Chromatin dynamics in host-gut microbiome interactions in the mouse colon

Mia Mosavie

A thesis submitted for the degree of

Doctor of Philosophy in Biological Sciences

Department of Life Sciences

University of Essex

October 2022

Impact of COVID-19

The COVID-19 pandemic greatly impacted my research. It prevented me from exploring exciting preliminary results with the SCFA – DSS treatment experiments. This part of my research was carried out at the University of Campinas (Unicamp) in Brazil, I was due to return in May 2020 to expand on the research I had already started in 2019. However due to travel restrictions across the globe in 2020, I was unable to return to Brazil. Students from the collaborating lab in Unicamp were also prevented from entering their lab so any attempts to have the mice experiments completed on my behalf and have mice material sent to the UK for analysis was also unfeasible. This has resulted in my Chapter 3 left with an unfortunate lose end.

National lockdown and the restrictions put in place within the University of Essex prevented me from travelling to Colchester and from entering the lab as often as I normally would have to complete experiments which overall slowed down my progression. My antibiotic ChIP samples in Chapter 4, were due to be sent off for sequencing in March 2020, but this was delayed until July 2020 pushing back the analysis of the data.

The downstream effect of the COVID-19 restrictions caused further delays in completing wet lab experiments with delays to reagents and chemicals being available for delivery. A ChIPseq experiment that was planned for the antibiotic treatment part of my PhD, required anti- H3K9ac and anti- H3K9cr antibodies, however it was not until June 2021 that these antibodies were able to be delivered from the US.

To mitigate the delay in analysing my own work, I reanalysed data from a past student's ChIPseq dataset with antibiotic treatment. Using different statistical analyses, I discovered a difference in our results, and I was able to compare them with my own ChIPseq dataset. This work has been included in my thesis.

Statement of originality

This thesis is the result of my own and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

It is not substantially the same as any that I have submitted, or is being concurrently submitted for a degree or diploma or other qualification at the University of Essex or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or is being concurrently submitted for any such degree, diploma or other qualification at the University of Essex or any other University or similar institution except as declared in the Preface and specified in the text.

Acknowledgements

I would like to thank several people who have helped me throughout my PhD and complete my thesis. Firstly, I would like to thank Dr Patrick Varga-Weisz for his guidance, patience, and supervision. I would also like to thank my lab colleagues, Salma El Sahhar, Mariane Font Fernandes, Sarah Oliveira and Mariana Portovedo for their kind help, guidance, and space to allow me to air my grievances when experiments refused to cooperate. I am forever indebted to Dr Juri Kazakevych for patiently and thoroughly teaching me all the secrets to ChIPseq. A big thank you to my assessor, Dr Radu Zabet, who provided guidance and advice whenever needed and was always very supportive.

I would like to also thank my friends at Essex who made my day to day enjoyable and was always ready for a coffee break – Nonny, Ola, Jess and Bhavana. A special thanks to the fellowship, a group of strong supportive women that have kept me going throughout, given me invaluable guidance and reminded to keep laughing – Anuja, Hannah, Madeleine, Nicki, Rachael, Sabine and Shirin. Thank you to my good friend Tass, for making sure I would take regular and adventurous breaks over the years and kept me moving and healthy.

Finally, a very special thank you to my family, my parents Hossein and Diana, and my sisters, Ranna and Hana for always being there, allowing me to go through the emotional rollarcoaster that comes with writing a thesis and for celebrating each and every win with me.

Acknowledgement of assistance received during the course of thesis

1. **Training in techniques and laboratory practice and mentoring**
   - Patrick Varga-Weisz for scientific advice and support.
   - Radu Zabet for scientific advice and support.
   - Juri Kazakevych for guidance and help with ChIP-sequencing and qPCR analysis, and with handling and dissecting mice.
   - Mariana Portovedo for guidance and help with handling, treating and dissecting mice, and isolating intestinal epithelial cells for experiments.
   - Mariane Font Fernades for guidance and help with handling and treating mice, and preparing western blot samples from tissue.
   - Simon Andrews for training and guidance with R.
   - Tony Marco for training and guidance with R and BASH.

2. **Data obtained from a technical service provider (e.g. DNA sequencing, illustrations, simple bioinformatics information etc)**
   - ChIP-sequencing (Smarcad1) was performed by BGI, Hong Kong, China
   - ChIP-sequencing (Antibiotic treatment) was performed by Novogene, Cambridge, UK.
   - Downstream sequencing processing (adaptor trimming and mapping) was carried out by the Babraham Bioinformatics Facility.

3. **Data produced jointly**
   - SCFA-DSS treated mice experiments were monitored for the daily measurements supported by Mariane Font Fernades and Sarah Oliveria.
   - Mice dissection, colon harvesting and IEC isolation were performed together with Mariana Portovedo, Mariane Font Fernades and Sarah Oliveria.
   - H3K18ac ChIPseq was performed with Mariane Font Fernades.
   - Extra antibiotic-treated IEC samples prepared by Sarah Oliveria.

4. **Data or material provided by someone else**
   - RNA-seq data was performed by Juri Kazakevych and Claudia Stellato.
   - Kcr ChIPseq data from antibiotic treatment was performed by Rachel Fellows.
   - Rachel Fellows for providing R scripts.

Abstract

This study aimed to gain a deeper understanding of the epigenetic interactions, functions and mechanisms of the host colon cells that are influenced by the gut microbiome. With the potential to contribute toward understanding the development of inflammatory bowel disease (IBD) and its co-morbidities and the increased risk of colorectal cancer (CRC), this knowledge, could facilitate and develop prevention strategies and possibly treatment.

The interactions between the chromatin dynamics and the gut microbiome are very complex and sophisticated. There are many pathways that facilitate the smooth running of the gut and protect it from external invaders and internal mutations. I have been able to gain some insight into the beneficial effect of SCFAs with DSS-induced colitis, suggesting a new potential pre-treatment for IBD patients through the use of crotonic acid supplemented into their diets. I have been able to describe the overall effects of the microbiome on histone modifications, specifically histone crotonylation, giving insight into the involvement crotonylation may have with tumour promotion and suppression, and regulation of immune responses. I have also established some of the roles the chromatin remodeller Smarcad1 has on histone modifications, predominantly influencing the activity of H3K9ac and H3K9cr, highlighting links to the maintenance of the gut and potential regulation of tumour development and suppression. Finally, I have been able to recognise a potential link between Smarcad1 and the microbiome with the notable changes Smarcad1 has on H3K9cr and the significant increase of this histone modification with recolonised germfree mice. Understanding the functions of these host-microbiome interactions has provided potentially new pathways of investigation to develop prevention measures and potential therapeutic targets for IBD and CRC.

Contents

List of Abbreviations and Acronyms

ac - Acetylation

ACSS2 – Acyl-CoA synthestase short chain family member 2 (Also known as AceCS1)

ATP - Adenosine triphosphate

bp - Base pair

BRD - Bromodomain

Bt - Butyrate

CD - Crohn's disease

CHD - Chromodomain helicase DNA-binding

ChIP – Chromatin immunoprecipitation

ChIPseq – ChIP-sequencing

CoA – Coenzyme A

Cr - Crotonate

cr - Crotonylation

CRC - Colorectal cancer

DESeq2 - Differential gene expression analysis 2

DNMTs - DNA methyltransferase

DPF – Double plant homodomain finger

DSS - Dextran sulphate sodium

EDTA - Ethylenediaminetetraacetic acid

EGTA - Egtazic acid

ES - embryonic stem cells

EtOH - Ethanol

FDR - False discovery rate

GF - Germ free mice

GO – Gene ontology

GPCR - G protein coupled receptor

H3K18 - Histone 3 lysine 18

H3K9 - Histone 3 lysine 9

HAT – Histone acetyltransferase

HBSS - Hank's balanced salt solution

HCl - Hydrochloric acid

HCT – Histone crotonyltransferase

HDAC – Histone deacetylase

HDCR – Histone decrotonylase

HDM - Histone demethylases

HMT - Histone methyltransferase

HPTM – Histone post-translational modification

IBD - Inflammatory bowel disease

IECs - Intestinal epithelial cells

IL-10/6 - Interleukin10/6

INO80 - Inositol auxotrophy 80 complex

ISWI - Imitation switch

Kac - Lysine acetylation

kbp - Kilo base pair

Kbr - Lysine butyrylation

Kcr – Lysine crotonylation

KO - Knockout

Kpr - Lysine propionation

LPS - Liposaccharides

MACS - Model-based analysis of ChIPseq

MAMPS - Microorganism-associated molecular patterns

me3 - Trimethylation

MS - Mass spectometry

NaCl - Sodium chloride

PBS – Phosphate buffered saline

PCR – Polymerase chain reaction

PHD – Plant homeodomain

PIC – Protease inhibitor cocktail

PMSF - Phenylmethylsulfonyl fluoride

PVDF - Polyvinylidene fluoride

qPCR – Quantitative polymerase chain reaction

RIPA - Radioimmunoprecipitation assay

RNA-seq – RNA-sequencing

RT – Room temperature

SCFA – Short-chain fatty acid

SDS - Sodium dodecyl sulfate

SWI/SNF - Switching defective/sucrose non-fermenting

TBS – Tris-buffered saline

TGFβ - Transforming growth factor-β

TNF - Tumour necrosis factor

TSS – Transcription start site

UC - Ulcerative colitis

WT - Wild type

β-Me - β-Mercaptoethanol

Chapter 1. Introduction

1.1 Inflammatory bowel disease, colitis and colon cancer

Inflammatory bowel disease (IBD) is a group of disorders that include Crohn's disease (CD) and ulcerative colitis (UC). Both diseases cause chronic inflammation of the intestine presenting with similar symptoms. Colorectal inflammation can result in diarrhoea, bloody mucus stool, abdominal pain, fever, fatigue and anaemia. Some patients may also develop extra-intestinal symptoms involving joints, eyes, skin and liver which can ultimately reduce the overall quality of life. CD can occur anywhere along the digestive tract, causing unpredictable flare-ups resulting in inflammation and patchy lesions. It can affect all layers of the colon wall causing wall thickening and fissures to develop. Chronic Crohn's can progress to form fibrotic strictures causing a "cobblestoning" effect on the inner wall of the lumen from reoccurring flare-ups damaging the wall repeatedly (Fig. 1.1.1.). Other manifestations include creeping fat[1] wrapping around the outer layer of the colon and small intestines. UC however, is restricted only to the colon and rectum and presents as a consistent inflammation along the entire colon. This causes continuous lesions, but only affects the mucosal layer of the bowel[1] which can develop into pseudopolyps, abscess formation and a loss of crypt structure (Fig. 1.1.1) predisposing the gut to colon cancer. The damage to the muscularia propria results in colonic dilation and a loss of haustra[2] – the loss of the segmented structure aiding food to move along the colon during contraction. The pathogenesis of UC involves genetic, immunological and microbiological factors[3].

In a normal gut, the luminal surface of the small intestine and colon is formed of a single layer of intestinal epithelial cells. Its role is to absorb nutrients while also acting

as a barrier to protect the body from damaging toxins and pathogenic microorganisms. The intestinal epithelium together with the lamina propria are two layers that make up this protective barrier. The epithelium is supported and nourished by the lamina propria, a thin layer of connective tissue that also gives the epithelium the ability to adhere to the tissue beneath it. The muscularia mucosae is a muscle layer that is separated from the submucosa and is situated outside of the lamina propria. Cell division, maturation, and migration occur in the proliferative cells, also known as crypt cells, at the base of the intestinal glands. These epithelial cells undergo a continuous renewal process in which they leave the crypt and travel to the outer part of the colon before eventually undergoing apoptosis and released into the intestinal lumen[4].

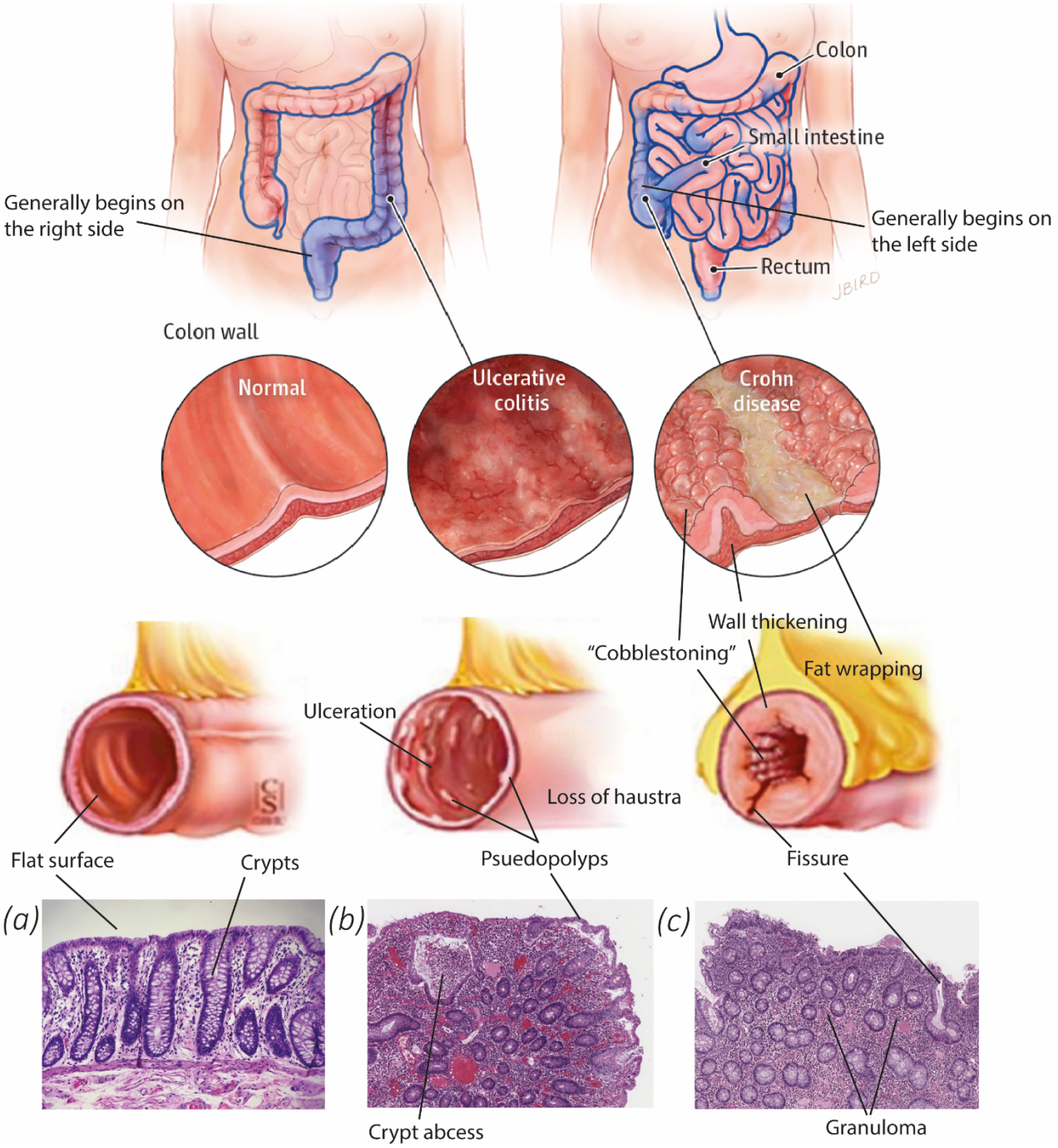


Figure 1.1.1 Clinical presentation of ulcerative colitis and Crohn's disease

Ulcerative colitis typically begins in the rectum and may extend continuously to involve the entire colon. It usually affects only the inner layer of the bowel wall and results in ulcerations, pseudopolyps and a loss of haustra on the structure of the colon. Crohn's disease mostly commonly involves the end of the small intestine and beginning of the colon and may affect any part of the GI tract in a patchy pattern. Crohn's can affect all layers of the bowel wall[5]*, resulting in fissures, thickening of the wall and a "cobblestoning" effect on the inner wall. Ulcerative colitis*[6]*.* a) *Normal histology of colonic mucosa. "Normal colonic mucosa shows a*

flat surface, evenly spaced crypts, and delicate connective tissue in the lamina propria with a mild degree of mononuclear inflammatory cells (H&E stain, 200X). There is no evidence of basal lymphoplasmacytosis or neutrophilic inflammation"[7]*. Chronic active ulcerative colitis (b) histology image show the mucosa layer separation resulting in pseudopolyps and crypt abscess formation. Cohn's (c) histology image reveal fissuring ulcer and granuloma formations. Figure adapted from Levine et al. (2012), Arias (1972) and Chang et al. (2018)*

A recent study from the University of Nottingham[8] shows that more than half a million people in the UK (1 in 123) are living with IBD. This number increases to 1 in 67 for people over the age of 70. In terms of percentage of the population, this statistic puts the UK second to the US with people living with colitis and Crohn's globally. This problem is growing. UC can develop at any age but the peak incidence is between 15 and 25 years with a second peak between 55 and 70 year olds. The concern with this growing number is that people receiving treatment for IBD need specialist care from clinicians including doctors, gastroenterologists, nurse specialists, dietitians and psychologists to help them manage the impact of their condition. The NHS's lifetime costs in treating these conditions and the number of patients currently living with IBD are unlikely to have reached their peak in the UK. The situation is no better outside of the UK, in 2015 the CDC estimated that in the US 1.3% of adults are living with IBD and according to the European Federation of Crohn´s and Ulcerative Colitis Associations (EFCCA), 10 million people are living with the disease worldwide. These figures are on the rise; over a 10-year study between 2006 – 2016 cases in the UK have risen by 33%[9] which are only expected to increase further with increasing life expectancy.

People that suffer from IBD are significantly impacted in their quality of life. Although there are many treatment options available there are no known cures. Currently, 70% of Crohn's sufferers and 30% of people with ulcerative colitis will require surgery at least once in their lifetime. There are treatments available to manage the symptoms

of the disease, such as corticosteroids and mercaptopurine or azathioprine. Both azathioprine and mercaptopurine belong to a group of drugs called thiopurines which are immunosuppressants. Thiopurines work by suppressing inflammation and slowing down the body's immune response. Both drugs are used to help maintain remission in patients with moderate to severe Crohn's and Colitis, minimising the need for corticosteroids[10]. However, patients do not always respond adequately to these treatments and a number of medications are recommended by the NICE guidelines. These are given to treat the most moderate to severe cases of ulcerative colitis in adults, for sub-acute manifestations of ulcerative colitis infliximab is given, and adalimumab for the treatment of severe ulcerative colitis and moderately to severely active ulcerative colitis[11]. Both infliximab and adalimumab are administered intravenously and are in a class of medications called tumour necrosis factor-alpha (TNF-alpha) inhibitors. They work by blocking the action of TNF-alpha, a substance in the body that causes inflammation. Although they are both effective in reducing inflammation in patients with ulcerative colitis, infliximab can come with some unwelcome side effects including, weight gain, hair loss, headaches, abdominal pain, nausea, joint pain, fatigue, eye problems and depression. Adalimumab is the alternative and although is effective for the induction and maintenance of remission in patients with moderate-severe ulcerative colitis refractory to conventional therapies[12], but as it is an immunosuppressant, it can lower the ability of your immune system to fight infections or make an infection worse. It may increase your risk of bacterial, fungal, or viral infections that may spread throughout the body.

Patients can have flare-ups and symptoms can disappear for months or years before another episode occurs, but approximately 50% of UC patients will relapse at least once a year. The main concern with UC is the comorbidities that are linked to it. It can

cause further complications such as primary sclerosing cholangitis[13] (inflamed bile ducts), bowel/colon cancer, osteoporosis and toxic megacolon[14] (swelling of the colon caused by trapped gas). Patients with severe UC classified under the Truelove and Witts severity index[15,16] are defined by having more than 6 bowel moments daily and are also systemically ill with tachycardia[13], fever, anaemia or a raised erythrocyte sedimentation rate. Acute severe UC is potentially life-threatening and requires emergency hospitalisation.

Patients with chronic UC are also at risk with at a significantly increased rate of developing colorectal cancer (CRC), due to the duration, extent and severity of the inflammation. The comorbidities such as the presence of inflammatory pseudopolyps, coexistent primary sclerosing cholangitis, and a family history of CRC are all major risk factors[17]. The colonic pseudopolyps are the most common complication in UC (Fig. 1.1.1) caused by forming mucosal tags after severe ulceration and disrupted mucosal integrity[18]. This mucosal repair from chronic inflammation can develop into pseudopolyps in 10-62%[19] of all UC cases. These patients are considered at intermediate risk of colorectal cancer[20].

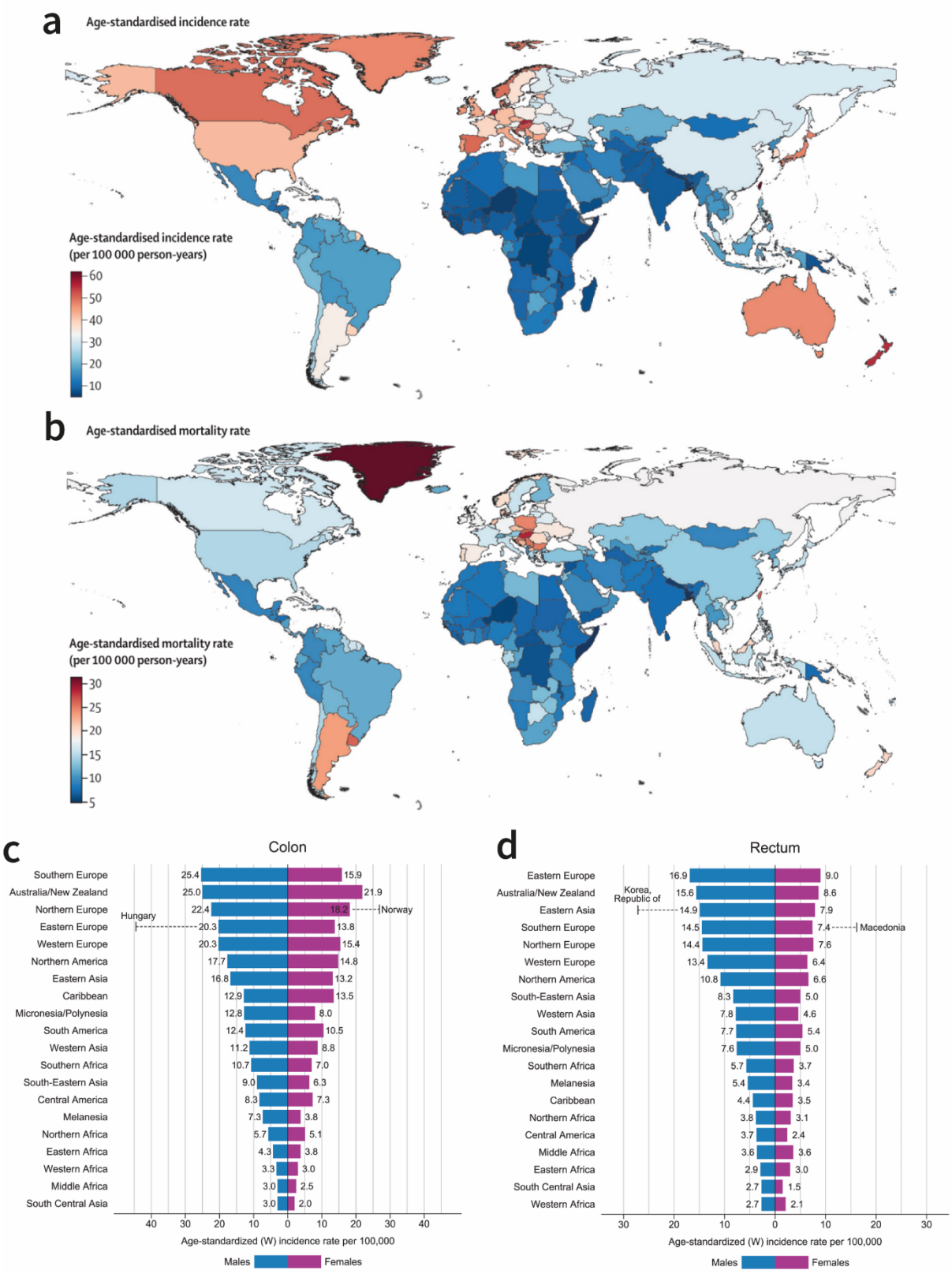


Figure 1.1.2 Geographical distribution of age-standardised rates CRC 2019

a) *Age-standardised incidence rate.* ***b)*** *Age-standardised mortality rate. Data source: Global Burden of Diseases, Injuries, and Risk Factors Study 2019*[21]*. Rates for cancers of the* ***c)*** *colon*

*and **d)** rectum are shown in descending order of the world (W) age-standardized rate among men, and the highest national rates among men and women are superimposed. Source: GLOBOCAN 2018[22]. Figure adapted from Sharma et al. (2022) and Bray et al. (2018)*

1.1.2 Colorectal cancer, global prevalence and risk factors

Globally CRC is the fourth most commonly diagnosed cancer and the third most deadly. The risk of developing CRC over a lifetime in the UK is 1 in 15 for men and 1 in 19 for women[23]. Incidences and mortality in low and middle-income countries are steadily on the rise[24,25] due to their social and economic development they are also adopting more of a "western" lifestyle. Trends in highly developed countries are stabilising or decreasing, but remain the highest in the world[25] (Fig. 1.1.2). The highest incidence rates are found in Europe, Australia/New Zealand, North America and East Asia (Fig 1.1.2a). Colon cancer incidence rates rank Hungary and Norway first among males and females respectively (Fig. 1.1.2c) rates in Uruguay are raised in both males and females also. Similar regional patterns exist for the incidence of rectal cancer, while the greatest rates for both males and females are seen in Macedonia and the Republic of Korea, respectively (Fig. 1.1.2d). Most of Africa and Southern Asia appear to have low incidence rates for both colon and rectal cancer. Despite the developed countries leading the table on high incidence rates, they have the economic backing to adopt best practices in cancer treatment and management, keeping the mortality levels relatively low. The main concern will be that if the increasing rates continue to rise globally, the low to middle-income countries will suffer greatly.

