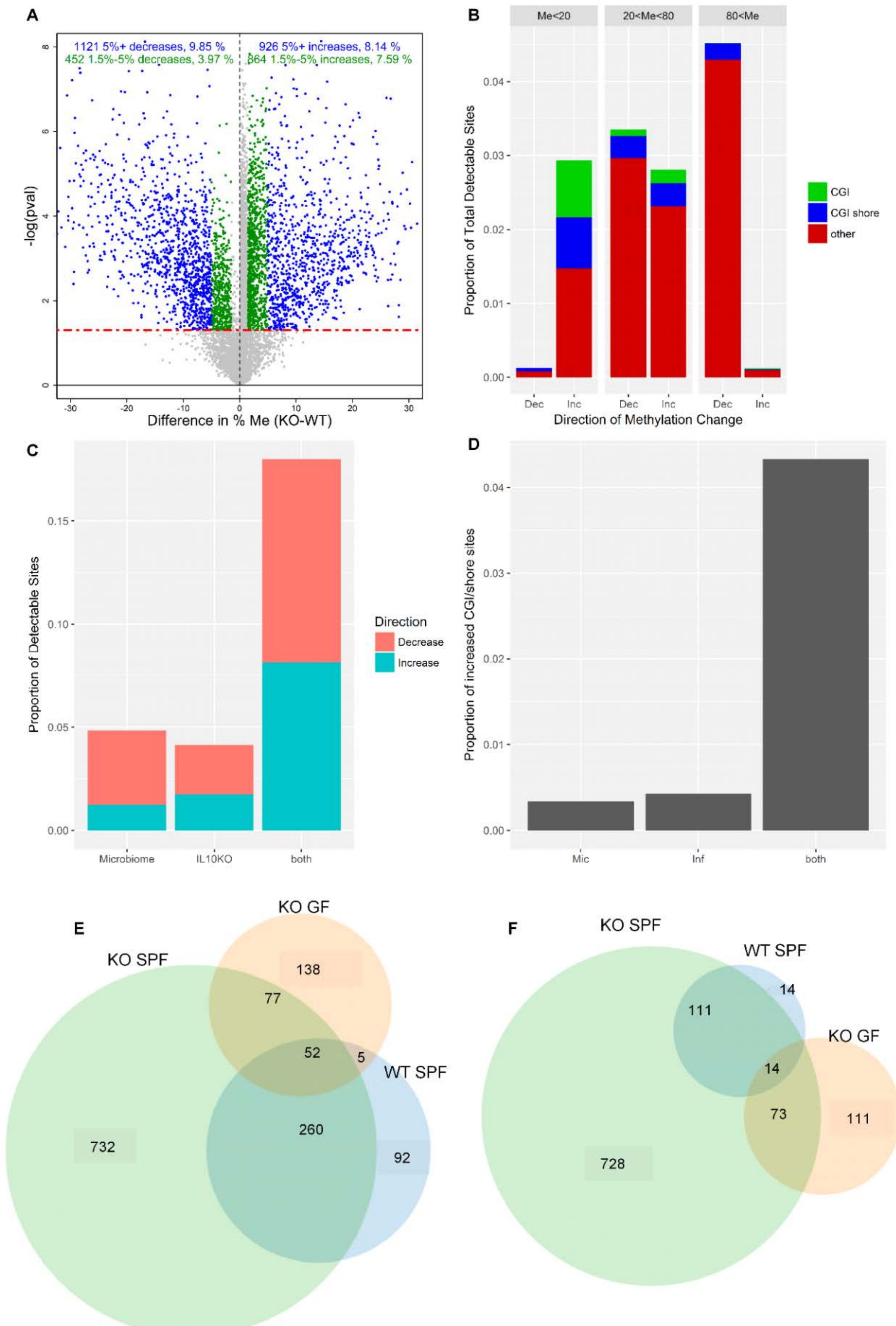

A
1121 5%+ decreases, 9.85 %
452 1.5%-5% decreases, 3.97 %
926 5%+ increases, 8.14 %
864 1.5%-5% increases, 7.59 %
-log(pval)
Difference in % Me (KO-WT)
B
Me<20
20<Me<80
80<Me
Proportion of Total Detectable Sites
Direction of Methylation Change
Dec
Inc
CGI
CGI shore
other
C
Proportion of Detectable Sites
Direction
Decrease
Increase
Microbiome
IL10KO
both
D
Proportion of increased CGI/shore sites
Mic
Inf
both
E
KO GF
KO SPF
WT SPF
138
77
52
5
260
732
92
F
KO SPF
WT SPF
KO GF
14
111
14
73
111
728

Figure 4: The microbiome and inflammation synergize to have a large effect on DNA methylation. A) Volcano plot analysis showing methylation differences between WT/GF and IL10KO/SPF mice. The x-axis shows the difference in average methylation between WT/GF and IL10KO/SPF mice for a given site. The y-axis is the negative log(10) of the p-value, which was determined with a student t-test. All sites above the red line are significant at p=0.05. Green sites change between 1.5% and 5% and blue sites at a magnitude of 5% or greater. B) Bar chart showing the proportion and type of CpG sites that change at least 5% between SPF and GF mice in 3a. Shore indicates the sites are not in CGIs, but within 2000 bp of them. C) Bar chart showing the proportion of total detectable sites that experience methylation change of at least 5%. GF WT to SPF WT (left), GF WT to GF IL10KO (middle), or GF WT to SPF IL10KO (right). D) Bar chart showing the proportion of sites located in CGIs or CGI shores that change under each condition. GF WT to SPF WT (left), GF WT to GF IL10KO (middle), or GF WT to SPF IL10KO (right). E) Venn diagram showing overlap between sites that decrease at least 5% from WT GF mice. P-values for all overlaps are statistically significant at p=0.05 and are listed below. KOGF-WT SPF, p=2.8x10^-28. KO GF-KO SPF, p=7x10^-59. WT SPF-KO SPF, p=6.4x10-241. triple overlap, p=6.6x10^-74. F) Venn diagram showing sites that increase at least 5% from WT GF mice. P-values for all overlaps are statistically significant at p=0.05 and are listed below. P-values for all overlaps are statistically significant at p=0.05 and are listed below. KOGF-WT SPF, p=2.1x10^-8. KO GF-KO SPF, p=6.3x10^-43. WT SPF-KO SPF, p=63.7x10^-121. triple overlap, p=4.5x0^-22.

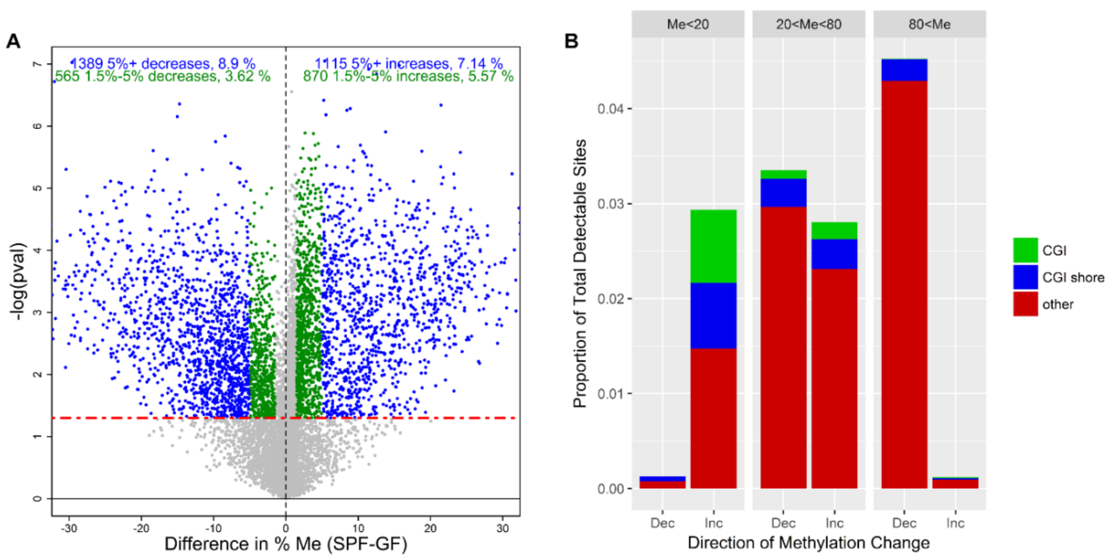


Figure 5: The effect of the microbiome in the presence of inflammation. A) Volcano plot analysis showing methylation differences between IL10KO/GF and IL10KO/SPF mice. The x-axis shows the difference in average methylation between SPF and GF mice for a given site. The y-axis is the negative log(10) of the p-value, which was determined with a student t-test. All sites above the red line are significant at p=0.05. Green sites change between 1.5% and 5% and blue sites at a magnitude of 5% or greater. B) Bar chart showing the proportion and type of CpG sites that change at least 5% between SPF and GF mice in S3a. Shore indicates the sites are not in CGIs, but within 2000 bp of them.

