Impact of Dietary Bioactive Components in Chronic Diseases Using Nutrigenomics, Nutriepigenomics, and Metagenomics Approaches

Bijaya Prasad Upadhyaya

Follow this and additional works at: http://openprairie.sdstate.edu/etd

Part of the Medicine and Health Sciences Commons

Recommended Citation
IMPACT OF DIETARY BIOACTIVE COMPONENTS IN CHRONIC DISEASES USING NUTRIGENOMICS, NUTRIEPIGENOMICS, AND METAGENOMICS APPROACHES

BY

BIJAYA PRASAD UPADHYAYA

A dissertation submitted in partial fulfillment of the requirements for the

Doctor of Philosophy

Major in Nutrition, Exercise and Food Sciences

South Dakota State University

2016
IMPACT OF DIETARY BIOACTIVE COMPONENTS IN CHRONIC DISEASES USING NUTRIGENOMICS, NUTRIEPIGENOMICS, AND METAGENOMICS APPROACHES

This dissertation is approved as a creditable and independent investigation by a candidate for the Doctor of Philosophy degree and is acceptable for meeting the dissertation requirements for this degree. Acceptance of this dissertation does not imply that the conclusions reached by the candidate are necessarily the conclusions of the major department.

Moul Dey, Ph.D. 
Dissertation Advisor

Matthew Vukovich, Ph.D. 
Head, Department of Health and Nutritional Sciences

Dean, Graduate School
ACKNOWLEDGEMENTS

I express my sincere gratitude and deep appreciation to my advisor Dr. Moul Dey for her continuous support throughout my Ph.D. study and the research. Her knowledge, patience, and motivation have immensely helped me in all aspects of my academic and professional life. I am always grateful to her giving me the opportunity to work in multiple projects that has tremendously helped me widen my research from various perspectives.

In addition to my advisor, I am thankful to other members of my Ph.D. committee: Dr. Matthew Vukovich (Head, Department of Health and Nutritional Sciences), Dr. Padmanaban Krishnan, Dr. Feng Li, and Dr. William Gibbons for their insightful comments, encouragement and support in all directions. I am also thankful to Dr. Chunyang Wang (Associate Dean- Education and Human Sciences) for his critical evaluation during my Ph.D. final oral exam.

I am also grateful to my research collaborators from South Dakota State University: Dr. Bonny Specker, Dr. Lacey McCormack, Dr. Jeffrey Clapper, Dr. David Knudsen as well as Dr. Michelle Baack (Sanford Research, Sioux Falls, SD) and Dr. Ali Reza Fardin-Kia (US Food and Drug Administration, College Park, MD) for their timely support and research. Their specific contributions have been acknowledged throughout the text.

My sincere thanks goes to my lab members: Dr. Dan Wang and Dr. Sailendra Nichenametla (former postdocs), Yi Liu, Robert Juenemann, Thamer Aljutaily, Sumadhuri Pamarthi, Jonathan Kahler, and Rania Mabrook (visiting scholar) for their stimulating discussions inside and outside the lab meetings. It is my pleasure to have that
opportunity to work as a team for last 4 years. In particular, I am thankful to Shivon Barbari, an undergraduate student-worker, for his exceptional help in data analyses.

I also acknowledge MGP Ingredients Inc. (Atchison KS) for research resources, and Second Genomics and Epigentek for their paid sequencing services in an efficient manner.

Finally, I wholeheartedly thank my friends and family: my parents, Baikuntha Prasad Upadhyaya and Bhawani Upadhyaya; my wife, Sushma Dhungana; and my son, Sion Upadhyaya for their never ending love, sacrifice and encouragement. Without their support, I could not have completed my Ph.D. study and research.
# TABLE OF CONTENTS

ABBREVIATIONS ........................................................................................................ viii

LIST OF FIGURES ........................................................................................................ x

LIST OF TABLES ........................................................................................................ xii

ABSTRACT .................................................................................................................... xiii

CHAPTER 1 . INTRODUCTION AND DISSERTATION ORGANIZATION ........ 1

1.1. Nutrigenomics and Nutriepigenomics in Health................................................. 1

1.2. Gut Microbiota and Metabolic Diseases............................................................... 2

1.3. Fiber diet, resistant starch and Metabolic Syndrome......................................... 3

1.4. Short-Chain Fatty Acids, Gut Epithelial Integrity and Mucosal Immunity ........ 5

1.5. High-Fat Diet, Gestational Diabetes and Fetal Programming ......................... 6

1.6. Phenethyl Isothiocyanate, Cancer-Stem Cells and Cervical Cancer .............. 7

1.7. Dissertation Organization ................................................................................... 9

CHAPTER 2 . IMPACT OF DIETARY RESISTANT STARCH TYPE 4 ON
HUMAN GUT MICROBIOTA AND IMMUNOMETABOLIC FUNCTIONS .... 13

2.1. Abstract ............................................................................................................... 13

2.2. Introduction ....................................................................................................... 14

2.3. Materials and Methods .................................................................................... 15

2.4. Results .............................................................................................................. 24

2.5. Discussion ....................................................................................................... 34

2.6. Conclusion ....................................................................................................... 37

2.7. Future Directions ............................................................................................ 37
CHAPTER 3. DIETARY RESISTANT STARCH TYPE 4-ASSOCIATED BUTYRATE MODULATES LYSINE 27 TRIMETHYLATION OF HISTONE 3 IN THE PROMOTER OF NUCLEAR FACTOR-KAPPA-B1 IN INTESTINAL EPITHELIUM .................................................. 51

3.1. Abstract .................................................................................................................. 51
3.2. Introduction ............................................................................................................ 52
3.3. Materials and Methods ....................................................................................... 54
3.4 Results ................................................................................................................... 61
3.5. Discussion ............................................................................................................ 65
3.6. Conclusions .......................................................................................................... 68
3.7. Future Directions ............................................................................................... 69

CHAPTER 4. MATERNAL HIGH-FAT DIET AND GESTATIONAL DIABETES AFFECT HISTONE MODIFICATIONS ON CARDIO-METABOLIC GENE PROMOTERS IN OFFSPRING ................................................................. 76

4.1. Abstract ............................................................................................................... 76
4.2. Introduction .......................................................................................................... 77
4.3 Methods ................................................................................................................. 79
4.4. Results .................................................................................................................. 83
4.5. Discussion ............................................................................................................ 89
4.6. Conclusion .......................................................................................................... 94
4.7. Future Directions ............................................................................................... 95
CHAPTER 5. PHENETHYL ISOTHIOCYANATE INHIBITS ALDEHYDE DEHYDROGENASE ACTIVITY AND TARGETS REACTIVE OXYGEN SPECIES-MEDIATED APOPTOSIS IN HUMAN CANCER STEM CELLS ..... 108

5.1. Abstract ................................................................................................................................. 108
5.2. Introduction ............................................................................................................................. 109
5.3. Materials and Methods ........................................................................................................... 112
5.4. Results ................................................................................................................................. 118
5.5. Discussion ............................................................................................................................. 121
5.6. Conclusion ............................................................................................................................ 124
5.7. Future Directions .................................................................................................................. 124

CONCLUSIONS .......................................................................................................................... 131

REFERENCES ............................................................................................................................. 132
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCL2:</td>
<td>Chemokine (C-C motif) ligand 2</td>
</tr>
<tr>
<td>CF:</td>
<td>Control Flour</td>
</tr>
<tr>
<td>ChIP:</td>
<td>Chromatin Immunoprecipitation</td>
</tr>
<tr>
<td>CRC:</td>
<td>Colorectal cancer</td>
</tr>
<tr>
<td>CSC:</td>
<td>Cancer stem cell</td>
</tr>
<tr>
<td>CVD:</td>
<td>Cardiovascular disease</td>
</tr>
<tr>
<td>DMEM:</td>
<td>Dulbecco’s modified eagle’s medium</td>
</tr>
<tr>
<td>DMSO:</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>FBS:</td>
<td>Fetal bovine serum</td>
</tr>
<tr>
<td>GCMS:</td>
<td>Gas chromatography-mass spectrometry</td>
</tr>
<tr>
<td>GDM:</td>
<td>Gestational diabetes</td>
</tr>
<tr>
<td>GSH:</td>
<td>Glutathione</td>
</tr>
<tr>
<td>H3Ac:</td>
<td>Histone 3 acetylated</td>
</tr>
<tr>
<td>H3K4me3:</td>
<td>Histone 3 trimethylated at lysine 4</td>
</tr>
<tr>
<td>H3K27me3:</td>
<td>Histone 3 trimethylated at lysine 27</td>
</tr>
<tr>
<td>HDACs:</td>
<td>Histone deacetylases</td>
</tr>
<tr>
<td>HDL:</td>
<td>High-density lipoprotein</td>
</tr>
<tr>
<td>HF:</td>
<td>High fat</td>
</tr>
<tr>
<td>IFNγ:</td>
<td>Interferon-γ</td>
</tr>
<tr>
<td>IL-:</td>
<td>Interleukin-</td>
</tr>
<tr>
<td>ITCs:</td>
<td>Isothiocyanates</td>
</tr>
<tr>
<td>LDL:</td>
<td>Low-density lipoprotein</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>LPS:</td>
<td>Lipo-polysaccharide</td>
</tr>
<tr>
<td>MetS:</td>
<td>Metabolic syndrome</td>
</tr>
<tr>
<td>NaBu:</td>
<td>Sodium butyrate</td>
</tr>
<tr>
<td>NFκB:</td>
<td>Nuclear factor-kappaB</td>
</tr>
<tr>
<td>PEITC:</td>
<td>Phenethyl isothiocyanate</td>
</tr>
<tr>
<td>PBS:</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>ROS:</td>
<td>Reactive oxygen species</td>
</tr>
<tr>
<td>RS4:</td>
<td>Resistant starch type 4</td>
</tr>
<tr>
<td>SCFAs:</td>
<td>Short chain fatty acids</td>
</tr>
<tr>
<td>SEM:</td>
<td>Standard error of mean</td>
</tr>
<tr>
<td>SFN:</td>
<td>Sulforaphane</td>
</tr>
<tr>
<td>STAT:</td>
<td>Signal transducer and activator of transcription</td>
</tr>
<tr>
<td>STZ:</td>
<td>Streptozotocin</td>
</tr>
<tr>
<td>T2DM:</td>
<td>Type 2 diabetes</td>
</tr>
<tr>
<td>TC:</td>
<td>Total cholesterol</td>
</tr>
<tr>
<td>TLRs:</td>
<td>Toll-like receptors</td>
</tr>
<tr>
<td>TJPs:</td>
<td>Tight-junction proteins</td>
</tr>
<tr>
<td>TNF-α:</td>
<td>Tumor necrosis factor-α</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1.1</td>
<td>Nutrigenomics</td>
</tr>
<tr>
<td>1.2</td>
<td>A dynamic extrinsic and intrinsic factors affecting gut epithelial permeability.</td>
</tr>
<tr>
<td>2.1</td>
<td>Study design</td>
</tr>
<tr>
<td>2.2</td>
<td>Phylum-level gut microbial composition of participants at baseline</td>
</tr>
<tr>
<td>2.3</td>
<td>Separation of the microbiome post intervention in RS4 and CF groups</td>
</tr>
<tr>
<td>2.4</td>
<td>Differential gut microbial composition after RS4 intervention at the species level</td>
</tr>
<tr>
<td>2.5</td>
<td>Representative chromatograms showing the overlay of retention times of butyl esters of SCFA fragment from standard and biological sample.</td>
</tr>
<tr>
<td>2.6</td>
<td>Effects of RS4 on fecal SCFA</td>
</tr>
<tr>
<td>2.7</td>
<td>Associations between gut microbiota and host biological parameters after RS4 and CF interventions</td>
</tr>
<tr>
<td>2.8</td>
<td>Intra-associations within bacterial species that were correlated with metabolic functions or SCFAs in an RS4-specific manner</td>
</tr>
<tr>
<td>2.9</td>
<td>Associations among host biological parameters in the RS4 group</td>
</tr>
<tr>
<td>3.1</td>
<td>Effects of NaBu on inflammatory gene expression in human intestinal epithelial cells</td>
</tr>
<tr>
<td>3.2</td>
<td>Concentration-dependent effects of NaBu on NFκB1 in human intestinal epithelial cells.</td>
</tr>
<tr>
<td>3.3</td>
<td>Effects of RS4 on feed intake and body weight in KK.cg- a+/a mice</td>
</tr>
</tbody>
</table>
Figure 3.4. Effects of RS4 on mice cecal weight, butyric acid levels and histone modification .......................... 73

Figure 4.1: Study timeline ....................................................................................................................... 96

Figure 4.2: Physiological characteristics pups and their dams under study ............................. 96

Figure 4.3: Peaks detection in ChIP sequencing ................................................................. 97

Figure 4.4: Differential histone modification profiles ............................................................. 98

Figure 4.5. Peaks overlapping any 5kb upstream region of known exons ...................... 100

Figure 5.1. Sphere culture enriches ALDH$^\text{hi}$CD44$^\text{hi}$ hCSCs population from parental HeLa cells ................................................................. 125

Figure 5.2. PEITC inhibits ALDH activity in HeLa and hCSC .............................................. 126

Figure 5.3. PEITC induces ROS in hCSCs ................................................................................. 127

Figure 5.4: PEITC inhibits proliferation of hCSCs and induces apoptosis ...................... 128

Figure 5.5. PEITC inhibits transcriptional factor Sp1 and multidrug resistance protein ABCB1 in hCSC ................................................................. 129
LIST OF TABLES

Table 1.1. Types of resistant starch and their food sources .............................................. 12

Table 2.1. Baseline characteristics of 20 participants ....................................................... 48

Table 2.2. Estimated nutrients intake at baseline and at the end of intervention periods 49

Table 2.3. Means of biological parameters at baseline and at the end of intervention periods .................................................................................................................. 50

Table 3.1. Primer sequences used for qRT PCR and ChIP assays ....................................... 74

Table 3.2. Composition of experimental diets per 100 grams ............................................. 75

Table 4.1. Diet-specific differential binding of peaks in respective histone modification in genomic and 5kb upstream region .................................................................................. 101

Table 4.2. List of “metabolic process” candidate genes with differentially bound peaks .......................................................................................................................... 102

Table 4.3. List of “metabolic process” candidate genes (high fat diet) with gene ontology enrichment .................................................................................................................. 104

Table 4.4. Metabolic process candidate qtls obtained from the candidate genes (high fat diet specific) .................................................................................................................. 106

Table 5.1. Number of tumors in treatment groups in NOD/SCID mouse model .......... 130
ABSTRACT

IMPACT OF DIETARY BIOACTIVE COMPONENTS IN CHRONIC DISEASES USING NUTRIGENOMICS, NUTRIEPIGENOMICS, AND METAGENOMICS APPROACHES

BIJAYA PRASAD UPADHYAYA

2016

Chronic diseases that include seven of the top ten leading causes of premature deaths in the United States demand more than eighty percent of all health care cost. Although typical Western diets--high fat (HF), low fiber and low vegetables-- are frequently associated with increased risk for chronic diseases like obesity, metabolic syndrome (MetS), type 2 diabetes, cardiovascular diseases (CVD), and cancer, there is a lack of mechanistic understanding of a causal relationship between diet and disease phenotypes, particularly in a clinical setting. This dissertation research aimed to address this research gap through the effects of wheat fiber-derived resistant starch type 4 (RS4) in MetS, maternal HF diet on offspring cardiovascular health, and cruciferous vegetables-derived phenethyl isothiocyanate (PEITC) in cancer prevention, using next generation “omics” technologies and molecular nutrition approaches.

RS4 flour (30%) enriched differential abundance of 71 bacterial operational taxonomic units in fecal samples obtained from Hutterite participants having MetS. RS4 not only showed butyrogenic effects in humans but also increased cecal mass and butyrate concentration in mice, signifying the role of dietary fiber in bowel health.
Butyrate induced tri-methylated lysine 27 on histone 3 (H3K27me3) in the promoter of nuclear factor-kappa-B1 (NFκB1) contributing to epigenetic repression of intestinal inflammation in mice. A genome-wide differential chromatin landscape of histone modifications (H3Ac, H3K4me3, and H3K27me3) was observed in offspring due to their maternal HF (40% kcal) diet, implicating the increased risks for obesity, diabetes, and CVDs beyond one generation. PEITC inhibited cancer stem cell (CSC) marker aldehyde dehydrogenase 1 (ALDH1) with concomitant induction of oxidative stress and apoptosis in cervical CSCs, conferring improved therapeutic outcomes in cancer patients.

This dissertation presents a body of translational research work on diet-gene interaction, which is expected to contribute to the fast-growing nutrigenomics knowledgebase and inform about (1) diet-responsive gut microbiota and metabolites, (2) fetal-reprogramming to prevent adult-onset of chronic diseases, and (3) chemo-preventive strategies to improve therapeutic outcomes of cancer by minimizing recurrence and metastasis. Together, these findings will contribute to future dietary recommendations that are expected to be more precise than at current times.

**Keywords:** Butyrate, Cancer stem cells, Clinical research, Gestational diabetes, Gut microbiota, High fat diet, Histone modifications, Metabolic syndrome, Phenethyl isothiocyanate, Preclinical disease model, Resistant starch type 4
CHAPTER 1. INTRODUCTION AND DISSERTATION ORGANIZATION

1.1. Nutrigenomics and Nutriepigenomics in Health

Since the time of Hippocrates (460-370 BC) food is considered a medicine in maintaining health and wellness. The Human Genome Project (1990-2003) followed by various post-genomic research initiatives along with the advent of modern biotechnological tools has opened numerous opportunities to establish the validity of this notion experimentally, and to explain how environmental factors like diet alter the expression and/or structure of an individual’s genetic makeup. This new era of molecular nutrition is unfolding in terms of individual response to nutrients due to own genetic variation (i.e., nutrigenetics), or in terms of the effects of food bioactives on the regulation of gene expression (i.e., nutrigenomics) [1]. Hence, nutrigenomics is a discovery science, which aims to understand how diet-gene interactions influences metabolic and disease pathways, how this regulation is disturbed in the early phase of diet-related diseases, and how dietary interventions prevent the individuals at risk from such diseases [2] (Figure 1.1).

Furthermore, nutriepigenomics deals with the elucidation of how biologically active dietary components modulate the landscape of a gene position in chromatin or chromosome, without changing the DNA sequence, and regulate the turning “on” and “off” the genes epigenetically [3]. However, many common diseases, such as obesity, cancer, diabetes, and cardiovascular diseases are polygenic in nature, which means that they arise from the dysfunction in a cascade of genes. Therefore, dietary intervention to prevent the onset of such diseases is a complex and challenging goal, and requires a thorough understanding of molecular nutrition, genetics, physiology, pathology,
biochemistry, bioinformatics, and wide array of “omics” technologies [4]. Powerful new parallel sequencing technologies have allowed for large scale genomic analysis of complex microbial populations, once unidentifiable due to inability to culture in vitro. These metagenomic approaches allow for both qualitative and quantitative microbial population analysis in both health and diseases.

1.2. Gut Microbiota and Metabolic Diseases

The human gut is colonized by trillions of microorganisms; including hundreds of different species of bacteria, viruses and fungi that outnumber the host cells by 10 fold [5]. These gut microbiota not only directly associates with the food that we eat [6, 7] but also involve in triggering immune cells underlying a single layer of gut epithelium. There exists a mutually beneficial association between the resident gut microbiota and a healthy host [8]. Commensal bacteria co-evolved with their hosts and hence are not pathogenic under normal or healthy conditions. The dominant residents of human gut are Gram-negative strict anaerobes, including Bacteriodes, Bifidobacterium, Fusobacterium, Eubacterium, Peptostreptococcus, and Atopobium [9]. However, under specific conditions, this ecological homeostasis cannot be maintained, leading to pathologic imbalance of gut microbiota, termed dysbiosis. Dysbiosis triggers a multitude of chronic human disease conditions, such as inflammatory bowel diseases (IBD) [10], intestinal cancers [11], obesity [12], type 2 diabetes [13], atherosclerosis [14], and alteration of bone metabolism [15]. Hence, the three-way interaction among the diet, gut microbiota and host play a critical role in shaping intestinal as well as extra-intestinal health and diseases.
The microbiota being an essential component in gut microenvironment plays a critical role in the pathogenesis of intestinal as well as extra-intestinal diseases. As most of the bacteria reside in the human gut in close proximity to intestinal epithelium, their composition and metabolism are directly influenced by the foods ingested. Given the ability to impact on the fine balance between health and disease, it is reasonable to tailor novel strategies to prevent bowel inflammation and associated comorbidities through dietary interventions (Figure 1.2).

1.3. Fiber diet, resistant starch and Metabolic Syndrome

While dietary fiber is associated with a lower risk of obesity and cancer [16], diets high in refined carbohydrates are associated with significant increases in obesity and metabolic syndrome (MetS) [17]. Based on International Diabetes Federation (IDF) criteria, MetS is identified with central obesity (≥102 cm in men and ≥88 cm in women) plus any two of the following four parameters:

1. Raised blood pressure: systolic BP ≥130 mm Hg or diastolic BP ≥85 mm Hg or any medication for hypertension,

2. Raised triglycerides: ≥150 mg/dl (1.7 mmol/l) or history of medication for triglycerides abnormality,

3. Reduced high-density lipoprotein (HDL) cholesterol: < 40 mg/dl (1.03 mmol/l) in males and < 50 mg/dl (1.29 mmol/l) in females or history of medication for HDL abnormality, and

4. Raised fasting blood glucose level: ≥ 100 mg/dl or previously diagnosed with type 2 diabetes mellitus [18].
Considering a typical Western lifestyle with high-fat and high-carbohydrates in diets with less physical activity, the chances of acquiring MetS is even higher. According to American Heart Association and National Health and Nutrition Examination Survey 1999–2010, almost 34% of American adults are affected by MetS that multiply a person's risk for heart disease, diabetes and stroke [19].

Prebiotics are non-digestible and selectively fermentable complex oligosachharides or other ingredients that confer health benefits due to selective changes in composition or activity of the gut microbiota [20]. Digestive fermentation of complex carbohydrates and fibers results in the formation of a number of SCFAs, which positively modulate colonic health through several mechanisms [21].

The major source of carbohydrate in the human diet is starch. There are five types of RS: physically inaccessible RS1 found in partly milled grains and seeds; resistant granules RS2 found in raw potato, green banana, some legumes, and high-amylose starches; retrograded RS3 found in cooked and cooled potato, bread, and cornflakes; and chemically modified RS4 as etherized, esterified, or cross-bonded starches [22] along with a newly reported amylose-lipid complex RS5 (Table 1.1). RS4 resists digestion in the small intestine due to its chemical modification, and passes to the colon where it is subjected to gut microbiota-mediated fermentation [23]. It has been shown that dietary interventions with RS4 attenuated various risk factors for metabolic syndrome (MetS) characterized by several conditions such as insulin-resistance, dyslipidemia, abdominal obesity, and hypertension that ultimately lead to atherosclerosis and cardiovascular disease (CVD) [24, 25]. Diets high in resistant starch (RS) profoundly affect the types of fecal bacteria, including species related to Ruminococcus.
bromii, which can contribute to starch degradation and SCFA production [26]. An increased consumption of RS produces more butyrate, which is associated with higher insulin sensitivity and improved glycemic control [27], suggesting the role of gut microbiota in promoting the host health [24].

RS has certain advantages over other dietary fibers in terms food applications. For instance, RS has a white, tasteless powder-like texture with a small particle size, rendering it more desirable for a variety of bakery products, pasta products, cereals, and snacks, etc. to increase the dietary-fiber contents and improve the quality of the final products [28]. This forms the basis for objective of chapter 2 to investigate how RS modulates gut microbiota and its metabolite SCFAs in individuals with signs of MetS.

1.4. Short-Chain Fatty Acids, Gut Epithelial Integrity and Mucosal Immunity

Gut epithelial cells are tightly packed with paracellular tight-junction proteins (TJPs) to act as a defensive barrier between luminal microbes and underlying immune cells. Only some antigen presenting cells (APCs) are well trained to constantly sample the luminal microbes that distinguish luminal microflora from pathogens, without alarming mucosal immunity [29]. However, in case of pathogenicity, these APCs recognize the pathogens based on their pathogen associated molecular pattern and signal to respective toll-like receptor (TLR) cascades, resulting into inflammation and associated immune response [30]. Therefore, it is most necessary to maintain the gut barrier function to control the chronic low-grade inflammation.

A recent study found a strong association between lower fiber intake, and subsequent lower production of SCFAs with early tumor development in patients at high risk of CRC [31]. Further, butyrate oxidation is the primary energy source for human
colonocytes and is strongly anti-inflammatory by inhibiting production of inflammatory cytokines and inhibition of inflammatory pathways like NFκB activation [21].