The manifestation of CRC can occur from many sources, from lifestyle choices, obesity, sedentary routines, and diet – red meat, alcohol and tobacco consumption (Fig. 1.1.3). These risk factors can drive the development of CRC through a heterogeneous group of diseases including IBD which can arise through various

molecular pathways that result in colon epithelial cells acquiring a chain of genetic and epigenetic mutations, influencing the susceptibility to cancer[23,24,26].

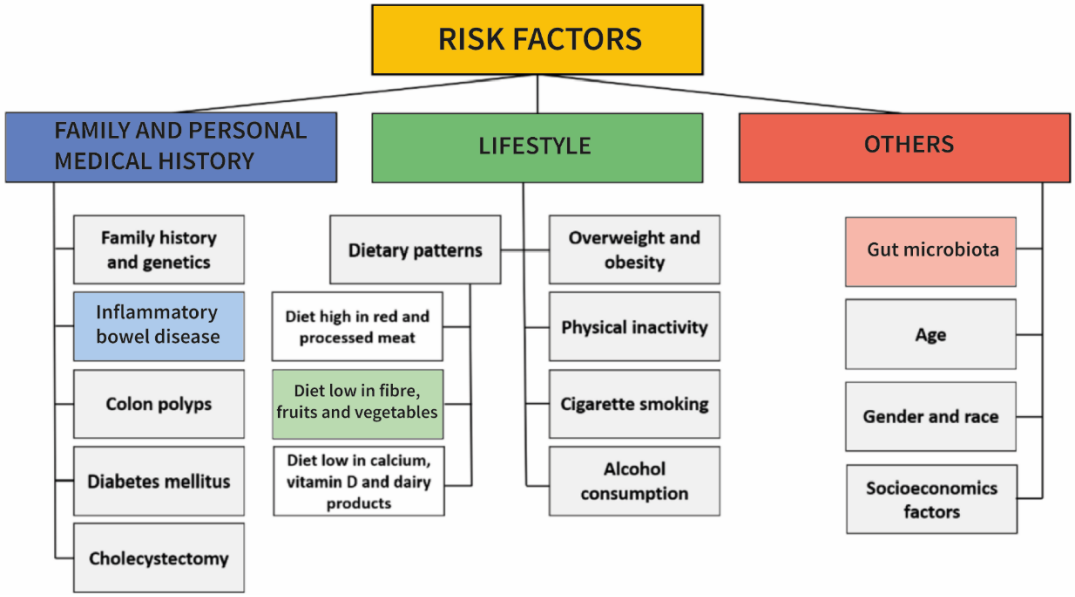


Figure 1.1.3 Risk factors of CRC

Risk factors linked to CRC[27]. Boxes coloured are areas linked to epigenetic influences on CRC and areas this thesis will focus on.

Understanding these molecular pathways that drive the disease will enable targeted therapies and tailored treatment regimens for tumours with specific mutation profiles to be established in routine therapies[23]. There are a number of molecular pathways that lead to CRC, the sequential accumulation of additional genetic events, including mutations in the oncogene *KRAS*, as well as in the tumour suppressor genes *DCC*, *SMAD4*, and *TP53*, drives tumour progression and the development of CRC[28]. Chromosomal instability pathway, the changes in the number of chromosomes and genomic amplifications accounts for approximately 60-85% of patients with CRC[23,29,30]. The CpG island methylator phenotypic cancer is the second most common pathway, accounting for about 35% of patients, followed by the mutator phenotype linked to Lynch syndrome, which accounts for 5% of patients[31].

Due to the vast array of mutations that can drive the development of CRC, it is difficult to determine a single design for successful molecular therapy. Colon cancer is however highly treatable and often curable, the primary form of treatment is to have surgery curing 50% of patients. Ultimately the best way to tackle colon cancer is to take measures in primary prevention and maintain a healthy gut environment.

1.2 Gut Microbiome

The gut microbiota plays a major role in maintaining homeostasis in the intestine and influences human physiology, metabolism, nutrition and immunity. It is estimated that around 1000 bacterial species live in the gut with variations depending on health, age, geographic location and diet of the individual[32]. Intestinal integrity is vital in maintaining gut function and a loss of this can increase permeability which can lead to the passage of harmful bacteria and toxins reaching other organs. The microbiota can affect human health by producing either harmful metabolites associated with disease or beneficial metabolites that protect against disease[33].

Dysbiosis of the gut microbiome caused by changes to the diet, environmental factors and antibiotics has been shown to alter both the transcriptome and proteome of intestinal epithelial cell, triggering an immune response and contributes to the development of conditions such as IBD, Crohn's disease, and UC[34]. Alterations in the microbiome are also associated with other host diseases such as obesity[35], diabetes, cardiovascular disease[36,37] and colon cancer[1,38,39] .

A mature microbiome generally consists of two main bacterial phyla, *Firmicutes* and *Bacteriodetes*[38], and at lower-levels, *Actinobacteria, Proteobacteria, Verrucommicrobia* and *Fusobacteria*. Different parts of the gastrointestinal tract can affect microbial colonisation depending on transit time, pH, oxygen exposure, nutrition

availability, host secretions (including bile and digestive enzymes), mucosal surfaces, and interactions with the immune system[40,41].

Diet can greatly influence diversity and function of the gut microbiota. Species within both butyrate-producing *Firmicutes* and *Actinobacteria* respond to changes in carbohydrate intake[42]. A resistant starch enriched diet can result in an increase in *Ruminoccaceae*[43], and a diet enriched with wheat bran increases *Lachnospiraceae*[44]*.* In low-carbohydrate diets however, a significant reduction in *Firmicutes (mainly Roseburia spp.* and *Eubacterium rectale)* and *Actinobacteria* (such as *Bifidobacterium spp.* and *Collinsella aerofaciens*) has been seen[45]. An interesting study by David *et al.* (2013)[46], explored the different responses diet can have on the gut microbiota population and diversity. Diet composed of plant based products rich in high fibre and low in protein and fat did not change the microbiota diversity, whereas animal based diet high in fat and protein and zero fibre resulted in an increase in bile-tolerant microorganisms (*Alistipes, Bilophila,* and *Bacteroides*) and decreased the levels of *Firmicutes*[46].

A highly diverse microbiota is a good indicator of a healthy individual. A study comparing a diverse bacterial gut population to a less diverse one, showed that individuals with a less diverse population tend to be dominated by the pro-inflammatory *Bacteroides* spp. and a decrease in the anti-inflammatory phyla, *Firmicutes*[47]*.* These individuals are more prone to having obesity and are associated with IBD and metabolic syndromes which could lead on to CRC[40].

1.2.1 Antibiotic treatment and germ free

Antibiotic treatment greatly impacts the diversity of the gut microbiota. This depletion of the microbiota results in the loss of some important taxa and causes major metabolic

shifts. The reduction in the microbiota also increases the susceptibility to be colonised by opportunistic pathogens such as *C. difficile*[48], causing overall harm and epithelial damage.

Epigenetically, when the gut microbiota has been depleted from antibiotic treatment a very interesting response occurs[49]. The colonocytes activate a number of genes through various epigenetic mechanisms to counteract this loss of function and energy producing metabolites[49]. The cascade of events that respond to this loss in the microbiota can result in dramatic downstream effect in the production of short chain fatty acids (described in Ch. 1.3). This reduces the concentration of metabolites that produce substrates and enzymatic regulators for epigenetic modifications such as histone modification, DNA methylation and non-coding RNA expression. This alteration to histone modifications as well as other epigenetic functions, can affect gene expression in counteracting this loss of microbiota and in maintaining the cellular functions of development and metabolism within the gut. Using antibiotics is a useful tool to investigate the involvement of the microbiota in gene regulation.

Mice that have been bred in sterile conditions to produce a germ free gut show interesting genomic differences to those with a normal microbiota and those that have had their microbiota depleted through antibiotic treatment[50]. A study showed that the non-pathogenic bacteria can modulate gene expression in germ-free mice, suggesting that bacteria such as the species present in the gut microbiota could recruit changes in the epithelium of genetically normal and abnormal hosts[51]. The gut microbiota can influence epigenetic mechanisms such as histone post translational modifications (HPTM) and in turn alter gene expression and regulation[52]. Germ-free mice can be used to help identify which of these HPTMs are noticeably affected by the presence

or absence of the microbiota and could shed further light on the host gut – microbiota interactions and the key players in regulating the gut and maintaining homeostasis.

1.2.2 Dextran Sulphate Sodium treatment on the gut

Dextran sulphate sodium (DSS) treatment is a method to artificially create a colitis response within the gut of the mice[3]. It is thought to be critically dependant on the innate immunity. DSS is a charged polymer that when consumed through drinking water diminishes the mucus layer in the colon, which in turn causes the epithelial barrier to become weakened and leaky. Microbial population of the lumen are directly exposed to the epithelium which causes activation of inflammatory cytokines, initiating an immune response inducing colitis. Studies have shown that DSS treatment in mice with a depleted microbiota from antibiotic treatment did not cause an immune response confirming the importance of the microbiota in the maintenance and homeostasis of the gut[53].

1.3 Short Chain Fatty Acids

The benefit of having a healthy and diverse microbiota in terms of host epigenetics, is its ability to produce an incredibly diverse range of metabolites, including small organic acids called short chain fatty acids (SCFAs) through the digestion of complex carbohydrates[38,54,55]. These SCFA provide a source of energy for epithelial cell metabolism and intestinal homeostasis[52,56,57]. Most microorganisms prefer carbohydrates over proteins which results in the concentration of SCFA highest in the proximal colon where most substrates for formation are available and reduces towards the distal colon[54,58].

1.3.1 Production of SCFAs

SCFAs can be produced naturally in the liver[38] but production is largely based in the colon with the help of specific colon bacteria. The three most common SCFAs are butyrate, acetate and propionate[58]. Although in humans butyrate is the least abundant SCFA (~60% acetate, 25% propionate and 15% butyrate)[59], with a combined concentration of 50-150mM in the colon, butyrate plays an important role in gut homeostasis. The most important Gram-positive anaerobic bacterial groups that produce butyrate are *Faecalibacterium prausnitzii* (in the *Clostridial* cluster IV) and *Eubacterium rectale/Roseburia* spp. (in the *Clostridial* cluster XIVa) and *Bacteriodes uniformis*. Each group typically accounts for ~5-10% of total bacteria detectable in faecal samples of healthy adult humans. They are important in symbiosis and host gene function. Butyrate is functionally involved in the activation of G-coupled receptors, inhibition of histone deacetylases and acts as the primary energy source for the epithelial cells[54], it also has anti-inflammatory properties, enhancing the intestinal barrier function and mucosal immunity[33]. It has been found to contribute towards obesity, but the effect of endogenous butyrate on the gut-brain axis requires further investigation[54].

Most enteric bacteria create acetate, the most prevalent SCFA, as a by-product of fermentation, although acetogenic bacteria, such as *Blautia hydrogenotrophica*, may also produce it from H_2 and CO_2 or from formate through the Wood-Ljungdahl pathway[60] (Fig. 1.3.1a). These acetogenic bacteria can manufacture three molecules of acetate from one molecule of glucose, while non-acetogenic anaerobes, which make up the majority of the microbiota, must dispose of reducing equivalents by producing various products in addition to (or instead of) acetate, such as succinate, propionate, butyrate, formate, D-lactate, L-lactate, and ethanol[40] (Fig. 1.3.1).

Propionate and butyrate are produced mainly by the breakdown of carbohydrates from the groups of enteric bacteria. *Bacteroidetes* and certain *Negativicutes (Firmicutes)* generate propionate mostly via the succinate pathway (such as *Phascolarctobacterium succinatutens*, *Dialister* spp. and *Veillonella* spp.)[40] (Fig. 1.3.1e). Under certain environmental conditions, succinate is a metabolic by-product for some bacteria, but specialised succinate users like *P. succinatutens* convert the majority of the generated succinate into propionate (Fig. 1.3.1C). The other two pathways that lead to the formation of propionate are the acrylate pathway that uses lactate and the propanediol pathway (Fig. 1.3.1f) which is more abdundant than the acrylate pathway (Fig. 1.3.1d) and uses deoxyhexose sugars[61]. This pathway is found in *Firmicutes* (including *Roseburia inulinivorans* and *Ruminococcus obeum*) and in the Proteobacteria, *Salmonella enterica.* The succinate pathway appears to be the main source of propionate as judged by the correlation between the amount of propionate in the total faecal SCFA and the relative abundance of Bacteroidetes[40,61].

The pathway to produce butyrate can either be through the butyryl-CoA:acetate CoA-transferase route used by *Firmicutes*, or through phosphotransbutyrylase and butyrate kinase route (Fig. 1.3.1c). The most predominant gut microbiota species uses the butyryl-CoA:acetate CoA-transferase route and requires acetate to produce butyrate[62] , some *Lachnospiraceae* species however, can use both lactate and acetate to produce butyrate[62]. Propionate and butyrate are only produced by a small number of anaerobes, and they come from various sources: *Rosburia inulinivorans* creates propionate from fucose and butyrate from glucose, whereas *Coprococcus catus* produces propionate from lactate and butyrate from fructose (via the acrylate pathway)[61]. The balance between the SCFAs and the presence of the diverse microbiota are important to achieve a stable ecosystem within the gut.

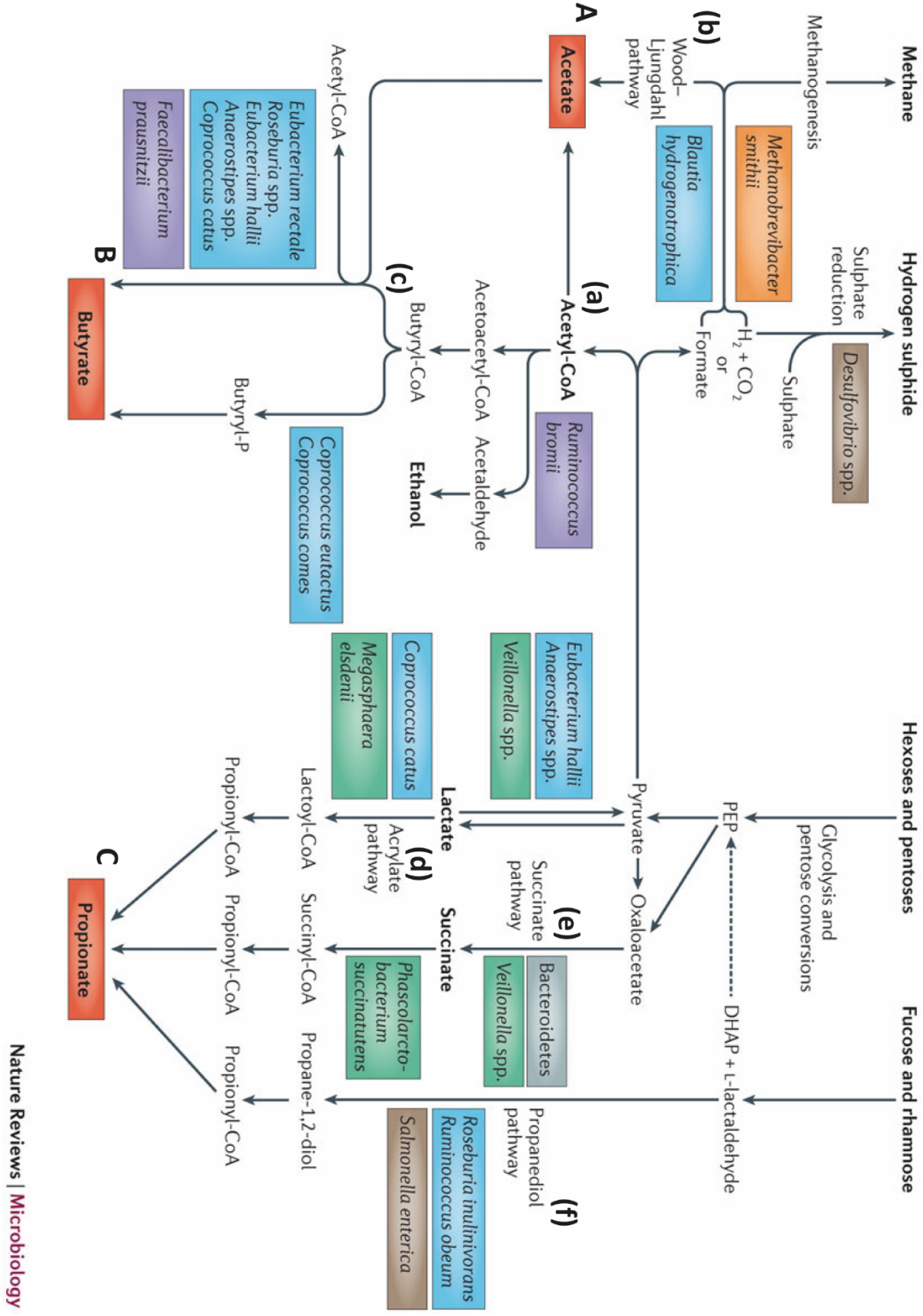


Figure 1.3.1 Pathways of biosynthesis of SCFAs by metabolised bacteria in the microbiota.[40]

A) *Acetate, one of the three primary short-chain fatty acids (SCFAs; illustrated in red), is generated by many intestinal bacteria from pyruvate via acetyl-CoA* ***(a)*** *as well as by acetogens like Blautia hydrogenotrophica via the* ***(b)*** *Wood-Ljungdahl route.* ***B)*** *Several Firmicutes convert two acetyl-CoA molecules into butyrate, and the last enzymatic step is often catalysed by the* ***(c)*** *butyryl-CoA:acetate.* ***C)*** *The main pathway for producing propionate is the* ***(d)*** *succinate pathway, which is utilised by certain Firmicutes to create propionate from lactate or succinate and by Bacteroidetes from carbohydrates. Two other propionate formation pathways are found in some gut bacteria:* ***(e)*** *the acrylate pathway, which utilises lactate, and the* ***(f)*** *propanediol pathway, which uses deoxyhexose sugars (such as fucose and rhamnose). Pathways that are involved in hydrogen metabolism and ethanol production are also shown. The bacterial species that are shown are based on studies of cultured isolates of dominant species and metagenomic analyses and are thus not exhaustive. Archaea are shown in orange, Bacteroidetes are shown in grey, Lachnospiraceae (Firmicutes) are shown in blue, Ruminococcaceae (Firmicutes) are shown in purple, Negativicutes (Firmicutes) are shown in green and Proteobacteria are shown in brown. DHAP, dihydroxyacetonephospate; PEP, phosphoenolpyruvate. Figure adapted from Louis et al. (2014).*

1.3.2 SCFAs and host cell interactions

These bacterially derived SCFAs are easily absorbed by the gut lumen, but they each have a differing effects on the host cell metabolism[63]. Butyrate is the mainly used as an energy source by gut epithelial cells, with minimal concentration levels found in the systemic circulation[40,64]. Propionate is primarily processed in the liver, and only acetate reaches relatively high levels in peripheral blood[40,58,63,65].

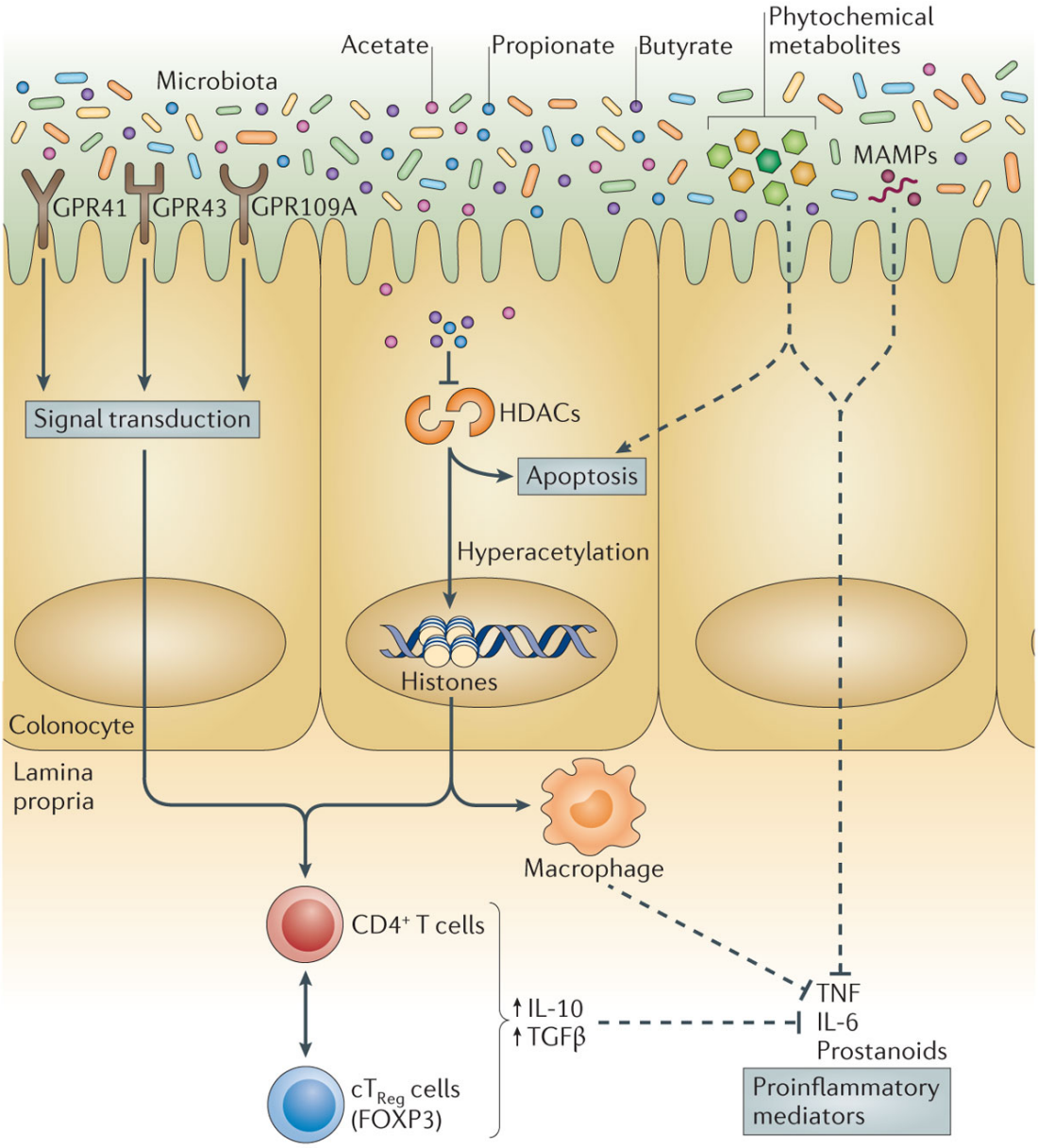
Intracellular butyrate and propionate but not acetate, inhibit the activity of histone deacetylases (HDACs) in colonocytes and immune cells[40], this then encourages hyperacetylation of histones as well as some transcription factors and proteins involved in signal transduction, resulting in a down-regulation of pro-inflammatory cytokines such as interleukin-6 (IL6) and IL12 in colonic macrophages[66] (Fig. 1.3.2). In mice, colonic regulatory T cells (cT_{Reg} cells) have been found to be regulated by

SCFAs, which may have significant anti-inflammatory effects[67]. Recent research demonstrates that butyrate and propionate stimulate the development of regulatory T cells that express the transcription factor FOXP3, which can play a critical role in regulating intestinal inflammation[68,69]. Butyrate is thought to increase histone H3 acetylation in the promoter and enhancer regions of the FOXP3 gene, resulting in enhanced FOXP3 expression[67].