The synergy data compared WT/GF mice to mice that had both IL10KO and a microbiome (ie. 2 hits). We confirmed these data by comparing mice that differed by only 1 hit: GF IL10KO versus SPF/IL10KO (Figure 5a and b) and SPF/WT versus SPF/IL10KO (described later). In both cases, the synergy was obvious with greater than 10% of sites changing in methylation.

To examine further whether the effects of the microbiome and inflammation were skewed towards hypermethylation or hypomethylation, the ratio of proportion of sites increasing versus decreasing sites was examined. Alone, the microbiome caused hypomethylation in 8.9% of sites, and hypermethylation in 1.4% of sites, giving a ratio of 6.4. Inflammation caused hypomethylation in 6.3% of sites, and hypermethylation in 1.8%, giving a ratio of 3.5. The microbiota-inflammation combination caused 10% of sites to hypomethylate and 8% of sites to hypermethylate, giving a ratio of 1.2. Taken together, these data indicate that individually, the effects of microbiota or inflammation are skewed towards hypomethylation. In contrast, when both conditions are present, the ratio of hypermethylation to hypomethylation is roughly equal. This corroborates our data indicating the CGI sites are hypermethylated during the simultaneous introduction of inflammation and a microbiota.

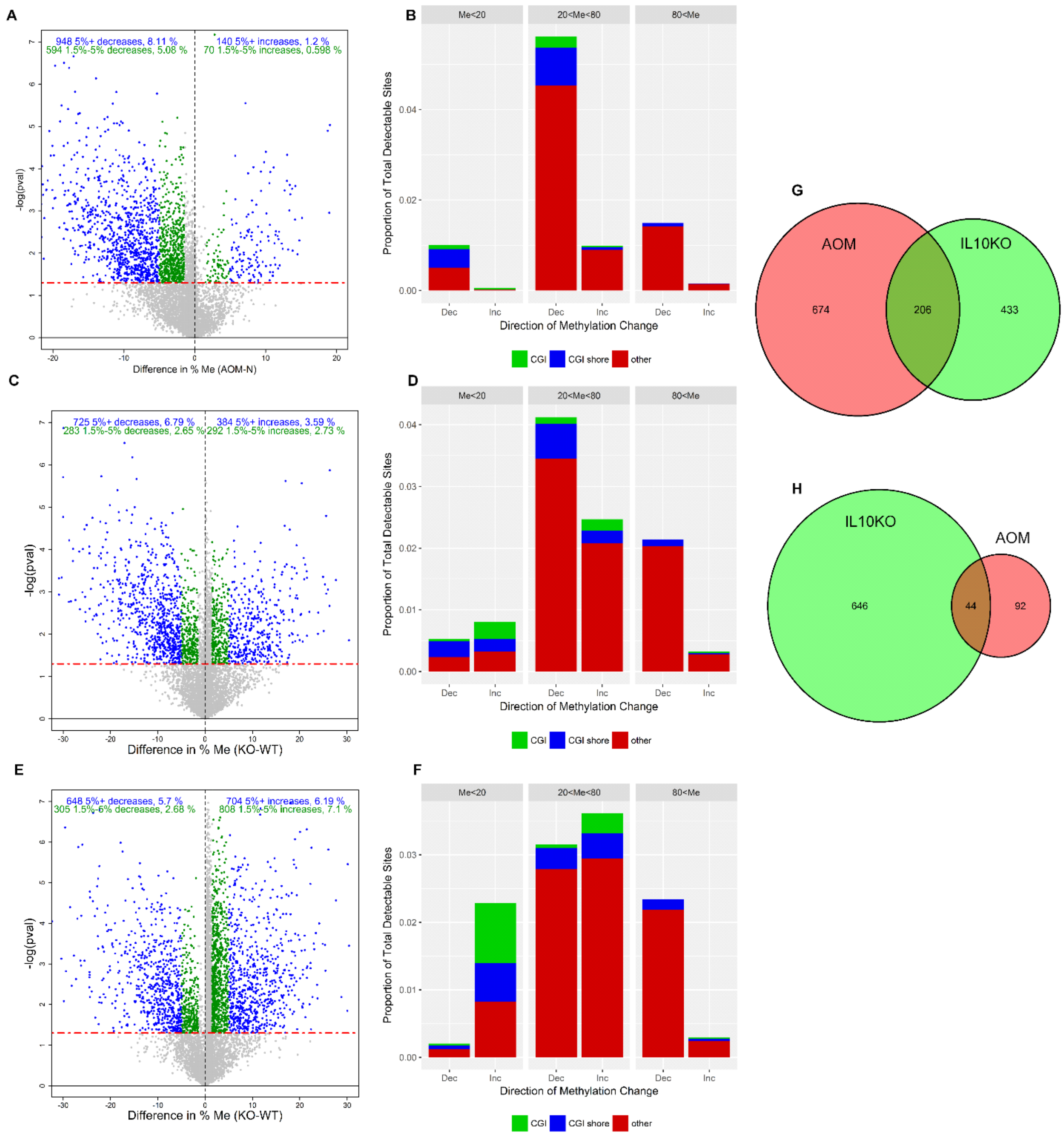


Figure 6: AOM is a potent hypomethylating inflammatory agent. A) Volcano plot analysis showing methylation differences between WT/SPF and WT/SPF+AOM mice. The x-axis shows the difference in average methylation between normal and AOM mice for a given site. The y-axis is the negative log(10) of the p-value, which was determined with a student t-test. All sites above the red line are significant at p=0.05. Green sites change between 1.5% and 5% and blue sites at a magnitude of 5% or greater. B) Bar chart showing the proportion and type of CpG sites that change at least 5% between SPF and GF mice in 4a. Shore indicates the sites are not in CGIs, but within 2000 bp of them. C) Volcano plot analysis showing methylation differences between WT/SPF and KO/SPF mice. D) Bar chart showing the proportion and type of CpG sites that change at least 5% between SPF and GF mice in 4c. E) Volcano plot analysis showing methylation differences between WT/SPF and IL10KO/SPF+AOM mice. F) Bar chart showing the proportion and type of CpG sites that change at least 5% between SPF and GF mice in 4e. G) Venn diagram showing overlap between sites that increase after treatment with AOM and sites that increase at least 5% after IL10KO in WT SPF mice, out of 11255 total sites. P=8.6x10^-15. H) Venn diagram showing overlap between sites that increase after treatment with AOM and sites that decrease at least 5% after IL10KO in WT SPF mice, out of 11255 total sites, p=2x10^-39.