SCFA butyrate is the product of bacterial carbohydrates and fiber polysaccharides fermentation in the human gut [32, 33]. Butyrate has both trophic and bioactive functions on host cells. Besides correcting the altered ratio of gut bacteria [34], butyrate can modulate intestinal and colonic inflammation by protecting against the potential pro-inflammatory molecules due to its ability to modulate gene expression [35], which could lead to an anti-inflammatory effect and inhibit the progression of low-grade chronic inflammation. Epigenetically, butyrate has been linked to histone deacetylase (HDAC) inhibiting properties in several cell lines and is considered to be a promising treatment for cancer. Butyrate might mitigate low-grade chronic inflammation by suppressing the transcription of inflammatory genes through H3K27 tri-methylation [36] or acting as an HDAC inhibitor [37] in colon cancer cells. This serves the basis of our objective of chapter 3 that dietary RS4 improves the luminal effects of butyrate on inducing histone modifications (e.g. H3K4Me3) by regulating chemokines (e.g. CCL2) and transcription factors (e.g. NFκB).

1.5. High-Fat Diet, Gestational Diabetes and Fetal Programming

Stimuli such as over nutrition, physical inactivity, and aging can result in chronic low-grade inflammation, leading to insulin resistance and diabetes in genetically or metabolically predisposed individuals [38]. Obesity epidemics is rising world wide [39], and alarmingly 34.4% of US women are obese in their child-bearing stage [40], resulting the higher incidence of gestational diabetes mellitus (GDM). GDM is one of the most
common complications in pregnancy that occurs when high fasting glucose (> 92–125 mg/dL) associated glucose intolerance is first diagnosed at any time of pregnancy, or in between 24 and 28 weeks in particular [41]. Of note, 17.8% (9.3-25.5%) of overall pregnancies in the US are affected by GDM [42].

Over nutrition or maternal high fat (HF) diet during the pregnancy is another critical factor to shape the adult onset of metabolic diseases in the offspring [43, 44]. In GDM, a food plan focused on controlling carbohydrate often adds up more fat in the diet [45] in order to meet the caloric needs of a rapidly growing fetus. Since heart constantly demands a high-energy source to maintain its contractile activity, heart prefers to utilize all fuel sources including fatty acids [46]. Hence, the quantity as well as quality of the dietary fat during pregnancy plays a major role in cardiometabolic health outcomes in offspring above and beyond the pregnant mother [47].

Altogether, the interaction between fetal and maternal environment and modulation of gene expression that starts very early in life and that can pass across generations due to epigenetic modifications such as DNA methylation and covalent posttranslational histone modifications [48]. This forms the rationale of chapter 4 that maternal HF diet affects streptozotocin (STZ)-induced gestational diabetes to alter the metabolic programming in the offspring using a neonatal heart tissues in suitable rat model.

1.6. Phenethyl Isothiocyanate, Cancer-Stem Cells and Cervical Cancer

Cancer statistics shows that cervical cancer is the fourth most common cancer in women, and the seventh overall, with an estimated 528,000 new cases in 2012
worldwide, while it remains the most common cancer in women in Eastern and Middle Africa. Interestingly, 87% of cervical cancer deaths occur only in the less developed regions [49], possibly due to less screening techniques.

Cancer stem cells (CSCs), a subset of cancer cells (0.1-30%) with self-renewal and differentiation potential, have been implicated in tumor initiation, maintenance, metastasis, and recurrence of a number of cancer types [50, 51]. To identify and isolate CSCs from within a heterogeneous population of cancer cells, recent studies have shown that aldehyde dehydrogenase family 1 (ALDH1) is a CSC marker that associates with tumor malignancy and self-renewal properties of stem cells in different tumors, including breast [52], colon [53], and cervical [54, 55] cancers.

PEITC is a dietary bioactive compound derived from common Brassicaceae family cruciferous vegetables, such as watercress, broccoli, cabbage, and cauliflower [56]. PEITC possesses anti-inflammatory [57, 58] and chemopreventive activity against various cancers, including colon [59], breast [60], cervical [61, 62] cancers. PEITC has no apparent toxicity, as revealed from safety studies in rats and dogs, even when administered in high doses determined by NOEL (no-observed-adverse-effect-level) [63]. Remarkably, undergoing clinical trials in the US has elucidated the protective role of PEITC against lung cancer (NCT00691132). This is the basis of our working hypothesis under chapter 5 that PEITC selectively targets chemo-and-radiotherapy resistant cancer initiating CSC) in a suitable mouse model.
1.7. Dissertation Organization

The whole dissertation research has yielded five research papers as described in subsequent chapters. Two of them are published, one is accepted for publication, and the remaining two papers are in preparation. These chapters are preceded by a general “Introduction”, while specific literature review is provided within the specific chapter, followed by a general “Conclusion”. All the figures and tables belonging to the same chapter are placed at the end of each chapter. The references are cited in a numbered fashion as they appear in the text and are compiled altogether at the end of this document.
Figure 1.1. Nutrigenomics

A model emphasizing nutrient-gene interaction responsible for health and disease status (modified from Peter Gillies [2]).
Figure 1.2. A dynamic extrinsic and intrinsic factors affecting gut epithelial permeability.
Epithelial monolayer separates luminal gut microbiota and their metabolites from underlying mucosal immune cells. Tight-junction proteins (TJPs) hold adjacent epithelial cell together to prevent paracellular bacterial translocation into the gut mucosa. The disequilibrium or dysbiosis determines the fate of epithelial cells whether to activate mucosal immune system through microbial pattern recognition receptor called toll-like receptors (TLRs) or to activate enterocyte apoptosis to turnover. Dietary bioactive components like PEITC and dietary fiber like RS4 modulate gut epithelial integrity maintaining the fine balance between health and disease.
Table 1.1. Types of resistant starch and their food sources (modified from Topping and Clifton [22])

<table>
<thead>
<tr>
<th>Types of Resistant Starch</th>
<th>Examples of Occurrence</th>
</tr>
</thead>
<tbody>
<tr>
<td>RS1; physically inaccessible</td>
<td>Partly milled grains and seeds</td>
</tr>
<tr>
<td>RS2; resistant granules</td>
<td>Raw potato, green banana, some legumes, and high-amylose starches</td>
</tr>
<tr>
<td>RS3; retrograded</td>
<td>Cooked and cooled potato, bread, and cornflakes</td>
</tr>
<tr>
<td>RS4; chemically modified</td>
<td>Etherized, esterified, or cross-bonded starches (used in processed foods)</td>
</tr>
<tr>
<td>RS5; amylose-lipid complex</td>
<td>Stearic acid-complexed high-amylose starch</td>
</tr>
</tbody>
</table>

RS: Resistant starch
CHAPTER 2. IMPACT OF DIETARY RESISTANT STARCH TYPE 4 ON
HUMAN GUT MICROBIOTA AND IMMUNOMETABOLIC FUNCTIONS


2.1. Abstract

Dietary modulation of the gut microbiota impacts human health. Here we investigated the hitherto unknown effects of resistant starch type 4 (RS4) enriched diet on gut microbiota composition and short-chain fatty acid (SCFA) concentrations in parallel with host immunometabolic functions in twenty individuals with signs of metabolic syndrome (MetS). Cholesterols, fasting glucose, glycosylated hemoglobin, and proinflammatory markers in the blood as well as waist circumference and % body fat were lower post intervention in the RS4 group compared with the control group. 16S-rRNA gene sequencing revealed a differential abundance of 71 bacterial operational taxonomic units, including the enrichment of three Bacteroides species and one each of Parabacteroides, Oscillospira, Blautia, Ruminococcus, Eubacterium, and Christensenella species in the RS4 group. Gas chromatography–mass spectrometry revealed higher fecal SCFAs, including butyrate, propionate, valerate, isovalerate, and hexanoate after RS4-intake. Bivariate analyses showed RS4-specific associations of gut microbiota with host metabolic functions and SCFA levels. Here we show that dietary RS4 induced changes in the gut microbiota are linked to its biological activity in
individuals with signs of MetS. These findings have potential implications for dietary guidelines in metabolic health management.

**Keywords:** 16S-rRNA gene sequencing, gut microbiota, resistant starch type 4, short chain fatty acids, bowel health

### 2.2. Introduction

Metabolic syndrome (MetS) encompasses co-morbidities like obesity, dyslipidaemia, hypertension, insulin resistance, and hyperglycaemia, which increase the risk of cardiovascular diseases [64], the leading cause of death in the US [65]. Although it is possible to manage these co-morbidities with dietary/lifestyle changes and/or therapeutic interventions, the overall prevalence of MetS is rising worldwide [65, 66]. One contributing factor could be non-adherence to healthy dietary practices beyond the short term, as convenience and taste remain the strongest determinants of food choices [67, 68]. In that context, our recent double-blind study allowed participants to maintain their habitual dietary practices during the intervention without adapting to any change [25]. This was possible due to the neutral and adaptable organoleptic properties of the resistant starch (RS) type 4 (RS4), a wheat-derived food-ingredient, used for the intervention.

Different chemical properties contribute to the functional differences between five RS types, particularly in terms of fermentability and its influence on the microbiota in the gut [69][24]. Fermentable carbohydrates such as RS4, may potentially increase colonic short-chain fatty acid (SCFA) production [70]. However, there remains a dearth of well-controlled intervention studies that have comprehensively examined the influence of RS4
on host physiological, gut microbiome, and SCFA changes [71-73]. One study reported influence of RS4 in healthy individuals, unrelated to any metabolic condition [24]. Other published studies regarding metabolic health benefits of RS4 involved animal models [74-76]. However, due to various metabolic adaptations in MetS patients, they may not always benefit from the information generated in healthy humans or animal models [77]. Taken together, there is a critical need for well-designed studies in individuals with MetS that systematically connects the influence of a functional and adaptable food ingredient on the gut microbial community, bacterial metabolites like SCFAs, and host metabolic functions.

In our trial, the RS4-group had an improved lipid profiles and body composition [25]. Since RS4 is indigestible, we hypothesized that the health benefits of RS4 are derived from its ability to influence the gut microbial community structure, which may, in turn, be linked to altered bacterial fermentation and SCFA production [22]. Therefore, here, in selected participants with MetS, we retrospectively examined the changes in the microbiota composition and the SCFA production in the gut, measured the concentrations of three circulatory adipocytokine markers, and estimated the macronutrient and caloric intake during the intervention period. Also, the host anthropometric and metabolic parameters were reanalysed in this cohort to show microbe-microbe and host-microbe interactions.

2.3. Materials and Methods

Subjects
Participants in this study included a subset of the parent cohort consisting of 20 individuals who originally participated in a dietary intervention with RS4 [25]. Briefly, this study was conducted in an adult population, consisting of both males and females, from two Hutterite colonies in eastern South Dakota, USA. The Hutterites are a culturally homogeneous Caucasian population of Central European ancestry. The parent trial is registered in clinicaltrials.gov (NCT01887964). From 40 participants who had signs of metabolic syndrome at baseline, 26 subjects submitted stool samples at all four data collection time points. Out of 26, 20 subjects (10 from each of the two colonies) were included in the current investigation as their fecal samples were adequate to carry out both sequencing and short chain fatty acid (SCFA) analyses (Figure 2.1). The selected cohort included 12 females and eight males, aged 32-77 (Table 2.1). Exclusion criteria included pregnancy, lactation, long-term antibiotic therapy, immune compromised state, cancer, and other conditions that would affect the ability to provide informed consent or comply with the protocol.

Test Diets

Participants consumed control flour (CF) and resistant starch type 4 (RS4) flours as described in Nichenametla et al. [25]. RS4 flour was made by substituting 30% (v/v) of the CF with RS4 (Fibersym RW, MGP Ingredients Inc., Atchison, KS). The sequence of flour consumption was randomly assigned to the colonies. Dietary intake was assessed by using a 3-day (2 weekdays and 1 weekend day) semi-quantitative food-frequency questionnaire, customized to include foods that Hutterites frequently consume, mostly consisting of items cooked from raw ingredients. Evidence show that the Hutterites
consume more fresh produce in summer and preserved or frozen food in winter resulting into a diet-influenced shifts in microbiome [78]. Since the study period did not overlap with winter months, the impact of seasonal variation on gut microbiome was not considered. Diet analysis was performed using Nutritionist Pro (Axxya Systems, Redmond, WA, USA) to obtain macro-nutrient breakdown of the consumed diets. Diet information was collected using a self-administered semi-quantitative food frequency questionnaire and was analysed using Nutritionist Pro (Axxya Systems, Redmond, WA, USA). Anthropometric, blood lipids, glycaemic parameters, and blood pressure measurement data were obtained from the parent study and reanalysed in this sub-cohort [25].

**Study design**

The original trial [25] was conducted as a two-period, total 26-week long double blind (participants–investigators), placebo-controlled, cluster cross-over intervention (Figure 2.1a). Each intervention period (CF and RS4) was for 12 weeks with a washout time of two weeks in between. All data and bio specimens were collected on-site from both colonies at baseline, 12, 14, and 26 weeks except body composition analyses, which were carried out at baseline, 12 and 26 weeks. A modified CONSORT flow chart is provided (Figure 2.1b). The study was conducted under free-living conditions, and no dietary restrictions were imposed. Gastrointestinal symptoms and stool consistency were scored based on a one-on-one interview with the participants [25]. Current and past medical information were obtained at each visit, which included medication or dietary supplement use, menopausal status, medical conditions, and hospital stays. Seven-day
physical activity-recall questionnaires were administered by trained staff to assess any atypical physical activity.

**Anthropometric measurements**

Height, weight, waist circumference, and blood pressure were measured at the beginning and at the end of each intervention period following standard procedures that were previously described [79]. Briefly, height was determined using a stadiometer and recorded to the nearest 0.5 cm. Body weight was measured on an electronic scale (Seca GmbH & Co., Hamburg, Germany) with the subjects in light clothes and no shoes and was recorded to the nearest 0.1 kg. Waist circumference was measured at the navel using a tension-sensitive, non-stretching Gulick tape and recorded to the nearest 0.5 cm. Blood pressure was measured with a digital sphygmomanometer in a sitting position. Body composition (fat mass and fat-free mass) was determined by total-body, dual-energy, x-ray absorptiometry (DXA) scanning (Hologic QDR Discovery A, Waltham, MA) as described previously [80]. Fat mass is presented as % body fat in paper.

**Blood biochemistry**

Overnight fasting blood was collected by venepuncture in vacutainer tubes (BD Biosciences, Franklin Lakes, NJ). Fasting blood glucose, total cholesterol (TC), Low-Density-Lipoprotein cholesterol (LDL), High-density-Lipoprotein cholesterol (HDL), non-HDL cholesterol (non-HDL), triglycerides were determined using the Cholestech LDX point-of-care analyzer (Alere Inc, Waltham, MA), lipid profile and glucose cassettes (Lipid Profile GLU, Alere Inc, Waltham, MA), following the manufacturers’
instructions. Postprandial glucose was determined 2 h after either breakfast or the noon meal by using a FreeStyle Freedom Lite blood glucose meter (Abbott Diabetes Care Inc, Alameda, CA), following the manufacturer’s instructions. Participants typically consumed a food item made with test flour during the meal prior to the postprandial glucose test. Glycosylated hemoglobin (HbA1C) was determined by an enzyme-based colorimetric assay using the Human HbA1C kit (Crystal Chem, Downers Grove, IL), following the manufacturer’s instructions. Each sample was run in duplicate, and a batch control was used to account for inter-assay variation. HbA1C levels were expressed as the percentage of total hemoglobin. Two markers of inflammation, interleukin 6 (IL6) and tissue necrotic factor-α were determined (in duplicate) in serum using Human ELISA Ready-SET-Go kits, following the manufacturer’s instructions (eBioscience, San Diego, CA). Plasma adiponectin levels, an indicator of metabolic functions, were detected by Human Adiponectin Radioimmunoassay following the manufacturer’s instructions (Linco Research, St. Charles, MO). Inter- and Intra-assay coefficients of variance were 5.0% and 11.7%, respectively.

**Fecal SCFA analysis by gas chromatography-mass spectrometry (GC-MS)**

SCFAs were derivatized to their corresponding butyl esters (SCFABE) followed by GC-MS analyses using a GC-MS 5977A and HP-5MS UI capillary column from the same manufacturer (Agilent, Wilmington, DE, USA). The data were expressed in mg/gm of fecal sample. To prepare fecal samples for SCFA analyses, 800-1000 mg fecal sample was homogenized and added to a glass tube containing 1 ml of internal standard (2-ethylbutyric acid in 1-butanol, 1 mg/mL). Organic solvent hexane and derivatizing agents,
hydrogen chloride 1-butanol or Boron trifluoride 1-butanol (both from Sigma-Aldrich, St. Louis, MO), were then added with vortexing between each addition. Boron trifluoride 1-butanol and hydrogen chloride 1-butanol were used to create butyl-esters for SCFA detection. Samples were then sonicated and placed in water bath (90-100 °C) for 1h. Once cooled to room temperature, water and additional hexane were added to the samples and centrifuged at 3000g for two min. The organic layer (~2 mL) was transferred into a sampling vial with 150 µL inserts, finally adding a pinch (~10 mg) of anhydrous sodium sulphate to remove the residual water. Fecal extracts were stored at -20 °C until further analysis.

GC-MS analyses were performed with an GC-MS 5977A (Agilent Technologies, Wilmington, DE, USA) and an HP-5MS UI capillary column (30 m x 0.25mm, 0.2 µm thickness, Agilent, Wilmington, DE, USA). Hydrogen was used as carrier gas at 1.9 mL/min constant flow and a typical injection volume was 1 µL in the split mode (1:10). The separation of SCFA butyl esters was achieved using an oven temperature program as follow: initial elution temperature was 55 °C for 4 min then 5 °C/min to 120 °C and then 20 °C/min to 220 °C for 10 min. The selective mass detector was operated in the ‘single ion monitoring and scan’ (SIM/Scan) mode. The source was maintained at 150 °C and the electron energy was 70 eV.

Community structure analysis of gut microbes

Stool DNA was extracted using the QIAamp DNA Stool Mini Kit (QIAGEN, Valencia, CA) according to the manufacturer’s instructions. All samples were quantified via the Qubit® Quant-iT dsDNA Broad-Range Kit (Invitrogen, Life Technologies, Grand
Island, NY). Sample DNAs were sent to Second Genome (South San Francisco, CA) for metagenome sequencing, and operational taxonomic unit (OTU) identification. Briefly, samples were enriched for bacterial 16S V4 rDNA region, DNA was amplified utilizing fusion primer designed against the Illumina (San Diego, CA) flow cell adapters and indexing barcodes. Each sample was PCR-amplified with two differently bar coded V4 fusion primers. For each sample, amplified products were concentrated using a solid-phase reversible immobilization method for the purification of PCR products and quantified by electrophoresis using an Agilent 2100 Bioanalyzer® (Agilent Technologies, Santa Clara, CA). Sequencing was carried out using Illumina MiSeq platform following standard protocols for 250 cycles with custom primers designed for paired-end sequencing. Using Quantitative Insights Into Microbial Ecology or QIIME [81] and generated custom scripts (Second Genome), sequences were quality-checked and demultiplexed to determine exact matches to the supplied DNA barcodes. Resulting sequences were then searched against the Greengenes reference database of 16S rRNA gene sequences, clustered at 97% by uclust (closed reference OTU picking). The longest sequence from each OTU was then assigned taxonomic classification via Mothur's Bayesian classifier, trained against the Greengenes database clustered at 98%. For unidentified Greengenes OTUs, we cross referenced with closest hits from NCBI 16S rRNA database with query cover (>90%), identity (>87%), and E value (<0.01). Raw sequences are deposited in NCBI sequence read archive (SRA, accession number SRP035338), belonging to BioProject accession number PRJNA308315.
Statistical analyses

Data analyses were performed using two approaches. Effects of CF and RS4 interventions were determined by comparing end-point data, while effects of RS4 intervention alone was presented by comparing the data before and after RS4 intervention. Linear mixed effects models (SAS MIXED procedure) were used to compare the effects of RS4 and CF on biological parameters. All models included variables for colony and season, where colony was a surrogate for randomization sequence and season was a surrogate for crossover treatment period. Because the crossover design used only two clusters (colonies), and any carryover effect might be confounded by cluster sampling effect (colony effects), thus a carryover effect could not be independently evaluated in the mixed models. Covariates for age, sex, baseline value of the outcome variable, change in total calorie intake were evaluated for their association with the response variables, and covariates that significantly affected at least one outcome variable were retained in the final model as fixed effects. The final models included season, colony, sex, age, and initial baseline value of the response variable as fixed effects.

When appropriate data were presented for either all-participants or stratifying participants based on missing data or outliers in microbial changes as our observations concur with a prior report of large inter-subject variations in response to RS4 [24]. For microbiome analyses one subject was excluded (n=19) because of *Lactobacillus* probiotic supplement use during the RS4 intervention period. For multivariate analysis PhyCA-Stats™ software package (Second Genome Inc, South San Francisco, CA) was utilized. Differences in the bacterial taxa and host phenotypes among time points were determined
by paired t-test or student’s t-test as appropriate. If the data were not normally distributed, values were subjected to logarithmic transformations as indicated.

Two approaches for sampling normalization were employed to account for uneven sequencing depth during data analyses on metagenome sequencing. First, taxa were filtered to those present in at least one of the samples, selecting 55,079 sequences from each sample before calculating community-wide dissimilarity measures. This facilitated calculations of microbial relative abundances (Figures 2.2, 2.4b, 2.6b, 2.7, and 2.8), and principal coordinate analyses (Figure 2.3). Second, the samples were normalized to 1 million counts. Thus the shift in microbial profile between the samples or groups was calculated based on per million sequences, which facilitated calculations of 71 differentially abundant OTUs or species (Figures 2.4a and 2.4c).

We used various R-packages (http://www.r-project.org/) for downstream data analyses on gut microbiome. Package edgeR (http://www.bioconductor.org/packages/release/bioc/html/edgeR.html) [82] which has the distinct benefit of increasing the sensitivity of detecting differentially abundant features, was utilized to control for inter-subject variability because of a model-based normalization to adjust for varying sequencing depths. This approach of identifying differentially abundant taxa between treatments borrows from the RNA sequencing literature testing for differential gene expression. These methods have recently been demonstrated to effectively translate to microbiome analyses [83], offering a robust method with options for paired sample comparisons even for experiments with minimal levels of biological replication. To correct for multiple comparisons, a false discovery rate correction was used to correct for p-values (adjusted p is represented as q). In order
to generate correlation matrices and heat maps, several R-packages, namely Hmisc (http://CRAN.R-project.org/package=Hmisc) [84], corrplot (http://CRAN.R-project.org/package=corrplot) [85], gplots (http://CRAN.R-project.org/package=gplots) [86], and ecodist (https://CRAN.R-project.org/package=ecodist) [87] were used. Significant variation in microbiome baseline values between intervention-groups were ruled out using a Permutational Multivariate Analysis of Variance that also confirmed the effectiveness of the washout period. Two-dimensional Principal Coordinate Analysis (PCoA) followed by Adonis test was used for plotting dissimilarity values so that the similar data points get closer to each other. Statistical analyses were carried out using Sigma Plot software (Systat Software, Inc., San Jose, CA) and the data were presented as means±S.E.M, unless otherwise noted. A $p$ value of 0.05 or less was considered significant, while a $p$ value of 0.05 to 0.09 was considered trend or approaching significance.

2.4. Results

Baseline physiologic, metabolic, and microbiome characteristics of the study participants

All twenty participants who had signs of metabolic syndrome at baseline and submitted adequate stool samples at four data collection time points were included in the current investigation (Figure 2.1), which allowed for comparison of the gut microbial and SCFA profiles before and after the interventions and also between the endpoints of the RS4 and CF (control) interventions. Potential adverse gastrointestinal side effects from the interventions were not evaluated in this cohort since none were observed in the parent
cohort [25]. Baseline characteristics of 20 participants are summarized in Table 2.1.

Taxonomic classification of a total of 55,079 sequences (present in at least one of the samples) were sorted into 5,949 OTUs, of which ~78% were associated with the phylum Firmicutes and ~9% with the phylum Bacteroidetes (Figure 2.2).

**Washout was effective in restoring microbiome characteristics**

Before switching the RS4 and CF diets in the cross-over study design, all the participants were supplied with CF during the 2-week washout period in order to avoid the potential carry-over effects of the RS4 intervention. For endpoint comparison between the RS4 and CF groups, it was necessary to check for a consistent baseline prior to each treatment period. Using permutational multivariate analysis of variance for distance matrices, no significant differences were observed among the starting microbiomes of the RS4 and CF groups (data not shown), which also confirmed that the two-week washout was effective and that any differences observed post-intervention are due to the intervention itself.

**Macronutrient intake pattern did not vary during the study**

Variation in macronutrient intake and total calories consumed can potentially influence the gut microbiota [88], thereby confounding the effects of the intervention. Although a large number of food options are offered at each meal, Hutterites have relatively small interpersonal differences in diet due to common meal planning, kitchen, and dining practices. No significant differences in overall macronutrients and caloric intake were observed between the baseline and post-intervention time periods, with the
exception of dietary fiber (Table 2.2). Dietary fiber intake, analysed separately from total carbohydrate intake, was significantly higher in the RS4 group ($p<0.001$), due to RS4 being classified as a prebiotic dietary fiber (Table 2.2). The average calories (~1,774 Kilocalories) consumed at baseline were estimated to come from carbohydrate (~49%), protein (~17%), and fat (~34%). These values fall within the Dietary Reference Intakes (DRI) for macronutrients, which are 45–65%, 10–35%, and 20–35% for carbohydrate, protein, and fat, respectively [89]. Of particular interest, saturated fat (12.6%, DRI<10%) and cholesterol (415 mg, DRI<300 mg) intakes were higher, while daily fiber intake was lower (18 g at baseline, DRI 20–30 g) than recommended in the participants studied.