SCFAs found within the gut can interact with important receptors in the host cells (Fig. 1.3.2). The G protein coupled-receptor 41 (GPCR41), GPCR43 and GPCR109A are expressed on the surface of colonocytes. The butyrate interaction with GPCR109A encourages an anti-inflammatory response by promoting regulatory T cells (T_{Reg} cells) and IL10 producing T cells. They also block nuclear factor-κB (NF-κB) activation and trigger apoptosis, a mechanism unrelated to HDAC inhibition. Acetate and propionate are also thought to play a significant role in triggering anti-inflammatory effects through modulating T_{Reg} cells via interactions with GPCR43. The tumour-suppressor genes GPCR43 and GPCR109A may mediate some of the cancer-preventive benefits of propionate and butyrate linked to high fibre intake. However, butyrate has a number of properties that are anti-tumorigenic including proliferation inhibition and selectively inducing apoptosis of CRC cells. These effects are likely due to HDAC inhibition resulting in changes in transcriptional regulation. Some studies suggest the GPCR109A binds to the co-activator p300 to induce histone 3 lysine 18 (H3K18) hyperacetylation, resulting in chromatin remodelling and activates transcription to act as a tumour repressor in colon cancer[70].

In addition to having an impact on host cells, SCFAs have anti-inflammatory properties that may also help maintain the balance of the gut microbiota (Fig.1.3.2). An

enlightening study by Chang *et al.* (2014)[66] demonstrated the importance and the dramatic effect butyrate concentration has on immune regulation in the gut through HDAC inhibition and regulation. They showed that with high levels of butyrate within the gut, an anti-inflammatory response could be a result of down regulation of immune function. In contrast, low butyrate levels set off a pro-inflammatory state that alters the gut microbiota by eradicating potential pathogens and restoring butyrate-producing species[66].



Nature Reviews | Microbiology

Figure 1.3.2 Anti-inflammatory and anti-apoptotic effects of colonic bacteria and their metabolites that are thought to mitigate colorectal carcinogenesis.

SCFAs - butyrate, propionate and acetate, can be recognized by receptors (such as the G protein-coupled receptors (GPCRs) GPCR41, GPCR43 and GPCR109A) on the surface of colonocytes. SCFAs are also transported into host cells, resulting in the inhibition of histone deacetylase (HDAC) activity by butyrate and propionate, causing hyperacetylation of histones. HDAC inhibition and GPCR signalling results in an increase in total colonic regulatory T cell (cT_{Reg}) numbers and the production of the anti-inflammatory cytokines interleukin-10 (IL-10) and transforming growth factor-β (TGFβ). HDAC inhibition also promotes apoptosis of colorectal cancer (CRC) cells. Dashed lines represent other possible anti-inflammatory compounds whose processes have not yet been determined. As an illustration, a number of phytochemicals produced by microbial transformation have been found to have anti-inflammatory properties, and both in vitro and in vivo studies suggest that they inhibit pro-inflammatory mediators like tumour necrosis factor (TNF), interleukin-6 (IL-6), and prostanoids. Finally, certain commensal bacteria's microorganism-associated molecular patterns (MAMPS) may support anti-inflammatory signalling. – Louis et al. (2014).

1.3.3 Crotonic acid

Another SCFA not yet mentioned is crotonate (crotonic acid), a SCFA moiety that occurs in the gut in very low concentrations. Crotonic acid is a natural product found in *Coffea Arabica* (coffee plant), *Daucus carota* (wild carrot) and *Croton tiglium* known as purging croton. It has a role as a plant metabolite and is a conjugate acid of a crotonate. *Croton tiglium*, also called jamaal gota in India where it is found, is used as herbal medicine to treat irritable bowel syndrome and aid in constipation.

Not much is known about crotonate and it functions within the gut as an energy source. It has however been recognised in the promotion of histone crotonylation in vitro. The addition of crotonic acid in a cell culture media converted to crotonyl-CoA by the enzyme ASCC2 (Acyl-CoA synthetase short-chain family member 2)[55]. Depletion of ACSS2 leads to reduced levels of histone Kcr, suggesting that SCFAs like crotonate might be an endogenous source of non-acetyl acyl-CoAs[71,72]. It has been suggested

that low nutrient conditions favour non-acetyl acylation such as histone crotonylation by activating ACCS2 through deacetylation[73].

While we know the vast influences butyrate has on gut and immune response through the modifications of histones such as histone acetylation, the effect of crotonate within the gut are yet to be explored. Given the holistic approaches using crotonate from plants in dealing with bowel issues, and the *in vitro* effects of crotonate on histone crotonylation, known to contribute towards gut homeostasis and immune response, it would be insightful to know how crotonate would respond within the gut when a colitis is induced. Butyrate has proven to provide protection and a beneficial response[74], it would in interesting to know whether crotonate does the same.

1.4 Epigenetics

Epigenetic modifications refer to the chemical or a physical change that regulates how genes are read and expressed without altering the DNA sequence. This regulation of gene accessibility is both heritable and reversible[75].

Determining cellular characteristics and functions involves both epigenetic changes and genetic regulation. Chromatin and the manner in which it is organised in the nucleus are important elements that can influence responses induced[1] by the gut microbiota. The epigenetic changes and chromatin organisation regulate diverse functions that include maintenance of homeostasis in the intestinal epithelium, the development and differentiation of immune cells, and the modulation of responses generated by the immune system to defend against potential pathogens[1].

Epigenetic changes can influence the function of these regulatory elements and the genes they regulate[76,77]. There are a number of different epigenetic mechanisms;

DNA methylation, chromatin remodelling and histone modifications, specific ATP-dependent remodelling and non-coding RNAs. DNA methylation silences gene expression by directly adding a methyl group to cytosine in the DNA sequence, chemically modifying DNA without altering the underlying DNA sequence. Chromatin remodelling and specific ATP-dependent remodelling both regulate expression by tightening and loosening chromatin structure[78]. In addition to the epigenetic alterations, the ATP-dependent remodelling complexes function to orient nucleosomes, promote the exchange of histones, and/or encourage long-distance connections between various DNA segments[79]. The final epigenetic mechanism are the non-coding RNAs which attach to complementary sequences resulting many processes such as gene silencing and heterochromatin formation, RNA interference, transcriptional regulation and chromatin remodelling[80].

All these epigenetic changes cooperate with one another to influence gene expression. However, the enzymes or processes involved with each of these modifications are specific to their mechanisms. DNA methyltransferases (DNMTs) regulate the DNA methylation in cells. Chromatin remodelling is primarily carried out through covalent histone modifications by enzymes such as histone acetylases (HATs), histone deacetylases (HDACs), methyltransferases, and kinases. HATs and HDAC complexes, regulate the transcriptional activity of genes by determining the level of acetylation of the amino-terminal domains of the nucleosomal histones associated with them[79]. Non-coding RNAs - including siRNAs, miRNAs, piRNAs, and long non-coding RNAs - epigenetically modify the genes by inducing transcriptional gene silencing[81].

These changes may affect how genes are expressed, either through activation or repression of transcription. Gene expression can however be disturbed as a result

of interruptions in the aforementioned mechanisms, which could lead onto the promotion of inflammation and possibly contribute to IBD pathogenesis. IBD is genetically complicated, and more than 200 IBD susceptibility loci have been found as a result of genome-wide association studies[82,83]. Many of these polymorphisms are thought to contain transcriptional regulatory elements as many of them have been found within non-coding regions of the genome[76,77].

1.4.1 Histone modifications

DNA within the nucleus of the cell are comprised of nucleosome structures. These structures consist of 147bp of DNA wrapped around 8 histone proteins. These spools of DNA and protein are made up of two core histone each; Histone 2A (H2A), H2B, H3 and H4[1]. Each nucleosome is separated from each other by 10 – 60 bp linker DNA and forms the complex known as chromatin (Fig. 1.4.1). When the nucleosomes are condensed, transcription is repressed. Chromatin in this form is known as heterochromatin. Transcription is activated when the chromatin is open in a relaxed structure, known as euchromatin[84]. The N- and C-terminal tails of the histones are responsible for mediating the folding of the chromatin. Lysine, arginine, serine and threonine, are targeted amino acids present on the tails by enzymes influencing post-translational modifications (HPTM) and altering the accessibility of gene binding[1] by transcription factors and RNA polymerase II machinery. The process of compacting chromatin is controlled by the modifications of the lysine such as acetylation, deacetylation, methylation, demethylation, phosphorylation, ubiquitination, sumoylation, crotonylation, ADP-ribosylation and deamination[1]. Depending on the site and type, the acylation can generate diverse local chromatin structures and functional outcomes with the potential to define disease-specific gene expression mechanisms[78,84].

Three histone modifications were of greater interest in my studies, histone acetylation, crotonylation and methylation. These have been shown to play significant roles in determining gene expression and transcription factor activation. I will go into further detail of these modifications, and I have focused on them throughout my thesis.

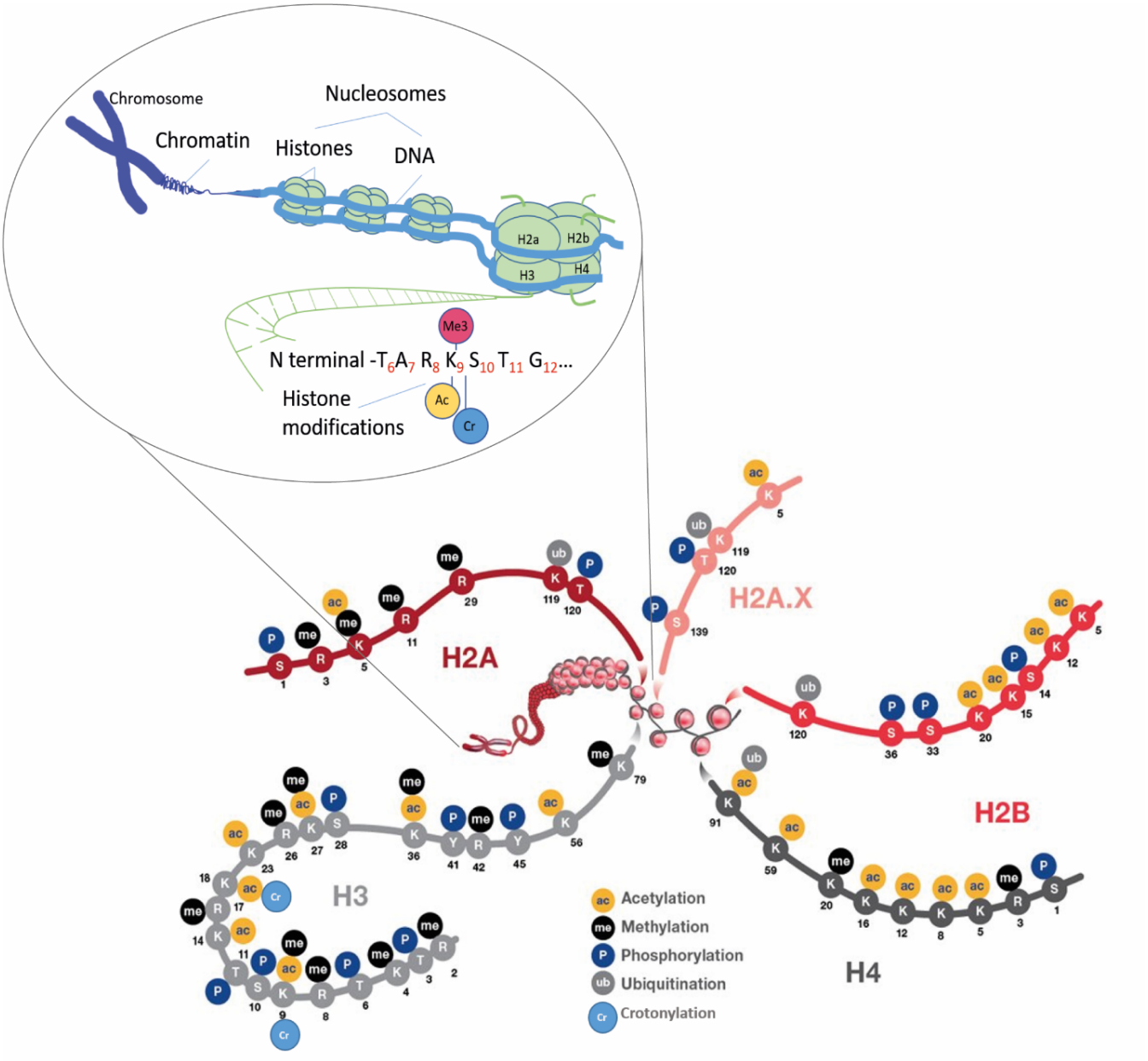


Figure 1.4.1 Chromatin and histone modification structure

Chromatin is made up of tightly wrapped DNA and proteins forming chromosomes. The nucleosome is the basic building block of chromatin and consists of approximately 147 bp of DNA coiled around an octamer of histones. The histone octamer contains two copies of each

of the core histones, H2A, H2B, H3, and H4. The N-terminal region of each core histone is unstructured when crystallized and therefore is likely to be a highly dynamic structure. These histone tails protrude out from the globular center of the nucleosome where they may interact with nuclear factors. The N-terminal tails are subject to a variety of posttranslational modifications, including acetylation, methylation, phosphorylation, ubiquitination and crotonylation. These modifications affect the binding of proteins to the histone tails and thus regulate the nature of the protein complexes that will associate with a region of chromatin.

1.4.1.1 Histone acetylation

In recent years there have been major discoveries in the study of histone lysine acylations. However one of the earliest identified by Allfrey *et al.* in 1964 was histone acetylation[85]. The function of histone acetylation has been studied in great detail over the years and has been found to play key roles in the transition and maintenance of chromatin switching from an open, relaxed state to a compact, folded state and back again. The opening of chromatin can be linked to active genes and is catalysed by the neutralisation of positive charges on the histones by the transfer of an acetyl group from the cofactor acetyl-coenzyme A (acetyl-CoA) to histone lysine residues[84] (Fig. 1.4.2). The majority of individual lysines in H3 and H4 tails are acetylated (Fig. 1.4.1), and this has been seen to be positively correlated with gene transcription. Histone acetylation can typically take place in two different states. The first of these states, hyperacetylation, is when all or the majority of the N-terminal lysines on histones are acetylated. Whereas, N-terminal lysines that are hypoacetylated frequently contain no or only one acetyl group[86]. The degree to which lysine residues on H3 and H4 are acetylated has been the subject of much investigation. It has been shown that the N-terminal tail of H3's lysines at positions 9, 14, 18, and 27 are the most essential positions for acetylation[87].

This process of acetylation of histones are modulated by two opposing enzymes; histone acetyltransferase (HAT)[55] which are responsible for adding acetyl groups to the lysine also known as “writers” and deacetylation, the reverse reaction where an acetyl group is removed using a histone deacetylase (HDAC) enzyme, known as the “erasers” (Fig. 1.4.2). Both enzyme families can influence transcription levels by directly recruiting transcriptional activators or repressors to upstream regulatory regions of genes and by performing global wide functions in the genome[84]. A third process that slots in the middle of the two mentioned are the “readers”. Certain domains, most commonly bromodomains and others such as PHD and YEATS interact with the histone modifications, “reading” them which leads to gene regulation, e.g. increased transcription.

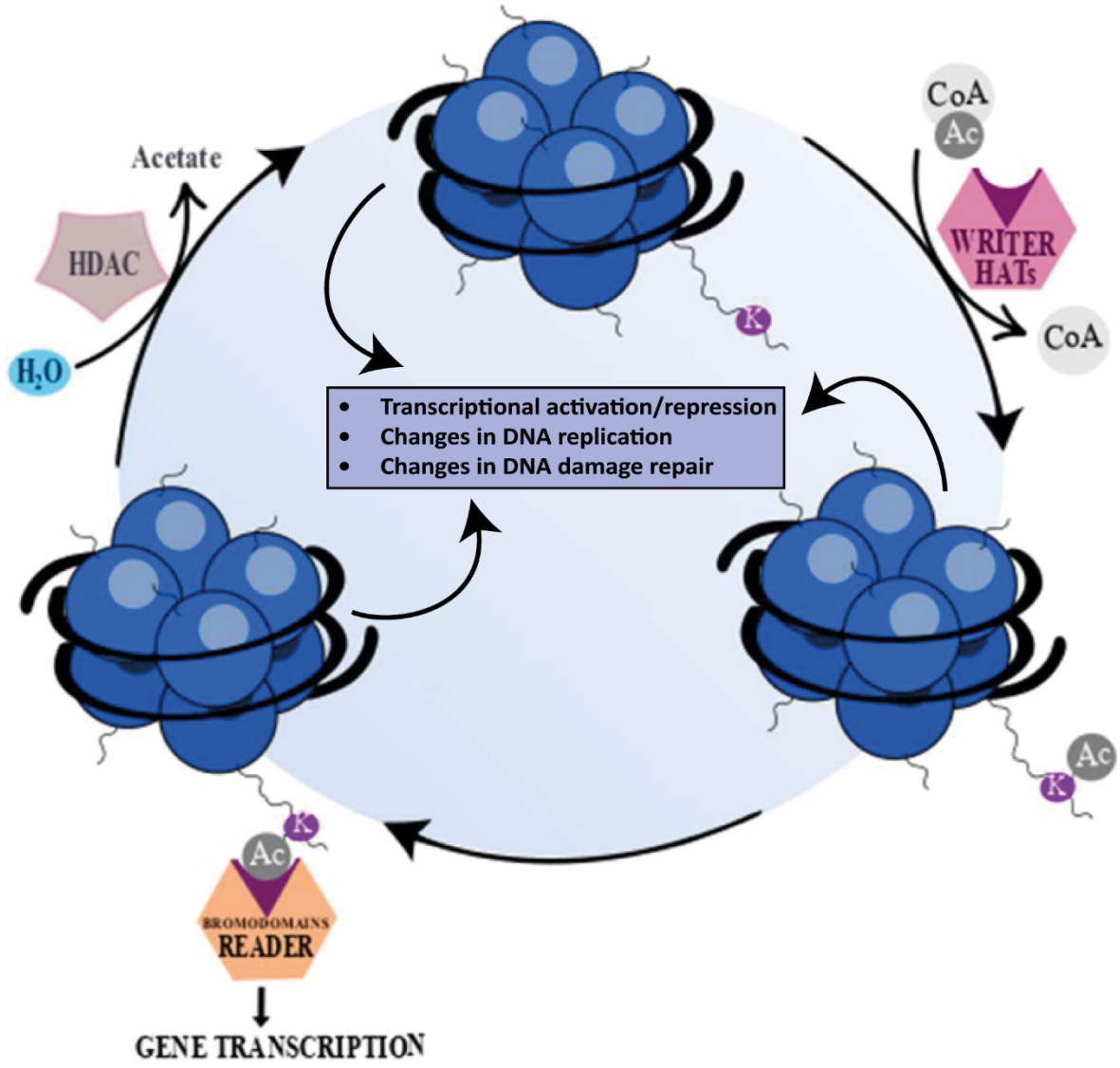


Figure 1.4.2 Writers, readers and erasers[88]

"Writers" (histone acetyltransferases (HATs)) (pink) catalyze the transfer of an acetyl group to a histone lysine (K) or arginine residue. "Erasers" (histone deacetylase proteins (HDACs)) (gray) catalyze removal of an acetyl group from acetylated lysine or arginine. "Readers" (orange) possess specialized domains which are able to recognize and interact with certain histone modification. Figure adapted from Hałasa et al. (2019).

1.4.1.1.1 Histone acetylases and deacetylases (HAT and HDAC)

HATs ("writers") are made up of a wide range of proteins. Approximately 30 HATs have been found in humans so far. Based on their subcellular location, HATs are largely divided into Type A and Type B classes. Type A HATs are present in the nucleus, whereas Type B HATs are found in the cytoplasm. While Type B HATs acetylate freshly generated histones and regulate the nucleosome structure, Type A HATs are involved in histone acetylation associated to transcription in chromatin [89–91].

Based on their catalytic domain, Type A HATs are further divided into three families[91]. The p300/CBP-associated factor (CREB- binding protein) family, Gcn5-related N-acetyltransferase (GNAT) family with Gnc5 and ELP3 as members of the family[92,93] and finally the MYST family. MYST is named after its founding members, MOZ (Monocytic Leukemic Zinc Finger), Ybf2 (Sas3), Sas2 and Tip60[71,94]. HATs also comprise a large number of steroid receptor co-activators[92].

The development of cancer can be linked to the dysregulation of HATs[95,96]. Depending on where it is located, p300/CBP can either act as a tumour suppressor or an oncogene during the cell cycle transition from the G1 to the S stage[95,97]. The cell cycle is inhibited and apoptosis is induced by the selective suppression of p300. In melanoma cells, it suppresses the response to DNA damage[96]. HATs from the MYST family play a variety of roles in stem cell formation and operation. Currently, HATs are being researched as drug targets to control tumour growth because they are reversible regulators and contribute in cell cycle progression[95].

As for histone deacetylases ("erasers"), in humans, 18 HDACs that have been classified based on their sequence homology to yeast proteins, into four classes: class I (HDAC1,2,3, and 8); class II (HDAC4,5,6,7, 9 and 10); class III (Sirtuins:

SIRT1,2,3,4,5, 6 and 7) and class IV (HDAC 11)[98]. HDACs of classes I, II, and IV are Zn^{2+} dependent enzymes with comparable functional processes, whereas HDACs of class III are NAD^{+} dependent[98]. However, several epigenetic regulators, including HDACs, have been shown to be involved in a crosstalk mechanism which can influence both the development of cancer tumours[99] and provide cancer specific therapeutic targets.

Class I HDACs are present in all tissues and have been linked to genome stability and are involved in the response to DNA damage[100]. HDAC1 and 2 have been linked to cellular processes like proliferation and apoptosis. Cell differentiation and development are linked to HDACs 4,5,7, and 9[100]. While Sirtuins generally play a role in cellular metabolism and DNA repair, HDAC11 controls the activation of interleukins[100–102]. HDACs are vital in the regulation of lymphoproliferation, inflammation and autoantibody production[103].

An imbalance in the equilibrium of histone acetylation linked to HDAC deficiency can have a significant effect on T cell function, apoptosis, cell cycle arrest and has been associated with tumorigenesis and cancer progression[104].

Over the years there has been a number of studies that have investigated the genetic response to histone acetylation and deacetylation leading to cancer progression and inducing gene silencing in tumour suppressor genes. Since the identification and characterisation of these writers, readers and easers revealing its interplay with cellular metabolism, Kac has become one of the most studied and best characterised histone PTMs.

One of the lesser researched acylations, histone crotonylation has piqued interest due to its similar behaviours to Kac. The extent at which Kcr and Kac functionally differ from each other within the cells has yet to be fully answered.

1.4.1.2 Histone crotonylation

Longer chain lysine acylations such as crotonylation (Kcr), butyrylation (Kbr) and propionation (Kpr) (Fig. 1.4.3) have been characterised relatively recently. Similar to acetylation, these unique modifications provide histones with a variety of chemical environments. They appear to concentrate near transcriptional start sites, and are correlated with gene activity[87]. In 2011, Tan *et al.*, identified lysine crotonylation within the N-terminal tail[105]. Using Mass Spectometry (MS) they were able to identify 67 novel PTMs sites which included 28 Kcr sites. This non-acetyl histone modification has become increasingly appreciated as an important epigenetic mark occurring at broadly in all the core histones[106]. Crotonylation and acetylation occur at the same sites in the histones (but cannot co-occur on the same lysine residue) and also occur on the ε-amino group of lysine[107], but are distinct by the planar orientation and the four-carbon length of the crotonyl group (Fig. 1.4.3).