3.4: AOM is a potent hypomethylating inflammatory agent

Azoxymethane is an agent commonly used to induce inflammation and tumorigenesis in mouse models of colorectal cancer. We sought to determine if the AOM-induced inflammation and IL10 deficiency-induced inflammation differed in their effects on DNA methylation profiles. In order to understand the effects of AOM and IL10KO on DNA methylation, we examined the effects of AOM alone, the IL10KO alone, and both in combination in the presence of a microbiome. A volcano plot analysis showed that AOM has a hypomethylating effect on DNA methylation (Figure 6a). Approximately 8% of sites decreased in average methylation as a result of AOM, as opposed to only 0.6% of sites increasing in average methylation. It is interesting to note that in both sites with 80%+ average methylation and 20% or less average methylation, AOM caused overall decreases in average methylation (Figure 6b). In addition, AOM does not seem to sites in target CGI or CGI shores and is most likely to cause a decrease in methylation of sites with medium levels of baseline methylation (20%<Me<80%, Figure 4b). By contrast, IL10KO in SPF mice led to 5.7% of sites increasing in methylation and 6.2% of sites decreasing in methylation (Figure 6c). About 2% of sites were located in CGIs or CGI shores that increased in methylation (Figure 6d). Thus, AOM and IL10 deficiency both induce hypomethylation but, in the presence of a microbiome, only IL10 deficiency` induces substantial CGI hypermethylation, pointing to potentially different mechanisms for the inflammation-methylation link. Indeed, there was relatively little overlap between sites changed by IL10KO or AOM (Figures 6g and 6h). Finally, we also examined mice exposed to both AOM and IL10KO (Figures 6e and 6f) in these mice; there seems to be antagonism in between the two exposures, leading to fewer hypermethylated sites, possibly due to the dominant hypomethylating effect of AOM.

In contrast to the IL10KO, the effect of AOM does not seem to be not dependent on the microbiome; Figures 7a shows a volcano plot analysis of the effects of AOM+IL10KO on GF mice (comparison of WT/GF and KO/GF+AOM mice. This volcano plot is similar to Figure 6a (the effect of AOM alone in the presence of the microbiome).

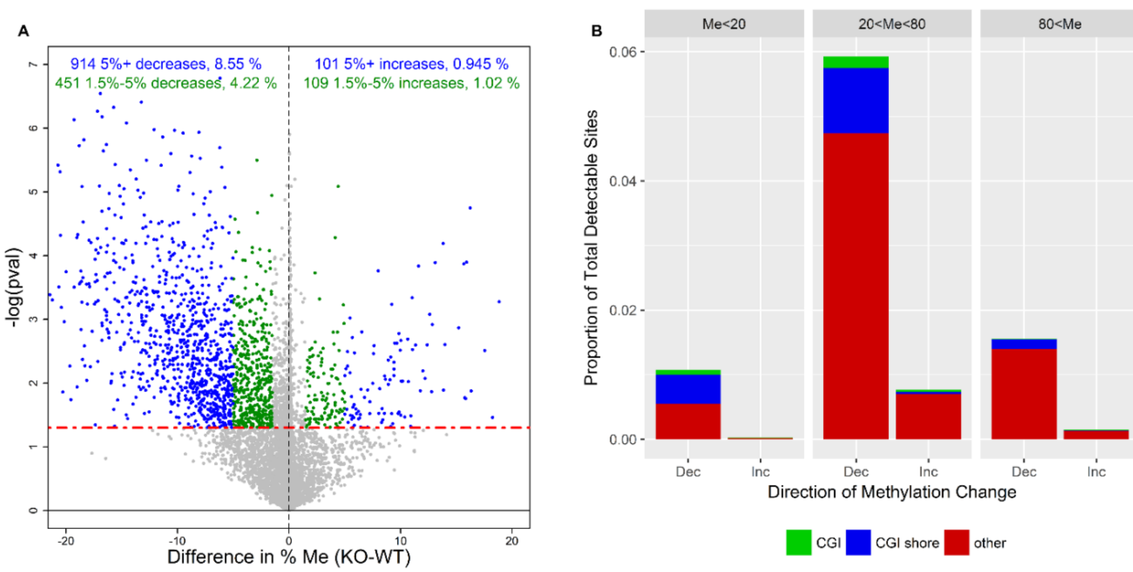
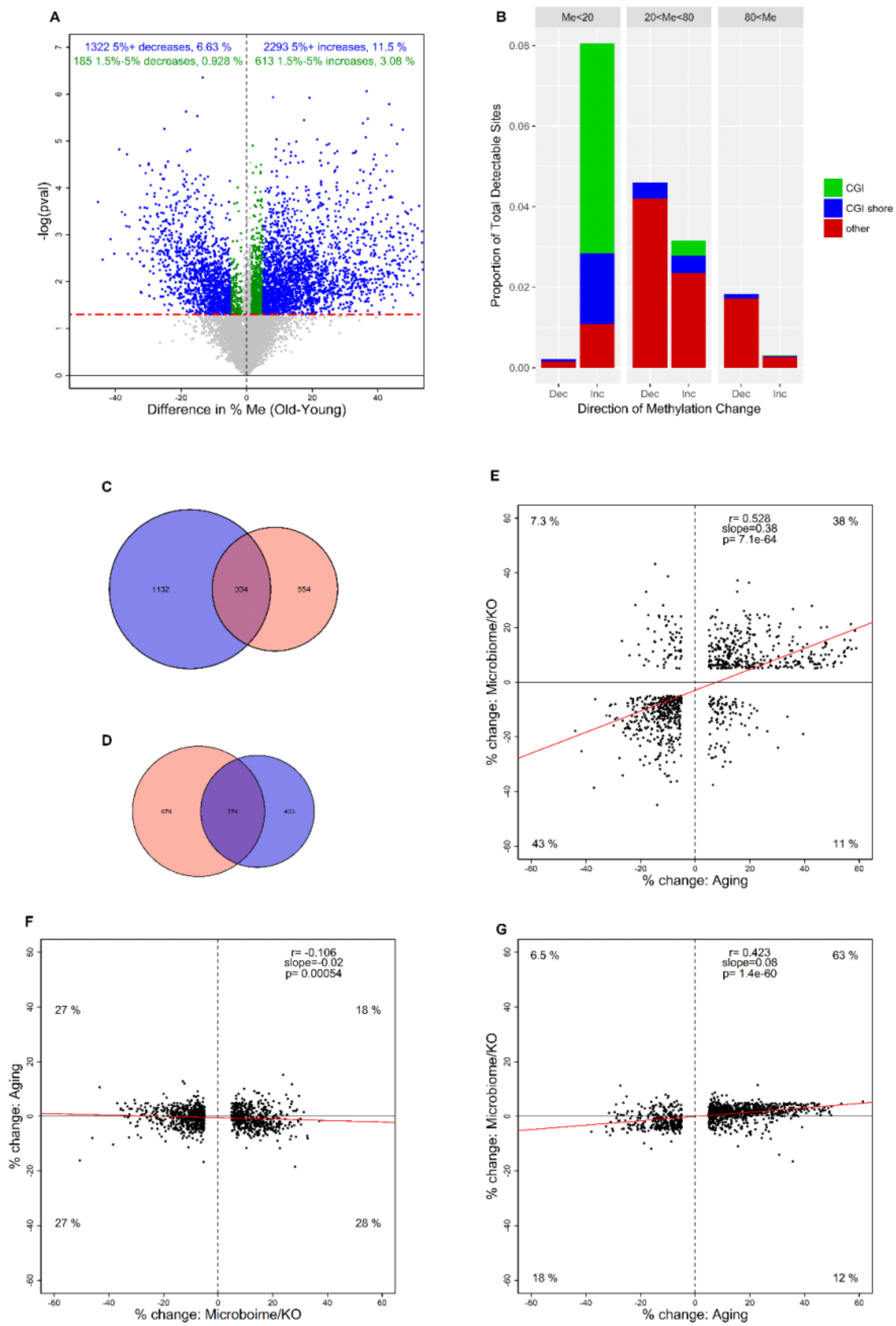


Figure 7: The effect of AOM and IL10KO in GF mice. A) Volcano plot analysis showing methylation differences between WT/GF and IL10KO/GF+AOM mice. The x-axis shows the difference in average methylation between WT and KO+AOM mice for a given site. The y-axis is the negative log(10) of the p-value, which was determined with a student t-test. All sites above the red line change are significant at p=0.05. Green sites change between 1.5% and 5% and blue sites at a magnitude of 5% or greater. B) Bar chart showing the proportion and type of CpG sites that change at least 5% between SPF and GF mice in S4a. Shore indicates the sites are not in CGIs, but within 2000 bp of them.