**Differential post-intervention effects of the RS4 diet compared with the CF diet on the gut microbiota**

The current understanding is that, in studies without a proper control group, inter-individual variation in gut microbial composition in adults frequently offsets the smaller changes induced by dietary interventions [71]. To address this problem, we compared microbial composition and abundance post RS4 compared with post CF intervention. Three taxa, all unclassified species of Firmicutes, differentially shifted between CF and RS4 treatments in eight male participants. Similarly, a differential effect of RS4 was observed in 16 Firmicutes taxa, with the most numerous genus being *Enterococcus*, which was significantly enriched after CF intake in 12 female participants (data not shown). No distinct trend for Firmicutes to Bacteroidetes ratio was observed in male or female participants (data not shown). The dominance of Firmicutes and Bacteroidetes was consistent with previous results, as reported in Hutterite [78] and other populations.
The Shannon diversity index was not associated with the age of the participants \((r=-0.2, \ p>0.05, \text{ data not shown})\). Likewise, the total diversity of the microbiota assessed from the Shannon diversity index did not significantly change after either CF or RS4 interventions (data not shown).

Principal coordinate analysis showed 26% and 13% variations on axes 1 and 2, respectively, indicating a major shift between the two groups \((p=0.01, \text{ Figure 2.3})\). The RS4 diet differentially modified 71 microbial OTUs \((q<0.05)\), including enrichment of four each of \textit{Ruminococcus} and \textit{Blautia}, two each of \textit{Bacteroides} and \textit{Oscillospira}, and one \textit{Parabacteroides} OTUs (data not shown). Of the 71 OTUs, 65 belonged to the phylum Firmicutes. The three Bacteroidetes OTUs all increased in abundance with RS4 relative to the CF treatment, while OTUs belonging to Firmicutes had a mixed response (data not shown). At the species level, some species were significantly enriched in the RS4 group, including three \textit{Bacteroides} species \((>121.2 \text{ fold, } q<0.05)\) along with \textit{Blautia glucerasea} \((2497.1 \text{ fold, } q<0.001)\), \textit{Christensenella minuta} \((2.4\times10^6 \text{ fold, } q<0.001)\), \textit{Eubacterium oxidoreducens} \((7723.2 \text{ fold, } q<0.01)\), \textit{Oscillospira} spp. \((2528.4 \text{ fold, } q<0.01)\), \textit{Ruminococcus lactaris} \((1.2\times10^5 \text{ fold, } q<0.001)\), and \textit{Parabacteroides distasonis} \((8642.2 \text{ fold, } q<0.001)\), while some were significantly decreased in abundance in this group, including pathogenic \textit{Enterococcus casseliflavus} \((-13603.2 \text{ fold, } q<0.001)\) and \textit{Streptococcus cristatus} \((-229.7 \text{ fold, } q<0.05)\) (Figure 2.4a). Although the enrichment fold changes for some of the bacterial species appear very high, their relative abundance in the whole microbial community could be low. This is due to the commonly used sampling normalization approach based on per million sequences to remove any bias due to varying sequencing depth (details in Methods). Overall, trends showed that
Bacteroidetes OTUs were increased in the RS4 group, leading to an overall lowering of the average Firmicutes-to-Bacteroidetes (F:B) ratio in the RS4 group from 14.6 at baseline to 12.9, but increasing to 19.2 post CF diet (Figure 2.4b). The lower F:B ratio is frequently perceived as an indicator of a leaner phenotype, although the previously reported results are not always consistent [91]. Firmicutes and Bacteroidetes are two major phyla and the species composition within each may vary widely in a given subject. It is possible that both phyla include species that may be characteristic of a particular phenotype. Therefore, a species level composition may represent a body weight phenotype more precisely than a broad estimation of F:B ratio.

Impact of RS4 on gut microbiota composition compared before and after RS4 intervention

Firmicutes species from Clostridium cluster XIVa account for almost 60% of the mucin-adhered microbiota [92]. A general observation was that species from Clostridial cluster XIVa, but not cluster IV, were enriched by RS4 supplementation of the diet. At the species level (Figure 2.4c), RS4 consumption increased the abundance of *Bifidobacterium adolescentis* (90.5 fold, q=0.087) and *Parabacteroides distasonis* (1180.2 fold, q<0.001) but not *Ruminococcus bromii* (−3.2 fold, q>0.05), *Fecalibacterium prausnitzii* (−1.2 fold, q>0.05), or *Dorea formicigenerans* (1.1 fold, q>0.05), which confirmed the previous report [24]. Novel observations include an RS4-induced increase in *Christensenella minuta* abundance (119.7 fold, q=0.038, 97% query coverage, 88% identity and E<0.001 in NCBI-BLAST) as well as in several OTUs in the family Ruminococcaceae and genus *Bacteroides*. At the species level, *Bacteroides ovatus*
(37.6 fold, \( q=0.087 \)), *Ruminococcus lactaris* (2866.7 fold, \( q<0.001 \)), *Eubacterium oxidoreducens* (3.3 \( \times \) 10^5 fold, \( q<0.001 \)), *Bacteroides xylanisolvens* (47.8 fold, \( q=0.037 \)), and *Bacteroides acidifaciens* (92.4 fold, \( q=0.038 \)) were enriched after RS4 intervention.

**RS4 consumption altered fecal SCFAs linked to specific gut microbes**

Acetate was the most abundant SCFA, accounting for over 60% of total SCFAs before and after the interventions in both RS4 and CF groups. The individual proportions of the SCFAs, butyric (69.5%, \( p=0.03 \)), propionic (50.2%), valeric (44.1%), isovaleric (20.3%), and hexanoic (19.2%) acids increased post intervention from baseline in the RS4 group (\( p<0.05 \), Figures 2.5 and 2.6a) but not in the CF group (data not shown). A 24.6% decrease in isobutyric acid in the RS4 group was observed. A Pearson correlation analysis showed a potential link between significant changes in the gut microbiota composition induced by RS4 and altered SCFA levels (Figure 2.6b). Acetate and butyrate levels were correlated (\( p<0.05 \)) with *Ruminococcus lactaris* (r=0.54) and *Oscillospira* spp. (r=0.41). Total SCFAs were correlated with the abundance of two species: *Methanobrevibacter* spp. (r=0.43) and *Ruminococcus lactaris* (r=0.52). Propionate and isobutyrate levels were linked to *Methanobrevibacter* spp. (r=0.65 and r=0.79, respectively), *Eubacterium dolichum* (r=0.42 and r=0.43, respectively), *Christensenella minuta* (r=0.39 and r=0.59, respectively), and *Ruminococcus lactaris* (r=0.59 and r=0.40, respectively), of which the latter two were increased by RS4 (Figure 2.4a, 2.4c, and 2.6b). Interestingly, these associations of SCFAs with specific gut microbiota were not observed after CF intervention (data not shown). To our knowledge, prior studies with RS4 have not reported significant SCFA changes in human fecal samples.
Impact of RS4 intervention on circulatory adipocytokines

In obesity, macrophages infiltrate adipose tissue and secrete proinflammatory cytokines such as IL6 and TNF-α [93]. Also, adiponectin is released by adipocytes in the blood and has important roles in lipid and glucose metabolism [94]. Reduced adiponectin levels are associated with various aspects of metabolic dysfunction [95]. Compared with baseline, IL6 decreased by 38% ($p=0.04$), and adiponectin levels increased by 20% ($p=0.002$) in the RS4 group, while TNF-α did not change significantly. Both TNF-α and adiponectin concentrations were lower post RS4 diet compared with post CF diet ($p=0.08$ and $p=0.02$ respectively, Table 2.3). To our knowledge, this is the first report showing changes in adipocytokines, which help determine progression to cardiovascular aberrancies [95, 96], in response to RS4 intake in humans.

Impact of RS4 consumption on body composition, lipids, and glucose metabolism

Individuals had lower % body fat ($p=0.05$) and lower non-high density lipoprotein (non-HDL, $p=0.003$), HDL ($p=0.005$), and total cholesterol (TC, $p<0.001$) post RS4 consumption compared with post CF consumption (Table 2.3). A trend was observed for lower waist circumference ($p=0.06$), glycosylated hemoglobin (HbA1C, $p=0.08$), and fasting blood glucose ($p=0.09$) following RS4 consumption compared with CF consumption. It is likely that response variation among participants to an RS4 diet contributed to these higher $p$-values. Varying responses to dietary interventions among individuals are frequently reported [97]. Changes in fasting glucose and HbA1C were more pronounced in this cohort (−8.6% and −1%, respectively) compared with the parent
cohort (−4.2% and no decrease, respectively). Attenuation of % body fat combined with a smaller waist circumference indicates a potential reduction in central obesity in these individuals. Although significant, these changes were modest, as measures of body composition do not change rapidly in adults and can take several months to years to show a larger change. Waist circumference, TC, HDL, and non-HDL were also reduced in the RS4 group compared with baseline (all, \(p<0.05\)). No significant effects of RS4 were observed on blood pressure or triglyceride levels in either group (Table 2.3). The average lipid and glycaemic profiles were apparently within normal limits, likely due to prescribed medication usage for various metabolic dysfunctions (Table 2.1).

**Inter-associations between gut bacteria and metabolic functions**

Multiple novel gut microbial associations with metabolic functions were observed post intervention in the RS4 group compared with the CF group (Figure 2.7). We propose that the associations detected post RS4 diet, but not post CF diet, could be induced by RS4. However, several associations were common to both groups, lacking specific response to RS4 enrichment. RS4-specific inverse correlations were observed between TC and the abundances of *Bacteroides plebeius* \((r=-0.46)\), *Blautia producta* \((r=-0.49)\), and *Prevotella stercorea* \((r=-0.45)\; \text{all, } p<0.05\). Although the abundances of *Parabacteroides distasonis* and *Oscillospira* spp. were enriched post RS4 compared with post CF intervention, their negative association with TC, low-density lipoprotein (LDL), and non-HDL were not RS4-specific \((\text{all, } p<0.05)\). In another instance, while RS4 did not significantly alter the abundance of *Fecalibacterium prausnitzii*, an RS4-specific negative correlations between this species and body mass index (BMI, \(r=-0.45\) and %
body fat ($r=−0.56$) were observed (all, $p<0.05$). An RS4-specific correlation between adiponectin and *Bacteroides ovatus* ($r=0.79$, $p<0.01$), *Bacteroides uniformis* ($r=0.56$, $p<0.05$), and *Bacteroides acidifaciens* ($r=0.82$, $p<0.001$) was observed (Figure 2.7). RS4 intake did not significantly enrich *Methanobrevibacter* spp. and *Eubacterium dolichum*, but these bacteria were correlated with weight and BMI (Figure 2.7) as well as with SCFA levels (Figure 2.6b) in an RS4-specific manner. *Methanobrevibacter* spp. ($r=−0.45$), *Ruminococcus gnarus* ($r=−0.56$), and *Prevotella stercorea* ($r=−0.45$) were negatively correlated with LDL ($p<0.05$), while *Blautia producta* ($r=−0.44$) and *Prevotella stercorea* ($r=−0.50$) were negatively associated with TC and non-HDL (all, $p<0.05$) in an RS4-specific way.

**Intra-association within gut microbes**

Little is known about how the relative abundance of one microbial species may influence the presence of another species within the gut ecosystem, particularly in response to RS4 consumption. To evaluate this question, we examined intra-association and clustering among those bacteria that showed RS4-specific association with SCFAs and metabolic features. Three *Bacteroides* species that showed a positive correlation with adiponectin and *Prevotella stercorea*, which associated with TC, LDL, and non-HDL, were clustered together (Figure 2.7 and 2.8). In general, a higher association within Bacteroidetes species or within Firmicutes species was observed, although there were exceptions. One example is *Bacteroides plebius*, which correlated with the Firmicutes member *Blautia producta* ($r=0.98$, $p<0.001$), and both of these were negatively linked to RS4-induced changes in TC or non-HDL. *Christensenella minuta* tended to associate
with *Ruminococcus lactaris* \((r=0.58, p=0.02)\), and both were enriched after RS4 intervention. Both *Christensenella minuta* and *Ruminococcus lactaris* also clustered with *Methanobrevibacter spp.* and *Eubacterium dolichum* (both \(p<0.05\)), but not with *Ruminococcus torques* and *Oscillospira spp.*, although all of them were associated with one or more SCFAs (2.6b and 2.8).

**Association among metabolic features**

Variation in response to RS4 among participants was observed, and the pattern is similar for various host metabolism parameters (response variation data not shown). This pattern may be due to the known link between high levels of circulating lipids and glucose with lower blood adiponectin levels [98] and the correlation of up-regulated IL6 and tissue necrotic factor-\(\alpha\) with the pro-inflammatory state in obesity [99]. The clustering of metabolic dysfunctions and CVD risk factors in adults has been observed in epidemiological studies and in the clinical setting [100]. In line with that, we observed consistent intra-associations among parameters of metabolic dysfunction within our data set independent of dietary changes (Figure 2.9). TC and non-HDL, but not HDL, correlated more closely with each other \((r=0.95, p<0.001)\). Similar correlations were observed among various anthropometric measures, such as weight, BMI, and waist circumference, as well as fasting glucose \((p<0.05)\). Fasting glucose correlated with IL6 level \((r=0.51, p=0.07)\), which in turn was associated with diastolic blood pressure \((r=0.78, p<0.01)\), systolic blood pressure \((r=0.52, p=0.07)\), and waist circumference \((r=0.60, p=0.03)\). Triglyceride concentrations were positively associated with weight \((r=0.82, p<0.001)\), waist circumference \((r=0.60, p=0.03)\), and BMI \((r=0.64, p=0.02)\), while
negatively correlated with HDL ($r=-0.65$, $p=0.02$) and adiponectin ($r=-0.45$, $p=0.12$) but had little apparent link with TC and LDL.

2.5. Discussion

Our present study provides for the first time a microbiome signature in response to RS4 consumption in subjects with MetS. Among the comorbidities of MetS, numerous studies associated a dysbiosis with obesity and type 2 diabetes [101], but less is known about the role of the microbiota-diet interactions in hypercholesterolemia [102]. A handful of in vitro, animal, and healthy volunteer studies suggested that the microbiota affects lipid metabolism [103-105], but a clear understanding is lacking. Reduction in plasma TC and non-HDL and HDL cholesterol after RS4 consumption was consistently observed in the larger parent cohort [25] as well as in the present sub-cohort when compared with the control group and with baseline. In addition, a novel link between TC and non-HDL with three bacterial species was observed in an RS4-specific manner. The effect of an RS4 diet on Parabacteroides distasonis without any context of lipid metabolism was previously reported [24] and further confirmed by our results. Parabacteroides distasonis was augmented post-RS4 compared with the post-CF and baseline. This species showed a correlation with TC and non-HDL in both intervention groups and clustered with species belonging to both the Bacteroidetes and Firmicutes phyla. While RS4 also lowered HDL, such a reduction is not always associated with increased cardio-vascular disease risk, while lowering of TC and non-HDL, which includes LDL, remains critical [106].
Another novel RS4-induced microbial enrichment involved *Christensenella minuta* in the family Christensenellaceae, which was identified in human faeces only recently [107]. This species was reported as being heritable, based on host genetics, and abundant in healthy, lower-BMI individuals, and its addition reduced microbiome-mediated weight gain in germ-free mice [108]. *Christensenella minuta* was enriched post RS4-diet compared with post CF-diet and the baseline, and correlated in an RS4-specific manner with higher propionate, isobutyrate, valerate, and isovalerate concentrations, similar to that of *Methanobrevibacter* spp. in the Methanobacteriaceae family. This result is in line with a previous report that this bacterium co-occurs with Methanobacteriaceae members, and both together produce SCFAs [108]. Moreover, *Christensenella minuta* was reported to augment *Oscillospira*, a bacterium that was enriched and associated with higher isovalerate levels post-RS4 but not post-CF. However, contrary to a prior suggestion [108], our results present evidence that *Christensenella minuta* is amenable to dietary intervention. We also observed a positive correlation of this species with weight and BMI, which was, however, non-specific to RS4 intervention. Future investigations may address whether *Christensenella minuta*-mediated energy harvest from SCFAs is utilized differently by individuals with MetS than by the lean and healthy individuals involved in prior studies [107, 108].

Complex carbohydrate or glycan availability is a major factor in shaping the gut microflora [109]. Since many of the starch-degrading enzymes are represented as a starch-utilization system (SUS) in ~20% of the genome of Bacteroidetes species [109-111], enrichment of *Bacteroides ovatus* and *Parabacteroides distasonis* after RS4 intervention in this study was as expected. Further confirmation with Carbohydrate-
Active Enzymes (CAZy) database (http://www.cazy.org/) [112] supported that Bacteroidetes species are dominant starch-degrading bacteria after RS4 intervention.

The strengths of this study include statistically significant observations of prebiotic RS4-induced changes across microbial composition, fecal SCFA levels, and host immunometabolic functions, all of which are relevant to the underlying physiology of individuals with pre-existing metabolic dysfunctions. Human dietary interventions conducted within natural settings and reporting statistically significant outcomes that are consistent across a broad range of metabolic health measures are rare [113]. One exploratory human intervention study attempted to show the prebiotic diet concept for combating obesity by undertaking a similar comprehensive investigation, but utilized a parallel design and a different kind of non-RS prebiotic diet. In that study, the changes in the microbiota and some metabolites, although not including SCFA, were observed, but without any concomitant improvement in metabolic functions [114]. This difference could be due to RS4 having a different prebiotic impact on such parameters or due to the crossover design of our study, which allowed each participant to serve as his/her own control, minimizing the influence of confounding variables on treatment effects. In addition, a stringent FDR at 0.05 helped minimize false positives for gut microbial changes. Furthermore, novel associations among host metabolic functions and species-level composition of the gut microbiota were observed, many of which also coincided with intra-species clustering, indicating the possibility of a synergistic (within bacterial species) as well as super-system interactions (symbiotic host–microbe systems) linked with metabolic functions. However, we note that fecal SCFAs represent <5% of the total SCFAs that are typically excreted, while the major portion of SCFAs is efficiently
absorbed in the intestinal lumen [103]. Also microbial communities in fecal samples may potentially exclude gut residents that are not shed. Nevertheless, fecal specimens are the most practical samples obtained from a human dietary intervention study for assessing the gut microbiome.

2.6. Conclusion

In conclusion, this study provides evidence that dietary RS4 supplementation selectively changes the gut microbial and metabolite environment as well as associated host metabolic functions. To our knowledge, this is the first holistic study that investigated the effects of the fermentable fiber, RS4, on the gut microbial ecology, functional metabolites like SCFAs, and physiological responses in the host in one well-designed study in a free living Caucasian cohort with signs of MetS. The findings support the perceived role of the microbiota–host interaction in nutritional therapies with important implications for dietary guidelines for individuals with metabolic disorders, a major public health concern of the present day.

2.7. Future Directions

1. To compare the gut microbial profiles from persons with or without MetS, which would help delineate the signature microbiome in MetS.

2. To identify the responders and non-responders to RS4 in terms of some selected gut microbiota or SCFAs, or parameters.

3. To identify the sources of health promoting metabolites, either from food, microbiota, or host itself.
Figures and Legends

(a) Time line of the placebo-controlled, crossover, dietary intervention with resistant starch (RS4) and control flour (CF). Stool and blood samples were collected before and after each treatment period (indicated by arrows).

(b) Trial profile and numbers of participants in the study.

**Figure 2.1. Study design**

127 subjects screened

- 86 participants, met inclusion criteria, signed consent form and enrolled in the parent trial (NCT01887964)
- 83 completed the trial. Primary outcome of the trial reported (Nichenumetla et al., 2014)
- 40 participants had metabolic syndrome at baseline (screened using International Diabetes Federation criteria). Most participants were on prescription medications for at least one sign of metabolic syndrome
- 26 submitted fecal sample at all time points
- 20 participants selected for this retrospective study based on adequate fecal sampling
- 19 included in microbial and short chain fatty acids data analyses. One subject was excluded due to probiotic supplement use during the study to avoid potential confounding effects. The n was further adjusted when using paired analysis, if any data point was missing
Figure 2.2. Phylum-level gut microbial composition of participants at baseline
Mean percentage of total bacteria in the major phyla (n=19).
Figure 2.3. Separation of the microbiome post intervention in RS4 and CF groups

Two-dimensional principal coordinate analyses (PCoA) based on the weighted UniFrac distance between samples, given the abundance of 5,831 taxa present in at least one sample (n=19). Axes 1 and 2 explain 26% and 13% of the variation, respectively ($p = 0.01$).
Figure 2.4. Differential gut microbial composition after RS4 intervention at the species level.

(a) Relative abundance of bacterial species (log$_{10}$ fold change) in the RS4 group compared with the CF group post intervention (n=19). Significant compositional variation between the two groups before the intervention was previously ruled out. (b) The Firmicutes/Bacteroidetes ratio after intervention (n=14). The dotted line represents this ratio at baseline. (c) Abundance of major bacterial species (log$_{10}$ fold change) before and after RS4 treatment. #, the closest hit from the NCBI 16S rRNA database cross-referenced with the OTU from the Greengenes database. *Adjusted $p (q) \leq 0.05$, **$q \leq 0.01$, ***$q \leq 0.001$, § $q \leq 0.09$ (trend/approaching significance), n=19.
Figure 2.5. Representative chromatograms showing the overlay of retention times of butyl esters of SCFA fragment from standard and biological sample. Butyl esters of acetic acid (1), propionic acid (2), iso-butyric acid (3), butyric acid (4), iso-valeric acid (5), valeric acid (6), internal standard (7), and hxanoic acid (8). * represents the di-butyl ether as a side product of butylation that did not co-elute with the sample analytes. One of the collaborators, Dr. Ali Reza Fardin-Kia helped with technical and instrumental aspects of GC-MS analyses.
Figure 2.6. Effects of RS4 on fecal SCFAs
(a) SCFA abundance before and after RS4 intervention (*p≤0.05, n=19). (b) Positive correlation of six bacterial species with increased SCFA levels in an RS4-specific manner (p<0.05). Pearson coefficients are shown on heat map. #, the closest hit from the NCBI 16S rRNA database cross-referenced with the OTU from the Greengenes database. †, species either significantly enriched or approached significance in the RS4 group. One of the collaborators, Dr. Ali Reza Fardin-Kia and Robert Juenemann helped with technical and instrumental aspects of GC-MS analyses.
Figure 2.7. Associations between gut microbiota and host biological parameters after RS4 and CF interventions

(a) Heat map showing Pearson’s r values (all, \( p < 0.05 \)). Black rectangular borders indicate an association present only post RS4 intervention. \#, the closest hit from the NCBI 16S rRNA database cross-referenced with the OTU from the Greengenes database. †, species either significantly enriched or approached significance in the RS4 group, \( n=15 \).
Figure 2.8. Intra-associations within bacterial species that were correlated with metabolic functions or SCFAs in an RS4-specific manner

Heat map showing Pearson’s r-values, corresponding to the size of the circle (n=19). The black border indicates clustering of species (*p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001, shown only in the upper triangle). #, the closest hit from the NCBI 16S rRNA database cross-referenced with the OTU from the Greengenes database. †, species either significantly enriched or approached significance in the RS4 group.
Figure 2.9. Associations among host biological parameters in the RS4 group
Heat map showing Pearson’s r-values, corresponding to the size of the circle (n=13, *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001).
Table 2.1. Baseline characteristics of 20 participants

<table>
<thead>
<tr>
<th>Criteria</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
</tr>
<tr>
<td>&lt;50</td>
<td>5</td>
</tr>
<tr>
<td>≥50</td>
<td>15</td>
</tr>
<tr>
<td><strong>BMI</strong></td>
<td></td>
</tr>
<tr>
<td>&lt;30</td>
<td>5</td>
</tr>
<tr>
<td>≥30</td>
<td>15</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>8</td>
</tr>
<tr>
<td>Female</td>
<td>12</td>
</tr>
<tr>
<td><strong>Medication for</strong></td>
<td></td>
</tr>
<tr>
<td>No medication</td>
<td>8</td>
</tr>
<tr>
<td>Type 2 diabetes</td>
<td>4</td>
</tr>
<tr>
<td>Heart diseases</td>
<td>2</td>
</tr>
<tr>
<td>Blood pressure</td>
<td>11</td>
</tr>
<tr>
<td>Fiber supplement</td>
<td>2</td>
</tr>
<tr>
<td>Probiotic supplement</td>
<td>1</td>
</tr>
<tr>
<td>Digestive support</td>
<td>3</td>
</tr>
<tr>
<td>Cholesterol lowering</td>
<td>5</td>
</tr>
</tbody>
</table>

n: number of individuals
BMI: body mass index
Table 2.2. Estimated nutrients intake at baseline and at the end of intervention periods