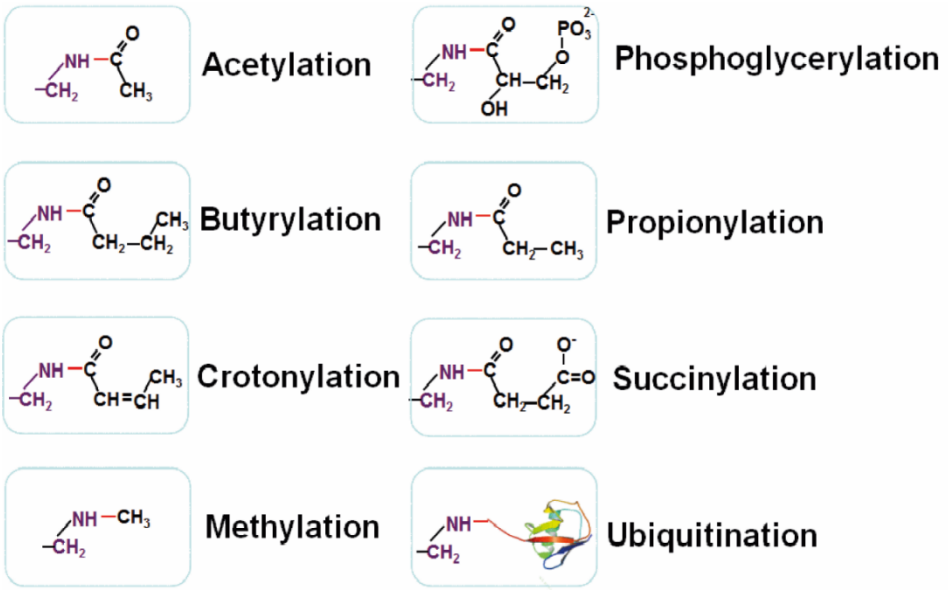


Figure 1.4.3 Chemical structure of PTMs

The lysine residue is in purple and the post-translationally added parts are in black or colourful ribbon for ubiquitination. Acetylation has a two-carbon chain on the end. The other acyl PTMs have an extension from this hydrocarbon end except for methylation which has just one hydrocarbon on the end of the lysine residue.

Crotonylation has been reported to be abundant near active gene promoters, boosting transcription and possibly enhancers to a greater extent than acetylation[108]. Crotonyl lysine and acetyl lysine chromosomal distribution studies indicate functional distinctions between the two, as they compete for sites along the histone N-terminal tails reflecting their cellular metabolism[105]. It has been suggested that these modifications are not functionally redundant from each other and have distinct roles in regulating gene function resulting in diverse biological outcomes[109].

The transcriptional co-activator p300 is a HAT and also possesses histone crotonyltransferase (HCT) activity regulating both H3K18cr and H3K18ac in cells[72,110]. The p300-catalysed histone crotonylation directly stimulates transcription and appears to be a more potent activator than histone acetylation[72]. But whether histone lysines are crotonylated or acetylated depends on the relative intracellular concentrations of crotonyl-CoA and acetyl-CoA, thereby linking cellular metabolism to gene expression[111]. It has also been suggested that the histone Kcr transferase activity of p300 is less efficient than Kac due to its structure and position in the substrate-binding tunnel, as the extended carbon length sterically restricts incoming substrates[112]. Kcr has been detected in many eukaryotes including some that do not possess the p300/CBP substrate, this would indicate that other enzymes are at work exhibiting HCT activity. Such as MOF, the evolutionary conserved HCT, can catalyse crotonylation at H3 lysine 4,9,18 and 23, and H4 lysine 8 and 12[106].

The class I HDACs have histone crotonylation “erasers” activity, all its members, HDAC1, 2, 3 and 8, exhibit histone decrotonylase (HDCR) activity *in vitro*[55,81]. To a lesser extent the sirtuins, Sirt1 and Sirt2[113] in mammalian cells and Sirt3 in *in* vitro show this activity. In embryonic stem (ES) cells HDAC1/2 have also shown “eraser” activity for both H3K18cr and H3K18ac overlapping at transcriptional start sites[114].

The proteins that are able to “read” histone crotonylation were identified among known Kac readers. These established acyl lysine readers, bromodomain double PHD finger, and YEATS (Yaf9, ENL, AF9, Taf14 and Sas5) have comparable binding capacities toward other short acyl chain modifications[71,115]. However the YEATS domain, Taf14 and double PHD finger proteins MOZ and DPF2[106] may be the preferred reader for Kcr. The crystal structure of AF9 YEATS domain was analysed in a complex with H3K9ac and H3K18cr by Li *et al.* (2016), they concluded that the YEATS domain preferred the longer and bulkier acyl groups[107]. Despite this preference towards Kcr over the other acyl groups[55], the YEATS domain nonetheless binds, albeit weaker, with other acyl lysine marks. Between Kcr and Kac the YEATS domain preference towards Kcr binding was 2-7 fold and several Kcr marks on H3K9, -K18, and -K27 can be bound by the AF9 YEATS domain. Taf14 and YEATS2 prefer H3K9Cr and H3K27Cr, respectively[116]. These preferences towards Kcr over Kac can be explored to further investigate the physiological and pathological role of histone Kcr.

Histone crotonylation has been greatly linked to active chromatin and identified as a positive regulator of transcription[107]. Kcr also has been shown to be associated with inflammatory gene expression: in macrophages stimulated with liposaccharides (LPS), p300 appeared to enhance these inflammatory responses linked to the presence Kcr and Kac at the promotor regions of the inflammatory genes. However, pre-treatment with crotonate, increasing the concentration of crotonyl-CoA and histone

crotonylation, enhanced dose-dependent site-specific H3K18cr of the inflammatory genes. This consequentially showed a decrease in H3K18ac[72].

Despite histone crotonylation being a relatively newly discovered acylation, it has been identified to have a diverse functional roles within the cell as well as physiological roles from signal-dependent gene activation to spermatogenesis and tissue injury[109].

In conclusion, histone crotonylation, can play a functionally different role in gene regulation from histone acetylation. The clear capacity of histone Kcr over Kac in the regulation of gene expression has contributed to recognising this modification as a crucial regulator for cellular signalling and tissue function. However, further research is required to clearly describe the regulation of histone crotonylation and its connections to different physiological and pathological processes.

1.4.1.3 Histone methylation

Histone methylation and demethylation can be linked to both activation and repression of the chromatin, depending on the histone modification. It is the transfer of one, two or three methyl groups from S-adenosyl-L-methionine to lysine or arginine catalysed by histone methyltrasferase (HMTs)[117]. The removal of these methyl groups is catalysed by histone demethylases (HDMs). The addition of mono-, di- and tri-methylation to the lysine has distinct functional implications[117].

When methylation occurs, specific genes within the DNA complexed with the histone can be activated or silenced. Methylation at residues H3K4, -K36 and -K79 are associated with active chromatin, whereas methylation at H3K9 and H3K27 mediates heterochromatin formation and is involved in gene silencing[118,119]. H3K27me3 are highly dynamic and very important in developmental processes, due to its reversible

nature. H3K9me3 are characteristically heterochromatic, but H3K9me2 is located more often at lowly expressed genes in euchromatin[120].

Histone methylation and its regulatory mechanisms are essential in mediating pathological processes in CRC. Under certain external conditions, aberrant histone methylation or demethylation can contribute to an array of processes that leads to carcinogenesis, proliferation, metabolic reprogramming, epithelial–mesenchymal transition, invasion, and migration[121]. One of the earliest initiators of CRC and drivers of tumour progression are mutations in the Wnt/ β-catenin signalling mediators[122]. In the absence of HMT, a reduction in H3K36me3 leads to increased activity of the Wnt/ β-catenin signalling pathway, which in turn promotes colorectal carcinogenesis. Another study demonstrating histone methylation directly stimulating the Wnt/ β-catenin signalling pathway and promoting the onset of CRC, found that the WNT3 promoter region had a reduction in H3K27me3 and a rise in H3K4me3[123]. High levels of H3K9me3 has also been found to be involved in cancer progression, specifically CRC. This is achieved possibly by promoting collective cell invasion[124] and motility[125]. High levels of H3K9me3, under hypoxic conditions can induce the negative regulator of p53, APAK, and can suppress p53-dependent apoptosis[126]. All these findings indicate that histone methylation can be linked to oncogenic functions, early research linked human cancer hallmarks with high expression of H3K9me3 and H4K20me3[125]. However, histone methylation is not just a marker for carcinogenesis, it also regulates many biological functions, including gene transcription, nucleosomal positioning, DNA replication and repair[125].

1.4.2 Functions of histone lysine modifications at specific residues

1.4.2.1 Histone 3 Lysine 9 (H3K9)

Modifications of the lysine residues of the N-terminal tails of histone H3 and H4 control the accessibility of chromatin regulating the promotion of gene transcription. Modifications at histone 3 lysine 9 (H3K9) have a number of roles. When H3K9ac histone is acetylated, genes are activated and when methylated (H3K9me) they are generally silenced. Since H3K9 is a residue that can be acetylated or methylated, many possibilities have been explored in terms of the progression of modification changes at H3K9, and what proteins coordinate modifications at H3K9 in various model systems such as embryonic development or gene expression models [127]. Chromosome condensation is a complex mechanism that ensures genome integrity. Since it can be acetylated and methylated at H3K9, histone deacetylation is required prior to H3K9 methylation[127] . It has been suggested that the deacetylation of H3K9 is a prerequisite of, and H3K9me3 is also required for chromosome condensation.

These pattern of histone modifiers can influence cellular processes across many cell types and one particular study discovered that H3K9ac and H3K9me3, as well as H3K4me3 expression was also regularly observed in CRC (75% of tumour samples, 77%, and 100% respectively). Lower levels of repressive histone mark (H3K9me3) were been detected in higher stages of tumours, but there have been no correlations between H3K9ac, H3K4me3, and HDAC1-3 [88,128].

1.5 Chromatin Remodelling

Chromatin remodellers are a group of proteins that use ATP to organise DNA and assemble nucleosomes for chromatin accessibly and nucleosome editing which involve the adding or removal of histone variants[129]. There are four subfamilies of ATP-dependent nucleosome-remodelling complexes, including imitation switch (ISWI), chromodomain helicase DNA-binding (CHD), switch/sucrose non-fermentable (SWI/SNF), and Inositol auxotrophy 80 complex (INO80), each are involved in the dynamic control of chromatin. Each subfamily has been tailored to more effectively accomplish certain chromatin objectives, such as chromatin assembly, access, or editing. These objectives are necessary for homologous recombination, DNA replication and repair, and regulation of gene expression.

The varied and crucial roles each remodeller uses ATP to mobilise DNA around the nucleosomes (Fig. 1.5.1). The key roles in regulating chromatin ranges from nucleosome assembly and spacing with the ISWI remodellers. CHD remodellers, associated with nucleosome sliding, eviction, spacing, and nucleosome assembly. The SWI/SNF family, which alters chromatin structure, participating in the sliding and eviction of nucleosomes[130,131]. The SWI2/SNF2-Related 1 Chromatin Remodelling Complex (SWR1) identifies nucleosomes that contain H2A and, in the presence of ATP and H2A.Z, substitutes H2A-H2B with the histone variant dimer H2A.Z-H2B in a histone exchange[132]. Finally the INO80 family which are closely related to SWR1, specializes in restructuring the nucleosome[129] while sliding nucleosomes along the DNA. They can also evict histone octamers from DNA.

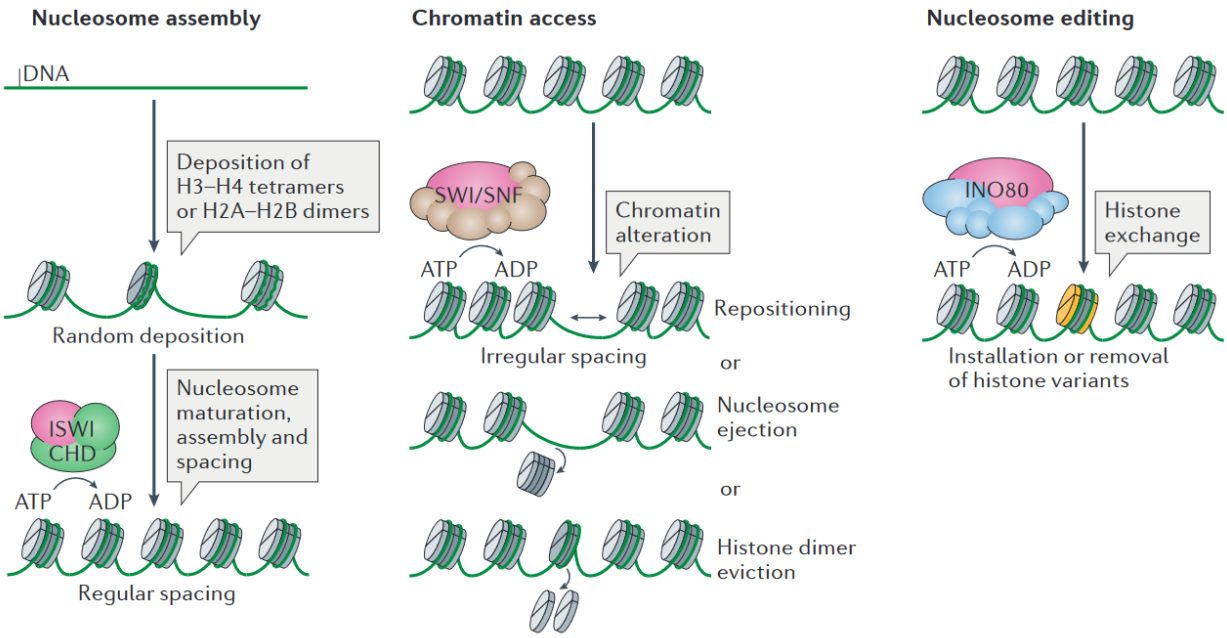


Figure 1.5.1 Chromatin remodelling functions.

"Functional classification of remodellers. The ATPase–translocase subunit of all remodellers is shown in pink; additional subunits are in green (imitation switch (ISWI) and chromodomain helicase DNA-binding (CHD)), brown (switch/sucrose non-fermentable (SWI/SNF)) and blue (INO80). ***Nucleosome assembly****: particular ISWI and CHD subfamily remodellers participate in the random deposition of histones, the maturation of nucleosomes and their spacing.* ***Chromatin access****: primarily, SWI/SNF subfamily remodellers alter chromatin by repositioning nucleosomes, ejecting octamers or evicting histone dimers.* ***Nucleosome editing****: remodellers of the INO80 subfamily (INO80C or Swr1 complex (SWR1C)) change nucleosome composition by exchanging canonical and variant histones, for example, and installing H2A.Z variants (yellow). This functional classification is a simplification, as INO80C, the ISWI remodeller nucleosome remodelling factor (NURF) and certain CHD remodellers can promote chromatin access." Figure adapted from Clapier et al. (2017)*[129].

1.5.1 Smarcad1

Smarcad1 (SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily-A containing DEAD/H box-1) is an ATP-dependent chromatin remodeller in the INO80 family[133,134] has appeared to be involved in histone eviction[134]. It plays a key role in gene regulation and maintenance[135]. The Smarcad1/Etl1/Fun30

family of chromatin remodelling factors is one of the most highly conserved families and is characterised by the inclusion of CUE-domains in addition to the helicase-like domains. Smarcad1 has a variety of roles in the regulation of genes, including gene silencing, maintaining heterochromatin and genome organisation and DNA repair[133]. One of the key functions of the chromatin remodeller is global deacetylation of histones H3 and H4, and the promotion of methylation. Through the recruitment of methyltransferase G9a/GLP, modifications of histones such as H3K9me3 can mediate gene silencing through heterochromatin markers[135,136 137]. Smarcad1 can influence heterochromatin through HDAC1 interactions and the transcriptional repressor KAP1[135].

Smarcad1 can also act as a gene promoter and form euchromatin. A study by Doiguchi *et al.* (2016)[138], showed a link between Smarcad1 and histone acetylases P300/CBP. They demonstrated Smarcad1's ability to enhance CBP-catalysed H2A K5 and K8 acetylation to promote transcription, supporting the idea that Smarcad1 is also a transcriptional activator[138].

Despite this remodelling factor being highly expressed in the intestinal epithelium, its tissue function is still poorly understood. However, mutations and deletions of Smarcad1 has been associated with many diseases including many carcinomas and some autosomal-dominant inherited diseases, e.g., Adermatoglyphia[139] and Basan Syndrome[136,140] (ectodermal dysplasia – abnormal development of skin, hair, nails, teeth and sweat glands, and an absence of fingerprints).

Our studies (Kazakevych *et al.* 2020)[141], within the Varga-Weisz group explored the role of Smarcad1 by generating a new conditional deletion model and focused on the role this factor plays within the intestinal epithelium. Analysis showed that a deletion

of Smarcad1 greatly affected the histone modification H3K9me3 linked to changes in chromatin assembly in the colon epithelium. This affect was seen over many sites including those genes that showed a reduction in gene expression when *Smarcad1* was knocked out.

This modified gene expression has also been reported to influence the intestinal epithelium-microbiome interactions resulting in a down regulation of a colitis response in a DSS-induced mouse model[141]. This study shows that highly conserved chromatin remodelling factors can result in colitis resistance within the cells, highlighting the impact the epigenome can have on anti-microbial defences.

The interaction found between Smarcad1 and the histone modification H3K9me3, a gene silencer, opened up the question towards the two previously mentioned gene activators, H3K9ac and H3K9cr. We know that in a study by Rowbotham *et al.* (2011)[135], SMARCAD1 regulates histone acetylation and H3K9 methylation during the cell cycle[135] and has shown in increase in acetylation of H3K9 upon SMARCAD1 depletion. This was also correlated with a decrease in di-and trimethylation of H3K9. Therefore, the question of how the knockout of Smarcad1 could affect the genomic localisation of H3K9ac and H3K9cr within the colon was investigated. It would be interesting find a link between Smarcad1 and these histone modifications to possibly light the way for future therapeutic targets on IBD and CRC.

1.6 Summary points

- IBD and CRC are a growing cause of concern worldwide.
- Gut microbiota plays an important role in both maintaining a homeostatic environment and facilitating an appropriate immune response.
- Antibiotic treatment can deplete the microbiota resulting in altered functions of the gene expression and histone modification responses. It is a good method to investigate host-microbiota interactions.
- DSS treatment can induce a colitis response in the gut. A good method to explore the epigenetic mechanisms during IBD.
- SCFAs can be generated by microbial fermentation and are metabolites with broad function in host metabolism and immunity.
- SCFAs are precursors of histone lysine acylations suggesting that SCFAs may function through regulating histone acylation and epigenetic responses.
- Histone acylations are modulated by the cellular metabolism of SCFAs.
- Histone lysine acylations identify transcriptionally active genes and have a role in a variety of physiological processes, including spermatogenesis, tissue damage, and stress brought on by metabolism.
- Histone acetylation and crotonylation are activators in chromatin regulation.
- Histone methylation acts as both repressor and activator depending on the lysine residue that is modified.
- Smarcad1 plays a dynamic role in chromatin assembly and gene expression within the colon.

1.7 Aims and hypothesis

The aim of this study will be to understand the link between chromatin dynamics and intestinal biology. The focus will be to explore how histone acylations promote gene expression and identify the connection between histone modifications, metabolism and the microbiota in cells from the colon intestinal epithelium.

My project will be focusing on the effect of epigenetic changes on histone modifications from 3 different angles.

I want to know whether the SCFA, crotonate has the same protective effect on the gut as butyrate. We know that butyrate can influence the histone modifications to maintain a homeostatic environment within the gut, and to provide an energy source to regulate immune responses and inflammatory activation from gut alterations. Crotonate has shown great potential in facilitating similar responses as butyrate considering its interaction with histone crotonylation. To investigate this, I will compare butyrate and crotonate effects on DSS induced colitis in mouse models.

I want to know how much of an influence the microbiota has on histone modifications and gene expression. We know that without the gut microbiota the downstream effect on epigenetic responses can be altered greatly. Histone acetylation and crotonylation both use the same protein receptors to control the function of the histones, and this may be linked to the SCFA generated by the microbiota, however but the functions of the two histone modifications have been suggested to vary greatly from each other. Therefore, focusing on global histone modifications of Kcr and site-specific histone modifications such as, H3K18 and H3K9, I want to explore the different functional, metabolic and biological processes that may be activated or supressed due to the

presence of acetylation or crotonylation. I also want to know the degree of enrichment and impact of each modification may have on a normal gut compared to a microbiota-depleted gut through antibiotic treatment.

Finally, I want to investigate whether the chromatin remodeller, Smarcad1 plays a role in controlling gene expression and immune response. We know that Smarcad1 can greatly influence histone tri-methylation at the H3K9 residue, supressing many genes from being expressed. What is yet to be investigated are the influences of Smarcad1 in H3K9ac and H3K9cr two activators with potentially different responses to Smarcad1. Through Smarcard1 KO the changes in gene regulation can provide a valuable insight into the interactions between with histone modifications and Smarcad1.

I will start by focusing in on the SCFAs and the phenotypic responses towards them within the colon, I will then focus on a genomic level and look at the effects of microbiota depletion on selected histone modifications and then lastly, I will focus on the Smarcad1 and its influence on histone modifications and gene regulation.

Chapter 2. Materials and Methods

2.1 Mice

C57BL/6 adult male mice (6-8 weeks) were obtained from Multidisciplincary Centre for Biological Investigation at the University of Campinas (CEMIB/UNICAMP). The animals were housed in the Animal Facility SPF of the Immunology area (DGEB/IB/UNICAMP) in a room with air conditioning, temperature and humidity control and light-dark cycles, with access to conventional chow (Nuvital) and water ad libitum. All animals were acclimatized and divided into experimental groups of 4-6 animals.

Mouse ethics for the SCFA/DSS experiments, were obtained with the approval of the Ethics Committee on Animal Use of the University of Campinas. For the Smarcad1 mouse experiments, mice were obtained from Babraham Instiute Biological Service Unit. The genotype of the KO mice was Lgr5-GFP-Sm-F-CE6:Hom-Vil-cre. Experimental protocols were approved by the Babraham Research Campus.

The rational to use the C57BL/6 were that they are the most used inbred strain in research. They are genetically identical within each strain, removing any genetic difference that could impact the research. The decision to use male only mice was to ensure the hormonal cycle in female mice would not influence the growth or behavioural variation in feeding, which could have thrown off the results.

2.1.1 SCFA pre-treatment

Prior to treatment the mice body weight was measured. The bedding containing their faeces from each cage was gathered, mixed and redistributed evenly to ensure a balanced exposure to their environments and microbiota.

All cages were given the same volume of food which is regular chow made with soybean meal, ground corn, fish meal and animal by-products[142].The mice were subjected to one week pre-treatment with 150 mM crotonate or butyrate. At the same time, the control groups were given 150mM Sodium chloride (NaCl) mixed into their drinking water to increase same osmolality of drinking waters.

2.1.2 DSS-induced colitis treatment

After a full 7 days of treatment, all groups were given 1.5% dextran sulfate sodium (DSS) (molecular weight 36,000-50,000; MP Biomedical's, cat no. 02160110-CF) added to their drinking water ad libitum. A control group received water without DSS throughout the experiment. Every day for 12 days the weight of the mice was measured and the amount of food and water consumed was monitored, too. In two other experimental models, 2.5% and 1% DSS was added to the drinking water and the effects of DSS and rate of sickness was monitored.

The clinical score was measured for each mouse over the experiment's duration to assess the colitis progression. The score was calculated following Kim et al. (2012), a point system that combines the weight loss, stool consistency and blood in stool and rectal bleeding of the mice. The point system is defined as, weight loss; 0 (no loss), 1 (1-5%), 2 (5-10%), 3 (10-20%), and 4 (>20%); stool consistency; 0 (normal), 2 (loose stool), and 4 (diarrhoea); and bleeding: 0 (no blood), 1 (haemoccult positive), 2 (haemoccult positive and visual pellet bleeding), and 4 (gross bleeding, blood around anus). Weight loss by more than 20% were considered as the point of no return and mice were put down.

2.1.3 Antibiotic treatment

Mice were treated with a 3 day course of antibiotics, 200µl per day. Antibiotics were administered by oral gavage. On the third day, 6 hours after the final gavage the mice were euthanized and dissected.

Table 1 - The antibiotic cocktail concentration.

Antibiotic	Conc (mg/ml)
Gentamicin	5
Metronidazole	5
Ampicillin	5
Neomycin	5
Vancomycin	2.5

2.2 ChIP-sequencing and ChIP-quantitative PCR

2.2.1 Mouse colon chromatin preparation

Mice colons were collected in both University of Campinas (Unicamp), Brazil and the Babraham Institute, Cambridge, UK. If collected in Unicamp, mice were anaesthetized using a super dose of 100 % isoflurane, followed by cervical dislocation. If collected at the Babraham Institute, animals were sacrificed by exposure to CO_2 followed by cervical dislocation.