3.5: The microbiome and inflammation affect the same CpG sites subject to age-related methylation drift

It has been well documented that methylation drift occurs as a result of aging; this drift is characterized by a global loss of DNA methylation while sites located in CGI promoters become hypermethylated[7]. To compare this to microbiome effects, we generated a volcano plot of DREAM data for old (aged 30 months) and young (aged 4-5 months) mice (Figure 8a). Of 19,934 detectable sites, 7.3% decreased in methylation, and 12.6% increased. Next, we determined the distribution of affected sites located in CGIs, CGI shores or other sites; we found that hypermethylated CGIs account for 7% of all detectable sites (Figure 8b).

We then analyzed the overlap between sites that change in methylation due to aging and sites that change due to the synergistic effect of the microbiome and inflammation. To accomplish this, we generated Venn diagrams comparing DREAM data from young and old mice to DREAM data from WT GF and SPF IL10KO mice. Of 11,041 total sites analyzed, 1466 sites increased in average methylation due to aging and 888 sites increased due to the combination in inflammation and the microbiome. There was an overlap of 334 sites (Figure 8c, p=1.7x10^-25). There were 777 sites that decreased in average methylation due to aging, and 1050 that decreased due to the combination of inflammation and the microbiome. It is interesting to note that this is the only situation in which the magnitude of the effect of the microbiome/inflammation was stronger than that of aging. There was a 374-site overlap (Figure 8d, p=1x10^-76). This corroborates previous data demonstrating partial overlap between inflammation-related and age-related methylation[36], and implicates the microbiome in this process.



A
1322 5%+ decreases, 6.63 %
2293 5%+ increases, 11.5 %
185 1.5%-5% decreases, 0.928 %
613 1.5%-5% increases, 3.08 %
-log(pval)
Difference in % Me (Old-Young)
B
Me<20
20<Me<80
80<Me
Proportion of Total Detectable Sites
Direction of Methylation Change
Dec
Inc
CGI
CGI shore
other
C
1132
334
554
D
E
7.3 %
r= 0.528
slope=0.38
p= 7.1e-64
38 %
43 %
11 %
% change: Microbiome/KO
% change: Aging
F
r= -0.106
slope=-0.02
p= 0.00054
27 %
18 %
27 %
28 %
% change: Aging
% change: Microboime/KO
G
6.5 %
r= 0.423
slope=0.08
p= 1.4e-60
63 %
18 %
12 %
% change: Microbiome/KO
% change: Aging

Figure 8: The microbiome and inflammation modify the same CpG sites subject to aging-related methylation drift. A) Volcano plot analysis showing methylation differences between young and old mice. The x-axis shows the difference in average methylation between young and old mice for a given site. The y-axis is the negative log(10) of the p-value, which was determined with a student t-test. All sites above the red line are significant at p=0.05. Green sites change between 1.5% and 5% and blue sites at a magnitude of 5% or greater. B) Bar chart showing the proportion and type of CpG sites that change at least 5% between SPF and GF mice in 5a. Shore indicates the sites are not in CGIs, but within 2000 bp of them. C) Overlap between sites that that increase at least 5% during aging and from IL10KO+microbiome, 11041 total sites, p=1.7x10^-25. D)Overlap between sites that that decrease at least 5% during aging and from IL10KO+microbiome, 11041 total sites, p=1x10^-76. E) Scatterplot of average change with age (x-axis) to average change by microbiome and inflammation (y-axis). Showing all sites that change at least 5% with age and microbiome/inflammation. Percentages in the corners indicate the percent of sites that fall within each quadrant. Red line is best fit. Pearson R=0.528, p= 7.1x10^-64, slope=0.38. F) Scatterplot of average change with microbiome/inflammation (x-axis) to average change with age (y-axis). Showing sites that change with only microbiome/inflammation. R=-0.106, p=0.00054 slope=-0.02) G) Scatterplot of average change with only age (x-axis) to average change by microbiome and inflammation (y-axis). Showing sites that change at least 5% with only microbiome/inflammation, R=0.423, p=1.4x10^-60, slope=0.08.

We have been using a threshold of 5% change in average methylation to consider a site affected by inflammation, the microbiome, or aging, and Figures 8c and 8d only display the overlap of affected sites at this magnitude. To investigate aging-microbiome/IL10KO overlap in sites that may change at a lower magnitude, we split sites into several groups, and generated scatterplots comparing a site's average percent methylation change with age, to change caused by the microbiome-inflammation combination. Sites were divided into the following groups: sites that change in both aging and inflammation aging (threshold of 5%), sites that change in only age, and sites that change only with the microbiome-inflammation combination. All sites (figure 9a) and sites that do not change (Figure 9b) were also investigated as controls. Figures 8e shows all sites that change at the 5% threshold in both aging and with the microbiome-inflammation combination, and plots the average change with age on the x-axis and average change with a microbiome/inflammation on the y-axis. A slope of 0.38 and a Pearson correlation coefficient of r=0.53 (p=7.1x10^-64) is observed. Interestingly, when looking at sites changed in only aging (Figure 8g), the correlation coefficient is also positive and significant (Figure 5f, r=0.42, p=1.4x10^-60). However, the slope is much lower (0.08). This was not seen for sites that change only after exposure to the microbiome and inflammation (Figure 8f, r=-0.11, p=0.0005); while this is statistically significant, the slope was near zero (slope=-0.02).

Taken together, these data indicate that nearly all sites that change with aging are affecting by the microbiome/inflammation combination, but there are many sites that are affected by the microbiome but not by aging. We characterized these unique sites separately and found that they are primarily non-CGIs (Figure 9c), especially when compared to sites that change in both aging and the microbiome/inflammation (Figure 9d). There is significant overlap in the CpG sites whose methylation is altered at least 5% with age or by simultaneous microbiota-inflammation addition (Figures 8c and 8d). Additionally, of sites that change only with age (Figure 8f), 63% are located in the top right quadrant of the scatterplot, indicating that the average change with the IL10KO/microbiome combination and with age are greater than zero, with a correlation coefficient of 0.42. This indicates that there are sites that change with age that are vulnerable to slight changes by the microbiome, which were not detectable in the Venn diagram analysis. Interestingly, the converse is not seen; there are not sites that change greatly with the microbiome that change slightly with age. This may indicate that the presence of a microbiome and inflammation may accelerate the effects of aging, but the effects of the microbiome and inflammation are not enhanced in aged mice.