<table>
<thead>
<tr>
<th>Nutrients</th>
<th>BL</th>
<th>Post CF</th>
<th>Post RS4</th>
<th>( p ): post CF vs post RS4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caloric intake (kcal/d)</td>
<td>1774±154</td>
<td>1528±121</td>
<td>1716±128</td>
<td>NS</td>
</tr>
<tr>
<td>Protein (g/d)</td>
<td>76±7</td>
<td>72±5</td>
<td>62±4</td>
<td>NS</td>
</tr>
<tr>
<td>Carbohydrate (g/d)</td>
<td>218±21</td>
<td>220±19</td>
<td>212±18</td>
<td>NS</td>
</tr>
<tr>
<td>Total fat (g/d)</td>
<td>68±7</td>
<td>60±5</td>
<td>53±5</td>
<td>NS</td>
</tr>
<tr>
<td>Cholesterol (mg/d)</td>
<td>415±40</td>
<td>442±38</td>
<td>407±35</td>
<td>NS</td>
</tr>
<tr>
<td>Saturated fat (g/d)</td>
<td>25±3</td>
<td>22±2</td>
<td>21±2</td>
<td>NS</td>
</tr>
<tr>
<td>Monounsaturated fat (g/d)</td>
<td>26±3</td>
<td>22±2</td>
<td>19±2</td>
<td>NS</td>
</tr>
<tr>
<td>Polyunsaturated fat (g/d)</td>
<td>10±1</td>
<td>9±1</td>
<td>7±1</td>
<td>NS</td>
</tr>
<tr>
<td>Total dietary fiber (g/d)</td>
<td>18±2</td>
<td>16±2</td>
<td>27±2</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

\(^a\) Data are Least Square Means±SEM adjusting for age, sex, season, colony and baseline value analysed by linear mixed model; \( n = 18-20 \) due to missing data points.
BL: baseline; CF: control flour; RS4: resistant starch type 4
One of the collaborators, Dr. Lacey McCormack helped with data analyses
| Table 2.3. Means of biological parameters at baseline and at the end of intervention periods<sup>a</sup> |
|--------------------------------------------------|--------------------------------------------------|--------------------------------------------------|--------------------------------------------------|--------------------------------------------------|
| BL                                               | Post CF                                          | Post RS4                                          | *p:* Post CF vs Post RS4 | *p:* BL vs post RS4 |
| **Anthropometrics**                              |                                                  |                                                  |                     |                     |
| Weight (kg)                                      | 90.9±3.4                                        | 91.0±0.4                                         | 91.6±0.4             | NS                  | NS                  |
| BMI (kg/m<sup>2</sup>)                           | 32.8±1.1                                        | 32.8±0.1                                         | 32.7±0.1             | NS                  | NS                  |
| Waist (cm)                                       | 109.0±2.8                                       | 108.8±0.9                                        | 106.6±0.9            | **0.06**            | **0.02**            |
| Systolic BP (mm Hg)                              | 135.0±3.9                                       | 134.6±3.5                                        | 137.5±3.5            | NS                  | NS                  |
| Diastolic BP (mm Hg)                             | 73.7±2.2                                        | 68.6±2.0                                         | 73.3±2.0             | NS                  | NS                  |
| % Body Fat                                       | 37.0±1.8                                        | 37.7±0.3                                         | 37.3±0.3             | **0.05**            | NS                  |
| Fat-free mass (kg)                               | 58.8±3.0                                        | 58.8±0.3                                         | 58.9±0.3             | NS                  | NS                  |
| **Glycemic Variables** (mg/dL)                    |                                                  |                                                  |                     |                     |
| Fasting glucose                                  | 106.5±4.1                                       | 111.5±4.2                                        | 101.9±4.3            | **0.09**            | NS                  |
| Postprandial glucose                             | 113.5±11.8                                      | 124.3±7.3                                        | 114.3±7.5            | NS                  | NS                  |
| HbA1C (% of total Hb)                            | 5.89±0.3                                        | 5.81±0.1                                         | 5.75±0.1             | **0.08**            | NS                  |
| **Lipid Variables** (mg/dL)                       |                                                  |                                                  |                     |                     |
| Total cholesterol                                | 196.6±11.6                                      | 192.8±4.6                                        | 187.8±.9             | <0.001              | 0.01                |
| HDL cholesterol                                  | 43.6±3.3                                        | 44.1±1.3                                         | 39.8±1.3             | <0.01               | 0.001               |
| LDL cholesterol                                  | 122.7±10.1                                      | 117.4±5.6                                        | 118.0±6.1            | **0.06**            | 0.06                |
| NonHDL cholesterol                               | 153.1±11.8                                      | 148.4±4.6                                        | 147.5±4.9            | <0.01               | 0.03                |
| TC/HDL (ratio)                                   | 5.0±0.5                                         | 4.7±0.2                                          | 5.1±0.2              | NS                  | NS                  |
| Triglycerides<sup>b</sup>                        | 161.5±19.9                                      | 144 (119-176)                                    | 138 (110-173)        | NS                  | NS                  |
| **Blood Biomarkers**                             |                                                  |                                                  |                     |                     |
| IL6 (pg/mL)                                      | 1.3±0.2                                         | 1.0±0.2                                          | 0.8±0.2              | NS                  | **0.04**            |
| TNF-α (pg/mL)                                    | 7.9±4.2                                         | 9.9±1.2                                          | 6.0±1.3              | **0.08**            | NS                  |
| Adiponectin (µg/mL)                              | 8.3±1.5                                         | 10.8±0.4                                         | 10.0±0.4             | **0.02**            | <0.01               |

<sup>a</sup> Data are Least Square Means±SEM adjusting for age, sex, season, colony and baseline values.

<sup>b</sup> Geometric mean and confidence interval are given for log-transformed triglyceride endpoints.

One of the collaborators, Dr. Lacey McCormack helped with data analyses.

3.1. Abstract

Indigestible resistant starches (RS) are substrates for gut-microbial metabolism and have been shown to attenuate intestinal inflammation but the supporting evidence is inconsistent and lacks mechanistic explanation. We have recently reported dietary RS type 4 (RS4) induced improvements in immunometabolic functions in humans and a concomitant increase in butyrogenic gut-bacteria. Since inflammation is a key component in metabolic diseases, here we investigated the effects of RS4-derived butyrate on epigenetic repression of pro-inflammatory genes in vivo and in vitro. RS4-fed mice, compared to the control-diet group, had higher cecal butyrate and increased trimethylation of lysine 27 on histone 3 (H3K27me3) in the promoter of nuclear factor-kappa-B1 (NFκB1) in the intestinal tissue. The H3K27me3-enrichment inversely correlated with concentration dependent down-regulation of NFκB1 in sodium butyrate treated human intestinal cells. Two additional inflammatory genes were attenuated by
sodium butyrate, but were not linked with H3K27me3 changes. This proof-of-concept exploratory study presents a new opportunity for studying underlying H3K27me3 and other methylation modifying mechanisms linked to RS4 biological activity.

**Keywords:** resistant starch type 4, sodium butyrate, H3K27me3, short chain fatty acid, inflammation

### 3.2. Introduction

Intestinal gut microbiota utilizes prebiotic fibers like resistant starch (RS) as a substrate to produce short chain fatty acids (SCFAs) that is proposed to play an anti-inflammatory role in the maintenance of intestinal immune homeostasis [115], although conflicting results exist [116]. Consumption of a fiber-rich diet is nevertheless considered beneficial in several diseases such as obesity, inflammatory bowel disease, and metabolic syndrome (MetS), all of which are associated with gut microbial dysbiosis and dysregulated mucosal immunity [117]. In particular, inflammation is considered as a key pathophysiological player in MetS [118]. We previously reported that consuming RS type 4-(RS4)-enriched flour (30% v/v) as part of a routine communal diet has significant cholesterol lowering effects in healthy as well as in subjects with signs of MetS [25]. A retrospective follow-up of the study showed that participants who consumed RS4 had increased fecal SCFAs, particularly butyrate, in conjunction with increased butyrogenic microbiota in fecal samples [119]. Butyrate, as a major and representative form of SCFAs, has both trophic and bioactive functions on host cells by correcting the altered ratio of gut bacteria and protecting against the potential pro-inflammatory molecules [35]. It has been reported that increasing large-bowel butyrate supply has the potential to promote colonic
integrity and lower colonocyte risk [120]. In that context this proof-of-concept study aimed to explore if RS4 (chemically modified RS), particularly through butyrate production, epigenetically regulates inflammatory gene expression in the gut.

Histone modifications, especially histone acetylation and methylation, play a dominant role in epigenomic regulation of gene expression [121]. Butyrate is well known as a histone deacetylase (HDAC) inhibitor that alters chromatin structure through histone acetylation, which in turn affects target gene expression associated with the maintenance of gut immune homeostasis [115, 122]. Butyrate’s anti-inflammatory activities [115] result from both HDAC dependent and independent inhibition of transcription factor nuclear factor kappa B (NFκB) [123, 124] as well as suppression of the downstream pro-inflammatory chemokines and cytokines [125, 126]. However, knowledge is limited regarding the role of butyrate in modulation of histone tail methylation. Of particular interest, in the context of inflammatory gene expression down-regulation, is the trimethylation of H3 at lysine 27 (H3K27me3), an important and well-studied modification that is frequently associated with transcriptional repression and gene silencing [59, 127].

In the present study, the wheat-derived cross-linked RS4 is used which has been thought to be effective in attenuation in postprandial glucose and insulin levels in humans [72]. Since RS4 is insoluble, we first examined the in vitro effects of sodium butyrate (NaBu) on H3K27me3 enrichment of gene promoters in conjunction with mRNA and protein expression of selected inflammatory markers in a human intestinal epithelial cell line, followed by a validation of RS4-induced butyrate production and a concomitant H3K27me3 modifying effects in vivo in mouse colon tissue. Of the three selected
inflammatory markers studied, a transcriptional factor NFκB1, which belongs to the NFκB family of transcription factors, plays a central role in inflammatory signaling [128]. The remaining two inflammatory mediators represented a chemokine (C-C motif) ligand 2 (CCL2) that regulates migration and infiltration of monocytes/macrophages [125], and a T-helper type 2 (Th-2) cytokine interleukin-10 (IL-10), which diminishes the activity of Th-1 cytokine interferon gamma (IFNγ) and monocyte [126].

3.3. Materials and Methods

Materials

Dulbecco’s Modified Eagle’s Medium (DMEM) with 4 mM L-glutamine and 4.5 g/l glucose was purchased from HyClone (Logan, UT), fetal bovine serum (FBS), and trypLE were purchased from Invitrogen Gibco (Grand Island, NY). Dimethyl sulphoxide (DMSO), lipo-polysaccharide (LPS, from Escherichia coli, O55:B5), penicillin/streptomycin, hydrogen chloride-1-butanol, and NaBu (sodium salt of butyrate) were purchased from Sigma-Aldrich (St. Louis, MO), while IFNγ was purchased from R&D Systems (McKinley, NB). For western blot, antibody against β-actin was purchased from Santa Cruz Biotechnology (Santa Cruz, CA), antibody against NFκB1 was purchased from Millipore (Billerica, MA), and Dylight 800 anti-rabbit secondary antibody was purchased from Li-Cor Biosciences (Lincoln, NE). Anti-trimethyl-Histone H3 Lys27 and rabbit IgG negative control antibodies used for chromatin immuno-precipitation (ChIP) assay were purchased from Upstate Biotechnology (Billerica, MA). Enzymes used for ChIP assay, Micrococcal Nuclease (MNase) and Proteinase K were purchased from Cell Signaling (Beverly, MA) and Pierce (Thermo Scientific, Rockford,
IL), respectively. All other ChIP assay chemicals, aprotinin, DL-1, 4-Dithiothreitol (DTT), Nonidet-P40 (NP-40) were purchased from Thermo Scientific (Rockford, IL), while protein-A sepharose and sucrose were purchased from Sigma-Aldrich (St. Louis, MO). ChIP incubation buffer was purchased from Abcam (Cambridge, MA) and DNA purifying slurry for ChIP-graded DNA purification was purchased from Diagenode (Denville, NJ). 2-Ethylbutyric acid (99%) and 1-butanol (99.9%) used as internal standard in gas chromatography- mass spectrometry (GC-MS) were purchased from Acros Organics (Mullica Hill, NJ) and Alfa Aesar (Ward Hill, MA), respectively. Hexane (>97.0%) and sodium sulphate (granular, anhydrous, >99.0%) were purchased from Fisher Scientific (Hampton, NH), while the inert helium gas was purchased from Matheson (Sioux Falls, SD). Oligonucleotides were synthesized by IDT DNA Inc. (Coralville, IA).

**Cell culture and treatment**

All *in vitro* assays were conducted in human colon cancer cell line, SW480 (CCL-228, ATCC, Manassas, VA). Cells were grown in DMEM supplemented with 10% FBS, 1% penicillin (25 U/ml)/streptomycin (25 µg/ml) in a 95% air/5% CO₂-humidified atmosphere at 37 °C. This cell line was cultured as previously described [59]. Briefly, SW480 cells were pre-treated with IFNγ (10 ng/ml) or control medium for 12 h, treated with NaBu or DMEM (as a negative control) for 5 h and then stimulated with LPS (10 ng/ml) for 4 h. For every experiment, one positive control (cells treated with LPS) and one negative control (cells treated without LPS or any NaBu treatments) were included.
Two replicates were used for both the treatments and controls. LPS and NaBu were reconstituted in DMEM.

**Total RNA extraction, cDNA synthesis and quantitative polymerase chain reaction (qPCR)**

Total RNA was extracted from cells using TRIzol reagent (Invitrogen, Grand Island, NY), following the manufacturer's instructions. RNA was quantified spectrophotometrically by absorption measurements at 260 nm and 280 nm using the NanoDrop system (NanoDrop Technologies, Wilmington, DE). RNA was then treated with DNase I (Invitrogen, Grand Island, NY), following the manufacturer's guidelines to remove any traces of DNA contamination. The cDNAs were synthesized using 3 µg of RNA for each sample using the High-Capacity cDNA Reverse Transcription (RT) Kit (Invitrogen, Grand Island, NY), following the manufacturers’ protocol. Two microliters of each diluted sample was added to 0.5 µl of gene-specific primers and 12.5 µl of Power SYBR green PCR master mix (Invitrogen, Grand Island, NY), and the final volume was brought to 25 µl by adding sterile distilled water. PCR amplifications were performed on a MX3005P system (Stratagene, San Diego, CA) using one cycle at 50 °C for 2 min, one cycle of 95 °C for 10 min, 40 cycles of 15 s at 95 °C and 1 min at 60 °C, and the last cycle with 95 °C for 1 min, 55 °C for 30 s, and 95 °C for 30 s. RNA extraction, purification, cDNA synthesis and RT-PCR were performed as previously described [129] in duplicate. Gene-specific primers used in the current study are described in Table 1. Calculations of relative gene expression levels were performed using the $2^{-\Delta\Delta C_t}$ method [130]. The mRNA levels were normalized to a housekeeping gene, glyceraldehyde-3-phosphate
dehydrogenase (GAPDH), and expressed as a fold change relative to positive-control cells.

**Western blot analysis**

For immunoblot analyses, IFN$\gamma$-primed NaBu-treated human intestinal epithelial cells were activated with LPS for 4 h and harvested using RIPA lysis buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM Na$_2$EDTA, 1 mM EGTA, 1% NP-40, 1% sodium deoxycholate, 2.5 mM sodium pyrophosphate, 1 mM $\beta$-glycerophosphate, 1 mM Na$_3$VO$_4$, 1 $\mu$g/ml leupeptin). The Pierce BCA Protein Assay kit (Thermo Scientific, Rockford, IL) was used to determine protein concentration. Proteins (35-50 $\mu$g/lane) were separated by 12% SDS-PAGE and the products were electrotransferred to polyvinylidene difluoride membranes (Thermo Scientific, Rockford, IL). The membranes were blocked with 5% skim milk for 1 h, and incubated with primary antibodies at 4°C overnight (NF$\kappa$B p105/p50). On the next day, membranes were incubated with Dylight 800 anti-rabbit secondary antibody (Li-Cor Biosciences, Lincoln, NE) for 1 h, and washed 3 times in PBS/T (0.1% Tween20 in PBS) all at room temperature. After rinsing in PBS/T, blots were imaged and quantitatively analyzed using an Odyssey infrared imaging system (Li-Cor, Lincoln, NE). A loading control protein, $\beta$-actin, was used to calculate the relative expression of NF$\kappa$B p105 and p50 subunits.

**Native ChIP assay in cell culture**

ChIP assay was performed as we previously described [59]. Briefly, cells were lysed by suspending in a sequence of lytic and purification buffers. DNA fragments of
300-800 bp long were obtained by treating the nuclear pellet, obtained from lytic buffers, with MNase (Cell Signaling, Beverly, MA) in digestion buffer for 6 min. DNA fragments were immunoprecipitated with specific antibodies (anti-trimethyl-Histone H3 Lys27 and rabbit IgG) at 4 °C overnight. Immunoprecipitated DNA fragments were extracted using protein-A sepharose (Sigma-Aldrich, St. Louis, MO) and purified using DNA purifying slurry (Diagenode, Denville, NJ). The amount of purified DNA was estimated by nanodrop of 200 ng purified DNA template using SYBR green chemistry as described in RT-qPCR quantification section. Promoter-specific primers used in the current study are described in Table 1. Calculations are expressed as a percentage of the input DNA.

**Animal housing, diet and tissue collection**

All *in vivo* procedures were approved and conducted following the Institutional Animal Care and Use Committee (IACUC) of South Dakota State University guidelines. Six-week old male KK.Cg-A^{a/y} (Jackson Laboratories, Bar Harbor, ME) mice were grouped at random to consume either RS4 or control diets. To avoid confounding effects of sex on weight gain, which was a measurable outcome for the study, only male mice were considered. Also, this study was conducted prior to the implementation of federal policy for considering both male and female sexes in biomedical experiments (NOT-OD-15-102). Mice were caged in groups of three in a laboratory animal facility at an ambient temperature of 24-26 °C with 12-h light/dark cycle. During the first 3 weeks mice were allowed to acclimatize while consuming standard rodent chow, LabDiet® 5001 (LabDiet, St. Louis, MO). Animals were then switched to experimental diets (LabDiet, St. Louis, MO) formulated based on AIN 93 [131] either with 20% RS4 or control diet (CD).
Detailed composition of experimental diets is provided in Table 2. Mice were provided the experimental diets and water *ad libitum* for 12 weeks. Following CO₂ euthanasia, colon tissues and cecum samples were immediately collected, snap frozen in liquid nitrogen and stored at -80 °C until further use.

**Determination of cecal weight and butyrate analyses**

Cecal tissues were weighed and pooled together (>500 mg in total weight) to represent one sample and vortexed for 1 min with 1 ml of internal standard (2-ethylbutyric acid in 1-butanol, 0.25 mg/ml). 0.5 ml of hexane (organic solvent) and 2 ml of HCl-1-butanol (derivatizing agent) were then added to each sample followed by an additional 1 min vortexing and 5 min sonication prior to being purged with an inert helium gas. The tubes were immediately closed and each sealed container was incubated in water bath at 60 °C overnight in order to catalyze the derivatization of butyrate analyte. Upon being cooled to room temperature, 1.5 ml of additional hexane and 15 ml of deionized water were added. Samples were mixed by vortexing 1 min following each addition and then centrifuged at 3000 g for 2 min. The top organic layer (~2 ml) was transferred into a 5 ml graduated vial before blowing down with helium to one-fourth of the volume, thereby increasing the final concentration of internal standard from 0.25 mg/ml to 1 mg/ml. Finally, each sample was transferred into a sampling vial containing a 150 µl insert and ~10 mg of anhydrous sodium sulphate was added to remove the water content before running into the GC-MS.

SCFAs were derivatized to obtain the corresponding butyl-esters (SCFABE) prior to their gas chromatography (Agilent 7890A, Agilent Technologies, Wilmington, DE)
and mass spectrometry (5977A MSD, Agilent Technologies, Wilmington, DE, USA) using HP-5MS UI capillary column (30 m x 0.25mm, 0.2 µm thickness, Agilent, Wilmington, DE, USA). Using hydrogen as carrier gas at a constant flow of 1.9 mL/min, a typical injection volume (1 µl) was injected in the split mode (1:10). The separation of SCFABE was achieved using an oven temperature program as follow: initial elution temperature of 55 °C for 4 min, an increase by 5 °C/min to 120 °C followed by 20 °C/min to 220 °C for 10 min. The selective mass detector was operated in the ‘single ion monitoring and scan’ (SIM/Scan) mode with the source temperature (150 °C) and the electron energy (70 eV). The abundance of butyrate analyte was identified by acquiring ions of a specific m/z value. Finally, the data were analyzed and expressed as a unit of mg/g of cecal tissues used.

**Native ChIP assay in mouse colon tissues**

Colon tissues were collected from mice immediately after sacrifice, snap-frozen in liquid nitrogen and stored at -80 °C. On the day of assay, around 0.2 g of colon tissues were rinsed in cold PBS and homogenized on ice in a pre-chilled tube with 1 ml of ice-cold Buffer 1 (0.06 M KCL, 15 mM NaCl, 5 mM MgCl2, 15 mM Tris-HCl pH 7.4-7.6, 0.1 mM EGTA, 0.3 M sucrose, 180 µg aprotinin, 5 mM sodium butyrate, 0.1 mM PMSF, 0.5 mM DTT), until the homogenate was free of clumps. The homogenate was then filtered through four layers of muslin cheesecloth, pre-autoclaved and pre-moistened with Buffer 1. The filtered cells were treated in the same way as described in the Native ChIP assay in cell culture section. Promoter methylation of genes of interest was determined in
mice fed RS4 or CD. Promoter-specific primers used in the current study are described in Table 1.

**Statistical analysis**

Data are expressed as mean ± SEM. One-way analysis of variance (ANOVA) was used to determine the significance between groups followed by the post-ANOVA Dunnett test. Intergroup comparisons in Western blot, ChIP assay in mouse colon tissues and for SCFAs in mouse cecal tissues were determined using Student’s t-test. A bivariate analysis was carried out using correlation coefficient method and presented as Pearson’s r. The significance of each treatment was interpreted by comparison with the appropriate control, as described under specific methods. Experiments were repeated at least three times. A probability (p) value of 0.05 or less was considered to be significant.

3.4 Results

**NaBu attenuates expression of inflammatory mediators in a concentration dependent manner in intestinal epithelial cells**

LPS stimulated the production of pro-inflammatory mediators in IFNγ-primed human intestinal epithelial cells. In the cells treated with NaBu, compared to the positive control cells treated with LPS alone, we observed a concentration-dependent mRNA attenuation of CCL2 and NFκB1 (Figure 3.1a). The highest concentration of NaBu tested (1.25 mM) showed the strongest attenuation, lowering the relative mRNA expression of CCL2 and NFκB1 by 90% and 63%, respectively (both p < 0.001). We also observed the similar effects in case of cytokine IL-10, where NaBu (1.25 mM) inhibited IL-10 mRNA
by 78% \((p < 0.001, \text{Figure 3.1a})\). The untreated cells that served as negative control showed the highest level of IL-10, indicating the potential anti-inflammatory role of IL-10 in the absence of a strong immune challenge by LPS (Figure 3.1a). However, it is possible that IL-10 gained a pro-inflammatory characteristics when primed with IFN\(\gamma\) prior to LPS induction as was observed by Sharif et al [132] and hence subsequently responded to NaBu in a similar manner to that of the other two pro-inflammatory mediators. This point is further elaborated in the “discussion” section.

**NaBu enriched H3K27me3 in the promoter of NF\(\kappa\)B1 in intestinal epithelial cells**

For ChIP assay we chose the highest dose (1.25 mM) of NaBu due to its most suppressive effects as observed for the gene expression data (Figure 3.1a). Compared with a positive control (Figure 3.1b), we observed the enrichment of H3K27me3 on the promoter region of NF\(\kappa\)B1 \((0.47 \pm 0.23\% \text{ vs } 1.28 \pm 0.08\%, p < 0.05)\). The H3K27me3 enrichment pattern was not clear for the IL-10 promoter, possibly due to high intersample variation. For CCL2, a similar enrichment pattern to that of NF\(\kappa\)B1 was absent, indicating the possibility of another gene regulatory mechanism influencing its mRNA downregulation (Figure 3.1a). Since CCL2 is a downstream effector of NF\(\kappa\)B signaling [133], it is possible that the mRNA suppression of CCL2 is due to inhibition of NF\(\kappa\)B1.