2.2.1.1 Intestinal epithelial cell isolation on antibiotic treated mice

Antibiotic treated mice were prepared and processed at Unicamp in Brazil. The dissected colons were opened longitudinally and washed three times with ice cold HBSS without Ca^{2+}/Mg^{2+}. Colons were cut into 3-5 mm pieces, transferred to a 50 ml tube with 25 ml of 30 mM EDTA/HBSS and incubated on ice with shaking at 60 oscillations per minute for 1 h. Tubes were shaken vigorously by hand for 5 min

followed by an additional 10 min on ice with a fresh 20 ml solution of 30 mM EDTA/HBSS and further shaken for 5 min. Both solutions were poured through a 70 µM cell strainer to remove mucus and sub-mucosa. The extracted cells were pelleted at 480x*g* at 4 °C for 10 min and resuspended in 5 ml ice cold PBS and re-pelleted at 480xg at 4 °C for 5 min. Cells were fixed in 10 ml of PBS with 1% methanol free formaldehyde or paraformaldehyde for 10 minutes at RT with gentle agitation. After quenching with 0.125 M glycine, cells were spun at 480x*g* and 4 °C for 5 min. The cell pellet was resuspended in 1.2 ml PBS and transferred to a 1.5 ml tube before being spun again as before. Then the supernatant was removed and cell pellets flash frozen in liquid nitrogen. Samples were to be sent from Brazil to the UK on dry ice and stored at -80 °C until further processing.

2.2.1.2 Intestinal epithelial cell isolation from Smarcad1-KO mice

For *Smarcad1*-KO mice, colon epithelial cells were processed at Babraham Institute, the dissected colons were opened longitudinally and washed three times with PBS-EDTA (5mM)-sodium butyrate. Colons were cut into 3-5 mm pieces, transferred to a 50 ml tube with 25 ml of fresh PBS-EDTA (5 mM)- sodium butyrate and incubated at room temperature rocking at 20 rpm for 30 minutes. Tissue was washed in fresh PBS-EDTA- sodium butyrate and shaken vigorously for 5 min to release the colon crypt cells. Supernatant of the colon crypt cell solution was filtered through a 70 µM cell strainer to remove mucus and sub-mucosa. The extracted cells were pelleted at 400x*g* at 4 °C for 5 min. The pelleted cells were washed then resuspended in 1ml 1% formaldehyde /PBS-EDTA- sodium butyrate and fixed for 10 minutes at room temperature (RT) with low agitation. The reaction was quenched with glycine at a final concentration of 0.125 M for 5 minutes and the cells were pelleted at

2400xg 4 °C for 1 minute and washed three times with 1 ml of PBS- sodium butyrate. Final spin at 5000xg for 2 minutes and supernatant removed.

2.2.1.3 Chromatin preparation

Frozen pellets or fresh material were resuspended in colon ChIP lysis buffer (50 mM Tris-HCl pH8, 10 mM EDTA, 1% SDS) supplemented with protease inhibitors (5 mM sodium butyrate, 1 mM PMSF and 1x Protease inhibitor cocktail (PIC)) at a ratio (v/v) of 75-100 µl buffer to 10 µl approximate pellet volume. After incubation for 5 minutes on ice, samples were sonicated. For *Smarcad1*- KO experiments sonication was initially performed in 180 µl of lysis buffer and later the volume was tripled to 540 µl to ensure the ~1million cells were fully disrupted to obtain fragment size of 300-500 bp. Samples were placed in a water-cooled Bioruptor (Diagenode), high power, set to 45 cycles with 10 sec on and 80 sec off. A 30 µl sample of the sonicated chromatin was taken to for monitoring the sonication and kept on ice. The remaining sonicated chromatin was diluted in 1200 µl of RIPA ChIP buffer (10 mM Tris-HCl pH 7.5, 140 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 1 % (v/v) Triton X-100, 0.1 % SDS, 0.1 % Na-deoxycholate, 5 mM Na-butyrate) with freshly added protein inhibitors. Chromatin and RIPA ChIP buffer was mixed and spun at 16,200 xg for 10 minutes at 4 °C. The supernatant was carefully aspirated into a clean tube to chill on ice, leaving the remaining 100 µl of supernatant and pellet. Another 500 µl of RIPA ChIP buffer was added to the pellet, resuspended and spun at 16,200 xg for another 5 minutes at 4 °C. The supernatant was removed leaving behind ~60-100 µl in the tube and pooled with the first set of supernatant. The diluted chromatin supernatant was left on ice while the sonicated chromatin was checked.

The 30 µl sonicated test material was added to 90 µl of elution buffer (20 mM Tris-HCl pH 7.5, 5 mM EDTA, 50 mM NaCl and 1 % SDS) and 20 mg/ml 1µl proteinase K. These samples were incubated for 1 hour at 65 °C in a thermomixer with low agitation 300 rpm followed by purification with Qiagen PCR purification kit, eluting in 30µl. 500 ng was separated on a 1.5% agarose gel at 120V and 250 mA for 40 minutes to monitor fragment size distribution.

2.2.1.4 Chromatin immunoprecipitation

The chromatin DNA concentration was calculated by measuring the concentration of the sonication test using the Qubit high sensitivity assay kit. There is ~ 4 pg of chromatin/cell, this was multiplied by the 1 million cells in the mouse colon samples, giving a total of 4 µg of chromatin in each of the mouse samples. For acetylation and crotonylation the optimisation steps lead to doubling this amount to 8 µg per ChIP and for trimethylation 2.5 µg of chromatin.

Calculation of volume to be added to ChIP:

$$X\ (\text{Total chromatin in tube}) = \frac{\text{Qubit (chromatin conc.) x sonicated sample volume}}{\text{RIPA ChIP buffer (volume added to sample after sonication)}}$$

$$\text{Ac / Cr} = \frac{8\ \mu g}{X} \qquad \text{Me3} = \frac{2.5\ \mu g}{X} \qquad = \text{volume of chromatin added to each ChIP}$$

Chromatin material was added to tubes with A/G Sepharose Dynabeads (Invitrogen cat. no. 10001D) which had been previously incubated with the specific antibody for ChIP (H3K9me3 – ab8898 Abcam, H3K9ac – Millipore 06-942 and H3K9cr - PTM516) at a concentration of 1 mg/ml each for 2 h. The chromatin-antibody-bead mix was incubated overnight on a rotating wheel at 4 °C.

For the inputs, 3 % of the volume for each chromatin sample was added to a separate tube and kept at 4 °C. Following incubation, the material (not including input) was washed 2 times with 1 ml RIPA low salt buffer (20 mM Tris-HCl, pH 8.0, 2 mM EDTA, 150 mM NaCl, 1 % Triton-X-100, 0.1 % SDS) inverting samples to mix well and using a magnetic rack to separate the buffer from the chromatin bound magnetic beads. Chromatin was washed once more with RIPA buffer high salt (containing an additional 500 mM NaCl). This step was removed from the acetylation ChIP after a separate optimisation experiment took place; ~2 million fixed cells were split to 1 million each, and taken forward with the rest of the protocol, one with no change and one removing the high salt wash. The results showed a higher enrichment with the samples that did not undergo a high salt wash, so this was adopted into the protocol only for acetylation. All three ChIPs were then washed with TE buffer (10 mM Tris-HCl, 1 mM EDTA).

Samples including the input samples were resuspended in 200 µl of elution buffer (20 mM Tris-HCl pH 7.5, 5 mM EDTA and 50 mM NaCl) and incubated for 20 min at 65°C on a thermomixer at 1300 rpm. Afterwards, samples were treated with 2 µl of RNase (20 mg/ml) and 20 µl of 10 % SDS for 30 min at 37 °C. Samples were then incubated further with 2 µl of proteinase K (20mg/ml) for 2 h at 65 °C at 300 rpm. The samples were spun at 11,000 xg for 1 min to collect the supernatant. The DNA was extracted with phenol-chloroform and precipitated with 3 x volume 100 % EtOH, 1 / 10 volume 3 M KOAc pH 5.5.and 1 µl glycogen. DNA was incubated for 1 hour at -80 °C, spun down at max speed for 30 minutes at 4 °C. The supernatant was removed and the pellet was washed with ice cold 70 % EtOH. The pellet was spun down for 10 min at max speed and the DNA was left to dry before being resuspended in 20 µl ultra-pure water. Full step-by-step protocol can be found in Appendix 8.2.

2.2.2 ChIP quantitative PCR

2.2.2.1 Oligonucleotide primer design

Four types of target areas were focused on when designing the oligonucleotide primers. Type one targets were areas that showed changes in ChIPseq signal between WT and KO from a previous analysis from extracts of small intestine, type two targets were areas of no difference in signal between WT and KO - a positive control, type three were areas that gave a weak signal and target type four was anywhere where there were no reads – a negative control.

SeqMonk v1.44.0 was used to identify these areas of interest, by comparing the signal from Smarcad1 WT and KO of a previously ChIPseq analysis from extracts of the small intestine. The chromosome number was loaded into UCSC genome browser to locate sequence. Using the mouse assembly GRCm38/mm10, PCR amplicon length of 70-100 base pairs (bp) were selected to design primers with a target size of 2000 bp. As a guideline, genes that have been previously identified with Smarcad1 in the colon crypts with significant difference in MACS peaks enrichment between WT and KO were looked at first. The selected targets can be seen in table 2.

Table 2 – Gene targets for ChIP-PCR

	Target	Primer Sequences
Change between WT & KO	**Bglap3**	F 5' -CCTCAGGCAACGATGAGCAA
		R 5' -TGGCATGGAGCCTTTGAACT
	Aplp1/Kirrel2	F 5' -CTCATGGGAAAGCGGCCATA
		R 5' -GTTGATGACCAACCCCTCCA
	Snx18	F 5' -GCCACCGGTACACTAAGAACA
		R 5' -GCATGGATAGCACCCTTGGA
Positive Control	**Tiam1**	F 5' -CCTCAGATCAGCCAAGGACC
		R 5' -CTGTCCTTCAGGGACTCGTG
	PcdH17	F 5' -TCCACCAGATTAGATGTGTCCA
		R 5' -GCCCTGAAGAGGCACACTTA
	Sall3	F 5' -GAAGTCCGCCCACTTGAAGA
		R 5' -ACTAGCCTTGCTGTCCCAAT
Weak Signal	**Bambi**	F 5' -GGCCAAGCCTGAAATCTTGC
		R 5' -AAAGTCCCTGCGGGTTGAAG
	Mbd1	F 5' -GGAAACGCCGAGAGTCCTTT
		R 5' -GAACATTGCCCAGAGCACAC
	Sema7a	F 5' -GGGTCTCTGAGAAGGTCCCA
		R 5' -AAAGACAGGTGCACAGCTCA
	Lyz1	F 5' -AGTAAGCAGCGAGTTCAACGA
		R 5' -CAGTGTTTGTTACTGCAGACAT
	Dlk1	F 5' -TCCACCAGATTAGATGTGTCCA
		R 5' -GCCCTGAAGAGGCACACTTA
	Osr2	F 5' -AACCCAGTTTATGCGGAGGG
		R 5' -GGTGGGGGAGGAGAATTTGG
	Zeb1	F 5' -AAAGCAAGCAGGTACTGAAGAGT
		R 5' -ACACACATTAAAAGGCAGTGACAAA
	Snai2	F 5' -GAAAGCAGCTGGATTCGTGC
		R 5' -ACTAGCCTTGCTGTCCCAAT
Negative Control	**Intergenic 93/94**	F 5' -CAGCATTCCAGGAGGTTAGC
		R 5' -GTGCCTCATGTGCAGTCACT
	Klf11	F 5' -GAGAGGTGACAGGGTTTCCA
		R 5' -TGGCTTTATCTCGGGTTTTG

2.2.2.2 ChIP quantitative PCR and data analysis

Purified ChIP or input DNA were diluted with nuclease free water and SYBR® green PCR master mix (applied biosystems) with 0.3 ng DNA per reaction. The master mixes and 300 nM primer were loaded on a 384 well plate in triplicate. The run was performed

on a BioRad CFX384 qPCR system, with 10 minutes initial denaturation at 95 °C, 45 cycles of 15 seconds at 95 °C, one minute at 62 °C with detection and a meltcurve. To obtain ΔCq, triplicates were averaged, outliers removed and the Cq values subtracted by the input Cq value minus log2/100 (adjusting to from 1% to 100%). To get percentage input, 2^-ΔCq value was multiplied by 100 and the resulting value adjusted according to the different volumes used in the qPCR master mix. Fold enrichment was determined by dividing by at least one negative target and values were plotted in GraphPad Prism version 8.

2.2.2.3 ChIP PCR for antibiotic treated experiments

The single colon ChIP preparation and ChIP PCR protocol described above was used for antibiotic-treated IECs. During the sonication step the samples were sonicated with the bioruptor (Diagenode), high power, set to 30 cycles with 30 sec on and 30 sec off. The antibodies used for the ChIP was H3K18ac (Abcam ab1191). The oligo primers used in Table 3

Table3 – Gene targets for ChIP-PCR

	Target	Primer Sequences
	Klf11	F 5’ -GGTAAGTGACGGGGCTGAG
		R 5’ -GTCCGGACCAGAGCAAATC
	Klf13	F 5’ -GACATGGACACGAGGCACT
		R 5’ -GACGACTCCTGCGAGAGC
	Isg15	F 5’ -CGGTTTCCTTTTCCTACGGCAT
		R 5’ TACAGCTTTGACCCTTGAGAGC
	Brd2	F 5’ -TGGGTATGTAGGCACGAGGT
		R 5’ -CAAACACCTGACAGCCAATG
	cys1	F 5’ -GGGGTTGGAGACTAGCACTG
		R 5’ -GTCTCCCTCCTCCTCTCACC
	Slc4a5	F 5’ -GGGGCAGCTCTAGGTGTTCT
		R 5’ -TTAGCGCTGAAGCTCAAACA
Positive control	Casd1	F 5’ -GCTACCCTGGCCAACCTTAG
		R 5’ -TGTAAGGCCCATGTTCACCG
Negative control	Adam23	F 5’ -TGCTGGAATGGAGTGCTCAG
		R 5’ -GGGTACGGAGACCTCGTAGT
	March11	F 5’ -TCTCTCTCACGCACGACTTG
		R 5’ -CTGGGAACAAGCACTGGACT

2.2.3 ChIPseq and bioinformatics sequence data preparation

The samples for Smarcad1-KO experiment libraries were assembled and sequenced by BGI using a company specific platform similar to Illumina sequencing in a 50 bp single end sequencing run.

The samples for antibiotic treated experiment, libraries were assembled and sequenced by Novogene, using the NovaSeq 6000 Pair End150 bp platform.

Downstream processing was performed by Babraham Bioinformatics Facility using the clusterflow fastq_bowtie2 pipeline, comprising adapter and quality trimming using trim galore v0.6.1 for Smacard1 experiments and v0.6.6 for antibiotic treated experiments. Trimming mode was with default single end parameters for Smarcad1 experiments and paired end parameters for antibiotic treated experiments, followed by mapping to

the mouse GRCm38/mm10 genome assembly using bowtie2 v2.3.2 for Smarcad1 and v2.4 for the antibiotic treated.

2.3 Western blot

Protein was extracted from colonic tissues by maceration on dry ice and immediately denatured at 95 °C for 5 minutes with 2 x Laemmli buffer and 5 % β-mercaptoethanol. Tissue was sonicated with the bioruptor (Diagenode), 6 cycles of 30 seconds on, 30 seconds off set to high power. The sonicated samples were then spun down at 11000 xg for 10 minutes to separate the protein in the supernatant from insoluble debris. The protein concentration was determined using the EZQ protein quantitation kit. The EZQ quantitated results were analysed using ImageJ computer software to measure the concentration of the samples against the standard curve, each sample was diluted to a final concentration of 1 µg/µl. This was achieved by adding more 2 x Laemmli + 5 % B-Me.

About 5 µg of proteins from the respective samples quantified were pipetted into 12% SDS-PAGE gels (TruPAGE Precast Gels, Sigma-Aldrich). After the run, the proteins were transferred to a polyvinylidene fluoride (PVDF) membrane using a semi dry The Trans-Blot Turbo Transfer System (Bio-Rad) with transfer buffer containing 20% methanol. The membrane was incubated for approximately 1 hour in blocking solution (3% albumin; 10 mmol/L Tris, 150 mmol/L NaCl, 0.02% Tween 20) and for 12 hours with antibody, e.g., anti-histone H3 (1:40,000; ab1791). After three short and three longer (each ~ 15 min) washed in 10 mmol/L Tris, 150 mmol/L NaCl, 0.02% Tween 20, the membrane was incubated with the anti-Rabbit IgG-Peroxidase secondary antibody (1:10,000; A0545) for 1 hour. Detection was by chemiluminescence using

ECL western blot substrate (Thermo Scientific, USA). The developed bands of the anti-H3 proteins were quantified by means of optical densitometry in ImageJ.

H3 loading control was calculated by square rooting the pixels (protein density) captured from ImageJ, calculating the average from all samples, then using the average to calculate the percentage each sample would needs to be altered by. Use this percentage work out the loading volume of each sample. (Example can be seen in appendix table 8.1).

The western blots were repeated using the calculated volumes per well, and one of the following primary antibodies was used: anti-crotonylated lysine (anti-Kcr) (1:5,000; PTM-501), anti-H3K18cr (1:1,000; Cusabio/PTM517), anti-H3K18ac (1:1,600; ab1191), anti-H3K9ac (1:1,000; Millipore 06-942), anti-H3K9cr (1:1,000; PTM-516) anti-H3K9me3 (1:1,000; Abcam 8898) and anti-histone H3 (1:50,000; ab1791).

The relative density of each sample bound to its corresponding antibody was calculated by averaging each pixel area obtained from ImageJ. The percentage of density was calculated by sample over average pixel multiplied by 100. This percentage density of each sample at each antibody was divided by the percentage density of H3 to give the relative density. These values were plotted using the ggplot2 package on R.

2.4 Bioinformatics analysis

2.4.1 ChIPseq analysis

Mapped reads were imported into SeqMonk (v1.47.2) as bam files for data exploration and analysis. For the Smarcad1 data set, bam files were imported into the program as single end reads with duplicates removed. The antibiotic treated data set I imported bam files as paired end reads. Data groups of the samples were made by histone

modification marker and by type – WT or KO, i.e H3K9ac_WT, H3K9ac_ KO, and Input.

MACS peak analysis was with SeqMonk. MACS peaks were defined using ChIP samples and input samples. Fragment size was 300. The significance threshold was 1.0E-5. The deduplication step was skipped. Quantitated with Read Count Quantitation using All Reads, no log transformation and did not correct for total read count to perform analyses on raw data.

EdgeR analysis was performed within Seqmonk v1.47.2 to identify differential enrichment. EdgeR stats filter tested on probes from the MACS peaks probes where KO vs WT replicate groups had a significance below 0.05 after Benjamini and Hochberg correction. Existing probes were quantitated again by read count and corrected for total read count of all reads. Log transformation with a base of 2.0 with a lower limit of 1.0 was applied and data viewed on at scatterplot and box and whisker plot. Probes were also generated over TSS +/- 500 bp and running windows of 200 bp. Quantitations were normalised by match distributions of all data sets bar the inputs.

2.4.2 RNAseq analysis

For gene expression the libraries analysed by DESeq2 ($p < 0.5$) and normalised in SeqMonk. Annotated gene lists of up and down regulated gene expression were saved and imported into each histone modification library analysis SeqMonk file as an annotated track to identify and compare differentially expressed genes against differentially acylated genes.

2.4.3 Gene ontology analysis

Gene ontology analysis was performed using the Panther (http://pantherdb.org/) statistical overrepresentation test using a list of gene IDs that were exported from

SeqMonk. The raw counts in my list of genes were compared with the expected values of differentially expressed genes found in the reference list from Panther, this gave the frequency of occurrence of each assigned functional category. To focus on the list of biological process that could be most significant, and avoiding those that could be identified by chance, terms containing more than 1000 genes in the reference list were removed. Statistical analysis on the frequency of occurrence was performed by Panther using Fisher's exact test, and terms with an FDR of more than 0.01 were removed. The analysis was performed using the reference list provided by the online software, this will have had an effect on the probability of the terms generated, which could be misinterpreted as important findings. The information taken from the gene ontology analysis is used as biological knowledge, not numbers. The terms were grouped into related processes and plotted in R using the ggplot2 package.

2.4.4 Motif analysis

The ChIPseq MACS peaks from H3K18ac, H3K18cr and Kcr were exported in a .bed file format from SeqMonk and converted into fasta format using the rtracklayer package in R. Each peak sequence was either extended to or cut off at 500bp +/- from the centre of the peak, and converted in to a FASTA format to import into the MEME-suite program (v5.1.1) (https://meme-suite.org/meme/tools/meme-chip)[143]. Each set of ChIPseq peaks were analysed against the *Mus musculus* input database Homo sapiens Comprehensive Model Collection (HOCOMOCO) Mouse (v11 FULL)[144]. This database provides TF binding models (binding profiles or binding motifs) for mouse transcription factors. All the models were produced by the ChIPMunk motif discovery tool and supplied with the quality ratings from A (highest confidence) to D (low confidence). ChIPseq data for motif discovery was extracted from Gene Transcription Regulation Database (GTRD)[145] of BioUML platform. The models' IDs are based on

UniProt protein IDs. The FULL collection contains the CORE collection; primary binding motifs that robustly represent binding sites across multiple experiments plus all the high-quality alternative and lower-reliability binding models built from limited experimental data. The FULL collection contains models of A/B/C/D quality and includes alternative binding models for a single TF. This *Mus musculus* input dataset is provided by the MEME-suite program. The Markov background model is used to normalise the biased distribution of letters and groups of letters in the sequences to distinguish the motifs from the sequence background noise[146]. The minimum motif size selected was a width of 6 letters and the recommended background model was a 2nd order of model sequences, correcting for biases.

2.4.5 Statistical analysis

Statistical analysis on the effect of SCFAs in DSS induced colitis, used a one-way ANOVA for all analysis comparing each test group against the control group. Bonferroni multiple correction test was performed to control for type 1 error.

Venn diagrams were generated in R using the venneuler package, counting each gene or promoter name that contained a MACS peak. Considering the data was comparing two variables against two other variables in a 2x2 contingency table, Chi^2 statistical calculator was performed with Yates correction on all of the Venn diagram results using the website (https://www.socscistatistics.com/tests/chisquare/default2.aspx). Volcano plots were generated in R using ggplot2, using the EdgeR and DESeq data sets retrieved from SeqMonk.

Chapter 3. Protection from DSS-induced colitis by short chain fatty acids

Inflammatory bowel diseases (IBD) include two disorders; Crohn's disease (CD) and ulcerative colitis (UC). Both diseases cause chronic inflammation of the intestines and present with similar symptoms. Colorectal inflammation results in diarrhoea, bloody mucus stool, abdominal pain, fever, fatigue and a number of other symptoms reducing the overall quality of life. However, UC occurs only in the colon and presents as a consistent inflammation along the entire colon. The pathogenesis of UC includes a combined mechanism of genetic, immunological and microbiological factors[3].

Inflammation in the gut occurs when the intestinal epithelial barrier is disrupted and the luminal bacteria can enter the lamina propria[147]. Administration, usually through the drinking water, of dextran sulphate sodium (DSS) can replicate the clinical presentation of ulcerative colitis (UC) in animal models, allowing this inflammation to occur. It is a water-soluble sulphated polysaccharide that triggers an immune response and disintegrates the gut mucosal layer[147]. This increases the colon permeability by destroying the intraepithelial cell tight junctions altering the microbiota and reducing the mucosal barrier. This then allows the infiltration of bacteria into the systemic circulation, triggering an inflammatory response. Inflammation develops mediating an innate immune response activating macrophages and promoting pro-inflammatory cytokines.

Short chain fatty acids (SCFAs) such as acetate (63-43mM in the gut)[148], propionate, butyrate (50-70-mmol)[149], and the lesser researched crotonate, are all organic acids

produced in the intestinal lumen by bacterial fermentation of complex carbohydrates[54]. SCFA can help regulate sodium, calcium and water absorption. They can also lower colonic pH, which inhibits opportunistic pathogens[74] and promotes the favourable bacteria phyla, Bacteroidetes and Firmicutes[150] which make up ~60% of the microbiota population[74]. A diverse microbial population within the gut can protect the lumen from many diseases including inflammatory bowel disease (IBD) by promoting a protective mucosal barrier between the gut tissue and potentially harmful pathogens[74], and biochemically by breaking down the dietary fibres producing the SCFAs to act as a useful energy source in the maintenance and protection of the gut.