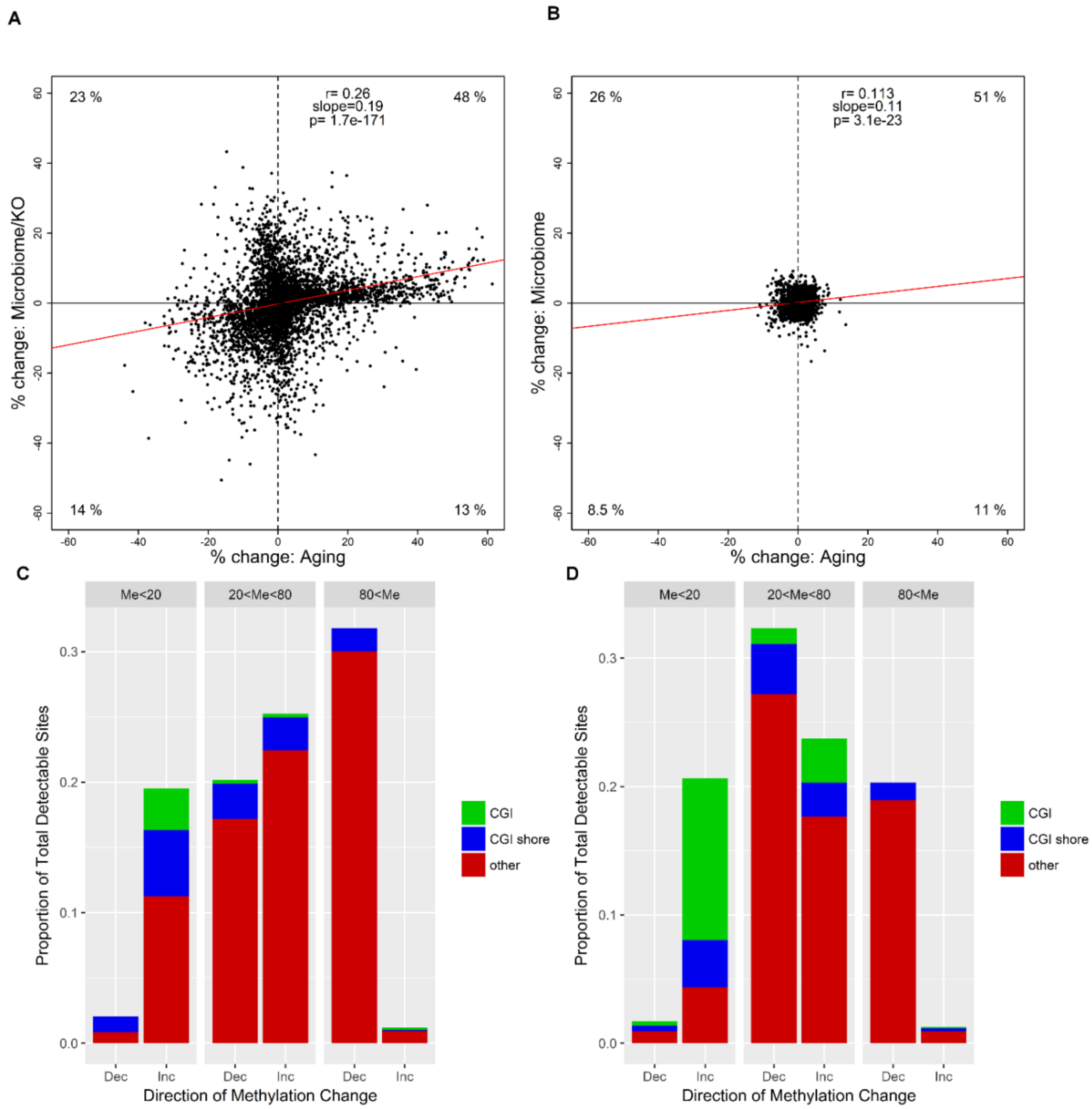


Figure 9: The microbiome and inflammation modify the same CpG sites subject to aging-related methylation drift. A) Scatterplot of average change with age (x-axis) to average change by the microbiota/inflammation (y-axis). Showing all sites. Percentages in the corners indicate the percent of sites that fall within each quadrant. Red line is best fit. Pearson R=0.261, p=1.7x10^-171, slope=0.19. B) Scatterplot of average change with age (x-axis) to average change by the microbiota/inflammation (y-axis). Showing sites that do not change with either the microbiome/inflammation combination or age. R=0.113, p=3x10^-23, slope=0.11). C) Bar plot showing distribution types of sites that change in with only the microbiome and inflammation, and not aging. D) Bar plot showing distribution types of sites that change in with both the microbiome and inflammation, and aging

CHAPTER 4

CONCLUSIONS AND DISCUSSION

4.1: Conclusions and Discussion

Our results demonstrate that, individually, a microbiome or inflammation via IL10KO cause moderate changes in DNA methylation profiles, with little or no changes in CGI methylation. In combination, these factors were able to cause more sites to change, and in greater magnitude. These changes were characterized by general global hypomethylation and CGI hypermethylation, changes that mimic aberrant methylation patterns observed in aging and cancer.

The microbiome alters DNA methylation

We showed that even in the absence of inflammation, the microbiota induces changes in DNA methylation, with approximately 5% of detectable sites changing in average methylation. Additionally, a previous study conducted in humans found that different microbial gut phyla was associated with different DNA methylation profiles[27]. However, the mechanism by which the microbiome-inflammation axis alters DNA methylation remains to be elucidated. One possibility is via production of metabolites, which are intimately linked to epigenetics; metabolites act as both cofactors and inhibitors for many epigenetic enzymes.

There are multiple possible mechanisms by which microbial metabolites may influence the host epigenome, including modulation of DNA methylation. The microbial metabolites acetate, propionate and butyrate are capable of inducing mass reprogramming of histone methylation and acetylation states[26]. These changes are characterized by increases in modifications associated with euchromatin (histone hyperacetylation). Interestingly, there is also a modest increase in H3K27 trimeythylation, a mark typically associated with gene repression and a heterochromatin state. Bacterial species producing these short-chain fatty acid metabolites are known to be decreased in cancer and inflammation. DNMTs have been shown to interact with modifications on histone tails[5], and as histone states change, DNA methylation patterns may as well. This mechanism could contribute to changing DNA methylation profiles as the microbiota composition shifts during inflammation. The resulting aberrant DNA methylation can then contribute to tumorigenesis. The initial inflammation that drives the microbiome composition change could be initiated by any trigger, including infection by a pathogen, further implicating microbes in this process. Additionally, the gut bacteria may cause alterations in host gene expression that affect the ability of host cells to uptake certain metabolites; for example it has been demonstrated that *E. coli* is capable of downregulating a protein responsible for butyrate uptake[55].

Aberrant DNA methylation may also be directly initiated by overproduction of oncogenic metabolites (oncometabolites). Loss of function mutations in citric acid cycle enzymes fumarate hydratase (FH) and succinate dehydrogenase (SDH) result in accumulation of fumarate and succinate respectively[23,44]. A gain of function mutation in isocitrate dehydrogenase (IDH) causes the enzyme to produce 2-hydroxyglutarate instead of α-ketoglutarate[23,44]. Each of these mutations are found in certain types of CIMP positive cancer[44], though they are not observed in colorectal cancer; FH mutations occur in hereditary leiomyomatosis and renal cell cancer (HLRCC), SDH mutations occur in paraganglioma (PGL) and pheochromocytoma (PCC) and IDH mutations are noted in gliomas, glioblastomas and acute myeloid leukemia (AML)[44]. 2HG binds to and inhibits α-ketoglutarate dependent TET and methyltransferase enzymes, causing DNA hypermethylation and tumorigenesis[44]. Succinate and fumarate have been shown to inhibit these enzymes in culture. In addition, there is overlap between the methylated genes associated with each metabolite, implying all three may act via this same mechanism[44]. Thus, bacteria may cause aberrant methylation directly via over-production of metabolites that inhibit TET enzymes. Additionally, the gut microbiota may over-produce precursors or derivatives of these compounds, driving host (or even other bacterial) metabolic reactions towards the production of the offending metabolites.

Finally, some mouse models have demonstrated bacteria may cause tumorigenesis via the induction of DNA damage. It has been suggested that DNA damage itself could lead to aberrant methylation[57]. GF IL10KO+AOM mice were mono-associated with *E. coli*, and aggressive colitis and tumorigenesis resulted[32]. Interestingly, when mice were mono-associated with another inflammation-inducing bacterial species, *Enterococcus faecalis*, they developed severe colitis but little to no tumorigenesis. Levels of colitis-associated inflammatory cytokines between the two mouse groups did not differ, indicating that the inflammation that resulted from the introduction of bacteria was not sufficient for tumor development. It was discovered that the *pks* pathogenicity island, which produces the compound colibactin and is found in certain strains of *E. coli* induces DNA damage leading to tumorigenesis. GF mice mono-associated with *E. coli* that lacked *pks* developed inflammation but did not develop tumors. AOM was also found to be required for the progression of severe dysplasia and tumorigenesis, as IL10KO mice associated with *pks*-containing *E. coli* but not AOM, developed only mild dysplasia, and no tumors. In addition, inflammation via IL10KO was found to be required for tumorigenesis; WT/GF mice mono-associated with *pks*-containing *E. coli* did not develop tumors, even when treated with AOM. This study demonstrates that while pathogenic bacteria and inflammation are both required for initiation of tumorigenesis; this microbe-inflammation is mirrored in this research project. Tumors were not caused by inflammation, but by DNA-damage induced by bacterial production of colibactin; however, inflammation seemed to "activate" the tumorigenic effects of *pks.*