**Concentration-dependent suppression of NF\(\kappa\)B1 protein and mRNA by NaBu inversely correlated with concentration-dependent enrichment of H3K27me3 in the NF\(\kappa\)B1 promoter**
We further examined the concentration-dependent effects of NaBu on NFκB1 mRNA and protein downregulations simultaneously with the determination of a concentration-dependent H3K27me3 enrichment on NFκB1 promoter region. As shown in Figure 3.2a, the protein expression of p105 as well as the active p50 subunits of NFκB1 showed a concentration-dependent attenuation ($p < 0.05$) in response to NaBu in the IFNγ-primed and LPS-induced human intestinal epithelial cells. Interestingly, the relative protein expressions of both the subunits dropped close to the level of the negative control when treated with the 1.25 mM NaBu (p105: $0.77 \pm 0.03$ vs $0.81 \pm 0.08$ folds and p50: $0.53 \pm 0.15$ vs $0.41 \pm 0.13$ folds). The suppression of both the subunits also negatively correlated with a simultaneous concentration-dependent upregulation of H3K27me3 modifications on the promoter region of NFκB1 (Figures 3.2b and 3.2c) (p105: Pearson’s $r = -0.98$, $p = 0.022$ and p50: Pearson’s $r = -0.97$, $p = 0.025$). Further, NaBu-associated H3K27me3 enrichment of NFκB1 promoter also inversely correlated with the relative NFκB1 mRNA expressions (Pearson’s $r = -0.95$, $p = 0.045$, Figures 3.1a, 3.2b, and 3.2c).

**Body mass and food intake in mice study**

We have previously reported that RS4-enriched diet may be effective for reducing pathophysiological consequences in humans with MetS [25]. Since KK.Cg-Aγ/a mice develops hyperglycemia, hyperinsulinemia, glucose intolerance and obesity by eight weeks of age, we in this study utilized KK.Cg-Aγ/a mouse model to mimic age-related MetS in humans [134], evaluating the effects of RS4 intake on butyrate production in the cecum and butyrate-associated epigenetic regulation of inflammatory genes in the
intestinal tissue. Since RS4 is a stealth ingredient [135], we incorporated RS4 in the mouse chow (20% v/v) and adjusted the CD to make it isocaloric with the addition of cellulose (Table 3.2). At the end of the 12-week feeding study, the cumulative food intake in RS4 group was around 35% higher when compared to that of CD group (560 ± 22 g vs 415 ± 26 g, $p = 0.004$, Figure 3.3a). However, the cumulative body weight increases in both groups were similar during the 12-week period (Figure 3.3b), suggesting a possible higher metabolic efficiency of RS4 diet than that of the control diet.

**Cecal butyrate increased after RS4 intake in mice (butyrogenic effect of RS4)**

We have observed increased fecal butyrate levels post RS4 consumption in humans with MetS [119], which was expected due to RS4 being a non-digestible fermentable fiber [25]. However fecal SCFAs represent <5% of the total SCFAs while the majority of SCFAs are efficiently absorbed in the intestinal lumen [103]. In contrast, the cecum butyrate concentrations from a MetS mouse model may represent a higher proportion of the SCFAs produced in the gut. The mice group receiving RS4 showed a trend for higher cecal weight compared to that of the CD group (322 ± 76 mg vs 135 ± 5 mg, $p = 0.07$, Figure 3.4a), which is in line with a previous report where fermentation of RS type 2 and RS type 3 increased the cecal weight in mice [136]. Using these cecal samples, GC-MS analyses showed that butyrate concentration was about two-fold higher (0.81 ± 0.12 mg/g vs 0.39 ± 0.05 mg/g, $p = 0.036$) in the RS4 group than the CD group (Figure 3.4a).
Dietary RS4 intake enriched H3K27me3 in the promoter region of NFκB1 in mouse intestinal tissue

After observing NaBu-mediated inhibition of histone modifications in vitro, we sought to examine the similar effects in the colon tissue from KK.Cg-A^v/a mice fed with 20% RS4 for 12 weeks. ChIP assay revealed enrichment of H3K27me3 modification on the promoter region of NFκB1 (1.4 fold, $p = 0.01$) after RS4 intervention compared with CD (Figure 3.4b). This result supports the concept that feeding a butyrogenic dietary fermentable fiber, such as RS4, enriches the epigenetic repressive mark (H3K27me3), potentially contributing to the amelioration of the gut inflammation, a key factor underlying many chronic disorders including MetS.

3.5. Discussion

Supplementation of diets with prebiotic fibers for potential mitigation of pro-inflammatory state in the context of chronic diseases, is an attractive area in public health research [117]. It was particularly proposed that the inhibition of inflammatory mediators through dietary enhancement of intestinal butyrate production has tremendous implications in the context of managing obesity-related metabolic diseases [137]. To our knowledge, this is the first report that dietary RS4 induced promoter-specific changes in the H3K27me3 enrichment in vivo likely by the way of increased colonic butyrate production. The in vitro observations further confirmed a concentration dependent repression of both NFκB1 mRNA and protein levels in response to NaBu with a concomitant enrichment in H3K27me3 levels.
Dietary fibers, which are otherwise indigestible to humans, undergo microbial fermentation in the hind-gut, particularly in the colon [138]. The colon is considered a metabolic organ where the mucosal epithelia absorb topical nutrients, such as SCFAs which play an important role in modulating gut immune system [139]. The gut microbiota is an essential component in the colonic microenvironment and an altered gut microbial composition or dysbiosis plays a critical role in the pathogenesis of intestinal as well as extra-intestinal diseases [140]. Butyrate acts as an energy source by epithelial cells [35] and inhibits genetic mutation [120] and oncogenic microRNA expression [141] in rectal biopsies associated diet-induced manipulation. In addition, in vivo research has investigated the anti-inflammatory role of SCFAs in regulating colonic regulatory T cells (cT\textsubscript{Reg} cells) in mice [142, 143]. Butyrate increases the expression of FOXP3, a transcription factor functioning in intestinal inflammation, by increasing activity of histone acetylation in its promoter and enhancer region [142, 144] and consequences downregulation of pro-inflammatory mediators [145]. As we [119] and others [24] have observed that RS4 mediates the modulation of butyrogenic gut microbiota, we hypothesized that the observed immunomodulatory effects of RS4 in humans are at least partially derived from the anti-inflammatory mechanisms of butyrate that is produced after microbial fermentation of the dietary RS4. To minimize use of animals, we first examined our hypothesis in an in vitro setting using a human intestinal cell line but since RS4 is insoluble [28, 146], we used sodium butyrate for the in vitro tests.

The cytokine IL-10 is predominantly anti-inflammatory in nature that suppresses the dendritic and macrophage cell functions [147]. The IL-10 may, however, acquire pro-inflammatory properties in inflammatory settings like endotoxemia, autoimmune diseases,
and graft transplantations [148-150]. The balance between pro- and anti-inflammatory activities of IL-10 is regulated by type I IFN [132]. Priming with type I IFNs (IFNα or IFNβ) also leads to enhanced cellular responsiveness to IFNγ [151], the latter acting as an enhancer of cellular responsiveness to LPS through the toll-like receptor 4 signaling pathway in human intestinal epithelial cells, including SW480 [152]. Therefore, in our study, it is possible that IL-10 showed a pro-inflammatory nature as the human intestinal epithelial cells were primed with IFNγ (10 ng/mL) and stimulated with the LPS (10 ng/ml), following Suzuki et al. [152]. Our results are also in line with the report of Saemann et al. that showed NaBu exerts its IL-10 enhancing properties only under the dose of 0.25 mM, while inhibits IL-10 beyond the concentration of 0.25 mM [153]. This finding further supports our observation of IL-10 inhibition by NaBu at the concentration range of 0.31 to 1.25 mM. Also of interest is that the normal physiologic level of SCFAs in portal blood is (0.38 ± 0.07 mM) [154], signifying the concentrations used in our study are likely within the physiologically relevant range.

Although congenic, age-matched, and randomized mice were supplied with isocaloric control diet, the cumulative food intake was consistently higher in the RS4 group. It is possible that palatability of RS4 and cellulose were different with RS4, imparting minimal alteration of the physicochemical and organoleptic properties of the final food products [28], the mouse chow in this case. Interestingly, in spite of higher food intake, but with expected similar basal metabolic rate and physical activity, the mice in the RS4 group did not have a significantly higher weight gain by the end of twelfth week. This is consistent with the previous findings by Gao et al., [27] where butyrate supplemented mice consumed higher feed intake and showed higher lean mass but lower
fat mass, thus preventing diet-induced insulin resistance and obesity through more energy expenditure. This is the limitation of our study that we could not examine the body mass composition in the mice under study. In addition, the dietary fibers may also slow down the digestibility of protein and fats in diets, thereby decreasing the metabolizable energy content [155]. Since we also previously reported that RS4 consumption improves body composition of humans [25], the current study further corroborates the previous findings in KK.Cg-Ay/a mice and supports the suitability of this mouse model for age-related MetS in humans as proposed by Kennedy et al [134].

The novelty of this research lies in the investigation of epigenetic mechanisms of butyrate on H3K27me3-enrichment that inversely correlated with concentration dependent down-regulation of a transcription factor NFκB1, which elegantly substantiates our previous findings of health promoting [25] and butyrogenic effects of RS4 [119]. Hence, RS4 and its derivative, butyrate, possesses promising clinical implications in the management of cardio-metabolic diseases in humans.

3.6. Conclusions

This exploratory study introduces a proof-of-concept that RS4 and its bacterial fermentation metabolite, butyrate, may function as an H3K27me3 modulator in conjunction with its suppressive effects on NFκB1 in the intestinal cells in the context of metabolic syndrome. Further characterization of prebiotic functional fibers-associated epigenetic alterations may provide effective strategies to mitigate the low-grade inflammation in the context of intestinal as well as extra-intestinal diseases associated with dysbiosis.
3.7. Future Directions

1. To include more of epigenetic marks, including both repressive and expressive marks, on more pro-inflammatory and anti-inflammatory mediators.

2. To expand this colon cancer cell model into other disease models

3. To precisely measure the metabolic efficiency of mice fed with RS4 so that we could explain why higher feed intake of RS4 did not result in high weight gain.
Figure 3.1. Effects of NaBu on inflammatory gene expression in human intestinal epithelial cells
(a) Concentration-dependent mRNA expression of inflammatory genes relative to a housekeeping gene, GAPDH. Values are expressed as mean ± SEM (n = 6), ***p < 0.001 compared with LPS treated (10 ng/ml) positive control. (b) Histone 3 lysine 27 trimethylation (H3K27me3) changes at promoter regions of inflammatory genes in cells treated with 1.25 mM NaBu for 5 h. Average percentage input ± SEM from each experiment (n ≥ 3) is plotted. *p < 0.05 compared with LPS treated positive control cells. Dr. Yi Liu helped with this figure.
Figure 3.2. Concentration-dependent effects of NaBu on NFκB1 in human intestinal epithelial cells.

(a) Representative immunoblots showing the suppression of total cellular NFκB1 subunits p105 and p50. Densitometric analyses showing relative protein expressions normalized to β-actin proteins and expressed as mean ± SEM (n = 3). *p < 0.05, compared with LPS (10 ng/mL) treated positive control. (b) Histone 3 lysine 27 trimethylation (H3K27me3) changes at promoter regions of NFκB1 in cells treated with different concentrations (0.31 to 1.25 mM) of NaBu for 5 h. Data points represent the input ± SEM (n ≥ 4). *p < 0.05, compared with LPS control. Dr. Yi Liu helped partly with this figure.
Figure 3.3. Effects of RS4 on feed intake and body weight in KK.Cg- Aγ/a mice
Cumulative food intake (a) and cumulative body weight gain (b) over a 12-week period in six-week old KK.Cg-Aγ/a mice fed with 20% RS4 (resistant starch) or control diet. Data points represent the mean ± SEM (n = 3), *p < 0.05, **p < 0.01. Dr. Yi Liu helped with this figure.
Figure 3.4. Effects of RS4 on mice cecal weight, butyric acid levels and histone modification
(a) Butyrate concentration (primary Y-axis on the left) per gram of cecal tissue (secondary Y-axis on the right) taken from KK.Cg-Aγ/a mice after 12 week of feeding RS4 (20%) or control diet. Data represent the mean ± SEM (n = 3), *p < 0.05. (b) Diet-induced differential tri-methylated histone 3 modifications on lysine 27 (H3K27me3) at promoter regions NFκB1 in the colon tissues of KK.Cg-Aγ/a mice fed with either RS4 (resistant starch, 20%) or control diet for 12 weeks. Average percentage input ± SEM from each experiment (n = 3) is plotted, **p < 0.01. Dr. Yi Liu helped partly with this figure.
<table>
<thead>
<tr>
<th>Gene</th>
<th>qRT PCR (SW480 cells)</th>
<th>ChIP Assay (SW480 cells)/ (Mice tissue)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CCL2</strong></td>
<td>F: 5’-agttctctgccgcccttct-3’</td>
<td>F: 5’-gtggtcagtcttggtatagtaatg-3’</td>
</tr>
<tr>
<td></td>
<td>R: 5’-gtgactggggcattggattg-3’</td>
<td>R: 5’-ctgctgagaccaaatagca-3’</td>
</tr>
<tr>
<td><strong>IL-10</strong></td>
<td>F: 5’-cataaatttaggtetccaaatc-3’</td>
<td>F: 5’-aagggcaatttaatccaaggtt-3’</td>
</tr>
<tr>
<td></td>
<td>R: 5’-aagggcggtgggtcagctat-3’</td>
<td>R: 5’-tttgtttctctaccttactgt-3’</td>
</tr>
<tr>
<td><strong>NFκB1</strong></td>
<td>F: 5’-acctgtacctctctcttttcttctc-3’</td>
<td>F: 5’-ttggcggcccccccaagag-3’/</td>
</tr>
<tr>
<td></td>
<td>F: 5’-agctcttttccccagatcttc-3’</td>
<td>F: 5’-cgatctgtgtagccgaga-3’</td>
</tr>
<tr>
<td></td>
<td>R: 5’-agctctttttccccagatcttc-3’</td>
<td>R: 5’-ggtttccccagatctttt-3’/</td>
</tr>
<tr>
<td></td>
<td>R: 5’-gctgggccaaagttcaatc-3’</td>
<td></td>
</tr>
<tr>
<td><strong>GAPDH</strong></td>
<td>F: 5’-agccacatgctcagac-3’</td>
<td></td>
</tr>
<tr>
<td></td>
<td>R: 5’-gccaatacgaccaatcc-3’</td>
<td></td>
</tr>
</tbody>
</table>
Table 3.2. Composition of experimental diets per 100 grams

<table>
<thead>
<tr>
<th>Ingredients</th>
<th>Control Diet</th>
<th>RS4 Diet</th>
</tr>
</thead>
<tbody>
<tr>
<td>MidsolTM 50&lt;sup&gt;1&lt;/sup&gt;</td>
<td>31.77</td>
<td>27.07</td>
</tr>
<tr>
<td>Fibersym RW&lt;sup&gt;2&lt;/sup&gt;</td>
<td>0.00</td>
<td>23.5</td>
</tr>
<tr>
<td>Powdered cellulose</td>
<td>18.80</td>
<td>0.00</td>
</tr>
<tr>
<td>Casein – vitamin free</td>
<td>14.00</td>
<td>14.00</td>
</tr>
<tr>
<td>Dextrin</td>
<td>13.67</td>
<td>13.67</td>
</tr>
<tr>
<td>Sucrose</td>
<td>8.82</td>
<td>8.82</td>
</tr>
<tr>
<td>Soybean oil</td>
<td>8.00</td>
<td>8.00</td>
</tr>
<tr>
<td>AIN 93M mineral mix</td>
<td>3.50</td>
<td>3.50</td>
</tr>
<tr>
<td>AIN 93M vitamin mix</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Choline Bitartrate</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>L-Cystine</td>
<td>0.18</td>
<td>0.18</td>
</tr>
<tr>
<td>t-Butylhyrodoquinone</td>
<td>0.0008</td>
<td>0.0008</td>
</tr>
</tbody>
</table>

<sup>1</sup> MidsolTM 50 is wheat starch with energy content 359.5 Kcal/100 g.
<sup>2</sup> Fibersym RW is resistant starch type 4 (85%, dry basis) with energy content 56.5 Kcal/100 g
CHAPTER 4. MATERNAL HIGH-FAT DIET AND GESTATIONAL DIABETES AFFECT HISTONE MODIFICATIONS ON CARDIO-METABOLIC GENE PROMOTERS IN OFFSPRING

Related publication: Upadhyaya, B. et al. Maternal high-fat diet and gestational diabetes affect histone modifications on cardio-metabolic gene promoters in offspring (2016) [In preparation]

4.1. Abstract

A nutrient-mediated fetal metabolic programming \textit{in utero} due to maternal high-fat (HF) diet and gestational diabetes mellitus (GDM) increases the risk of the adult-onset of obesity, type 2 diabetes mellitus, and cardiovascular disease (CVD) in offspring. However, the precise mechanism of excess circulatory lipids in cardiac tissue-specific epigenetic programming remains poorly understood. In this pilot study, we seek to investigate a genome-wide chromatin landscape due to histone modifications (H3Ac, H3K4me3, and H3K27me3) in rat offspring heart tissue in response to either maternal HF (40% kcal) diet or streptozotocin-induced GDM, or both using a previously validated rat model. We observed that gestational diabetes was associated with cardiac hypertrophy ($p<0.05$). Chromatin immunoprecipitation sequencing of offspring’s heart tissue revealed the differential peaks distribution (FDR<0.1) due to Ac, H3K4me3 and H3K27me3 histone modifications on the promoter regions mapped to rat genome. Downstream gene ontology analyses showed that out of 28 annotated genes, supposedly expressed (H3K4me3 gained peaks, FDR<0.1) in high fat-diet group, 15 genes were 1) annotated to
“metabolic process” in general and “positive regulation of cholesterol biosynthetic process” (FDR=0.0295) in particular; 2) overlapped with a total of 455 quantitative trait loci (QTL), primarily consisting of 66 QTL for blood pressure (19.34%), 24 QTL for body weight (5.27%), and 20 QTL for serum cholesterol (4.40%); 3) related to susceptibility or disease progression to several CVDs revealed through disease ontology analyses. Together, these findings show that maternal HF with or without GDM changes the cardiac epigenetic signature through histone modifications in rat offspring. This differential fuel-mediated epigenetic programming help assess the risks for obesity, diabetes, and CVDs beyond one generation, potentially leading to novel preventive and treatment strategies.

**Keywords:** Cardiovascular disease, ChIP sequencing, developmental programing, gestational diabetes, histone modifications, maternal high-fat diet

### 4.2. Introduction

Gestational diabetes mellitus (GDM), one of the most common complications in pregnancy, occurs when hyperglycemia (> 92–125 mg/dL, fasting glucose) associated glucose intolerance is first diagnosed at any time of pregnancy, or in between 24 and 28 weeks in particular [41]. 17.8% (9.3-25.5%) of overall pregnancies in the US are affected by GDM [42] and its occurrence parallels to the increasing obesity epidemics [39], which shows 34.4% of US women are obese in their child-bearing stage [40]. In addition to well-known risk factor of genetic susceptibility, over nutrition of the fetus in utero is another critical factor to shape the metabolic profile of offspring for the rest of their life [43, 44]. Siblings born to a diabetic mother possess a significantly higher risk of
developing diabetes later in life than in those born to the same mother before being diagnosed with diabetes [156]. Similarly, embryo transfer experiments in rats [157] further provided compelling evidence to support that adult metabolic consequences are caused by the diabetic environment in utero.

The food plan that tightly controls the blood glucose level is often the first recommended strategy to manage GDM. However, the problems may arise when dietary carbohydrates go down (from ~60% to ~40%) and the compensatory fats usually go up (from ~20% to ~40%) [45] in order to meet the caloric needs of a rapidly growing fetus. Since heart constantly demands a high-energy source to maintain its contractile activity that requires a turnover of entire ATP pool in every 10 seconds, it has the ability to utilize all fuel sources including fatty acids, which indeed is the primary energy source for the normal adult heart [46]. Hence, the quantity as well as quality of the dietary fat during pregnancy plays a major role in cardiometabolic health outcomes in offspring above and beyond the pregnant mother [47]. Offspring of obese or diabetic mother have significant risk of cardiovascular diseases (CVD) at birth [158, 159] and throughout the life [160]. Interestingly, since studies show that many women even with good glycemic control have infants with CVD [159, 161], mechanistic research, which links adult nutrigenomic consequences to fuel-mediated fetal programming due to exposure of excess circulatory lipids in utero [162], are critically needed especially in combination with GDM.

It has recently become apparent that posttranslational modification (PTM), which may be epigenetically inherited, are involved in gene expression or repression during developmental nutritional programing [163]. DNA methylation and histone modification, two major types of epigenetic marks that also crosstalk with each other, lead to a stable
long-term gene repression and a reversible short-term gene expression/repression, respectively [164]. These modifications alter gene expression by changing the conformation of the chromatin states into open or closed, hence modulating the accessibility of the promoter regions, upstream to the transcription start site, to bind with transcription factors and RNA Polymerase II before the gene transcription starts [165]. Only a few studies show maternal HF diet during pregnancy regulates rat offspring’s gene expression through histone modifications in liver-specific manner [166-168]. However, the genome-wide chromatin landscape in rat offspring heart tissue in response to maternal obesity and diabetes in pregnancy remains unknown.

Since we already reported that a maternal HF diet further impairs cardiac functions due to lipid droplet accumulation, metabolic disturbances, oxidative stress, and mitochondrial dysfunction in rat offspring born to diabetic dams [169], this follow-up (proof-of-concept) study seek to investigate three most widely studied histone PTMs, namely acetylated histone 3 (H3Ac, a gene activation mark), trimethylated lysine at position 4 of histone 3 (H3K4me3, a gene activation mark), and trimethylated lysine at position 27 of histone (H3K27me3, a gene repression mark), following prenatal exposure to maternal HF diet and late gestation diabetes, either individually or in combination, using the previously validated rat model.

4.3 Methods

Animal study and dietary intervention

Female Sprague-Dawley rats (Harlan Laboratories Inc., Indianapolis, IN) were fed either control diet (TD2018 Teklad, Harlan Laboratories, Madison, WI)) or high-fat
diet (TD95217 custom diet Teklad, Harlan Laboratories, Madison, WI) for at least 28 days prior to breeding and throughout the pregnancy (Figure 4.1), following the guidelines of Institutional Animal Care and Use Committee at Sanford Research. All rats were housed in a temperature-controlled, light-dark cycled facility with water and chow supplied ad libitum. 18% and 40% of calories came from saturated and monounsaturated fat content in case of control diet (CD) and high-fat diet (HF), respectively, as we recently reported [169]. To induce gestational diabetes on day 14, pregnant dams were injected intraperitoneally either with 65mg/kg of streptozotocin (STZ, Sigma Life Sciences, St. Louis, MO) or 0.09 M citrate buffer (CB) as a placebo. Maternal late gestation blood glucose levels were monitored twice daily by tail nick sampling to a One Touch Ultra meter (LifeScan Inc., Milpitas, CA). The offspring were delivered normally on gestation day 22. Finally, neonate rat hearts from 4 different groups- CD with CB (CC), CD with STZ (CS), HF with CB (HC), and HF with STZ (HS)- were collected and snap-frozen after euthanizing the rat pups soon after the delivery.

**Chromatin immunoprecipitation (ChIP) assay**

Five to six frozen heart tissues were pooled to make >250mg/sample and homogenized followed by chromatin isolation using P-2001 ChromaFlashTM Chromatin Extraction Kit (Epigentek, Farmingdale, NY). Chromatin was sheared using Episonic2000 Sonication System (Epigentek, Farmingdale, NY) in 300 µl of ChIP buffer, followed by a quality control of sheared chromatin using fluorescence quantification. 4 µl of sheared chromatin was purified to obtain DNA, which was eluted with 20 µl of water. DNA fragment (100-300 bp) was quality-checked using Bioanalyzer analysis. For each
ChIP reaction, 10 µg of chromatin, 3 µg of H3K4me3 polyclonal antibody or 3 µg of H3K27me3 antibody or 3 µg of H3 Acetyl polyclonal antibody (all from Epigentek, Farmingdale, NY) and 6 µl of protein A/G beads were used. ChIPed DNA was eluted in 20 µl of water. The H3K4me3, H3K27me3 and H3 Acetyl antibodies were tested and proven to be ChIP-grade.