Despite butyrate being low in abundance (~15% of total SCFAs found in human colons) compared to acetate (~60%) and propionate (~25%)[74,151], this SCFA has received significant attention due to its instrumental role in cellular metabolism, intestinal homeostasis, proliferation, differentiation and apoptosis regulation[74,150]. Butyrate is the preferred energy source for the gut epithelial cells and is a key nutrient in modulating biological responses such as inflammation regulation and promoting genomic stability[74]. It is a histone deacetylase (HDAC) inhibitor and binds to specific G protein-coupled receptors (GPCR) which can influence immune function and tumour suppression[54,152]. Many studies have noted that butyrate can function as a protective factor against colonic disorders[151], promoting the removal of dysfunctional cells and highlighting its potential protection against colon cancer[74].

HDACs are a class of enzymes that remove acetyl groups from ε-N-acetyl-lysine on the histones causing the DNA to wrap tightly. Butyrate has the greatest ability to inhibit this out of all the SCFAs and many anticancer mechanisms of butyrate have been linked to the inhibition of HDAC. This leads to cell proliferation inhibition, cell differentiation and apoptosis activation, and regulation of gene expression. With

inflammatory bowel disease where macrophages and T-lymphocytes are activated, HDAC inhibition also results in an anti-inflammatory response; butyrate down-regulates the pro-inflammatory effectors in lamina propria macrophages[66] as well as regulates the cytokine expression in T cells[150], suggesting that butyrate can protect the gut from IBD.

Crotonate, a SCFA that has not yet been fully explored but is of great interest in terms of the SCFA's effect on intestinal health and epigenetic responses. Recent research has shown that with crotonate supplementation histone crotonylation was increased in intestinal organoids[55], this HTPM can also be promoted by butyrate. Histone crotonylation is regulated by class I HDACs affecting gene expression and the epigenome suggesting that HDACs and the corresponding SCFA could play a similar role that butyrate plays in gut protection, metabolic function and immune response linked to the gut microbiota. However, crotonate does not appear as good a HDAC inhibitor as butyrate[55] where supplementation of crotonate promotes crotonylation, but not acetylation, while butyrate promotes both. This suggests these similar SCFA are functionally distinct.

Based on similar studies that have investigated butyrate as a protective element in DSS induced colitis. I wanted to determine the crotonate response and whether this SCFA would provide protection to the gut to the same degree.

3.1 Acute or chronic DSS induced colitis with crotonate

DSS induced colitis can be greatly altered by the concentration of the DSS administered, resulting in either an acute colitis response or a chronic one[153]. To determine the effects of DSS concentration and the level of protection that crotonate may have, mice were treated with 150 mM of crotonic acid in their drinking water for 7 days prior to a DSS induction. Studies have shown noticeable differences in the gut microbiota with a greater abundance of SCFAs-producing bacteria and lower α-diversity compared to the chronic colitis model[3].

3.1.1 Acute DSS induced colitis with crotonate

Over a course of 7 days, 150 mM of crotonate was given to mice (n=5) through their drinking water. The average water intake per mouse was measured daily to ensure they were drinking enough water and to note whether the test group were just as hydrated as the control group (n=8) on water only. Their food consumption was also measured daily along with their body weight. Over the 7 days, there was an average consistency of food and water intake (Fig. 3.1.1a) and a gradual increase in body weight which can be expected as these were young growing mice.

The control group was split into control – water only (n=4) and water plus DSS 2.5% (n=4). DSS 2.5% was added to the crotonate group (n=5) and all three groups were monitored over the course of 6 days. The water intake of DSS group initially jumped on the first day but levelled out on subsequent days (Fig. 3.1.1b). The food intake of all three groups appeared to also be consistent with each other on each day. The total weight loss of the DSS only group dropped to an average of 4g and the DSS plus crotonate group also dropped by 3g after 6 days of DSS. This drop of 20% and 13% in weight respectively, indicates crotonate may possess a potential protective ability

against DSS induced colitis, but these numbers are not significant enough to give a definitive answer.

After the experiment the mice were sacrificed and the colons were harvested for analysis. The cecum and the colon length were measured for each mouse and a stark difference was found in the size of colons from the mice that had been treated with DSS (Fig. 3.1.1c) against the control. Both the DSS only and the crotonate plus DSS treated mice had significantly shrunken colons and cecum. Although the crotonate group had only a very slightly larger cecum than the DSS only groups, their colons had both shrunk the same rate on average.

The rapid and noticeable effects of DSS on the mice's overall decline in health and body weight suggests that 2.5% DSS is far too concentrated and too severe for the mice. It did not allow time for any protective responses to occur within the gut from the crotonate treatment. A study by Cui *et al.* (2021)[154], induced acute and chronic colitis using a method of adding 2.5% DSS dissolved in drinking water given to mice ad libitum for 7 consecutive days followed by a 2-day tap water period for acute colitis. Chronic colitis was induced by giving the mice the same concentration of DSS for the same length of time, with an added 14 days of H_2O, before resuming with the DSS solution. They repeated this method two more times, to allow the mice to have a mini recovery period and to mimic a chronic colitis response. This method could have been adopted in my own experiments to determine whether the longer time period supplementing the mice with crotonate and 2.5% DSS intermittently, could have resulted in a greater protective response to chronic colitis instead of acute colitis.

However due to time constraints, another method was adopted instead by lowering the DSS concentration to 1%, which would reduce the severity of the inflammation

within the intestinal lumen and overall colitis response. This was to attempt to create a condition close to chronic colitis, and allow the mice some time to respond to the crotonate.

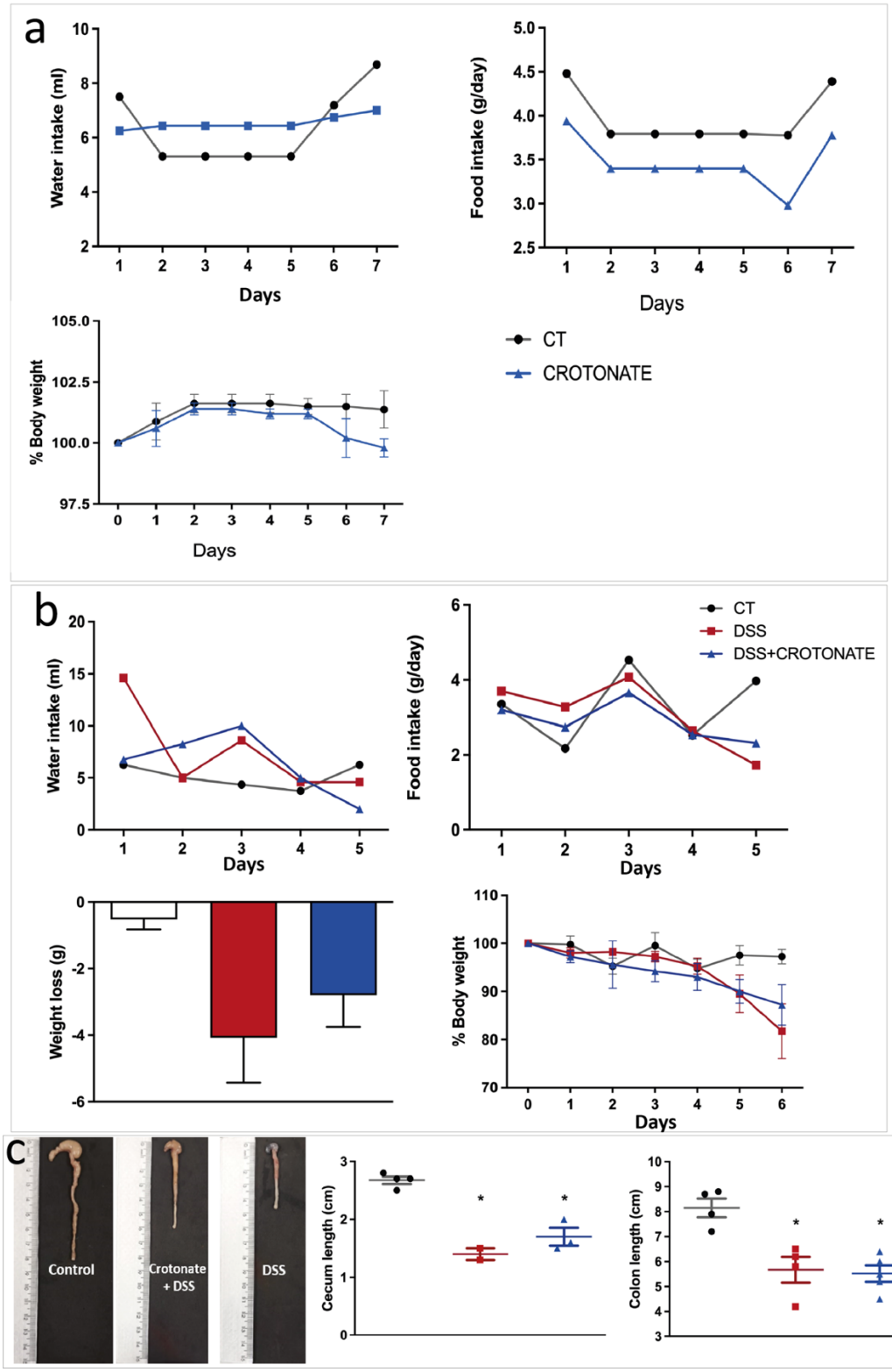


Figure 3.1.1 Crotonate treatment on acutely induced DSS colitis

Three mouse groups were analysed, control n=4, 2.5% DSS n=4 and 2.5% DSS plus crotonate n=5. ***a)*** *Water intake (ml/day), food take (g/day) and body weight percentage were measured over 7 days and the mean averages of each group was plotted on a line graph.* ***b)*** *the same mice were measured for 6 more days and the water intake (ml/day), food intake (g/day), total weight loss in grams and body weight percentage loss were all plotted on graphs.* ***c)*** *Pictured are 3 colons with cecums one from each group and presented to measure size difference, the cecum length (cm) and colon length (cm) were plotted on a scatter graph and the significance was measured.*

3.2 Chronic DSS induced colitis with crotonate

Mice (n=4) were treated with 150 mM of crotonate over 8 days. The control group of n=8 were given water only alongside the treatment group. The water and food intake were measured daily, to ensure consistent consumption between the groups (Fig. 3.1.2a). The water intake over the 8 days showed dramatically fluctuating volumes of water consumed each day in the crotonate treated group. However, the body weight of both groups stayed fairly consistent with each other.

The control group was again split into a water only group (n=4) and a 1% DSS only group (n=4). The crotonate group was also give 1% DSS and again the food and water intake was monitored each day along with the weight loss/gain of each group. This experiment took twice as long as the 2.5% DSS treated experiment before a noticeable response was seen from the DSS. Indicating a successful attempt at inducing a chronic colitis response. The water intake of each of the groups started out fairly consistent over the first 8 days after DSS treatment (Fig. 3.2.1b) until day 9, where the crotonate plus DSS group drank almost triple the amount the other two groups drank and remained slightly higher until the end of the experiment.

At day 8 and 9, two mice from the crotonate treated group were terminated due to extreme weight loss and sickness. This can explain the spike in the average water intake at day 9 as both mice were excluded from the graphs but the averages were affected by the smaller group number. It is possible that the mice drank too much water due to the increased osmolality of the crotonate water.

Over the course of the experiment the daily weight loss between the control group and the DSS group was very apparent. The total change in weight between the two groups was a 13 % difference by day 14. The crotonate plus DSS treated group although appeared to show a fair level of protection against DSS induced colitis with an average weight loss of only 2%, it is not a reliable result as two of the four animals in the group had been prematurely terminated skewing the average. The total weight loss of each mouse in each group can be seen in table 3.2.1.

Table 3.2.1 Total weight changes of mice

CT % (g)	DSS % (g)	DSS + CR % (g)
1 (0.6)	-9 (-2.9)	*-29* (-7.9)*
5 (1.1)	-8 (-1.9)	-5 (-0.9)
4 (1)	-7 (-2.6)	*-22* (-6.4)*
3 (0.9)	-8 (-2.2)	0 (0.5)

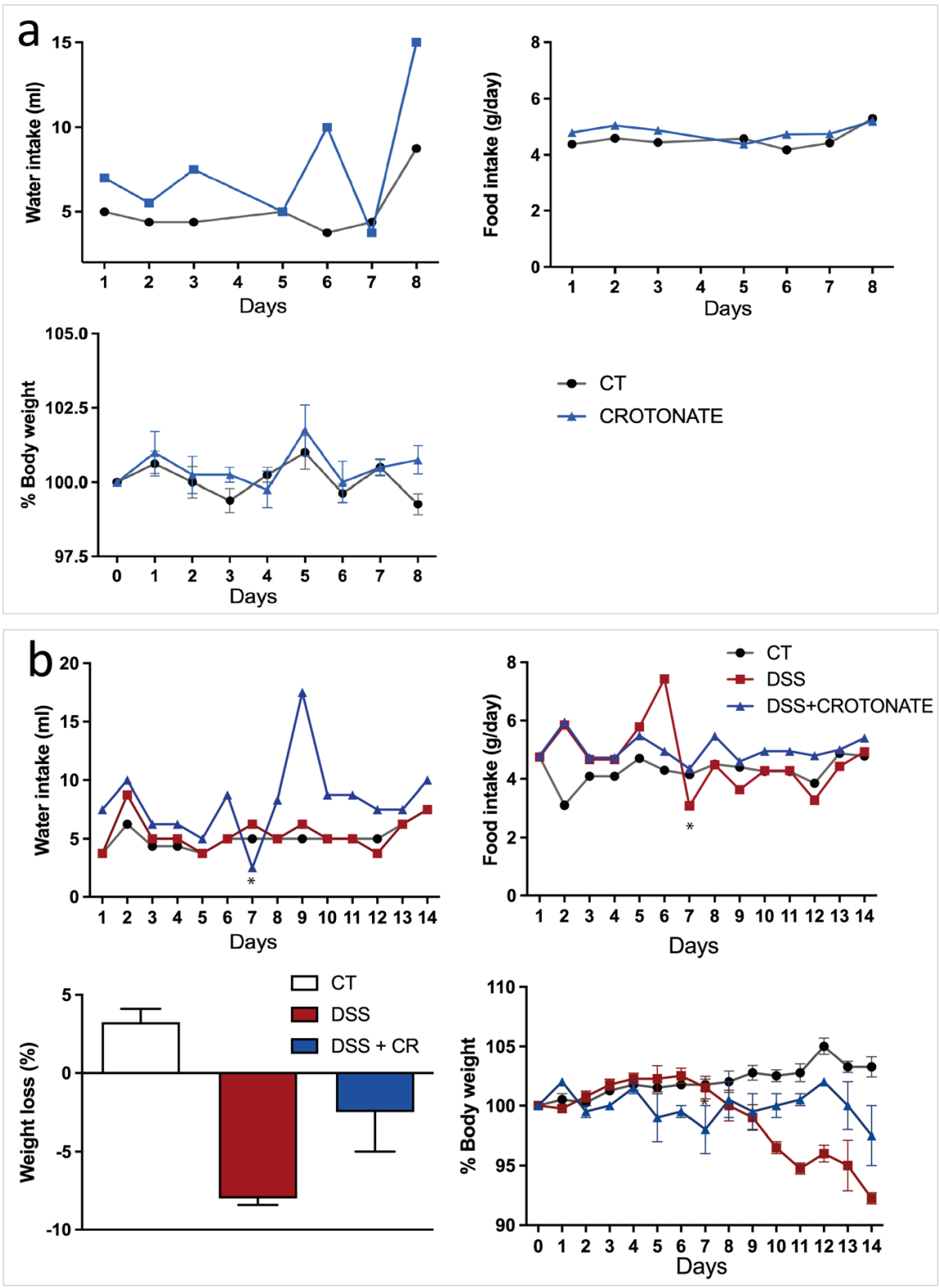


Figure 3.2.1 Crotonate treatment on chronically induced DSS colitits

Three mouse groups were analysed, control n=4, 1% DSS n=4 and 1% DSS plus crotonate n=4 (2 mice were euthanized due to rapidly reaching end point, indicated by the asterisk in ***b****).*

a) *Water intake (ml/day), food take (g/day) and body weight percentage were measured over 8 days and the mean averages of each group was plotted on a line graph.* ***b)*** *The same mice were measured for 14 more days and the water intake (ml/day), food intake (g/day), total weight loss in grams and body weight percentage loss were all plotted on graphs.*

3.3 Effects of crotonate and butyrate supplementation on DSS induced colitis

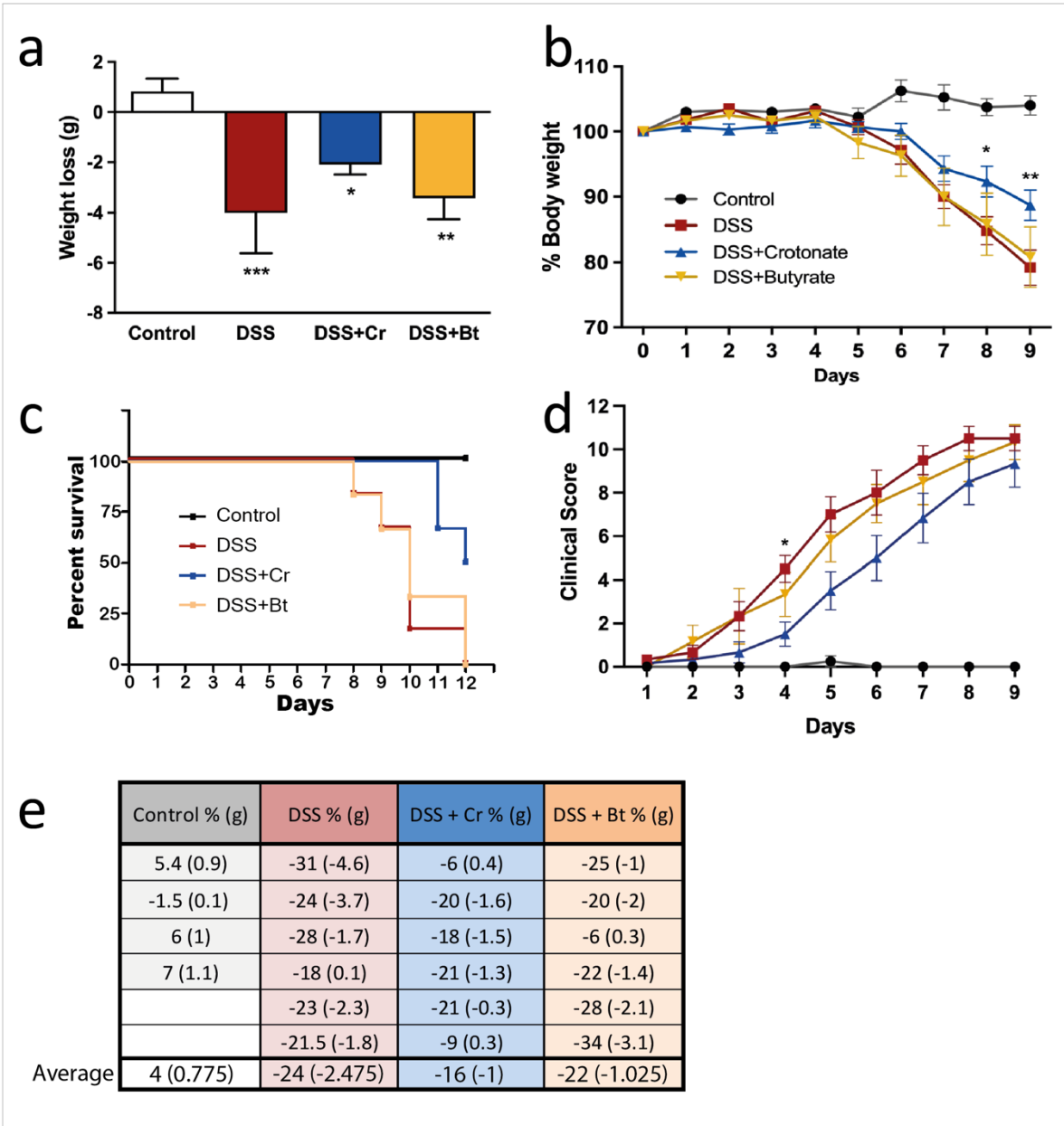
The purpose of the experiment was to determine the survival rate of mice while being treated with crotonate and butyrate. It has been established that butyrate can provide a great level of protection to the gut epithelium, providing a valuable energy source to stay healthy and regulate immune responses. From the previous potentially interesting results with crotonate and DSS induced colitis, it would be interesting to compare the two SCFAs and their protection levels relative to each other and compared with no SCFA treatment.

Two experiments took place in parallel, with a 24-hour gap between them. In groups of 3 mice per cage, (done in duplicate, two cages per group) mice (n=6) were subjected to one week pre-treatment with 150 mM crotonate or butyrate. While the control groups were given 150 mM NaCl mixed into their drinking water to increase osmolality and to similarly reflect the two SCFA treated waters.

After a full 7 days of treatment, the groups had 1.5 % DSS added to their drinking water. A control group of four mice in 150 mM NaCl, remained DSS free. Every day for 12 days the weight of the mice were measured and the amount of food and water consumed was monitored also. The clinical score for each mouse was taken daily.

The clinical score was measured over the duration of the experiment to assess the progression of the colitis. The score was calculated using the method by Kim et al.

(2012)[155], a point system that combines the weight loss, stool consistency and bleeding of the mice. The point system is defined as, weight loss; 0 (no loss), 1 (1-5%), 2 (5-10%), 3 (10-20%), and 4 (>20%); stool consistency; 0 (normal), 2 (loose stool), and 4 (diarrhoea); and bleeding: 0 (no blood), 1 (haemoccult positive), 2 (hemoccult positive and visual pellet bleeding), and 4 (gross bleeding, blood around anus). Weight loss by more than 20% were considered at the point of no return and mice were sacrificed.



e

	Control % (g)	DSS % (g)	DSS + Cr % (g)	DSS + Bt % (g)
	5.4 (0.9)	-31 (-4.6)	-6 (0.4)	-25 (-1)
	-1.5 (0.1)	-24 (-3.7)	-20 (-1.6)	-20 (-2)
	6 (1)	-28 (-1.7)	-18 (-1.5)	-6 (0.3)
	7 (1.1)	-18 (0.1)	-21 (-1.3)	-22 (-1.4)
		-23 (-2.3)	-21 (-0.3)	-28 (-2.1)
		-21.5 (-1.8)	-9 (0.3)	-34 (-3.1)
Average	4 (0.775)	-24 (-2.475)	-16 (-1)	-22 (-1.025)

Figure 3.3.1 The survival rate of DSS induced colitis mice with crotonate and butyrate treatment.

Four groups of mice were analysed, a control group (n=4) with 150 mM NaCl in the drinking water and no DSS, a 1.5% DSS plus 150 mM NaCl group (n=6), a 1.5% DSS plus 150 mM crotonate group (n=6), a 1.5% DSS plus 150 mM butyrate group (n=6). ***a)*** *Total weight loss in grams of each group was plotted and the p-values were calculated against the control group. DSS only group had the most significant weight loss, followed by butyrate and then crotonate.* ***b)*** *The body weight percentage over the course of 8 days were plotted in a line graph.* ***c)*** *The survival percentage of each group over 12 days was plotted. Mice were killed once they reached < 20 % weight loss or were allowed to die naturally whichever came first.* ***d)*** *The clinical score over 8 days while all the animals were alive. The clinical score calculation can be found in 3.3.2 Results.* ***e)*** *Table of final weight loss/gain percent and grams for each mouse after 12 days of DSS administration.*

Mice that were treated with 1.5% DSS only, became sick and died at a faster rate than those that had been treated with the SCFAs. The mice either died naturally or were killed once they reach a point of < 20 % weight loss. The rate of weight loss in the DSS only group was significantly greater than the control group and the DSS plus crotonate group with average total loss of 24% (Bonferroni's multiple comparisons test) (Fig 3.3.1e). The butyrate plus DSS group appeared to show minimal to no protection in preventing DSS induced colitis with an average weight loss of 22%. However, the very interesting finding was that the crotonate plus DSS group was significantly different compared to DSS only treated mice with a weight loss of 16% after 12 days of treatment and slightly significant compared to butyrate treated group (Fig 3.3.1e). When measuring the total weight loss between the groups (Fig 3.3.1a), Control vs DSS, DSS+Cr and DSS+Bt all three show a statistical significance with Control vs DSS being the most significant followed by Control vs DSS+Bt and then Control vs DSS+Cr as the least significant (Bonferroni's Multiple Comparison Test). Although the total difference between DSS and DSS+Cr was not statistically significant with a Paired t-test (P = 0.0911) (Fig 3.3.1a), the decline in body weight percentage each

day (Fig 3.3.1b) did show a significant difference in weight between DSS and DSS+Cr, and DSS+Bt and DSS+Cr on day 8 and 9 (Holm-Sidak's multiple comparisons test). The decline in body weight percentage in both DSS only and DSS+Bt were more dramatic compared to DSS+Cr as the experiment progressed.