Inflammation alters the DNA methylome

Previous research indicates that inflammation alters the DNA methylome[2,36,38], though the mechanism by which this occurs is unknown. The inflamed gut may exert different selective pressures on host intestinal epithelial cells than the non-inflamed gut. It is possible that the gut microbiota is constantly causing aberrant methylation change in host cells, but in the inflamed gut, the affected cells are selected for, and can thrive and divide. It has also been proposed that inflammation increases cell turnover rate, and inflammation-induced methylation may simply be a reflection of an increased rate of age-related methylation; however, we found the microbiome and inflammation induced methylation changes that were not observed in aging, and this mechanism cannot explain all of the inflammation-induced methylation change. It is also possible that direct interactions with inflammatory cytokines alter a cell's methylation profile[58]. In addition, the pathogenic bacteria *H. pylori* has been shown to cause aberrant DNA methylation via induction of an inflammatory response; the resulting methylation was not due to *H. pylori* itself[38].

We also demonstrated small-scale changes in DNA methylation profiles in germ-free mice by the interleukin-10 KO (IL10KO). GF mice are characterized by underdeveloped immune systems, though inflammatory responses within GF mice are known to occur[20]. Despite this, moderate methylation differences between WT/GF and IL10KO/GF mice were observed; these differences may be due to differences in WT and IL10KO immune function.

It is interesting to note that the inflammatory agent, AOM, and IL10-deficiency by knock-out both induce inflammation, but have different effects on DNA methylation patterns. We showed that AOM seems to cause primarily hypomethylation of sites, regardless of their baseline methylation levels; even sites with less than 20% methylation at baseline were further hypomethylated. This contrasts the effect of inflammation via IL10KO, which caused both hypermethylation at unmethylated sites, hypomethylation at highly methylated sites, and changes in both directions at intermediately methylated sites, though these changes were skewed towards hypomethylation. Furthermore, when comparing WT mice to IL10KO mice treated with AOM, middling amounts of hypo and hypermethylation were observed. This indicates that IL10KO and AOM may create opposing changes on DNA methylation, but more evidence is required to confirm whether that is the case. In addition, it is likely that the IL10KO and AOM induce methylation changes via different mechanisms; AOM does not appear dependent on the microbiome, whereas the IL10KO requires the presence of the microbiome to induce methylation changes.

Microbiome inflammation synergy

This research project also shows that the presence of both inflammation and the microbiome was required for the observed large-scale DNA methylation changes. It is known that inflammation leads to microbial dysbiosis[32]; the WT/SPF, IL10KO/SPF, and mice treated with AOM are highly likely to have different microbial compositions. The change in the environmental pressures of the gut in different mice shifts microbial community makeup, and may alter bacterial gene expression[59]. In this way, inflammatory status may lead to increased production of aberrant-methylation inducing metabolites, and/or decreased production of protective metabolites. Microbes also stimulate an inflammatory response, increasing production of cytokines and inducing colitis in the IL10KO model[32]. Additionally, diet is a known link between these three factors, correlating with microbial composition, metabolite production, and host inflammatory status[60].

Multiple Mechanisms of Microbiota-Inflammation combination: acceleration of age related methylation drift and effects on other sites

The existence of age-related methylation drift has been well documented[7]. Aberrant methylation is known to occur faster in tissue with high rates of cell turnover, indicating that replication errors in the maintenance of the DNA methylome may play a role in this process[2]. This research project demonstrated that age related methylation was characterized by strong hypermethylation of CGIs (4% of detectable CGI or CGI shore sites changed, accounting for 10% of all methylation changes observed, Figures 5b, 5c, 5d). Hypomethylation of thousands of sites were also observed, and our observations corroborate previous studies. GF mice are known to live longer than their SPF counterparts[61]; though a study on age-related methylation drift in GF mice has yet to be performed, it could be hypothesized that GF mice experience a slower rate of drift. We demonstrated that all age-related methylation changes are vulnerable to modulation by the microbiome-inflammation combination. As inflammation has been shown to accelerate age-related methylation in conventionally housed mice (thus, mice with a microbiome)[37], it is likely that the microbiota/inflammation combination accelerates age-related methylation drift. This is especially likely given the large degree of overlap between microbiota/inflammation associated and age associated methylation.

However, this may not be the sole mechanism by which the microbiome and inflammation may accelerate age-related methylation drift. As an organism ages, microbial dysbiosis is known to occur, correlating microbiota composition with age-related methylation drift. Additionally low-grade systemic inflammation[62], including an increase in circulating inflammatory cytokines, is observed in the aged intestine. The aged large intestine is also characterized by decreased integrity of the mucosal barrier in both rodent and non-human primate models. This decrease in barrier function has been associated with cell senescence in baboons, and is thought to be enhanced by TNF-α production[63]. Another recent study in mice demonstrated that age-associated inflammation and microbial dysbiosis drove the increased permeability in the intestinal barrier[62]. GF mice were protected from systemic inflammation, but developed it in response to cohousing with old, but not young, mice. Additionally, mice with “old” microbiota and heightened cytokine levels showed a diminished ability to kill pathogenic bacteria. Though there were some differences, both studies supported inflammatory cytokines as a mediator of decreased intestinal permeability. Additionally, IL10 decreases intestinal permeability[63], which may indicate that our IL10KO mice had high intestinal permeability, which allowed the microbiome to induce such large changes in DNA methylation. Butyrate has also been thought to decrease intestinal permeability, making the loss of butyrate producing microbes in inflammation even more consequential.

However, neither study can explain why age-associated microbial dysbiosis occurs. Interestingly, TNF KO mice in the aforementioned study did not undergo microbial dysbiosis with age[62], further indicating that inflammatory cytokines and inflammation may be responsible for age-associated microbial dysbiosis. It may be that over the course of an individual's life, they are exposed to more environmental inflammatory triggers, increasing the likelihood that inflammation will develop, causing microbial dysbiosis, which then causes increased intestinal permeability; once triggered, these three factors may even act in a positive feedback loop, with each one exacerbating the others. Taken together, this research may indicate that as an organism ages, the intestinal epithelium becomes increasingly vulnerable to invasion by pathogenic bacteria: decreased mucosal barrier function allows microbes to invade the epithelium, further triggering inflammation (and cell turnover) and exposing the host to microbial metabolites. However, it is important to note that intestinal permeability is increased due to inflammation, and that this phenomenon may be exacerbated in aged mice, but is not exclusive to them.

We demonstrated that all age-related methylation changes are vulnerable to modulation by the microbiome-inflammation combination, but the reverse was not true; there was a large subset of sites that changed only in the microbiome/inflamed mice, but not in aging. The large subset of non-age-related sites may indicate that inflammation and the microbiome cause methylation changes via multiple mechanisms. The sites that experience methylation changes from both microbiome/inflammation combination and age-related methylation drift are likely to be due to cell turnover, whereas the sites that change only in inflammation/microbiome combination may be due to cytokine and metabolite interaction with the intestinal epithelium.

Overall conclusions and model

Taken together, this research indicates that the microbiome results in moderate DNA methylation changes in the absence of inflammation. Inflammation and the microbiome synergize to create much broader effects on the host methylome, and it is only when both are present that CGI and CGI shore sites are affected. When age-related methylation drift was compared to the microbiome/inflammation induced methylation changes, it was found that all sites that change with age are vulnerable to a microbiota and inflammation, indicating that the microbiome and inflammation accelerate the epigenetic effects of aging. The converse was not true; there was a subset of sites changed only by a microbiome and inflammation. These different subsets of sites indicate the multiple mechanisms of action, which likely include, but may not be limited to: inflammation-induced increased rate of cell turnover, host cell exposure to inflammatory cytokines (the production of which may be triggered by microbes), and bacterial metabolites interacting with host epigenetic machinery. Figure 10 shows a model of these interactions.