**ChiP sequencing and analyses**

A ChIP-seq library was prepared using the EpiNext ChIP-Seq High Sensitivity Kit with amplification, DNA end polishing and adaptor ligation, following the manufacturer’s instruction. 10 nM of sample libraries were provided for next generation sequencing on a HiSeq 2500. The basic analysis of ChIP-Seq was based on the published protocol [170] utilizing Bowtie, version 1.0.0 [171] and MACS, version 2.1.0 [172]. Raw reads were quality checked using FastQC, version v0.10.1 [173] and mapped onto the rat RN5 genome sequence using Bowtie, version 1.0.0 [171]. The option of “-m 1” was activated so that only uniquely mapped reads were allowed to map. The mapping results in SAM were converted to BAM and sorted according to coordinates using samtools, version 0.1.19 [174]. Mapping results of each ChIP sample and the input sample were subjected to ChIP enriched peak calling. The option of “--broad” was activated to optimize the calling algorithm for broad binding regions. The called peaks are annotated to the nearest TSS (Transcript Starting Site) using ChIPpeakAnno, version 3.2.0 [175]. ChIP quality control was performed using ChIPQC, version 1.2.2 [176]. All raw and processed sequencing data are deposited to NCBI GEO database under the accession number GSE84831.
**Bioinformatics analyses**

Enriched peaks were called and then annotated to the nearest transcription start site of the genes. Peaks were visualized using Integrative Genomics Viewer (PMCID: PMC3346182). Genomic distribution of peaks were summarized in a heat map using Seqminer version 1.3.3 [177]. First, the genomic reference peak set (TSS.rat.Rnor_5.0.bed) was imported. Then the mapping results (BAM) were loaded; read densities were extracted and clustered (linear Kmeans clustering) using the default parameters. Differential binding analyses was carried out to determine the differential enrichment of cardiac-specific genes due to 4 different diet-conditions in all regions and in promoter region only. Promoter regions were further validated by having CpG islands using online database [178]. Likewise, a gene ontology analysis [179] was carried out to determine which biological processes that were triggered by 4 different diet-conditions. Enriched gene ontology analysis was carried out using online database GOEAST [180]. Rat genome database [181] was extensively used to annotate the gene, overlapping quantitative trait loci (QTL) and disease ontologies.

**Statistical analyses**

Statistical analyses were performed with Sigma Plot (Systat Software, Inc. San Jose, CA). To compare two diet-groups, a t-test was used. When the data were not normal, a Mann-Whitney U test was used. Principal coordinate analysis was used to show the difference among three histone modifications. The genome mapping statistics was extracted using Picard, version 1.90 [182]. Differential binding analyses was carried out
using a DiffBind Bioconductor R package [183]. Data were expressed as mean±SEM. A value of \( p<0.05 \) or FDR<0.1 was considered significant.

4.4. Results

**Maternal late gestation STZ-induced diabetes increases the offspring’s heart weight**

To analyze the effects of maternal late gestational diabetes on the heart weight of the offspring, we compared the average weight of the hearts obtained from the pups whose mothers were fed either control diet (C) or high fat (H) diet with (CS and HS) or without (CC and HC) streptozotocin (S)-induced diabetes. Pups from CS group had significantly higher average heart weight compared to their control CC (55.64±1.95 mg vs. 48.94±2.03 mg, \( p=0.030 \)). A similar effect of STZ was observed in high fat diet-fed mice groups combined with STZ (HS) when compared to diet-matched control group HC (52.92±1.62 mg vs. 44.91±2.30 mg, \( p=0.013 \)) (Figure 4.2a). In addition, pups from CS group had higher heart to body weight ratio when compared to control CC group (0.0089±0.0004 vs. 0.0071±0.0002, \( p=0.001 \)). Although we observed the similar trend in HS vs. HC groups, the heart to body weight ratio were not significantly different (\( p=0.118 \)). However, HS group had significantly higher heart to body weight ratio when compared to the CC control group (0.0082±0.0003 vs. 0.0071±0.0002, \( p=0.006 \)) (Figure 4.2b). Interestingly, these significantly higher heart weight in the offspring from the STZ groups corresponded to a trend of elevated maternal late gestation glucose levels induced by STZ when compared to CC and HC diet-matched control groups (305.30±92.01 mg/dL vs. 82.23±3.92 mg/dL and 295.42±39.27 mg/dL vs. 88.30±2.61 mg/dL,
respectively) (Figure 4.2c). Taken together, gestational diabetes is related to enlargement of offspring’s heart, perhaps due to fuel mediated energy homeostasis.

**Differential histone modification profiles in cardiac tissue of offspring related to maternal HF diet and GDM in pregnancy**

Since we recently showed that the maternal high-fat diet alters the cardiac functions in offspring of diabetic pregnancies [169], we further seek to investigate the intergenerational effects of high-fat diet and gestational diabetes epigenetically through histone modifications. Using ChIP-grade antibodies, ChIP assays followed by ChIP sequencing revealed the differential peak distribution in offspring’s heart tissue due to Ac, H3K4me3 and H3K27me3 modifications, which were responsive to control or HF diet with or without STZ induction during pregnancies (Figure 4.3). A total of around 41000, 73000 and 54000 significant peaks (FDR<0.1) were observed due to H3Ac, H3K27me3 and H3K4me3 histone modifications, respectively, in control diet group. When we mapped peaks to the rat genome, only 29%, 11% and 36% of the total peaks overlapped to the known exonic regions (data not shown), which shows that most of H3K27me3 peaks occurred in non-exonic and gene regulatory elements. These differential histone modification profiles were further validated with principal component analyses that showed 24% and 20% of variations were explained by principal component axes 1 and 2, respectively. Both being an expressive epigenetic mark, H3Ac and H3K4me3 modifications clustered together, separated from a repressive epigenetic mark, H3K27me3 (Figure 4.4a). However, the effects of diet were not so distinct within each category of histone modification. Since we carried out ChIP sequencing from the
offspring’s’ heart whose mothers, not the offspring themselves, were fed a specific diet, we observed only subtle differences in the genome wide histone marks for H3Ac, H3K4me3, and H3K27me3 (Figure 4.4b). Interestingly, in 5kb upstream regions, combination of diabetes and high fat diet in gestation (HS group) showed relatively higher and lower mean densities (tag/50bp) than the individual ones (CS or HC alone) in case of expressive marks (H3Ac and H3K4me3) and repressive mark (H3K27me3), respectively (Figure 4.4c). Additionally, when we compared the different histone modifications from a representative high fat-diet group, H3K4me3 showed a much higher mean density that was slightly correlated (Pearson’s r=0.29) with another expressive H3Ac mark but without any correlation (Pearson’s r=0.09) with a repressive H3K27me3 mark. Together, both the expressive histone marks, H3Ac and H3K4me3, functions similarly, which is different to that of H3K27me3.

**Diet-specific differential histone modifications in the gene promoter region**

In order to locate the histone modifications in the annotated genomic region, ± 5kb distances to the nearest transcriptional start site (TSS) of genes annotated to rat genome 5.0 were identified and plotted. Around 40% of all the peaks were enriched in the upstream (-5kb) of any known TSS, which was the consistent pattern for all four diet-groups. In case of H3Ac modification, we observed a decrease in the total number of peaks in 5kb upstream of any known TSS in diabetes-induced (CS group) by 52% and in high fat diet-fed (HC group) by 53%. However, we found the opposite trend where CS and HC groups had additional 33% and 73% of K4me3 specific peaks and 56% and 70% of K27me3 specific peaks, respectively (Figure 4.5). Interestingly, K4me3 modification
also showed that the high-fat and diabetes combined (HS group) had an additional 89% of peaks when compared to the control (CC group), which is already 9.5% more than its diet-matched control (HC group) (Figure 4.5). In case of H3Ac, differential binding analyses (FDR< 0.1) revealed that CS and HC groups had two and three significantly lost peaks, respectively, in 5kb upstream gene promoter region when both were compared to CC group (Table 1). H3K27me3 showed a mixed response: six peaks were significantly gained for each CS and HC group, while three and one peaks were significantly lost for CS and HC groups, respectively (Table 4.1 and Table 4.2). However, two and 28 peaks were significantly gained in CS and HC groups, respectively, when compared to CC group in case of H3K4me3. Out of these 28 annotated genes supposedly expressed in high fat-diet group, 15 genes were related to “metabolic process” in the gene ontology category of biological process, validating the critical role of HF diet in GDM due to fetal metabolic programming in utero.

**Maternal high fat diet aggravates gestational diabetes that associates with metabolic consequences in the offspring**

Since both high fat diet and diabetes closely link to each other, we hypothesize a common gene network regulating metabolic process. As expected, H3K4me3 enriched two closely related genes, heat shock protein 1a1 (Hspa1a) and Hspa1b, were found to be associated with of CS diet (6.54 fold, FDR=0.02) and HC diet (5.62 fold, FDR<0.001)(Table 4.2). Remarkably, other maternal high-fat diet (HC group) induced offspring’s H3K4me3 enriched genes were ATP synthase mitochondrial Fo complex, subunit C3, subunit 9 (Atp5g3, 6.10 fold, FDR=0.07); Cytochrome P450, family 4,
subfamily f, polypeptide 18 (Cyp4f18, 5.06 fold, FDR=0.01); autophagy associated transmembrane protein (El24, 2.79 fold, FDR=0.02); endoplasmic reticulum oxidoreductase beta (Ero1lb, 6.66 fold, FDR=0.002); farnesyl diphosphate synthase (Fdps, 1.9 fold, FDR=0.06); hexose-6-phosphate dehydrogenase (H6pd, 6.75 fold, FDR<0.001); P450 (cytochrome) oxidoreductase (Por, 6.16 fold, FDR=0.06); solute carrier family 11, member 2 (Slc11a2, 6.7 fold, FDR<0.001) and tripartite motif containing 63, E3 ubiquitin protein ligase (Trim63, 4.99 fold, FDR<0.001).

On the other hand, since H3K27me3 is a repressive mark, we considered the loss of peaks as a gain of gene function. We found several candidate genes on which H3K27me3 peaks were significantly lost in HF diet combined with GDM (HS group) when compared to GDM alone (CS group), suggesting the expression of those genes associated to further metabolic stress and diabetes. These candidate genes included ATP binding cassette subfamily B member 9 (Abcb9, 5.6 fold, FDR=0.02); Atp5g2 (5.92 fold, FDR=0.003); lysine demethylase 6B (Kdm6b, 5.37 fold, FDR=0.08); LIM domain binding 1 (Ldb1, 5.73 fold, FDR=0.01); nuclear receptor subfamily 4, group A, member 3 (Nr4a3), 3.41 fold, FDR=0.004); RuvB-like AAA ATPase 2 (Ruvbl2, 3.40 fold, FDR=0.05); Solute carrier family 44, member 4 (Slc44a4, 6.17 fold, FDR<0.001); steroid sulfatase (microsomal), isozyme S (Sts, 1.82 fold, FDR=0.08); urocortin 2 (Ucn2, 2.61 fold, FDR=0.06). Similarly, a significant gain of H3K27me3 peaks in two candidate genes were supposed to be repressed in HS group when compared to CS group. These candidate genes were adenylate kinase 3 (Ak3, 1.95 fold, FDR=0.09) and signal transducer and activator of transcription 5B (Stat5b, 3.76 fold, FDR=0.09). Taken together, maternal high fat diet not only shares with gestational diabetes some candidate
genes that affect the metabolic consequences in the offspring but also aggravates the gestational diabetes.

**Maternal high fat diet and gestational diabetes affect the genes and QTLs related to metabolic diseases in the offspring**

Some of the gene ontologies related to mitochondrial function were enriched among those 15 genes of interest in case of H3K4me3 due to high fat diet. Interestingly, both diabetic diet (CS) and high-fat diet (HC) groups showed H3K4me3 peaks gain in heat shock protein 1a and 1b in common. Enriched gene ontology analyses revealed that both Hspa1a and Hspa1b were enriched for “primary metabolic process” in HF-induced H3K4me3 modification (log2 odd ratio 1.38, FDR=0.06) (Table 3). Moreover, enriched gene ontology of “positive regulation of cholesterol biosynthetic process” (FDR=0.0295) was associated to farnesyl dipospho synthase (Fdps) and P450 (cytochrome) oxidoreductase (Por) (Table 4.3 and Supplementary Table 4.1). Despite the genes of interest belong to different chromosomes or genomic regions, we found that those 15 genes overlap with a total of 455 QTLs, primarily consisting of 66 QTLs for blood pressure (19.34%), 24 QTLs for body weight (5.27%), and 20 QTLs for serum cholesterol (4.40%) (Table 4). Particularly, Fdps overlapped with a total of 33 QTLs, out of which 22 QTLs (66.67%) were related to blood pressure. Atp5g3 overlapped with a total of 22 QTLs, out of which 9 QTLs (40.91%) were related to body weight. Similarly, Por and Ero1lb overlapped with a total of 12 and 19 QTLs, respectively, out of which 2 (16.67%) and 3 (15.79%) QTLs, respectively, were related to serum cholesterol (Table 4.4, suggesting the putative role of HF diet in hypercholesterolemia. Furthermore, disease
ontology revealed that Hspa1a gene was related to the disease progression of systolic heart failure and other cardiovascular diseases like coronary disease, atherosclerosis, and acute coronary syndrome. Hspa1b, a close relative of Hspa1a, was related to obesity, type 2 diabetes mellitus, and hyperlipidemia. H6PD was related to experimental diabetes mellitus and inborn errors on metabolism, further supporting the role of HF diet either alone or in combination with GDM on the fetal metabolic programming of adult-onset diabetes and cardiovascular diseases.

4.5. Discussion

A successful pregnancy should primarily be defined not only by the outcome at birth but also by the health status in later life [44]. A growing number of animal models and human epidemiological studies have shown that fetal metabolic programming play a crucial role in perpetuating the vicious cycle of cardiometabolic conditions through one generation to the other [184] but the molecular mechanisms remain poorly understood. To this context, we previously reported that maternal HF diet (40% kcal from fat) aggravates the gestational diabetes through lipid droplet accumulation, metabolic abnormalities, oxidative stress, and mitochondrial dysfunction in a rat model [169]. To our knowledge, this follow-up study is the first to further explain the heart tissue-specific epigenetic marks in the rat offspring born to diabetic and obese pregnancy, utilizing a genome-wide ChIP sequencing.

Using a rat model of gestational diabetes with or without maternal obesity [169], we introduced STZ-induced diabetes in pregnant rats at gestational day 14 (Figure 4.1), simulating GDM in humans that occurs in third-trimester of pregnancy [41]. This is
appropriate because the late GDM is less likely associated with congenital heart, which mostly occurs in early pregnancy [185, 186], minimizing the confounding cardiovascular outcomes due to maternal diet. HF diet, the second half of the intervention in this study, included more saturated (~13% kcal from fat) and monounsaturated (~17% kcal from fat) fatty acids but with similar polyunsaturated fatty acids (~8% kcal from fat) when compared to CD diet, showing a translational or feasible diet choice for typical Americans [187]. More recently paternal diet has also been shown to play a key role in the metabolic programming of the offspring [188]. In our study, female rats were bred with normal control diet-fed males, ruling out the possible effects of the paternal diet on the observed outcomes.

The fuel-mediated teratogenesis hypothesis shows the association between maternal hyperglycemia in pregnancy and obesity or cardiovascular complications in fetus [156]. In our study, late gestational maternal glucose was elevated, as expected, after STZ-induced diabetes (Figure 4.2c). However, the concomitant high level of maternal insulin that we previously observed [169] cannot freely cross the placenta to the fetus [189]. Therefore, in order to respond to this high glucose load the developing fetal pancreas produces additional insulin, causing fetal hyperinsulinemia as we and others previously reported [169, 190]. Interestingly, insulin resistance, a biomarker of type 2 diabetes, has also been associated with altered cardiac singling and the development of hypertension-associated left ventricular hypertrophy [190, 191]. This aligns with our observation that weight of the offspring hearts (Fig 2a) and heart to body ratio (Figure 4.2b) were significantly higher in STZ-induced diabetic group (CS) when compared to control group (CC).
Since histone modification, mostly of short-term and reversible in nature, usually precedes the DNA methylation that facilitates a long-term gene repression [164], here in this proof-of-concept study we focused only few well-studied representative histone marks: H3Ac, H3K4me3, and H3K27me3 using a validated ChIP-grade antibodies. These choices were based on the proposal that developmental specification is accompanied by a striking transition from a permissive chromatin state with widespread remodeling to a restrictive state with pervasive Polycomb repression [192]. Broadly, the oxidative stress-mediated regulation cascade is the common mechanistic link among the pathogenesis of diabetes, hypertension, and other related inflammatory diseases [193]. In our study, several H3K4me3 peaks, which are the gene activating marks, were differentially gained in the promoter regions of genes pertaining to metabolic stress regulation, including Atp5g3 associated with mitochondrial injury [194], Ei24 reported to have an effect on autophagy [195], H6pd associated with fuel metabolism and experimental diabetes mellitus [196], Erol1b associated with decrease insulin synthesis leading to diabetes mellitus [197], SLC11A2 increased with body fatness [198], CYP4F associated with oxidation of inflammatory lipid substrates [199], and Trim63 associated with several human heart conditions, such as hypertrophic cardiomyopathy and chronic heart failure [200].

Notably, Hspa1a and Hspa1b were the common candidate genes in HF and CS treated groups, showing a degree of common pathways of obesity and diabetes. Indeed, increased Hspa1a is not only linked with HF-associated oxidative stress in metabolic syndrome patients [201] but also correlates with HbA1c values in women with GDM [202]. This could be the reason why our disease ontology analyses linked Hsp1a1 and
Hspa1b with severe disease progression to systolic heart failure [203] and CVD [204]. Furthermore, Por and Fdps showed enriched “metabolic process” gene ontology in general and “positive regulation of cholesterol biosynthesis process” in particular (Table 3). Por is associated with microsomal P450s-mediated xenobiotic metabolism as well as cholesterol and bile acid synthesis [205], while Fdps is associated with cholesterol biosynthesis and hypercholesterolemia [206]. Interestingly, all of these annotated genes overlapped with “blood pressure” QTLs while most of them overlapped with “cholesterol synthesis” QTLs, which clearly demonstrates that the epigenetic cues already in place at birth affect the adult-onset diseases like atherosclerosis, hypertension, and CVD. These results further corroborate our previous findings that HF diet in GDM causes oxidative stress and mitochondrial dysfunction [169].

Since H3k4me3 (active) and H3K27me3 (inactive) modifications co-exist to form bivalent promoters in early life, promoters quickly respond to the developmental stimuli and change their histone modification state mainly through H3K27me3 demethylation [207]. Interestingly, when we compared the effects of HF diet on diabetic background (HS vs. CS groups), a lysine demethylase enzyme, Kdm6b, was associated with a significant H3K27me3 loss [208]. This suggests that several candidate genes with H3K27Me3, including Kdm6b, were likely upregulated when HF diet is provided in GDM. To this context, our candidate genes, ABCB9 (associated with type 2 diabetes [209]; ATP5G2 (associated with insulin secretion [210]; Ldb1 [211], NR4A3 [212] and SLC44A4 [213](all associated with β-cell and pancreatic islets functions); RUVBL2 (associated with insulin-stimulated GLUT4 translocation [214]; Sts (associated with HF diet and ob/ob models of obesity and type 2 diabetes [215]; and Ucn2 (associated with
heart failure patients [216], lost a repressive H3K27me3 mark and were likely expressed. Interestingly, in case of H3K27me3 peaks gain in the promoter regions of AK3 (associated with downregulation of insulin secretion [217]) and STAT5B (associated with negative regulation in adipogenesis [218] further validated the HF-induced gene repression of AK3 and STAT5B, leading to more insulin secretion and adipogenesis.

Hence, epigenetic processes represent a central underlying mechanism of developmental programming of T2D and CVD [219]. In a similar study to ours, rat adult offspring, exposed to a maternal HF diet in utero, at 6 months of age showed significantly higher body weights, greater adiposity, and increased triacylglycerol levels than that of control rats in spite of their similar weights at birth [220]. To this context, we also expect that it might take a longer time to manifest more phenotypic outcomes because there were no significant differences in the birth weights observed among offspring in different groups (data not shown). Indeed, we have already collected heart tissues of the weaned pups (3 weeks age) and adult rats (12 weeks age) from the same breeding mates kept in the same diet groups. Due to limited resources, only the neonatal hearts were analyzed for ChIP sequencing to report in this pilot study. Our prospective and comprehensive longitudinal study will also encompass whether the effects of post-natal dietary habits at different life stage of the offspring can alter the epigenetic landscape set at the birth, making them either less or more prone to adult-onset CVD. Since the fetal primordial germ cells, that determines the fate of F2 generation is also likely affected by intrauterine environment [219], studies highlighting the importance of HF diet and GDM in the context of intergenerational and transgenerational (F3 generation and beyond) metabolic programming [221] above and beyond the first
generation is imperative. Moreover, considering the gender-specific fetal response to the same environmental insults due to maternal as well as paternal diet would reveal more information on differential metabolic profiles due to sexual dimorphism and genomic imprinting. Besides, the functional microarray or RNA Seq would allow us to delineate the functionality of epigenetic marks annotated to the nearest transcription start sites. Therefore, characterizing these altered epigenetic marks at-birth by taking the surrogate tissues like placenta and cord blood with a minimal invasive technique could function as the first-step in developing a clinical biomarker based on the intra-uterine environment, which merits further study on screening the fetus at an increased risk of adult-onset chronic diseases.

4.6. Conclusion

This follow-up study of fuel-mediated cardiac dysfunction in offspring [169] is the first to further explain the role of maternal HF diet on heart tissue-specific differential histone modifications (H3Ac, H3K4me3, H3K27me3) in the rat offspring born to obese and/or diabetic pregnancy, utilizing a genome-wide ChIP sequencing and bioinformatics analyses. Identification of this reversible epigenetic landscape in utero and its association to risks for adult-onset of chronic CVDs provides an effective strategy for novel cardio-metabolic health management in next generation and future generation
4.7. Future Directions

1. To carry out ChIP sequencing and RNA sequencing on the same tissue so that the histone modifications and putative functionality can be compared at the same time.

2. To carry out ChIP sequencing of heart tissues collected from different age groups (e.g., 3 weeks or 12 weeks in rats) so that the changing epigenetic landscape can be monitored.

3. To carry out transgenerational metabolic programming study that involves feeding of mother or father or both and observing the epigenetic effects in 3 or more generations to come.
Figure 4.1: Study timeline
Female Sprague-Dawley rats were randomized into four groups and fed high fat diet (HF) or control diet (CD) for 28 days prior to their breeding and throughout the pregnancy. On gestational day 14, each HF or CD group was treated with intraperitoneal injection of Streptozocin (STZ, 65 mg/Kg) or citrate buffer (CB) as a control. Pups were normally delivered on gestational day 21 and sacrificed for their hearts collection. One of our collaborators, Dr. Michelle Baack developed and carried out the rat model used in this study.