The survival rate of each of the group shows that those mice treated with crotonate resulted in a better survival rate than those mice treated with butyrate. At the end of the experiment, 50% of mice treated with crotonate had survived whereas mice treated with butyrate did not make it to the end of the experiment Fig 3.3.1c). The butyrate group presented with a clinical score very similar to the 1.5% DSS only group (Fig. 3.3.1d). Despite the DSS+Cr treated group showing a clinical score trend lower than the other two test groups, the crotonate group only showed any statistical significance compared to the 1.5% DSS only group on day 4 (Tukey's multiple comparisons test). Overall, the crotonate group fared slightly better than both the 1.5% DSS only group and the DSS+Bt group in protecting the mice from an induced colitis response, but it must be noted that the clinical scores are marked by human interpretation which can introduce errors, albeit all the scores were measured by myself to ensure consistency. Together the results show exciting and promising results indicating the effectivity of crotonate in providing a level of protection the gut.

3.4 Discussion

Crotonate appears to have a significant effect on the protection of the gut when induced with 1.5% DSS. This finding has many potential avenues to explore, particularly as crotonate appeared more protective than butyrate,

The simplest explanation for this stark difference between the two SCFA responses could be that the group sizes are just too small to give a fair indication of how they

provide protection within the gut. However, there have been a number of animal studies that have shown SCFA treatment reducing weight gain[156] and any increases in energy expenditure and metabolism are influenced by these SCFAs[157]. In obese mice, oral treatment of butyrate have been shown to reduce body weight by increasing fat oxidation[156]. Another study with oral treatment using acetate, propionate and butyrate given to mice on a high fat diet showed increased insulin sensitivity and also a reduction in body weight[158]. These studies could suggest that in our experiments the butyrate could have contributed to the weight loss on top of the DSS effects on the body.

In normal intestinal epithelial cells (IECs), butyrate acts as a support element but in colonic carcinoma cells, it induces apoptosis. A study by Zeng & Briske-Anderson[159], showed a prolonged butyrate treatment *in vitro* with HT1080 tumour cells resulted in a dual effect of both pro- and anti-metastatic gene expression. The overall effect however showed the inhibition of the pro-metastatic and activation of the anti-metastatic genes. The concentration levels of butyrate is a factor in cell growth and migration. Low concentrations enhanced cell proliferation and high amounts inhibited the activation of tumour cell potential [159]. This supports another study highlighting the importance of butyrate concentration levels, Jiminez *et al.* (2017)[160] showed that administering butyrate at a concentration of 140 mM through rectal administration fared much better in clearing gut infection than mice given 80 mM of butyrate. They identified the beneficial effects a higher concentration of butyrate would provide to gut inflammation. This is interesting to note as the route of passage differs from our experiments; we gave the mice 150 mM butyrate or crotonate through drinking water which would have been absorbed in the small intestine to a degree once the SCFA

travelled through the digestive system and reached the colon. Perhaps the direct route to the colon would have shown us a different overall result between the two SCFAs.

In relation to this study, perhaps the duration of pre-treatment could have also affected the protective response of both SCFAs. Increasing the time by 10 days or even 14 days would allow more SCFAs to reach the colon and could dramatically change the response. It took 5 days for any of the groups to show a change in weight after DSS was administered and 8 days before the mice began to reach end point in the butyrate group and the DSS only group. This delay could be the window where the equilibrium between the microbiota and butyrate shifted due to the depletion of the microbiota from DSS treatment causing the mice to begin to die. This did not occur until a few days later for the crotonate group, perhaps crotonate doesn't rely on the microbiota population as much as butyrate does and therefore less effected by the depletion from the DSS disrupting the microbiota perhaps crotonate is able to function well at a low concentration for a longer period of time.

Absorption ability is also key in enabling the SCFAs to be functional. Perhaps crotonate is more absorbent than butyrate, so when the two SCFAs have been given as treatment at the same concentration levels, the greater positive effects of crotonate are apparent.

Following on from these results, the next steps to investigate would be the effect of the concentration of the SCFAs – an experiment with 2 crotonate groups, one 150mM and another 50mM, 2 butyrate groups with the same two concentrations and another 2 groups with NaCl and finally group with water only, all groups would be then given DSS after the pre-treatment and another set of groups with no DSS. The two SCFAs could provide different and/or better protection at these differing concentrations. An

interesting approach would be to analyse extracts from the colons from these mice by western blot to identify HPTM changes with the presence of crotonate and butyrate, ChIPseq on the intestinal epithelial colon cells to investigate the genomic effects linked to HPTMs, luminal content to measure the SCFAs levels and microbiota diversity, ELISA to measure the cytokines and finally take histology sections to visualise the extent of the damage the DSS would have on each test group. As these experiments would have been performed in Brazil, it has been difficult to return to carry these out and unfortunately, I have been unable to complete this line of investigation.

However, the exciting take-home message from the research so far, is that crotonate has outdone butyrate in its ability to protect the gut. This SCFA has lots of potential for future research in pre-treatment or treatment for UC and IBD, but more still needs to be understood in identifying the full extent of its mechanisms and protective abilities.

Chapter 4. Histone crotonylation links gene expression to the colon microbiota

The symbiotic relationship between the microbiota and host is essential for many aspects of host physiology, but the molecular mechanisms of this interaction are poorly understood. The microbiota is thought to have evolved in partnership with its host and plays an important role in digestion, immune development and defence against pathogens. Research into the microbiota and its effects has gained speed in recent years with rapidly improving techniques; commensal bacteria have been linked to many diseases such as obesity, diabetes and even depression [161]. As host epigenetic states can regulate diverse sets of cellular functions, it is essential to determine how the epigenome responds to the microbiota in the context of the intestine.

The microbiota produces many metabolites which have the capacity to modulate the host epigenetic state. Short chain fatty acids (SCFAs), including acetate, propionate and butyrate, are produced by the intestinal microbiota as a result of fermentation of dietary fibre and butyrate constitutes a major source of energy for intestinal epithelial cells[162]. SCFA molecules are chemically similar to important histone post-translational modifications (HPTMs), including histone acetylation, propionylation, crotonylation, butyrylation and β-hydroxybutyrylation. The abundance of these HPTMs, known collectively as histone acylations, have been shown to reflect the availability of SCFAs and their coenzyme A modified precursors in the cell [163–170].

SCFAs can act as modulators of chemotaxis, gene expression, proliferation, differentiation and apoptosis[152,171,172]. SCFAs have been found to have anti-cancer,

immunosuppressive and immune activating properties as well as effects on gut mobility and energy expenditure[66,171,173,174]The many effects of SCFAs may be because they can act through multiple pathways including provision of energy substrates, activation of G protein coupled receptors (GPCRs), stabilization of hypoxia inducible factor (HIF), inhibition of histone deacetylases (HDACs) and provision of precursor metabolites for histone acetyltransferases (HATs) [175,176]. SCFAs can affect cell types in different ways, for example butyrate is metabolised and acts as a HAT activator in normal colonocytes whilst in cancerous epithelial cells it is not metabolised due to the Warburg effect and act as a HDAC inhibitor, resulting in inhibition of proliferation [167]. HATs and HDACs have also been shown to catalyse the addition and removal of histone crotonylation and butyrylation [170,177–183].

Histone post translational modifications (HPTMs) are essential regulators of gene expression as they modulate the packing of chromatin and specifically recruit cellular machinery. Whilst histone acetylation has been extensively studied, longer chain acylations such as histone crotonylation have been identified only recently and are less well understood [184]. Histone crotonylation has been shown to activate transcription to a greater extent than histone acetylation, due to its bulkier side chain[185]. Histone crotonylation is poorly recognised by the acetyl-reading bromodomain containing proteins and is bound instead by the YEATS domain proteins and the PHD domain containing proteins MOZ and DPF2, which prefer crotonyl groups over acetyl groups, and in particular H3K9cr over H3K9ac[115,182,186–190]. This provides the potential for histone acetylation and crotonylation to link to distinct functions despite both being transferred to histones by p300/CBP. Histone crotonylation has already been linked to a number of diverse biological functions including spermatogenesis, cancer progression, cell cycle and DNA damage; and linked to

several disease states including Alzheimer's disease, HIV and kidney injury [170,183,184,191–195].

SCFAs are one of a number of microbially derived metabolites that have been shown to directly affect histone modifications in the intestinal epithelium. In a study by Boffa *et al.* (1992), supplementation of fibre to the diet of rats increased total SCFA levels and histone acetylation. Interestingly, butyrate levels were highest in the 5 and 10% wheat bran groups but cell proliferation was lowest compared to the 0 and 20% groups[196].

In another study by Krautkramer *et al.* (2016), microbial colonization of germ-free mice was shown to promote histone acetylation and SCFA supplementation of germ-free mice could partially match histone acetylation and methylation patterns of conventionalised mice[197]. Studies have shown phenotypic differences in germ-free mice compared to conventional mice, and some have seen a noticeable reduction in immune responses with germ-free mice[198]. Other studies have identified changes in gene expression within the gut epithelium of germ-free mice when recolonised with a microbiota harvested from conventional mice. Many of these differentially expressed genes have been associated with growth, apoptosis, cytoskeleton and immune response[51]. It has also been previously shown by Fellows *et al.* (2018) that histone H3 lysine 18 crotonylation (H3K18cr) is abundant in intestinal epithelial cells and that histone acetylation and crotonylation are globally reduced on antibiotic induced microbial depletion[170].

In this study, we show that global levels of many histone crotonylation marks, are responsive to depletion of the microbiota. We demonstrate that changes to crotonylation, on antibiotic treatment are linked to changes to gene expression,

suggesting that crotonylation rather than acetylation is involved in the regulation of these genes in response to the microbiota.

4.1 Gut microbiome and Germ-free gut

4.1.2 Differences in colon histone crotonylation and acetylation levels between germ-free and conventionalised mice

There have been a few studies focusing on histone acetylation and histone methylation with germ-free mice, but none explored histone crotonylation. In this study, germ-free mice colons were compared against germ-free-recolonised mice colons to identify any differences of global levels of various histone modifications. Histone-3 lysine-18 acetylation and crotonylation (H3K18ac, H3K18cr), histone-3 lysine-9 acetylation and crotonylation, and tri-methylation (H3K9ac, H3K9cr and H3K9me3) were all compared with pan-crotonylation (Kcr) and histone-3 (H3) used as a control by using western blot analysis of whole tissue extracts (Fig.4.1).

All of the analysed histone modifications except for H3K9me3 were detected at lower levels in germ-free colons than in the recolonised colon extracts (Fig 4.1b). H3K9me3 showed no change between the two sample groups indicating that its global levels are is not affected by the presence or absence of the microbiota. The relative intensity of the western blots were quantified and a two-tailed t.test was used. The most statistically significant difference measured was with H3K9cr (p-value of 0.0001) H3K18ac (p-value of 0.019) and H3K9ac (p-value of 0.05).

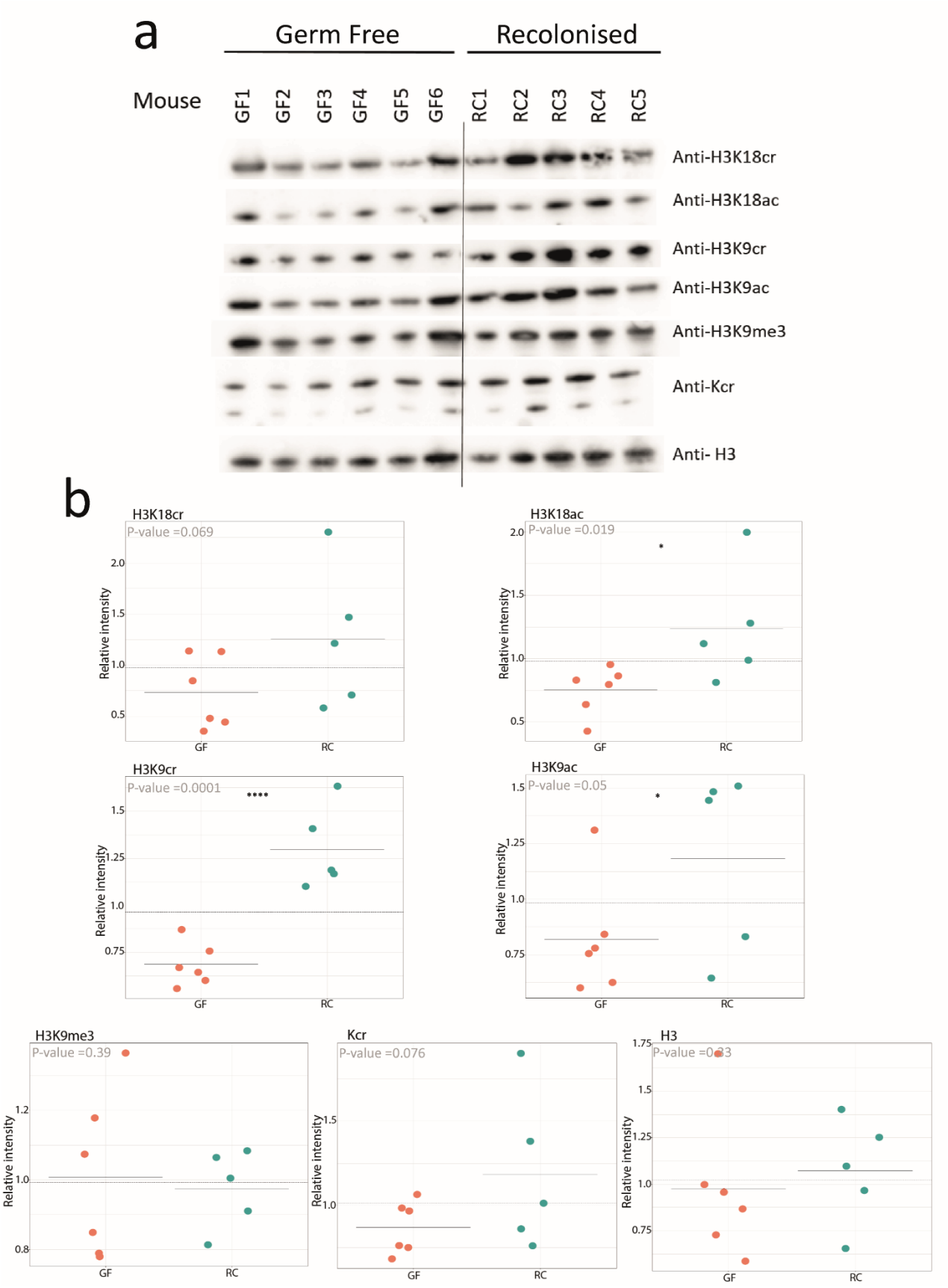

a
Germ Free
Recolonised
Mouse
GF1
GF2
GF3
GF4
GF5
GF6
RC1
RC2
RC3
RC4
RC5
Anti-H3K18cr
Anti-H3K18ac
Anti-H3K9cr
Anti-H3K9ac
Anti-H3K9me3
Anti-Kcr
Anti- H3
b
H3K18cr
P-value =0.069
H3K18ac
P-value =0.019
H3K9cr
P-value =0.0001
H3K9ac
P-value =0.05
H3K9me3
P-value =0.39
Kcr
P-value =0.076
H3
Relative intensity
GF
RC

Figure 4.1. Germ-free and recolonised gut microbiota can affect PTM abundance in the colon.

Whole colon extracts of Germ-free mice (GF), n = 5, and Germ-free mice colonised with microbiota from wild type mice (RC), n = 6, were investigated by Western blot using indicated antibodies. ***(a)*** *The mouse colon extracts were probed with antibodies against H3K18cr, H3K18ac, H3K9cr, H3K9ac, H3K9me3, Kcr and an anti-histone H3 was used as the loading control.* ***(b)*** *Quantifications of the Western blots. The intensity of bands were plotted relative to the anti-histone H3 quantification. The average percentage of each sample of the antibody in question was divided by the average percentage of the H3 from the corresponding sample. H3 samples were normalised to the first sample GF1. Two-way t test was performed and the p-values of each PTM are shown on the graphs. With a p-value threshold of 0.05 only two PTM were significantly affected by the recolonization of the gut microbiota.*

H3K18 and H3K9 can be acetylated or crotonylated and associated with many sites of gene activation. The methylation of H3K9 however is generally associated with gene silencing. The interesting finding here is that although H3K18cr show an increase when the colon has been recolonised with a microbiota, it is H3K9cr, H3K9ac and H3K18ac that show a significant increase between the germ-free colon and the recolonised colon. H3K9cr in particular is extremely significant which could suggest that the microbiota is heavily influential to the histone modification crotonylation, enabling it to bind effectively at the K9 residue or to facilitate its function to activate gene expression at a greater level than acetylation can at either residue, K18 or K9.

To understand the implications to these findings a more robust investigation needs to be carried out to look at the host chromatin – microbiota interactions on a genomic level. This has been addressed in chapter 4.2.

4.2 Effects of antibiotic treatment on the gut microbiome

Mice are treated with broad-spectrum antibiotics to deplete the gut microbiota and this can be rapidly applied to any genotype or mouse condition making such a depleted microbiota mouse model attractive. Germ-free mice however require sterile conditions from birth with continuous maintenance and checks for contamination throughout their lives. This can become very expensive to maintain and can make it impractical for large experiments in the long term.

Comparing the two methods biologically, germ-free mice are impaired in many aspects of development and early immunity, whereas antibiotics treatment in adult mice allows the studying of the role of the microbiota in cell function, signalling pathways and disease development[50] in a more controlled and consistent manner. However, antibiotic treatment studies also complements the GF approach as it reveals the effect of acute microbiota depletion.

To develop a deeper understanding of the effects of the microbiota on the gut and the HPTMs involved, antibiotic treated mice were used to focus on the changes in Kcr, H3K18ac, H3K9ac and H3K9cr using chromatin immunoprecipitation (ChIPseq), identifying genes that are differentially crotonylated or acetylated at these particular residues and those genes that are differentially regulated by gene expression when the microbiota has been depleted.

RNAseq data was collected by a past post doc, Dr Juri Kazakevych, ChIPseq data of Kcr was collected by a past PhD student Rachel Fellows, and ChIPseq data of H3K18ac was collected by myself and another PhD student Mariane Font Fernandes.

All analysis was performed by me, and Kcr ChIPseq analysis was a re-analysis of the data.

4.2.2 A large number of genes are dysregulated on antibiotic treatment of mice

Mice were given a cocktail of antibiotics over the course of three days. The antibiotic cocktail comprised of 5 mg/ml gentamicin, metronidazole, ampicillin, neomycin and 2.5 mg/ml vancomycin. Differential gene expression analysis was performed to determine how antibiotic treatment altered the regulation of genes in colon epithelial cells. Over 800 genes identified to be significantly different between the control group and antibiotic treated group (microbiome depleted group) (Fig. 4.2.1). DEseq2 analysis identified 558 significantly up regulated (red) and 305 significantly down regulated (green) genes. Some of the genes that have changed in expression after the microbiome has been depleted are indicated in Fig. 4.2.1b). Gene ontology analysis was performed on the differentially expressed genes revealing a larger number of significant terms linked to the genes that were up regulated on antibiotic treatment (Fig. 4.2.1c, left panel). Terms for several processes including metabolism, biosynthetic processes, electron transport and ATP synthesis were both highly significant and 5 to 20 fold enriched compared to the *Mus musculus* colon epithelium reference list. Additionally, purine ribonucleotide, carbohydrate, lipid, fatty acid and ATP metabolism were significantly enriched in the up regulated gene list. GO terms that occurred in the up regulated list gave “T cell mediated immunity”, “stimulus response” and “positive regulation of leukocyte mediated immunity”, this could indicate the changes in gut epithelium-environment interactions. Fascinatingly, “circadian rhythms” is also an enriched term, potentially illustrating the microbiota’s involvement in maintaining biological rhythms of gut functions[199]. For the genes that were down

regulated on antibiotics treatment, the terms identified were associated with “negative regulation of cell differentiation” (Fig. 4.2.1c, right panel). A number of genes were linked to this term including, *Sox9* (Transcription factor SOX-9), a transcription factor linked to stem cell maintenance and commonly overexpressed in solid cancers including colorectal cancer, *EphB2* (Ephrin type-B receptor 2), regulates bidirectional migration of intestinal precursor cells in the crypt-villus axis and an overexpression of EphB2 can inhibit colon cancer[194] and *MMP14* (Matrix metalloproteinase-14) gene associated with cell proliferation and high expression of MMP14 is associated with poor prognosis of colorectal cancer[200,201].

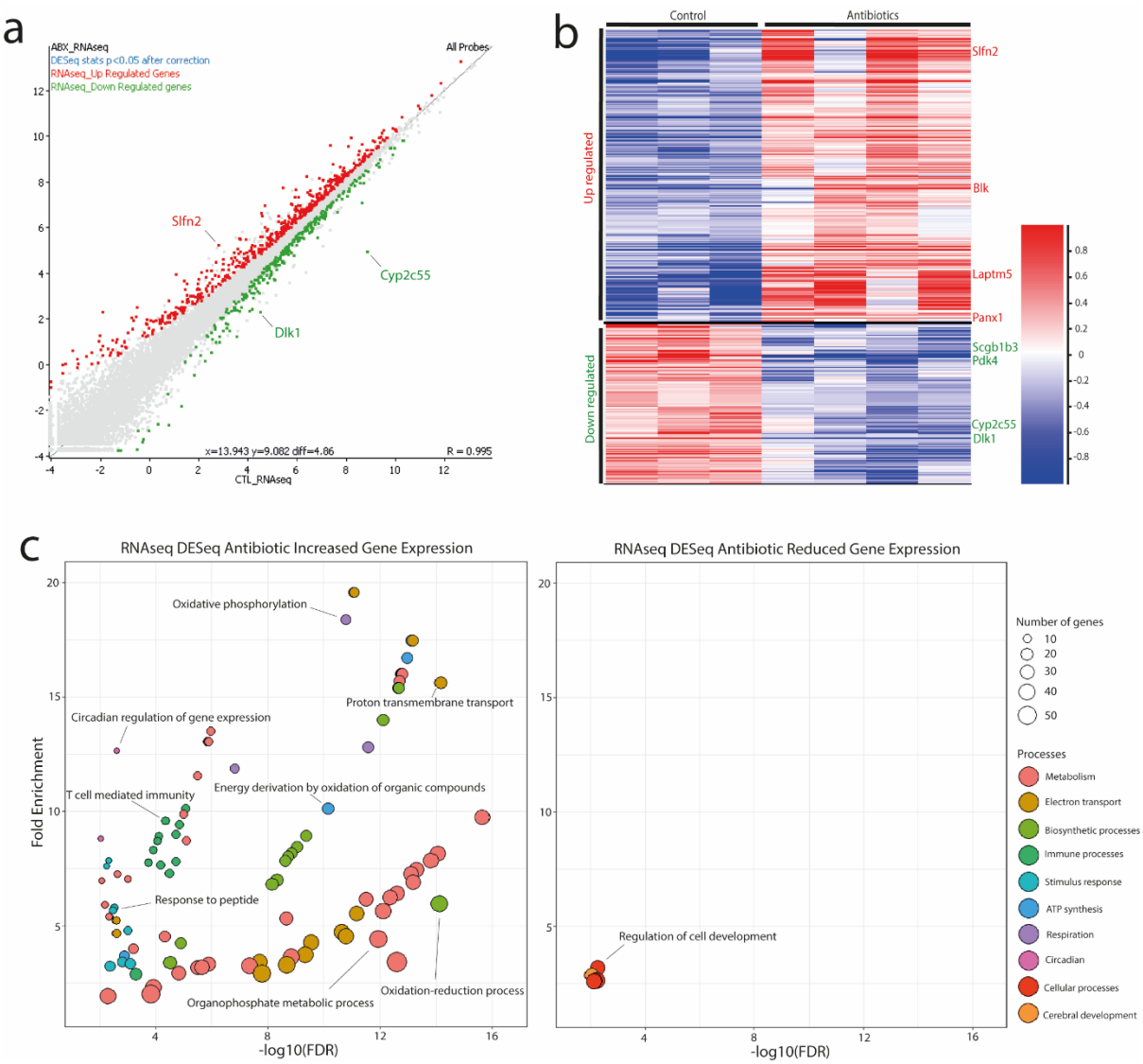


Figure 4.2.1. Antibiotic treatment results in substantial changes in gene expression in mouse colon epithelium

RNA-sequencing (RNA-seq) was performed on colon epithelial cells from antibiotic treated or control mice. DESeq2 analysis was performed. ***(a)*** *Scatter plot comparing gene expression of antibiotics treated colon epithelial cells to control after probes were generated using the SeqMonk RNA-seq pipeline and log2 transformed at a threshold of 1. Up regulated gene expression is highlighted in red and down regulated gene expression is highlighted in green, using a threshold p-value of 0.05 and adjusting for multiple testing.* ***(b)*** *Heatmap of differentially expressed genes. The two clusters, "Up regulated" and "Down regulated" have been split highlighting the change in gene expression from the control data set to the antibiotic treated data set. Some genes have been annotated at the location on the heatmap.* ***(c)*** *Gene ontology analysis of antibiotics treated up regulated (left) or down regulated genes (right) using Panther statistical overrepresentation test against a reference list of all expressed genes in the RNA-seq data set (15379 genes, excluding*

unobserved probes). Fold enrichment is plotted against significance. Terms are coloured based on functional categories, with size corresponding to the number of genes in the gene ontology term in the DESeq2 up or down regulated list.