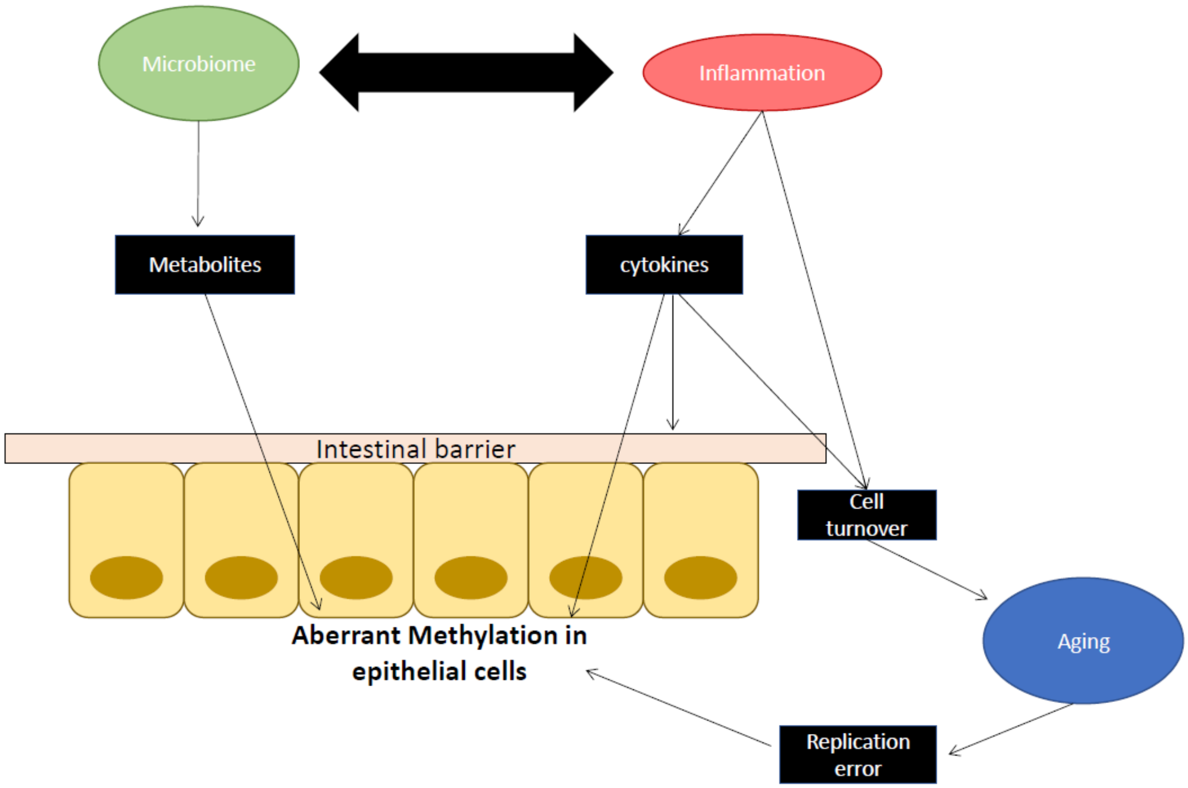


Figure 10: Model of interactions between the microbiome, inflammation, aging and aberrant methylation in the large intestine. The microbiome is able to cause inflammation, and in return, inflammation is able to shift microbiota composition. Inflammation produces inflammatory cytokines that increase intestinal permeability; this allows invasive microbes and microbial products to pass through the barrier and impact host cells. Cytokines may also directly promote methylation changes. Inflammation also increases cell turnover, which increases the rate of age related methylation drift. Errors in maintenance of methylome during cell division increases aberrant methylation. Finally, microbes can influence DNA methylation via production of metabolites, which interact with host epigenetic enzymes.

REFERENCES

1. Jones, P. A., Issa, J.-P. J. & Baylin, S. Targeting the cancer epigenome for therapy. *Nat. Rev. Genet.* **17,** 630–641 (2016).

2. Issa, J. Aging and Epigeneric Drift: a Vicious Cycle. *J. Clin. Invest.* **124,** 24–8 (2014).

3. Good, C. R. *et al.* A novel isoform of TET1 that lacks a CXXC domain is overexpressed in cancer. *Nucleic Acids Res.* **45,** 8269–8281 (2017).

4. Jelinek, J. *et al.* Conserved DNA methylation patterns in healthy blood cells and extensive changes in leukemia measured by a new quantitative technique. *Epigenetics* **7,** 1368–1378 (2012).

5. Baubec, T. & Schübeler, D. Genomic patterns and context specific interpretation of DNA methylation. *Curr. Opin. Genet. Dev.* **25,** 85–92 (2014).

6. Moore, L. D., Le, T. & Fan, G. DNA Methylation and Its Basic Function. *Neuropsychopharmacology* **38,** 23–38 (2013).

7. Maegawa, S. *et al.* Widespread and tissue specific age-related DNA methylation changes in mice. 332–340 (2010). doi:10.1101/gr.096826.109.332

8. Tahara, T. *et al.* Colorectal carcinomas with CpG island methylator phenotype 1 frequently contain mutations in chromatin regulators. *Gastroenterology* **146,** 530–538.e5 (2014).

9. Talens, R. P. *et al.* Epigenetic variation during the adult lifespan: Cross-sectional and longitudinal data on monozygotic twin pairs. *Aging Cell* **11,** 694–703 (2012).

10. Rubinstein, M. R. *et al.* Fusobacterium nucleatum Promotes Colorectal Carcinogenesis by Modulating E-Cadherin/??-Catenin Signaling via its FadA Adhesin. *Cell Host Microbe* **14,** 195–206 (2013).

11. Kostic, A. D., Xavier, R. J. & Gevers, D. The microbiome in inflammatory bowel disease: Current status and the future ahead. *Gastroenterology* **146,** 1489–1499 (2014).

12. Huttenhower, C. *et al.* Structure, function and diversity of the healthy human microbiome. *Nature* **486,** 207–214 (2012).

13. Le Chatelier, E. *et al.* Richness of human gut microbiome correlates with metabolic markers. *Nature* **500,** 541–546 (2013).

14. Louis, P., Hold, G. L. & Flint, H. J. The gut microbiota, bacterial metabolites and colorectal cancer. *Nat. Rev. Microbiol.* **12,** 661–672 (2014).

15. Arumugam, M. *et al.* Enterotypes of the human gut microbiome. *Nature* **474,** 666–666 (2011).

16. Bonder, M. J. *et al.* The effect of host genetics on the gut microbiome. *Nat. Genet.* **48,** 1407–1412 (2016).

17. Sears, C. L. & Garrett, W. S. Microbes, microbiota, and colon cancer. *Cell Host Microbe* **15,** 317–328 (2014).

18. Tomkovich, S. & Jobin, C. Microbiota and host immune responses: A love-hate relationship. *Immunology* **147,** 1–10 (2016).

19. Jain, N. & Walker, W. A. Diet and host–microbial crosstalk in postnatal intestinal

immune homeostasis. *Nat. Rev. Gastroenterol. Hepatol.* **12,** 14–25 (2014).

20. Olszak, T. *et al.* Microbial Exposure During Early Life Has Persistent Effects on Natural Killer T Cell Function. *Science (80-.).* **336,** 489–493 (2012).

21. Schwitalla, S. *et al.* Intestinal tumorigenesis initiated by dedifferentiation and acquisition of stem-cell-like properties. *Cell* **152,** 25–38 (2013).

22. Donohoe, D. R. *et al.* A gnotobiotic mouse model demonstrates that dietary fiber protects against colorectal tumorigenesis in a microbiota- and butyrate-dependent manner. *Cancer Discov.* **4,** 1387–1397 (2014).

23. Feinberg, A. P., Koldobskiy, M. A. & Göndör, A. Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. *Nat. Rev. Genet.* **17,** 284–299 (2016).