Figure 4.2: Physiological characteristics pups and their dams under study
Bar graphs showing the higher neonatal heart weight (a) and neonatal heart to body weight ratio (b) that corresponds to elevated maternal late gestation glucose levels (c) in Streptozotocin-induced diabetes groups. n= 10-12 (a and b), n= 2 (c). *p<0.05, **p<0.01, ***p<0.001 (t-test). CC, controls; CS, diabetes induced; HC, high-fat diet fed; HS, high-fat fed and diabetes induced. One of our collaborators, Dr. Michelle Baack analyzed the characteristics of dams and pups under study.
Figure 4.3: Peaks detection in ChIP sequencing
Representative genomic region of a negative strand (chromosome 7) near the transcriptional start site of Atp5g2 gene, illustrating 5kb upstream peaks in cyan (acetalyation), red (H3K427me3), blue (H3K4me3), and grey (Input control) colors. CC, controls; CS, diabetes induced; HC, high-fat diet fed; HS, high-fat fed and diabetes induced.
Figure 4.4: Differential histone modification profiles

a) Principal component analyses showing different profiles among three histone modifications. Axes 1 and 2 possess 24 and 20% of variations, respectively. b) Genome wide differential histone marks in H3Ac, H3K4me3 and H3K27me3 modifications due to different diets. Read density in +/- 5 kb of any known transcriptional start site for each marker are clustered and visualized in a heat map, each of which consists of four panels in the order of CC (control), CS (diabetes-induced), HC (high fat diet-fed), and HS (combined HC and CS) from left to right. From the vertical view, clusters with the central density profile are located at the top, followed by clusters with peaks in the -5 kb region (marked under the black lines), those with peaks in the +5 kb region, those with diffuse read density, and those without significant read density. c) Corresponding mean density (tag/50bp) plots for H3Ac, H3K4me3 and H3K27me3 modifications in +/- 5 kb of any known transcriptional start site due to different diets.
Figure 4.5. Peaks overlapping any 5kb upstream region of known exons

For each histone modifications of H3Ac (Ac), H3K4me3 (K4) and H3K27me3 (K27), peaks with 5kb upstream of any known TSS are counted and plotted as indicated. CC, controls; CS, diabetes induced; HC, high-fat diet fed; HS, high-fat fed and diabetes induced
Table 4.1. Diet-specific differential binding of peaks in respective histone modification in genomic and 5kb upstream region

<table>
<thead>
<tr>
<th>Histone Modification</th>
<th>CC vs CS (gained/lost peaks)</th>
<th>CC vs HC (gained/lost peaks)</th>
<th>CS vs HS (gained/lost peaks)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All region</td>
<td>Promoter Region</td>
<td>All region</td>
</tr>
<tr>
<td>H3Ac</td>
<td>0/32</td>
<td>0/2</td>
<td>5/103</td>
</tr>
<tr>
<td>H3K4me3</td>
<td>8/6</td>
<td>2/0</td>
<td>449/11</td>
</tr>
<tr>
<td>H3K27me3</td>
<td>186/82</td>
<td>6/3</td>
<td>309/114</td>
</tr>
</tbody>
</table>

CC, controls; CS, diabetes induced; HC, high-fat diet fed; HS, high-fat fed and diabetes induced
Table 4.2. List of “metabolic process” candidate genes with differentially bound peaks

<table>
<thead>
<tr>
<th>Modification</th>
<th>Diets</th>
<th>Genes</th>
<th>Ensemble ID</th>
<th>Fold</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>H3Ac</td>
<td>CC vs CS</td>
<td>Myrf</td>
<td>ENSRNOG00000028274</td>
<td>-5.53</td>
<td>1.59E-02</td>
</tr>
<tr>
<td>H3K4me3</td>
<td>CC vs CS</td>
<td>Hspa1a</td>
<td>ENSRNOG00000005647</td>
<td>6.54</td>
<td>2.49E-02</td>
</tr>
<tr>
<td></td>
<td>CC vs CS</td>
<td>Hspa1b</td>
<td>ENSRNOG00000005647</td>
<td>6.54</td>
<td>2.49E-02</td>
</tr>
<tr>
<td></td>
<td>CC vs HC</td>
<td>Atp5g3</td>
<td>ENSRNOG0000001596</td>
<td>6.10</td>
<td>7.63E-02</td>
</tr>
<tr>
<td></td>
<td>CC vs HC</td>
<td>Cyp4f18</td>
<td>ENSRNOG00000015751</td>
<td>5.06</td>
<td>1.27E-02</td>
</tr>
<tr>
<td></td>
<td>CC vs HC</td>
<td>Ei24</td>
<td>ENSRNOG00000030391</td>
<td>2.79</td>
<td>2.42E-02</td>
</tr>
<tr>
<td></td>
<td>CC vs HC</td>
<td>Ero1lb</td>
<td>ENSRNOG0000002609</td>
<td>6.66</td>
<td>1.96E-03</td>
</tr>
<tr>
<td></td>
<td>CC vs HC</td>
<td>Fdps</td>
<td>ENSRNOG00000043377</td>
<td>1.92</td>
<td>5.66E-02</td>
</tr>
<tr>
<td></td>
<td>CC vs HC</td>
<td>H6pd</td>
<td>ENSRNOG00000017523</td>
<td>6.75</td>
<td>7.39E-04</td>
</tr>
<tr>
<td></td>
<td>CC vs HC</td>
<td>Hspa1a</td>
<td>ENSRNOG0000005647</td>
<td>5.62</td>
<td>3.76E-05</td>
</tr>
<tr>
<td></td>
<td>CC vs HC</td>
<td>Hspa1b</td>
<td>ENSRNOG0000005647</td>
<td>5.62</td>
<td>3.76E-05</td>
</tr>
<tr>
<td></td>
<td>CC vs HC</td>
<td>Mapk8ip2</td>
<td>ENSRNOG00000032828</td>
<td>6.16</td>
<td>6.08E-02</td>
</tr>
<tr>
<td></td>
<td>CC vs HC</td>
<td>Napsa</td>
<td>ENSRNOG00000019854</td>
<td>6.18</td>
<td>5.66E-02</td>
</tr>
<tr>
<td></td>
<td>CC vs HC</td>
<td>Nfix</td>
<td>ENSRNOG0000002983</td>
<td>2.09</td>
<td>8.40E-02</td>
</tr>
<tr>
<td></td>
<td>CC vs HC</td>
<td>Por</td>
<td>ENSRNOG0000001442</td>
<td>6.16</td>
<td>6.08E-02</td>
</tr>
<tr>
<td></td>
<td>CC vs HC</td>
<td>Slec11a2</td>
<td>ENSRNOG00000019550</td>
<td>6.73</td>
<td>9.55E-04</td>
</tr>
<tr>
<td></td>
<td>CC vs HC</td>
<td>Trim63</td>
<td>ENSRNOG00000016543</td>
<td>5.00</td>
<td>9.51E-04</td>
</tr>
<tr>
<td></td>
<td>CC vs HC</td>
<td>Zim1</td>
<td>ENSRNOG00000015071</td>
<td>6.28</td>
<td>3.43E-02</td>
</tr>
<tr>
<td>H3K27me3</td>
<td>CC vs CS</td>
<td>Bmp8b</td>
<td>ENSRNOG00000033209</td>
<td>2.25</td>
<td>3.03E-02</td>
</tr>
<tr>
<td></td>
<td>CC vs CS</td>
<td>Gli1</td>
<td>ENSRNOG00000025120</td>
<td>2.77</td>
<td>6.51E-02</td>
</tr>
<tr>
<td></td>
<td>CC vs CS</td>
<td>Rps29</td>
<td>ENSRNOG00000032542</td>
<td>3.73</td>
<td>7.29E-02</td>
</tr>
<tr>
<td></td>
<td>CC vs CS</td>
<td>Snapc3</td>
<td>ENSRNOG00000010825</td>
<td>4.51</td>
<td>7.29E-02</td>
</tr>
<tr>
<td></td>
<td>CC vs CS</td>
<td>Ccdc126</td>
<td>ENSRNOG0000009322</td>
<td>-4.10</td>
<td>9.17E-02</td>
</tr>
<tr>
<td></td>
<td>CC vs HC</td>
<td>Dstyk</td>
<td>ENSRNOG00000021298</td>
<td>2.57</td>
<td>8.56E-02</td>
</tr>
<tr>
<td></td>
<td>CC vs HC</td>
<td>Lhpp</td>
<td>ENSRNOG00000017097</td>
<td>5.63</td>
<td>1.69E-02</td>
</tr>
<tr>
<td>Gene</td>
<td>ENSRNOG ID</td>
<td>Log2 Fold Change</td>
<td>p-value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>------------------</td>
<td>------------------</td>
<td>-----------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zfp444</td>
<td>ENSRNOG00000015517</td>
<td>-5.83</td>
<td>4.77E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ak3</td>
<td>ENSRNOG00000015273</td>
<td>1.95</td>
<td>9.12E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stat5b</td>
<td>ENSRNOG00000019075</td>
<td>3.76</td>
<td>8.66E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abcb9</td>
<td>ENSRNOG0000001082</td>
<td>-5.60</td>
<td>2.22E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atp5g2</td>
<td>ENSRNOG00000015320</td>
<td>-5.92</td>
<td>3.41E-03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B3gnt3</td>
<td>ENSRNOG00000018764</td>
<td>-5.79</td>
<td>7.71E-03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm6b</td>
<td>ENSRNOG00000037613</td>
<td>-5.37</td>
<td>7.80E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ldb1</td>
<td>ENSRNOG00000018468</td>
<td>-5.73</td>
<td>1.10E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metap1</td>
<td>ENSRNOG00000012947</td>
<td>-5.58</td>
<td>2.46E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nfic</td>
<td>ENSRNOG0000004505</td>
<td>-4.58</td>
<td>3.65E-03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nr4a3</td>
<td>ENSRNOG0000005964</td>
<td>-3.41</td>
<td>3.70E-03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ruvbl2</td>
<td>ENSRNOG00000020793</td>
<td>-3.40</td>
<td>5.43E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slc44a4</td>
<td>ENSRNOG000000878</td>
<td>-6.18</td>
<td>2.79E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Snapc3</td>
<td>ENSRNOG00000010825</td>
<td>-4.17</td>
<td>4.90E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sts</td>
<td>ENSRNOG00000032487</td>
<td>-1.82</td>
<td>7.95E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ucn2</td>
<td>ENSRNOG00000020655</td>
<td>-2.61</td>
<td>5.70E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uqcrq</td>
<td>ENSRNOG00000048174</td>
<td>-5.73</td>
<td>1.10E-02</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CC, controls; CS, diabetes induced; HC, high-fat diet fed; HS, high-fat fed and diabetes induced.
Table 4.3. List of “metabolic process” candidate genes (high fat diet) with gene ontology enrichment

<table>
<thead>
<tr>
<th>Gene</th>
<th>Enriched GO ID</th>
<th>Description of GO</th>
<th>Log₂ Odds Ratio</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atp5g3</td>
<td>GO:1901360</td>
<td>Organic cyclic compound metabolic process</td>
<td>1.96</td>
<td>0.0976</td>
</tr>
<tr>
<td></td>
<td>GO:0044238</td>
<td>Primary metabolic process</td>
<td>1.38</td>
<td>0.0636</td>
</tr>
<tr>
<td>Cyp4f18</td>
<td>GO:1901360</td>
<td>Organic cyclic compound metabolic process</td>
<td>1.96</td>
<td>0.0976</td>
</tr>
<tr>
<td></td>
<td>GO:0044248</td>
<td>Cellular catabolic process</td>
<td>2.87</td>
<td>0.0976</td>
</tr>
<tr>
<td></td>
<td>GO:0044238</td>
<td>Primary metabolic process</td>
<td>1.38</td>
<td>0.0636</td>
</tr>
<tr>
<td></td>
<td>GO:0016709</td>
<td>Oxidoreductase activity</td>
<td>6.13</td>
<td>0.0976</td>
</tr>
<tr>
<td>Ei24</td>
<td>GO:0030308</td>
<td>Negative regulation of cell growth</td>
<td>4.80</td>
<td>0.0726</td>
</tr>
<tr>
<td>Fdps</td>
<td>GO:0045542</td>
<td>Positive regulation of cholesterol biosynthetic process</td>
<td>8.68</td>
<td>0.0195</td>
</tr>
<tr>
<td></td>
<td>GO:1901360</td>
<td>Organic cyclic compound metabolic process</td>
<td>1.96</td>
<td>0.0976</td>
</tr>
<tr>
<td></td>
<td>GO:0044238</td>
<td>Primary metabolic process</td>
<td>1.38</td>
<td>0.0636</td>
</tr>
<tr>
<td>H6pd</td>
<td>GO:1901360</td>
<td>Organic cyclic compound metabolic process</td>
<td>1.96</td>
<td>0.0976</td>
</tr>
<tr>
<td></td>
<td>GO:0044238</td>
<td>Primary metabolic process</td>
<td>1.38</td>
<td>0.0636</td>
</tr>
<tr>
<td>Hspa1a</td>
<td>GO:0008180</td>
<td>COP9 signalosome</td>
<td>6.31</td>
<td>0.0976</td>
</tr>
<tr>
<td></td>
<td>GO:1901360</td>
<td>Organic cyclic compound metabolic process</td>
<td>1.96</td>
<td>0.0976</td>
</tr>
<tr>
<td></td>
<td>GO:0044248</td>
<td>Cellular catabolic process</td>
<td>2.87</td>
<td>0.0976</td>
</tr>
<tr>
<td></td>
<td>GO:0090083</td>
<td>Regulation of inclusion body assembly</td>
<td>8.09</td>
<td>0.0312</td>
</tr>
<tr>
<td></td>
<td>GO:0044238</td>
<td>Primary metabolic process</td>
<td>1.38</td>
<td>0.0636</td>
</tr>
<tr>
<td>Hspa1b</td>
<td>GO:0008180</td>
<td>COP9 signalosome</td>
<td>6.31</td>
<td>0.0976</td>
</tr>
<tr>
<td></td>
<td>GO:0043281</td>
<td>Regulation of cysteine-type endopeptidase activity</td>
<td>4.43</td>
<td>0.0976</td>
</tr>
<tr>
<td></td>
<td>GO:1901360</td>
<td>Organic cyclic compound metabolic process</td>
<td>1.96</td>
<td>0.0976</td>
</tr>
<tr>
<td></td>
<td>GO:0044248</td>
<td>Cellular catabolic process</td>
<td>2.87</td>
<td>0.0976</td>
</tr>
<tr>
<td></td>
<td>GO:0090083</td>
<td>Regulation of inclusion body assembly</td>
<td>8.09</td>
<td>0.0312</td>
</tr>
<tr>
<td></td>
<td>GO:0044238</td>
<td>Primary metabolic process</td>
<td>1.38</td>
<td>0.0636</td>
</tr>
<tr>
<td>Gene</td>
<td>GO:xxxxx</td>
<td>Process</td>
<td>Value</td>
<td>P-value</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
<td>----------------------------------------------</td>
<td>-------</td>
<td>---------</td>
</tr>
<tr>
<td>Mapk8ip2</td>
<td>GO:0044238</td>
<td>Primary metabolic process</td>
<td>1.38</td>
<td>0.0636</td>
</tr>
<tr>
<td>Napsa</td>
<td>GO:0044238</td>
<td>Primary metabolic process</td>
<td>1.38</td>
<td>0.0636</td>
</tr>
<tr>
<td>Nfix</td>
<td>GO:1901360</td>
<td>Organic cyclic compound metabolic process</td>
<td>1.96</td>
<td>0.0976</td>
</tr>
<tr>
<td></td>
<td>GO:0044238</td>
<td>Primary metabolic process</td>
<td>1.38</td>
<td>0.0636</td>
</tr>
<tr>
<td>Por</td>
<td>GO:0045542</td>
<td>Positive regulation of cholesterol biosynthetic process</td>
<td>8.68</td>
<td>0.0195</td>
</tr>
<tr>
<td></td>
<td>GO:0043281</td>
<td>Regulation of cysteine-type endopeptidase activity involved in apoptotic process</td>
<td>4.43</td>
<td>0.0976</td>
</tr>
<tr>
<td></td>
<td>GO:0044248</td>
<td>Cellular catabolic process</td>
<td>2.87</td>
<td>0.0976</td>
</tr>
<tr>
<td></td>
<td>GO:0044238</td>
<td>Primary metabolic process</td>
<td>1.38</td>
<td>0.0636</td>
</tr>
<tr>
<td></td>
<td>GO:0016709</td>
<td>Oxidoreductase activity</td>
<td>6.13</td>
<td>0.0976</td>
</tr>
<tr>
<td>Slc11a2</td>
<td>GO:0043281</td>
<td>Regulation of cysteine-type endopeptidase activity involved in apoptotic process</td>
<td>4.43</td>
<td>0.0976</td>
</tr>
<tr>
<td></td>
<td>GO:1901360</td>
<td>Organic cyclic compound metabolic process</td>
<td>1.96</td>
<td>0.0976</td>
</tr>
<tr>
<td></td>
<td>GO:0044238</td>
<td>Primary metabolic process</td>
<td>1.38</td>
<td>0.0636</td>
</tr>
<tr>
<td>Trim63</td>
<td>GO:0044248</td>
<td>Cellular catabolic process</td>
<td>2.87</td>
<td>0.0976</td>
</tr>
<tr>
<td></td>
<td>GO:0044238</td>
<td>Primary metabolic process</td>
<td>1.38</td>
<td>0.0636</td>
</tr>
</tbody>
</table>
### Table 4.4. Metabolic process candidate QTLs obtained from the candidate genes (high fat diet specific)

<table>
<thead>
<tr>
<th>Gene</th>
<th>QTL Total</th>
<th>Blood Pressure</th>
<th>Body Weight</th>
<th>Glucose Level</th>
<th>Insulin Level</th>
<th>Insulin Dependent Diabetes Mellitus</th>
<th>Non-insulin Dependent Diabetes Mellitus</th>
<th>Serum Cholesterol</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(Percent Total)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atp5g3</td>
<td>22</td>
<td>7</td>
<td>9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(31.82%)</td>
<td>(40.91%)</td>
<td>(4.55%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(4.55%)</td>
</tr>
<tr>
<td>Cyp4f18</td>
<td>14</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(21.43%)</td>
<td>(21.43%)</td>
<td></td>
<td>(14.29%)</td>
<td></td>
<td></td>
<td></td>
<td>(7.14%)</td>
</tr>
<tr>
<td>Ei24</td>
<td>19</td>
<td>7</td>
<td>1</td>
<td>2</td>
<td></td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>(36.84%)</td>
<td>(5.26%)</td>
<td>(10.53%)</td>
<td>(5.26%)</td>
<td></td>
<td></td>
<td></td>
<td>(10.53%)</td>
</tr>
<tr>
<td>Ero1lb</td>
<td>19</td>
<td>7</td>
<td>1</td>
<td>2</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>(36.84%)</td>
<td>(5.26%)</td>
<td>(10.53%)</td>
<td>(5.26%)</td>
<td></td>
<td></td>
<td></td>
<td>(15.79%)</td>
</tr>
<tr>
<td>Fdps</td>
<td>33</td>
<td>22</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(66.67%)</td>
<td>(9.09%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(6.06%)</td>
</tr>
<tr>
<td>H6pd</td>
<td>12</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(58.33%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(16.67%)</td>
</tr>
<tr>
<td>Hspa1a</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(33.33%)</td>
<td>(33.33%)</td>
<td>(33.33%)</td>
<td>(33.33%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hspa1b</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(33.33%)</td>
<td>(33.33%)</td>
<td>(33.33%)</td>
<td>(33.33%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mapk8ip2</td>
<td>12</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>(50.00%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(16.67%)</td>
</tr>
<tr>
<td>Napsa</td>
<td>16</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>(43.75%)</td>
<td>(6.25%)</td>
<td>(6.25%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(12.50%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Nfix</strong></td>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(25.00%)</td>
<td></td>
<td></td>
<td>(25.00%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Por</strong></td>
<td>12</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(50.00%)</td>
<td>(8.33%)</td>
<td>(8.33%)</td>
<td>(8.33%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(16.67%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Slc11a2</strong></td>
<td>8</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(62.50%)</td>
<td></td>
<td></td>
<td>(12.50%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Trim63</strong></td>
<td>19</td>
<td>7</td>
<td>3</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(36.84%)</td>
<td>(15.79%)</td>
<td></td>
<td>(5.26%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(21.05%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Zim1</strong></td>
<td>6</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(16.67%)</td>
<td>(33.33%)</td>
<td></td>
<td>(16.67%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(16.67%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 5. PHENETHYL ISOTHIOCYANATE INHIBITS ALDEHYDE DEHYDROGENASE ACTIVITY AND TARGETS REACTIVE OXYGEN SPECIES-MEDIATED APOPTOSIS IN HUMAN CANCER STEM CELLS

Related publications:


Upadhyaya, B. et al. Phenethyl isothiocyanate inhibits aldehyde dehydrogenase activity and targets reactive oxygen species-mediated apoptosis in human cancer stem cells (2016) [In preparation].

5.1. Abstract

Aldehyde dehydrogenase family 1 (ALDH1) is a cancer stem cell (CSC) marker that associates with tumor malignancy and self-renewal properties of stem cells in different tumors. Tumor cells also have higher load of oxidative stress than normal cells where as cancer stem cells (CSCs) are more stable because of the redox balance machinery. Mounting evidence has supported that further exposure of reactive oxygen species (ROS)-load will exhaust their antioxidant potential and induce even more apoptosis in CSCs compared to cancer cells. However, how phenethyl isothiocyanate (PEITC), a prooxidant and an inhibitor of liver ALDH2, selectively targets cervical CSCs has remained unknown. Hence, we seek to investigate whether PEITC could inhibit ALDH1 and cell proliferation and induce ROS in HeLa CSCs (hCSCs). Results showed
that sphere-culture method significantly enriched CSCs from HeLa cells \((p<0.05)\). Flow cytometry revealed that PEITC attenuated ALDH\(^{hi}\) cells in hCSCs \((p<0.05)\). ROS assay showed that PEITC induced oxidative stress in hCSCs, which was glutathione dependent. PEITC also inhibited a transcription factor Sp1 and multidrug resistance protein, P-glycoprotein in hCSCs \((p<0.05)\). Xenotransplantation study in a non-obese diabetic, severe combined immunodeficient (NOD/SCID) mouse model and histopathology showed anti-tumorogenic and anti-metastatic potential of PEITC in vivo. Taken together, this proof-of-concept study showed that PEITC effectively targets multi-drug resistant and metastatic hCSCs through inhibition of ALDH enzyme and Sp1 transcription factor and induction of oxidative stress and apoptosis in cervical hCSCs. Therefore, PEITC commonly found in cruciferous vegetables may offer a novel approach to target CSCs for improving therapeutic outcomes in cancer patients.

**Key words:** ALDH, cancer stem cells, oxidative stress, phenethyl isothiocyanate, Sp1 protein

### 5.2. Introduction

Cancer stem cells (CSCs), a subset of cancer cells (0.1-30\%) with self-renewal and differentiation potential, have been implicated in tumor initiation, maintenance, metastasis, and recurrence of a number of cancer types [50, 51]. Survival of CSCs from chemo or radiotherapy include expression of drug efflux transporters, expression of anti-apoptotic proteins, enhanced activity of repair enzyme, hypoxic stability, vascular niche, and dormancy [51]. To identify and isolate CSCs from within a heterogeneous population of cancer cells, several CSC putative surface markers have been proposed, such as
CD133\(^+\), CD24\(^-\) and CD44\(^+\), that often associate with aggressive cancer types and poor prognosis [222, 223]. Recent studies have also shown that aldehyde dehydrogenase family 1 (ALDH1) is a CSC marker that associates with tumor malignancy and self-renewal properties of stem cells in different tumors, including breast [52], head and neck [224], melanoma [225], hepatic [226], pancreatic [227], colon [53], lung [228], prostate [229], bladder [230] ovarian [231], endometrial [232] and cervical [54, 55] cancers.

Hence, CSCs-targeted therapy is of growing interest and concern given the fact that conventional chemo or radiotherapy cannot and does not eradicate the drug-resistant CSCs [51]. This is due to the activity of drug efflux transporter proteins or ATP binding cassette (ABC) transporters, such as P-glycoprotein (P-gp), breast cancer resistance protein (BCRP) and multidrug resistance-associated proteins (MRP), which usually overexpress in drug-resistant CSCs [233]. Particularly, P-gp having a broad substrate specificity is considered as a main impeding factor in the anticancer therapy [234]. Since P-gp protein is mediated through expression of MDR1 gene, activation of MDR1 promoter region gets activated with specificity protein (Sp1) [235]. Sp1 being a ubiquitous transcription factor for Sp1-regulated genes involved in cell proliferation and cancer development [236], there is a need of anti-cancer compound that does not easily get excreted out of the cell and target Sp proteins in CSCs.

Besides drug-resistance potential, CSCs possess another survival machinery of enhanced redox homeostasis. Due to increased oxidative metabolism and shortage of nutrient supply rapidly growing cancer cells have higher levels of reactive oxygen species (ROS) than normal healthy cells [237]. Hence, one established strategy to induce apoptosis and kill only cancer cells is to further increase the ROS levels, since the
threshold to resist elevated ROS levels is usually low in cancer cells when compared to normal cells. However, the challenges arise due to the fact that the CSCs possess lower redox stress due to high levels of anti-oxidant machineries [238], such as glutathione (GSH) and lipid peroxidation enzyme, aldehyde dehydrogenase (ALDH). ALDH enzymes irreversibly catalyze the oxidation of aldehydes to their respective carboxylic acids, thereby reducing the cellular oxidative stress [239]. Therefore, to exploit the similar effect of elevated ROS in CSCs, naturally occurring pro-oxidant dietary compounds is essential because of their no side effects even in the high doses.

Phenethyl isothiocyanate (PEITC) is a biologically active dietary compound present in cruciferous vegetables, including watercress, broccoli, cabbage, and cauliflower [56, 240]. Others and we have previously reported the anti-inflammatory [57, 58] and chemopreventive characteristics of PEITC against various cancers [59, 61, 240]. Because of its non-toxicity [63] and high efficacy among other isothiocyanates (ITCs) [241] PEITC is currently under phase II clinical trials for lung cancer (NCT00691132). PEITC depletes GSH and selectively inhibits proliferation of cancer cells due to elevated ROS-accumulation inside the cells [242, 243]. Interestingly, PEITC also inhibits liver ALDH enzyme in the context of ethanol administration in rats [244]. However, the pro-oxidant effect of PEITC in inhibiting ALDH1 in the context of CSCs and more specifically cervical CSCs is entirely unknown.

Apoptosis, or programmed cell-death, is essential to maintaining tissue homeostasis, and its impairment is implicated in many human diseases, including cancers [245]. Since we recently showed that PEITC targets CSCs through apoptosis [61], in this study we hypothesized that PEITC induces ROS-mediated apoptosis in CSCs and inhibits
a CSC marker ALDH1 and drug efflux transporter protein P-gp, using a cervical cancer HeLa sphere culture model *in vitro* and non-obese diabetic/severe combined immune deficit (NOD/SCID) mouse model *in vivo*.