24. Arpaia, N. *et al.* Metabolites produced by commensal bacteria promote peripheral regulatory T-cell generation. *Nature* **504,** 451–455 (2013).

25. Zheng, L. *et al.* Microbial-Derived Butyrate Promotes Epithelial Barrier Function through IL-10 Receptor–Dependent Repression of Claudin-2. *J. Immunol.* ji1700105 (2017). doi:10.4049/jimmunol.1700105

26. Krautkramer, K. A. *et al.* Diet-Microbiota Interactions Mediate Global Epigenetic Programming in Multiple Host Tissues. *Mol. Cell* **64,** 982–992 (2016).

27. Kuma;H *et al.* Gut Microbiota as an Epigenetic Regulator : Pilot Study Based on. *MBio* **5,** 1–4 (2014).

28. Aleksandrova, K., Romero-Mosquera, B. & Hernandez, V. Diet, Gut Microbiome

and Epigenetics: Emerging Links with Inflammatory Bowel Diseases and Prospects for Management and Prevention. *Nutrients* **9,** 962 (2017).

29. Kostic, A. D., Xavier, R. J. & Gevers, D. The microbiome in inflammatory bowel disease: Current status and the future ahead. *Gastroenterology* **146,** 1489–1499 (2014).

30. Ohkusa, T. *et al.* Fusobacterium varium. *Gut* **52,** 79–83 (2003).

31. Rubinstein, M. R. *et al.* Fusobacterium nucleatum Promotes Colorectal Carcinogenesis by Modulating E-Cadherin/??-Catenin Signaling via its FadA Adhesin. *Cell Host Microbe* **14,** 195–206 (2013).

32. Arthur, J. C. *et al.* Intestinal Inflammation Targets Cancer-Inducing Activity of the Microbiota. *Science (80-.).* **338,** 120–123 (2012).

33. Ho, J. T., Chan, G. C. & Li, J. C. Systemic effects of gut microbiota and its relationship with disease and modulation. *BMC Immunol.* **16,** 21 (2015).

34. Ridaura, V. K. *et al.* Gut Microbiota from Twins Discordant for Obesity Modulate Metabolism in Mice. *Science (80-.).* **341,** 1241214–1241214 (2013).

35. Bäckhed, F., Manchester, J. K., Semenkovich, C. F. & Gordon, J. I. Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. *Proc. Natl. Acad. Sci.* **104,** 979–984 (2007).

36. Hahn, M. A. *et al.* Methylation of polycomb target genes in intestinal cancer is mediated by inflammation. *Cancer Res.* **68,** 10280–10289 (2008).

37. Issa, J. J., Ahuja, N. & Toyota, M. Accelerated Age-related CpG Island

Methylation in Ulcerative Colitis Accelerated Age-related CpG Island Methylation in Ulcerative Colitis 1. 3573–3577 (2001).

38. Niwa, T. *et al.* Inflammatory processes triggered by Helicobacter pylori infection cause aberrant DNA methylation in gastric epithelial cells. *Cancer Res.* **70,** 1430–1440 (2010).

39. Issa, J. P. DNA methylation as a clinical marker in oncology. *J. Clin. Oncol.* **30,** 2566–2568 (2012).

40. De Zhu, J. The altered DNA methylation pattern and its implications in liver cancer. *Cell Res.* **15,** 272–280 (2005).

41. Langevin, S. M., Kratzke, R. A. & Kelsey, K. T. Epigenetics of lung cancer. *Transl. Res.* **165,** 74–90 (2015).

42. Tahara, T. *et al.* Examination of whole blood DNA methylation as a potential risk marker for gastric cancer. *Cancer Prev. Res.* **6,** 1093–1100 (2013).

43. Toyota, M. *et al.* CpG island methylator phenotype in colorectal cancer. *Med. Sci.* **96,** 8681–8686 (1999).

44. Yang, M., Soga, T. & Pollard, P. J. Oncometabolites: Linking altered metabolism with cancer. *J. Clin. Invest.* **123,** 3652–3658 (2013).

45. Siegel, R. L. *et al.* Colorectal Cancer Statistics , 2017. **67,** 177–193 (2017).

46. Lao, V. V. & Grady, W. M. Epigenetics and colorectal cancer. *Nat. Rev. Gastroenterol. Hepatol.* **8,** 686–700 (2011).

47. Dienstmann, R. *et al.* Consensus molecular subtypes and the evolution of precision

medicine in colorectal cancer. *Nat. Rev. Cancer* **17,** 79–92 (2017).

48. Zeki, S. S., Graham, T. A. & Wright, N. A. Stem cells and their implications for colorectal cancer. *Nat. Rev. Gastroenterol. Hepatol.* **8,** 90–100 (2011).

49. Tahara, T. *et al.* Fusobacterium in colonic flora and molecular features of colorectal carcinoma. *Cancer Res.* **74,** 1311–1318 (2014).

50. Suehiro, Y. *et al.* Epigenetic-genetic interactions in the APC/WNT, RAS/RAF, and P53 pathways in colorectal carcinoma. *Clin. Cancer Res.* **14,** 2560–2569 (2008).

51. Toyota, M., Issa, J. P. & U, J. J. I. CpG island methylator phenotypes in aging and cancer. *Semin Cancer Biol* **9,** 349–357 (1999).

52. Wu, N. *et al.* Dysbiosis Signature of Fecal Microbiota in Colorectal Cancer Patients. *Microb. Ecol.* **66,** 462–470 (2013).

53. Maekita, T. *et al.* High levels of aberrant DNA methylation in Helicobacter pylori-infected gastric mucosae and its possible association with gastric cancer risk. *Clin. Cancer Res.* **12,** 989–995 (2006).

54. Chang, M. S. CpG island methylation status in gastric carcinoma with and without infection of Epstein-Barr virus. *Clin. Cancer Res.* **12,** 2995–3002 (2006).

55. Kumar, A., Alrefai, W. A., Borthakur, A. & Dudeja, P. K. *Lactobacillus acidophilus* counteracts enteropathogenic *E. coli* -induced inhibition of butyrate uptake in intestinal epithelial cells. *Am. J. Physiol. - Gastrointest. Liver Physiol.* **309,** G602–G607 (2015).

56. Kostic, A. D. *et al.* Fusobacterium nucleatum Potentiates Intestinal Tumorigenesis and Modulates the Tumor-Immune Microenvironment. *Cell Host Microbe* **14,** 207–215 (2013).

57. Russo, G. *et al.* DNA damage and Repair Modify DNA methylation and Chromatin Domain of the Targeted Locus: Mechanism of allele methylation polymorphism. *Sci. Rep.* **6,** 33222 (2016).

58. Gasche, J. A., Hoffmann, J., Boland, C. R. & Goel, A. Interleukin-6 promotes tumorigenesis by altering DNA methylation in oral cancer cells. *Int. J. Cancer* **129,** 1053–1063 (2011).

59. Arthur, J. C. *et al.* Microbial genomic analysis reveals the essential role of inflammation in bacteria-induced colorectal cancer. *Nat. Commun.* **5,** 4724 (2014).

60. Vipperla, K. & O'Keefe, S. J. Diet, microbiota, and dysbiosis: a 'recipe' for colorectal cancer. *Food Funct.* **7,** 1731–1740 (2016).

61. Tazume, S. *et al.* ment . Both groups of food-restricted mice showed slower weight gain for 16 weeks and then remained at this lower weight throughout , gaining significantly less body weight than their full-fed counterparts (p < 0 . 01). Organ weights : Liver , kidneys ,. **40,** 517–522 (1991).

62. Thevaranjan, N. *et al.* Age-Associated Microbial Dysbiosis Promotes Intestinal Permeability, Systemic Inflammation, and Macrophage Dysfunction. *Cell Host Microbe* **21,** 455–466.e4 (2017).

63. Tran, L. & Greenwood-Van Meerveld, B. Age-associated remodeling of the

intestinal epithelial barrier. *Journals Gerontol. - Ser. A Biol. Sci. Med. Sci.* **68,** 1045–1056 (2013).