5.3. Materials and Methods

**Sphere cultures of hCSCs**

Following our previous work [61], human cervical HeLa cell line (CCL-2, American Type Culture Collection, Manassas, VA) was cultured and maintained in a T-25 flask with Dulbecco’s modified eagle’s medium (DMEM) containing 4 mM L-glutamine and 4.5 g/L glucose (HyClone, Logan, UT), supplemented with 10% heat-inactivated fetal bovine serum (Invitrogen, Grand Island, NY) and 1% penicillin (25 U/ml)/streptomycin (25 µg/ml) (Sigma-Aldrich, St. Louis, MO) in a 5% CO₂-humidified atmosphere at 37°C. HeLa cells were trypsinized with TrypLE (Invitrogen, Grand Island, NY) and then sub-cultured with a 1:5 splitting ratio when the cells reached about 90% confluency. From the parental HeLa cells CSCs (termed simply as hCSCs in the rest of the document) were cultured following a modified protocol described by Gu et al. [246]. Briefly, single-cell suspensions of HeLa cells (4×10⁴) were seeded into a 100-mm ultra-low attachment (ULA) petri dish (Corning Inc., Corning, NY) containing 8 ml of serum-free mammary epithelial basal medium (MEBM, Lonza, Allendale, NJ), supplemented with 1x B27 (Invitrogen, Grand Island, NY), 4 µg/ml heparin (Sigma-Aldrich, St. Louis, MO), 20 ng/ml hEGF, and 20 ng/ml hFGF (Invitrogen, Grand Island, NY). After an initial 5-day culture in suspension at 37°C, an additional 9 ml of sphere culture medium was added for another 5 days of culture. On day 10, spheres were harvested by
centrifugation at 500 $x$ g for 3 minutes, followed by washing with phosphate-buffered saline (PBS), trypsinization with TrypLE for 10 minutes at 37 °C, centrifugation at 500 $x$ g for 3 minutes, resuspension in 5 ml of hCSC culture medium, and counting with a hemocytometer. Both HeLa cells and hCSCs were used for flow cytometry experiments.

**Aldefluor assay**

The Aldelfluor Kit (STEMCELL Technologies, Vancouver, Canada) was used to determine the enzymatic activity of ALDH in HeLa and hCSCs. hCSCs were treated with PEITC (Sigma-Aldrich, St. Louis, MO) or dimethyl sulfoxide (DMSO, Sigma-Aldrich, St. Louis, MO) 24 hr prior to cell harvesting. Cells were trypsinized, washed with 1x PBS, and counted using a hemocytometer. Cells were then resuspended in Aldefluor buffer at a concentration of $2 \times 10^5$ cells/mL. Activated Aldefluor reagent (300 µM) was prepared following the manufacturer’s instructions. PEITC Cells were incubated with activated Aldefluor reagent and its control reagent, diethylaminobenzaldehyde (DEAB) at 37°C for 45 min, with intermittent shaking after every 15 min. Following incubation, all the tubes were centrifuged at 250 $x$ g for 5 min at 4°C. Aldefluor-stained cells were washed once in Aldefluor buffer, and maintained in 4°C until acquired through FL1 channel in flow cytometer (FACSCalibur, Becton, Dickinson and Company, San Jose, CA) as ALDH high and side scatter low phenotypes.

**ROS assay**

Cells were harvested, trypsinized, washed with 1x PBS, and counted using a hemocytometer before collecting in conical test tube. Using Cellular ROS Detection
Assay Kit (Abcam, San Francisco, CA), the cells were stained with 20 µM 2’, 7’ – dichlorofluorescin diacetate (DCFDA) in 1X Buffer supplied for 30 minutes at 37 °C following the manufacturer’s protocol. 2 x 10^5 hCSCs per sample were aliquoted into a six-well plate and incubated with 10 µM PEITC or DMSO control for 3, 6, or 9 hr at 37 °C. GSH (Sigma-Aldrich, St. Louis, MO) was used as an inhibitor of the ROS production and a pro-oxidant hydrogen peroxide (H_2O_2) was used as a positive control in the assay. The read signal was detected at Ex485 nm/Em535 nm using a flow cytometer.

**Flow cytometry**

Cells were washed with 2 ml of PBS, trypsinized with 1 ml of TrypLE, and resuspended in 1 ml of PBS, followed by immunostaining. Similarly, hCSCs were collected after 10 days of culture, trypsinized, and resuspended in 2 ml of PBS with a density of 1x10^6 cells/ml, followed by immunostaining. Cells were immunostained with anti-CD24–FITC (1:500 v/v, Millipore, Billerica, MA) or anti-CD44–FITC (1:500 v/v, Millipore, Billerica, MA) antibodies for 1 hour at room temperature. Immunofluorescence was measured using FACSCalibur cell analyzer (Becton Dickinson, San Jose, CA) with approximately 10,000 events in each sample. Propidium iodide/annexin V staining was performed according to the manufacturer’s instructions. Briefly, 5x10^5 cells were centrifuged and resuspended in 100 µl of 1x binding buffer (Invitrogen, Grand Island, NY). The cells were treated with 10 µM PEITC or vector control (DMSO) for a total of 24 h. The cells were then incubated with 5 µl of annexin V–FITC (eBioscience, Inc., San Diego, CA) and 5 µl of propidium iodide (eBioscience,
Inc., San Diego, CA) at room temperature for 5 minutes in the dark before acquiring at least 10,000 cells in a flow cytometer.

**Sphere-formation assay**

The hCSCs were enriched in spheres in serum-free medium. Sphere culture was carried out as previously described in the sphere culture section. Cells were treated with predetermined doses of 2.5, 5 or 10 µM of PEITC or DMSO as control. After 7 days incubation, photomicrographs of spheres were acquired under an inverted phase-contrast microscope (Olympus America Inc., Center Valley, PA), and the number of hCSCs was counted using a hemocytometer.

**Cell proliferation assay**

A standard colorimetric method (MTS assay) was used to determine the number of viable cells, which were harvested and counted with a hemocytometer before seeding into 96-well microplates at a density of 2×10^4 cells per well. Cells were cultured in DMEM supplemented with 100 U/ml penicillin, 100 µg/ml streptomycin, 5% heat-inactivated FBS, and 50 µM 2-mercaptoethanol. hCSCs were treated with three concentrations of PEITC (2.5, 5, and 10 µM). After 24 and 48 hours of incubation, 20 µl of CellTiter reagent was added directly to the cell-culture wells and incubated for 1 hour at 37 °C, followed by cell viability assessment using the CellTiter 96 AQueous One Solution Cell Proliferation Assay kit (Promega, Madison, WI), containing [3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium, inner salt; MTS]. The manufacturer’s instructions were followed, and treatments were
compared with vehicle control (DMSO-treated cells) at 490 nm in a BioTek Synergy H4 multimode plate reader (BioTek, Winooski, VT).

**Immunoblotting**

hCSCs \((1 \times 10^6)\) were seeded in each well of a 6-well plate and incubated overnight at 37°C in a 5% CO\(_2\) incubator. Old culture medium was replenished by culture medium containing 1.25, 2.5, 5 or 10-µM concentrations of PEITC for 3 hours. Cell harvesting and immunoblotting were carried out as we previously reported [59]. Briefly, cells were lysed in ice-cold RIPA buffer containing 150 mM NaCl, 50 mM Tris (pH 8.0), 10% glycerol, 1% Nonidet P-40 (NP-40), and 0.4 mM EDTA, followed by a brief vortexing and rotation for 30 minutes at 4 °C. Equal amounts (v/v) of cell lysates were separated by SDS-PAGE through a 12% separating gel, transferred to nitrocellulose membranes blocked with 5% non-fat dry milk, and double-probed overnight at 4 °C with mouse anti-human Sp1 (1:1000 v/v, Millipore, Billerica, MA) and rabbit anti-human β-actin (1:5000 v/v, Millipore, Billerica, MA) antibodies. Blots were then washed in PBS and further incubated with secondary antibodies, goat anti-rabbit horseradish peroxidase (1:5000 v/v, Millipore, Billerica, MA), for 1 hour at room temperature. Finally, after rinsing in Tween20 (0.1% in PBS), blots were incubated for 5 min in enhanced chemiluminescent substrate for horseradish peroxidase (SuperSignal™ West Femto, Thermo Fisher Scientific, Waltham, MA) and imaged with a ChemiDoc XRS+ Imaging System (BioRad, Hercules, CA), followed by a mean fluorescence intensity analysis of Sp1 levels.
**Tumorigenicity study in mice**

Animal studies were carried out in accordance with the guidelines of the Institutional Animal Care and Use Committee (IACUC), South Dakota State University (IACUC approval #12-087A). Sixteen female non-obese diabetic, severe combined immunodeficient (NOD/SCID, NOD.CB17-Prkdc<sup>scid</sup>/J) mice (Jackson Laboratories, Bar Harbor, ME), 5 weeks old, were randomly grouped into four groups (four mice per group) in specific pathogen-free (SPF) housing at a constant temperature of 24–26 °C with a 12-h:12-h light/dark cycle. All mice were allowed to acclimatize for 1 week and were provided with sterile food and water *ad libitum*. HeLa and hCSCs were cultured, trypsinized, washed, pre-treated with 10 µM PEITC where indicated, and resuspended in PBS at the concentration of 1×10<sup>7</sup> cells/ml before injecting into the mice. Each mouse was subcutaneously injected at both flank regions with one injection of PBS (100 µl, control group), HeLa (1×10<sup>6</sup>), hCSCs (1×10<sup>6</sup>), or hCSCs pretreated with 10 µM PEITC (1×10<sup>6</sup>). The cell number in each injection was consistent with the study previously carried out by Gu et al. [246]. All mice were routinely monitored for tumor formation, weight loss, pain, and distress. The mice were euthanatized with CO<sub>2</sub> asphyxiation 28 days post-treatment. The average tumor number per injection was calculated in each group.

**Statistical analysis**

Statistical analyses were carried out using Sigma Plot software (Systat Software, Inc., San Jose, CA). Analysis of variance (ANOVA) followed by a Dunnett’s post hoc test was used to compare multiple means. Statistical significance between the groups was
assessed by Student’s $t$ test when compared two groups receiving similar treatments. Data were expressed as means ± SEM. Experiments were repeated at least three times. The significance of differences between means is represented by asterisks: $^*p \leq 0.05$, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$.

5.4. Results

**Sphere culture enriches ALDH$^{\text{high}}$CD44$^{\text{high}}$ hCSCs**

Since others and we have shown that sphere culture enriches the population of sphere-forming CSC-like cells in HeLa cervical cancer cell line [61, 246], we further seek to quantify this enrichment in HeLa cells through a well-established cervical CSC marker, ALDH1 [55]. In order to use Aldefluor assay we first optimized the proper concentration of hCSCs and found that 2 x 10$^5$ cells/ml produced the optimum signal intensity when compared to 5 x 10$^5$ and 1 x 10$^6$ cells/ml (data not shown). After growing the cells in serum free sphere-culture medium, 25.40 % of ALDH$^{\text{hi}}$ in HeLa cells at day 0 significantly enriched ($p<0.001$) to 57.75 % at day 10 (Figure 5.1). On the contrary, HeLa cells grown for 10 days in complete growth medium with 10% fetal bovine serum (FBS) did not allow sphere formation and hence did not change ALDH$^{\text{hi}}$ cells (data not shown). These ALDH$^{\text{hi}}$ hCSCs were also confirmed to be CD44$^{\text{hi}}$ cells, further validating the sphere-culture mediated enrichment of hCSCs.

**PEITC inhibits the activity of ALDH1 in both HeLa cells and hCSCs**

PEITC inhibits ALDH2 in liver [244]. Since NCBI blast revealed that ALDH1 and ALDH2 share 68% of amino acid sequences, we hypothesized that PEITC could also
inhibit ALDH1 in CSCs. As expected, PEITC attenuated ALDH$^{hi}$ cells in HeLa cells when compared to DMSO control (15.82% vs 22.41%), using DEAB as a negative control for Aldefluor reagent Figure 5.2a). Similarly, PEITC (10 µM) also attenuated ALDH enrichment in hCSCs when compared to DMSO control (40.96% vs 56.71%), using disulfiram as a positive control (Figure 5.2c, 5.2d). We observed that both PEITC and disulfiram had similar ALDH1 inhibiting effects on hCSCs. Further, we sought to understand the dose response (1.25 – 10 µM) of PEITC on inhibiting ALDH1 in hCSCS. 1.25 µM PEITC inhibited ALDH$^{hi}$ cells by 20% ($p<0.01$), whereas 5 and 10 µM PEITC inhibited ALDH$^{hi}$ cells by around 40% and 65% ($p<0.001$), respectively (Figure 2e).

Taken together, PEITC inhibited ALDH1, further attenuating the population of hCSCs.

**PEITC induces ROS in hCSCs**

Being an electrophilic in nature PEITC tends to covalently interact with nucleophilic GSH, leading to ROS-induction in cells [240]. Since CSCs possess high level of GSH as defensive machinery, we hypothesized that PEITC could selectively induce more ROS in hCSCs. DCFDA ROS assay revealed that PEITC (10 µM) increased ROS production by 1.4 fold higher than that of DMSO control using H2O2 (50 µM) as a positive control (Figure 5.3). Interestingly, when the PEITC-treated cells were replenished with GSH (25 nM), the cells came back to the normal redox status, further supporting the ROS induction by PEITC in hCSCs.
PEITC inhibits proliferation of hCSCs and induces apoptosis

Sphere formation assay showed that the sphere-forming capacity of hCSCs was inhibited by PEITC in a dose dependent manner (Figure 5.4a). Moreover, the MTS cytotoxicity assay revealed PEITC significantly reduced the viability of hCSCs by around 20% and 70% when treated for 24 hr and 48 hr, respectively with a concentration of 5 µM ($p<0.05$) and 10 µM ($p<0.001$) (Figure 5.4b). Propidium iodide and Annexin V assay revealed that PEITC (10 µM) significantly increased ($p<0.01$) the early apoptotic cells when compared with DMSO control (6.42% vs 3.89%) (Figure 5.4c). Together, PEITC showed its pro-apoptotic and anti-proliferative properties in hCSCs.

PEITC inhibits Sp1 and P-gp proteins in hCSCs

In our previous study [61], when compared to parental HeLa cells hCSCs caused more efflux of Hoechst 33342 dye, which is a substrate of multi-drug transporter P-gp protein [247]. Hence, we hypothesized that PEITC might inhibit P-gp protein in hCSCs. Flow cytometric analyses revealed that PEITC (10 µM) inhibited P-gp positive hCSCs when compared to DMSO control (4.5% vs 9%) (Figure 5.5c, 5.5d). Since Sp1 transcription factor is involved in the binding of promoter region of MDR1 (P-gp) gene, we sought to understand if PEITC could inhibit Sp1 protein. Western blot followed by mean fluorescence intensity analyses showed that Sp1 protein was inhibited by PEITC in a dose dependent manner, using Sp1 inhibitor mithramycin as a positive control (Figure 5.5a, 5.5b). To confirm the correct band of Sp1 protein, we used HeLa nuclear extract as a positive control in a separate blot (data not shown).
PEITC inhibits CSCs in vivo and prevents metastasis in mice

In order to confirm the higher tumorigenic potential of hCSCs in vivo and its attenuation by PEITC, we carried out a xenotransplant NOD-SCID mouse model with two treatment groups and a negative control group. Normal body weight and activity among all the groups confirmed that there was no alteration in food intake of the mice due to tumor development (data not shown). Equal number of hCSCs (1x10^6) with 10 μM PEITC pre-treatment (hCSC+PEITC) developed lower tumor load when compared to without PEITC pretreatment group (hCSC) (Table 5.1). Interestingly, mice in hCSC group started forming the tumors in only about 2 weeks. Taken together, hCSCs initiated tumor formation and its tumorigenicity was attenuated by PEITC pre-treatment prior to xenotransplant.

5.5. Discussion

Absence of rigorous screening program has made the cervical cancer the second-most-fatal cancer in women worldwide, especially more prevalent in developing nations [248]. Since others and we have shown that PEITC can induce the death receptor-mediated extrinsic apoptosis pathway in human cervical cancer cell line and cervical cancer stem cells [61, 249], this study highlighted the intrinsic apoptotic pathway involving ROS and mitochondrial oxidative stress in cervical CSCs. Using aldefluor assay on hCSCs enriched through cervical HeLa sphere culture, here we showed for the first time that PEITC inhibited ALDH1, a marker of cervical cancer stem cells and induced ROS-mediated apoptosis in hCSCs through inhibition of Sp1 transcription factor.
20 years ago, Lindros et al. described PEITC as an inhibitor of mitochondrial ALDH in liver in the context of alcohol administration, with a similar effect as that of a known ALDH inhibitor, disulfiram [244]. This is due to accumulation of ethanol-produced toxic acetaldehyde [250], which cannot be converted into acetic acid due to lack of ALDH2 enzyme. Interestingly, this old concept did not get much attention until ALDH was later proposed as a novel CSC marker in a breast cancer [52]. Since ALDH is functionally associated with redox balance in CSCs, an inhibition of ALDH1 by PEITC could induce ROS in CSCs (Figure 5.3). Despite the paradoxical role of PEITC, an antioxidant in lower doses and pro-oxidant in higher doses in normal healthy cells, its pro-oxidant role in cancer cells is well studied [243]. Intriguingly, another ALDH2 inhibitor, disulfiram, has been associated with induction of ROS and inhibition of a CSC marker ALDH1 [251]. Moreover, sulforaphane, a closely related isothiocyanate to PEITC, has been already shown to inhibit ALDH1 in breast CSCs [252], further supporting the CSC-targeted chemotherapeutics potential of PEITC.

Previous study that showed PEITC modulate drug efflux transporter proteins like P-gp [253] in cisplatin-resistant gastric cancer cell line also supports our observation that PEITC inhibits P-gp in CSCs (Figure 5.5). This could be one of the explanations for PEITC-associated sensitization of other conventional drugs, which are mostly effluxed through P-gp transporter [254], which is mediated through multidrug resistance gene (MDR1). Therefore, PEITC, either alone or in combination, may be useful as the therapeutic strategy for overcoming multi-drug resistance.

Interestingly, promoter of the MDR1 gene gets activated with specificity protein (Sp1) [235], which may link to its significantly higher expression in colon CSCs than
parent cancer cells [255]. Mechanistically, higher expression of Sp1 and related Sp proteins in cancer is governed, in part, by microRNA-regulated transcriptional repressors like ZBTB proteins, which directly compete with Sp proteins to bind onto GC-rich promoter regions. In this context, recently it has been shown that PEITC induced glutathione (GSH)-dependent ROS in pancreatic cancer cells resulting a downregulation of miR-27a and miR-20a/miR-17-5p, which induced ZBTB10 and ZBTB4, respectively, and in turn inhibited Sp proteins [256]. Moreover, the Sp1 transcription factor is a target of histone acetylation associated with histone deacetylase (HDACs) inhibitors [236]. Since PEITC is also an HDAC inhibitor in prostrate cancer cells [257], Sp1 inhibition could be an important epigenetic mechanism of action for induced apoptosis due to PEITC and other ROS-inducing anticancer agents [256], further validating our findings in CSC-specific context that merit further study.

In this study PEITC also induced ROS in 3 hr, and the effect was attenuated when the cells were replete with GSH (Figure 5.3). Since cellular redox homeostasis is dependent on GSH and thiol-regulating glutaredoxin enzyme system [258], PEITC utilizes this redox machinery to bind and deplete GSH, leading to ROS-induced cell apoptosis [243, 256]. Strikingly, CSCs that have higher levels of GSH are relatively more sensitive to PEITC, which may also explain its selectivity toward CSCs compared with cancer cells and normal cells. 1.5–10-μM final concentrations of PEITC used in this study are achievable following oral administration in human [259] and have been already used in our prior studies to induce apoptosis in the SW480 colon cancer cell line [59] and cervical cancer stem-like cells [61]. Hence, targeting CSCs with a natural compound
PEITC have significant implications on minimizing tumor recurrence and associated health care costs.

5.6. Conclusion

This novel study demonstrated that PEITC inhibited CSC marker ALDH1 and induced ROS-mediated apoptosis, inhibiting the proliferation and sphere-formation potential of hCSCs, which correlated with the less tumor counts in vivo. Since PEITC also suppressed Sp1 transcription factor and multi-drug resistance protein P-gp, this study provides a substantial proof-of-concept that PEITC targets cancer stem cells using HeLa CSCs as a model. Hence, PEITC commonly obtained from cruciferous vegetables can effectively target chemo-and-radiotherapy-resistant CSCs, fighting against the cancer relapse and associated healthcare-cost burden.

5.7. Future Directions

1. To sort the ALDH^{hi} cells using FACS and carryout the in vitro experiment,

2. To inject CSCs in multiple concentrations to see the concomitant inhibitory effects PEITC concentrations in tumor inhibition in NOD/SCID mice

3. To identify the active binding site of ALDH1 enzyme and to establish the role of cysteine in activity of PEITC
Figure 5.1. Sphere culture enriches ALDH$^{\text{hi}}$CD44$^{\text{hi}}$ hCSCs population from parental HeLa cells

a) Representative FACS dot plots showing enrichment of ALDH$^{\text{hi}}$ cancer stem cells from day 0 to day 10 due to sphere culture in low anchorage dishes. The ALDH$^{\text{hi}}$ gated cells were mostly CD44 positive. b) Bar graphs representing ALDH enrichment in HeLa (Day 0) and hCSCs (Day 10). All data represent means ± SEM, ***p ≤ 0.001
Figure 5.2. PEITC inhibits ALDH activity in HeLa and hCSC
a) Representative FACS dot plots showing PEITC inhibits ALDH activity in HeLa cells (a) and hCSCs (c). b) Bar graphs showing the reduction of ALDH high cells in HeLa (b) hCSC, hCSC + PEITC, and hCSC + Disulfiram treatments, using Disulfiram as a positive control (d). e) Bar diagrams showing inhibition of ALDH high hCSCs by PEITC in a dose dependent manner. All data represent means ± SEM, *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001.
Figure 5.3. PEITC induces ROS in hCSCs
a) Representative FACS histograms showing PEITC induces ROS in hCSCs with H$_2$O$_2$ as a positive control. b) Overlay image of histograms showing ROS induced by PEITC can be replenished by exogenous glutathione c) Bar diagram showing the ROS induction by PEITC in hCSCs.
Figure 5.4: PEITC inhibits proliferation of hCSCs and induces apoptosis  
a) Representative photomicrographs showing attenuation of hCSCs sphere formation (day 7) by PEITC in a concentration-dependent manner (400-µm scale). b) Bar graphs showing concentration-dependent effects of PEITC on the viability of hCSCs after 24 hr treatment (n=6). The dotted line represents the baseline cell viability for DMSO, which served as a control to determine statistical significance. c) Bar graphs showing the percentage of apoptotic cells obtained from propidium iodide and Annexin V assay. All data represent means ± SEM, *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001. Dr. Yi Liu and Dr. Dan Wang helped with this figure.
Figure 5.5. PEITC inhibits transcriptional factor Sp1 and multidrug resistance protein ABCB1 in hCSC

a) Western blots showing PEITC inhibits Sp1 in a dose dependent manner and their corresponding bar graphs with relative Sp1 expression (b)  c) FACS histograms showing PEITC inhibits multi-drug resistance protein P-Gp (ABCB1). D) Bar graphs showing PEITC inhibits multi-drug resistance protein P-Gp (ABCB1). All data represent means ± SEM, *p ≤ 0.05.
Table 5.1. Number of tumors in treatment groups in NOD/SCID mouse model

<table>
<thead>
<tr>
<th>Cells Injected</th>
<th>Tumors at week 2</th>
<th>Tumors at week 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>HeLa</td>
<td>0/8</td>
<td>3/8</td>
</tr>
<tr>
<td>hCSC</td>
<td>1/8</td>
<td>4/8</td>
</tr>
<tr>
<td>hCSC + PEITC</td>
<td>0/8</td>
<td>1/8</td>
</tr>
</tbody>
</table>

NOD/SCID: Non-obese diabetic/severe combined immunodeficient
hCSC: HeLa cancer stem cell
PEITC: Phenethyl isothiocyanate
CONCLUSIONS

This dissertation study provides the evidence that dietary fiber RS4 supplementation selectively changes the gut microbial and metabolite environment as well as associated host metabolic functions in a free living Caucasian cohort with signs of MetS. These results backed-up by in vivo study using a suitable mouse model of MetS further supports the perceived anti-inflammatory role of bacterial metabolites, particularly butyrate, in attenuating intestinal inflammation epigenetically. This study is also the first to further explain the role of maternal high fat diet on heart tissue-specific differential histone modifications in the rat offspring born to obese and/or diabetic pregnancy, highlighting the fuel-mediated cardiac dysfunction in offspring that increase the adult-onset of CVDs. Further, PEITC commonly obtained from cruciferous vegetables effectively target chemo-and-radiotherapy-resistant cancer stem cells, fighting against the cancer relapse and associated healthcare-cost burden.

Further characterization of prebiotic functional fibers-associated epigenetic alterations may provide effective strategies to mitigate the low-grade inflammation in the context of intestinal as well as extra-intestinal diseases associated with dysbiosis. In addition, identification of the reversible epigenetic landscape in utero and its association to risks for adult-onset of chronic CVDs provides an effective strategy for novel cardio-metabolic health management in next generation and future generations. Hence, this study enhances the knowledgebase of nutritional therapies with important implications for dietary guidelines to decrease the high fat with an increase in fiber and vegetables consumption in our daily meal for all individuals in all life stages.
REFERENCES


107. Morotomi, M., F. Nagai, and Y. Watanabe, *Description of Christensenella minuta gen. nov., sp. nov., isolated from human faeces, which forms a distinct branch in*