4.2.3 Changes of crotonylation occur in the colon epithelium with antibiotic treatment on mice

In order to understand how the distribution of crotonylation is affected genome-wide, chromatin immunoprecipitation sequencing (ChIPseq) was performed on colon epithelium extracts from control and antibiotic treated mice using an anti-lysine-crotonylation (Kcr) antibody. This antibody was validated previously in studies investigating histone crotonylation [169,184,202]. This ChIPseq was performed by a PhD student in the Varga-Weisz lab and I followed up with the analysis. I found large numbers of MACS peaks changed significantly in crotonylation between control and antibiotics treated mice (Fig. 4.2.2). Using EdgeR analysis on the control and antibiotic treated sequence data sets, 1324 probes (genomic loci) were identified with significant changes in crotonylation when the gut microbiome was depleted. Of those, 328 genes were increased with crotonylation and were annotated up to +/-2kbp each side of the enriched probes, 614 genes showed a reduction in crotonylation. A list of differentially expressed genes highlighted in Fig. 4.2.1a, was overlapped with the differentially crotonylated genes to identify those that have changed in gene expression and crotonylation (Fig. 4.2.2a). The converse was also performed to show the differentially expressed genes highlighting those that have also changed in crotonylation levels (Fig. 4.2.2b). There are many overlapping genes between the differentially crotonylated and differentially expressed data sets (Fig. 4.2.2a, b). Plotting fold change in gene expression against fold change in histone crotonylation of these overlapping genes, reveals a positive correlation (R^2=0.48, Fig. 4.2.2c). *Schlafen 2* (*Slfn2*), which

is associated with negative regulation of cell growth [203], was the most significantly up crotonylated gene. *Lysosomal-associated protein transmembrane 5* (*Laptm5*) known to positively regulate proinflammatory signalling pathways[204], was also crotonylated by a fold increase of 1.7. *Cyp2c55* has anti-inflammatory and anti-tumorigenic effects[206], with antibiotic treatment this gene has a 0.9 fold reduction in crotonylation and 3.9 fold reduction in expression. However a few genes show opposing changes: on antibiotic treatment, *N-acetylgalactosaminyltransferase 6* (*Galnt6)* is significantly up crotonylated by 1.4 fold increase but expression is down. *Galnt6* has been found to majorly suppress the progression of colorectal cancer[205]. Gene ontology analysis of the genes that showed a loss in crotonylation, returned terms such as “regulation of transcription”, “regulation of biological quality” and “negative regulation of cellular processes” including “negative regulation of transcription” (Fig. 4.2.2e, right panel). This suggests that a loss of crotonylation could be associated with reduced gene expression. In the groups that showed an increase in crotonylation the only term enriched was “response to viruses” (Fig. 4.2.2e, left panel), suggesting an increased immune response with relatively increased crotonylation. The overlaps between up crotonylated and up regulated genes, and between down crotonylated and down regulated genes were highly significant relative to total expressed genes (Fig. 4.2.2d). The percentage of up regulated genes in gene expression where the genes also increased in crotonylation was 8%, and those genes that were down regulated in gene expression and crotonylation came to 18.8%. For those genes that showed a contrast between gene expression and crotonylation for example where there was an increase in gene expression and a reduction in crotonylation, no statistical signficance was seen. This suggests that crotonylation could be linked to changes in gene expression only in a subset of the differentially regulated genes.

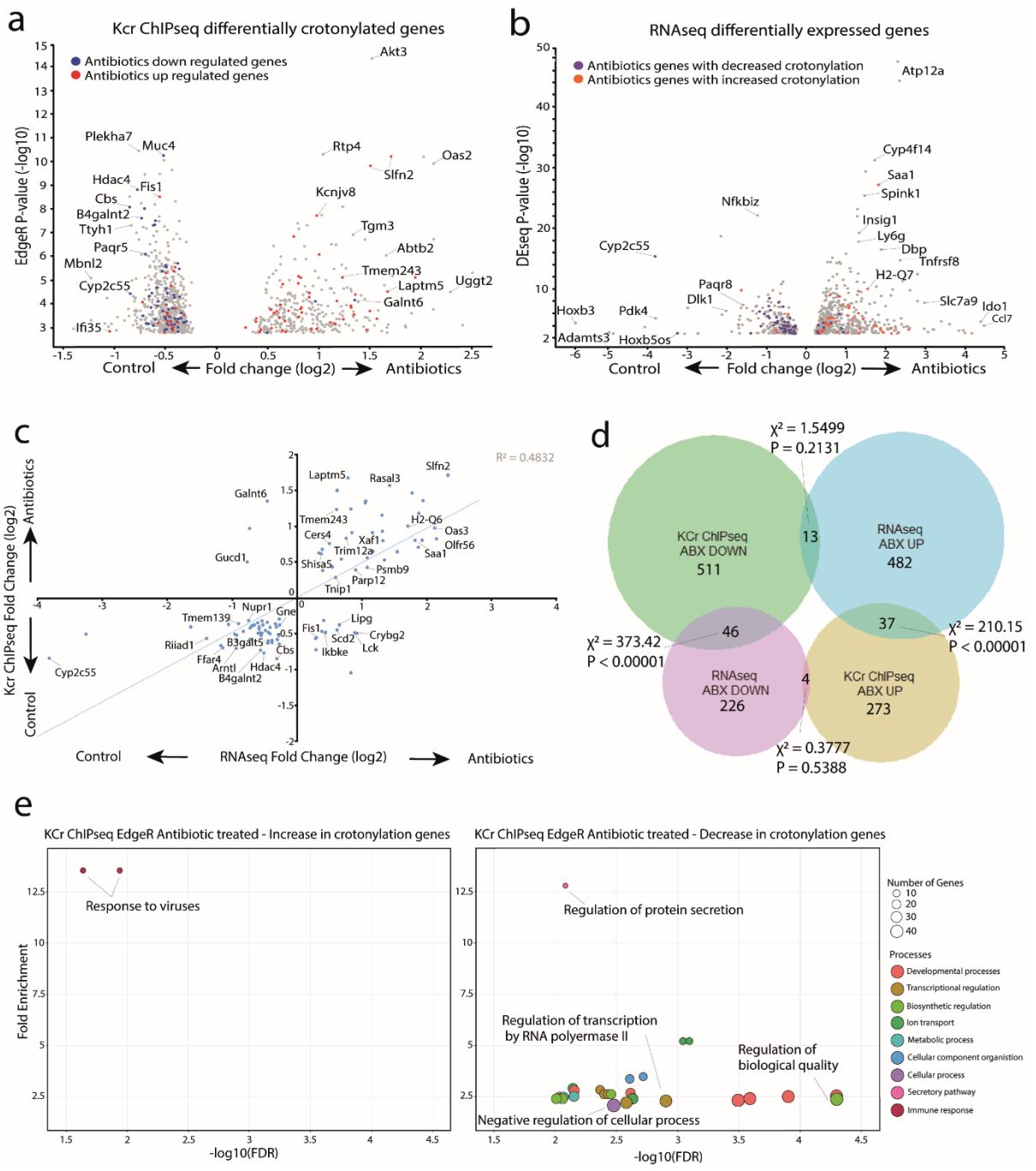


Figure 4.2.2. Changes in histone crotonylation correlate with changes in gene expression in a subset of genes on antibiotics treatment

RNA-seq and Kcr ChIPseq was performed on colon epithelial cells from control mice or mice treated with antibiotics. ***(a)*** *Volcano plot of genes containing MACS peaks that changed in crotonylation, as identified using EdgeR statistical testing of the Kcr ChIPseq data set, using a threshold p-value of 0.05 and adjusting for multiple testing. In cases where there was more than one MACS peak in a gene, only the MACS peak with the greatest fold change was*

included. Unannotated genes e.g. Gm20442 are not labelled. The genes with most significant changes in crotonylation and those with largest fold changes on treatment are labelled. Genes that increase or decrease in gene expression, as identified by RNA-seq, are coloured blue or red respectively. ***(b)*** *Volcano plot of differentially expressed genes identified using DEseq2 analysis of the RNA-seq data set, using a threshold p-value of 0.05 and adjusting for multiple testing. Genes that increase or decrease in histone crotonylation, as identified by Kcr ChIPseq are coloured purple or orange respectively.* ***(c)*** *XY plot showing the fold change of genes that change in both crotonylation and expression, with a trend line showing the relationship between the two.* ***(d)*** *Four-way Venn diagram showing the overlap between up-crotonylated, up-regulated, down-crotonylated and down-regulated genes on antibiotics treatment (ABX). The significance of overlap was tested using a 2x2 two-tailed Chi-square test with Yates correction and the total number of expressed genes of 34245, as determined from the RNA-seq data set excluding probes not annotated with a gene.* ***(e)*** *Gene ontology analysis of more crotonylated (left) or less crotonylated genes on antibiotics treatment (right) using Panther statistical overrepresentation test against a standard Mus musculus reference list. Fold enrichment is plotted against significance. Terms are coloured based on functional categories, with size corresponding to the number of genes in the gene ontology term in the DEseq2 up or down regulated list.*

4.2.4 Comparing global and local changes in histone crotonylation and acetylation.

To understand how the distribution of crotonylation and acetylation compares, I performed ChIPseq on colon epithelium extracts from control and antibiotic treated mice using antibodies against Histone H3K18 acetylation (H3K18ac) and compared with the ChIPseq results obtained using anti-Kcr (described above) and anti-H3K18cr (published, Fellows et al.). Two repeat ChIPseq experiments of antibiotic treatment on both histone modifications were performed independently and analysed. Spike-in controls were used to normalise the ChIP, using crotonylated drosophila and found no difference between KCr and H3K18ac. EdgeR analysis on the first experiment group "A", identified 8 genes that were differentially acetylated at the H3K18 modification (Fig. 4.2.3a). Three of these genes increased in enrichment with microbiota depletion (Fig 4.2.3b), Nicastrin (*Ncstn),* Ras Associated Domain Family (*Rassf4*) and Laminin

subunit alpha-3 (*Lama3). Ncstn* has potential to influence intestinal homeostasis through glycosylation towards its substrate the Notch receptor[207]. *Rassf4* gene is a part of a family of proteins that are tumour suppressors and can modulate apoptosis and duration of inflammation in many organs including the colon[208]. Lastly, *Lama3* is a gene that encodes of the α3 subunit of lamanin α5 protein, contributing towards basement membrane development protecting against colorectal cancer and plays a role in gene expression and regulating tissue development[209,210].

The remaining five genes decreased in enrichment (Fig. 4.2.3c), Cell Division Control 42 Protein 1 (*CD42ep1)*, Lectin Galactosidase Soluble 2 (*Lgals2),* MYB proto-oncogene like2 (*Mybl2),* Diacylglycerol Lipase Alpha (*Dagla)* and Beta-1,3-galactosyltransferase-5 (*B3galt5). CD42ep1* is a Rho GTPase that maintains cellular homeostasis through regulating proliferation, differentiation migration and polarity[211], and a reduction in enrichment of this gene can lead to intestinal crypt hyperplasia. *Lgals2* encodes the glycan-binding protein Galectin 2 (Gal2), which is predominantly expressed in the gastrointestinal tract and is downregulated in colon tumours[212]*. Mybl2* is a transcription factor and regulator of cell cycle progression, cell survival and cell differentiation. Without this transcription factor regulation, it can contribute towards cancer initiation and progression[213]. *B3galt5* gene was reduced in both H3K18 acetylation and pan-crotonylation (Fig. 4.2.3d), and codes for type II membrane-bound glycoproteins and also encodes the most probable candidate for synthesis of the type 1 Lewis antigens which are frequently found to be elevated in gastrointestinal and pancreatic cancers.

The repeating the ChIPseq experiment on H3K18ac, group “B” sequence analysis with MACS peaks and EdgeR analysis showed no significant changes between antibiotic

treated samples and controls. The combined the two datasets, group A and B resulted in no significant changes between antibiotic treated and control.

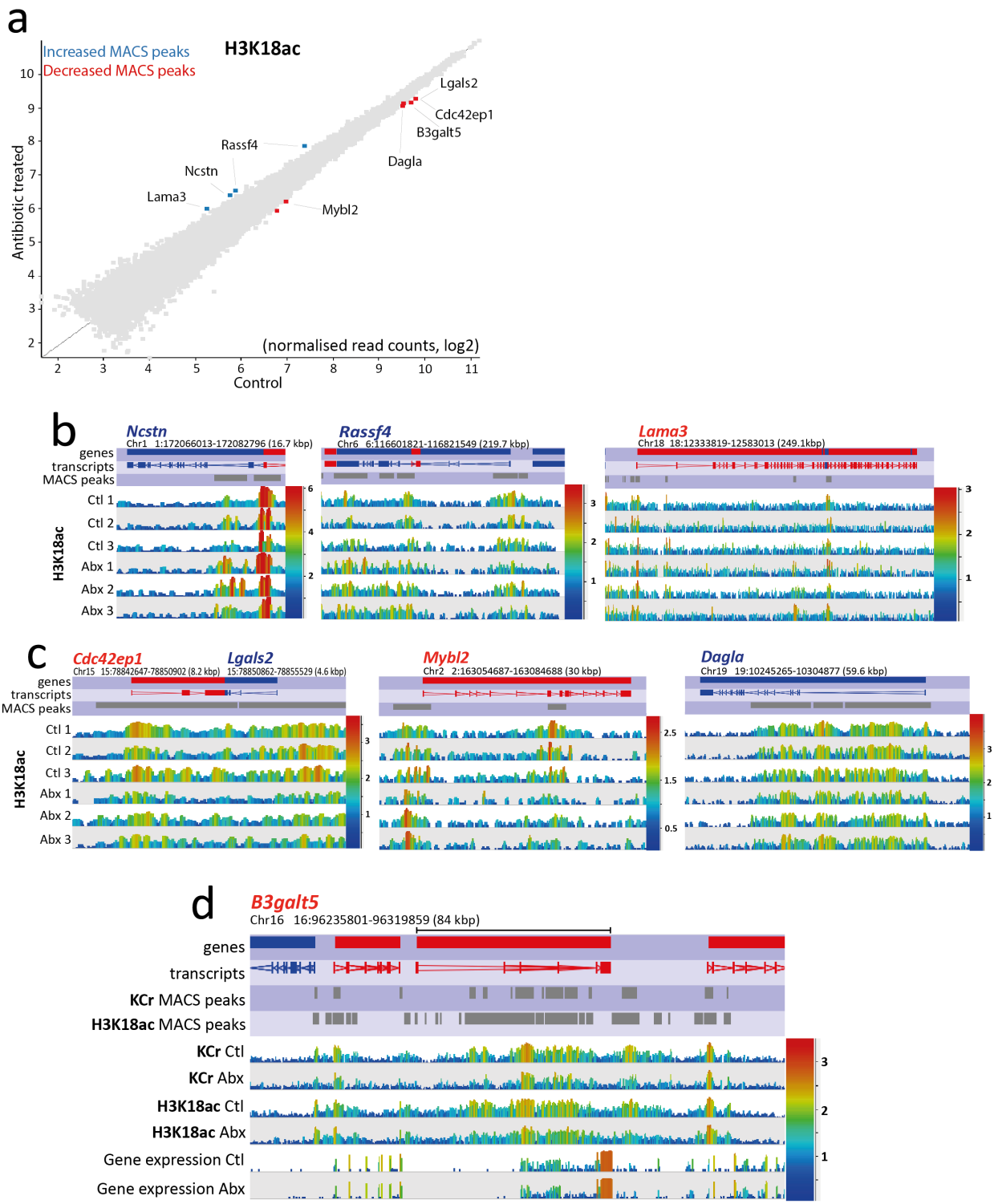


Figure 4.2.3. H3K18ac changes linked to gut microbiome depletion after antibiotic treatment

H3K18ac and KCr ChIPseq and RNA-seq was performed on colon extracts from control mice and mice treated with antibiotics to deplete their microbiota. ***(a)*** *Scatterplot of genes containing MACS peaks that changed in H3K18ac in one set of ChIPseq data (A), identified using EdgeR analysis ($p < 0.05$). In cases where MACS peaks did not overlap with a gene were not labelled. MACS peaks over 3 genes increased H3K18ac when the microbiome had been depleted, identified in blue. MACS peaks over 5 genes identified a decrease in H3K18ac, shown in red.* ***(b, c & d)*** *Selected regions showing changes with H3K18ac. The top annotation tracks indicate genes, transcripts and H3K18ac MACS peaks* ***(d)*** *and KCr MACS peaks. Read counts are shown in linear scale and colour coded. Chromosomal positions and window sizes in kbp are indicated.* ***(b)*** *Peaks in the antibiotic treated samples reveal regions with significant alterations in H3K18ac when the microbiome has been depleted.* ***(c)*** *Peaks in the control samples reveal regions with significant reductions in H3K18ac when the microbiome has been depleted.* ***(d)*** *Peaks located at the gene B3galt5 reveal with both H3K18ac and KCr there is a reduction in read count when the microbiome has been depleted. Gene expression correlates with this showing a small reduction when the microbiome has been depleted also.*

4.2.5 Genome-wide distribution of H3K18ac, H3K18cr and Kcr are highly overlapping in the colon epithelium

Histone crotonylation and acetylation varied in response to antibiotic treatment. Crotonylation changes were linked to the changes in gene expression while H3K18ac did not show consistent changes reflecting gene expression in my ChIPseq analysis. I compared the genomic distribution of Kcr, H3K18ac and crotonylated histone-3-lysine-18 residue (H3K18cr) to identify similarities or differences. The genome-wide profiles of Kcr, H3K18cr and H3K18ac demonstrated a striking similarity, despite some overall differences in enrichment which could potentially be technical (Fig. 4.2.4a). To identify the link of gene expression with each modification, the levels of expression were split by percentage and grouped into 20% “bins”. Analysis shows that both globally and locally, crotonylation has a greater level of gene expression at the 20-40% and 80-100% percentile “bins”, whereas with acetylation the gene expression is grouped mainly at the lower percentile “bin”, between 0-40% and 80-100% (Fig.

4.2.4b). Where there was a substantial overlap between the histone acylations, nearly all genes with peaks in the H3K18cr data set also contained peaks of Kcr or H3K18ac (Fig. 4.2.4c).

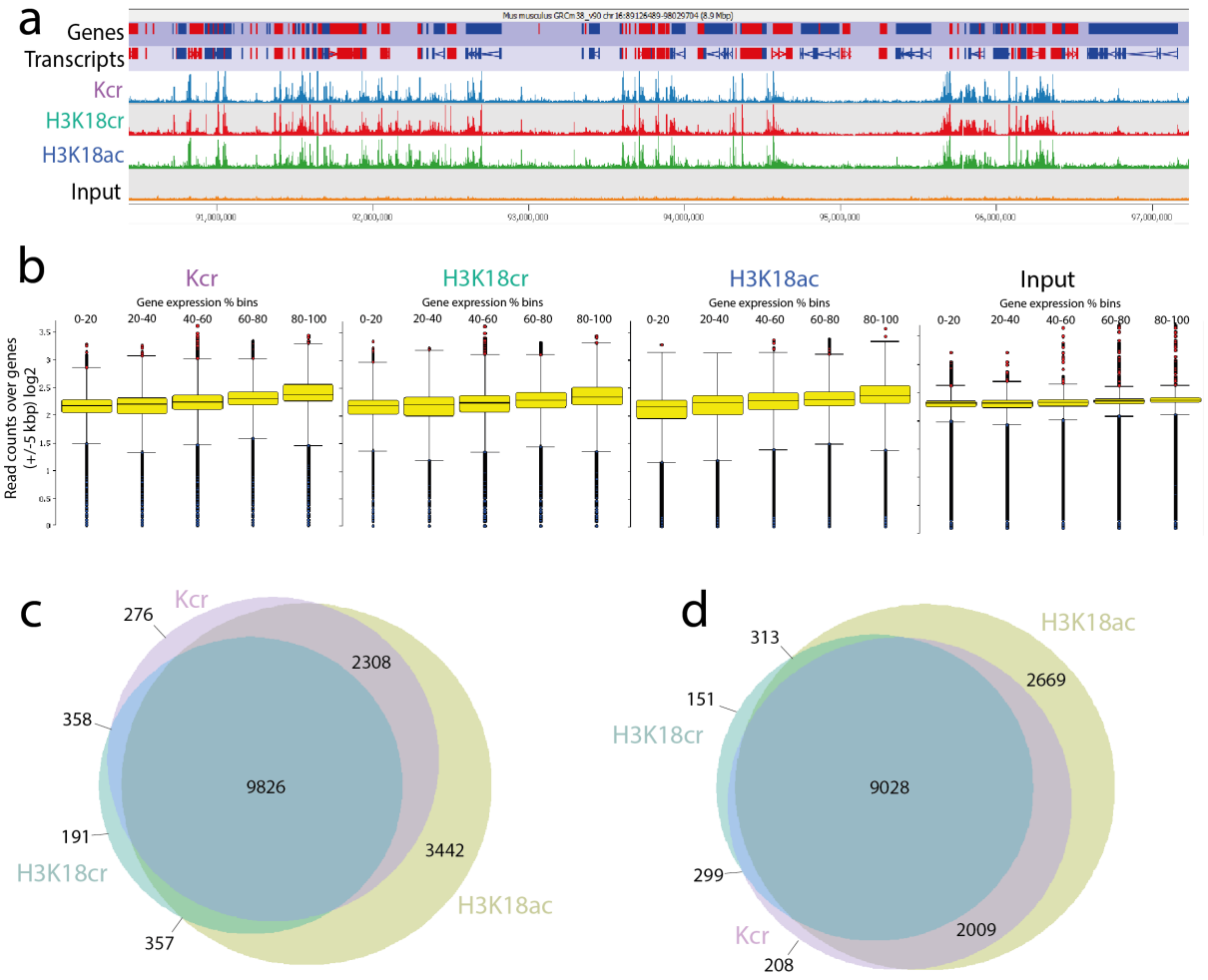


Figure 4.2.4. ChIPseq analysis of Kcr, H3K18cr and H3K18ac shows very similar genomic profiles in mouse colon epithelial cells

(a) *Browser view showing a representative profile of ChIP peaks. Running window quantitated by read count and match distributed probes are shown. Kcr is a combination of three biological replicates, H3K18cr is a combination of two biological replicates and H3K18ac is a combination of six biological replicates.* ***(b)*** *Number of reads overlapping genes relative to gene expression. Quantified by total read counts of largest data store, normalised with match distribution and log transformed. Genes were put into percentile bins according to their normalized RNA-seq read counts from lowly expressed (0-20 %) to highly expressed (80-100 %) and box plots generated.* ***(c)*** *Overlap between Kcr, H3k18cr and H3K18ac data sets.*

MACS peaks were generated against each data set and exported as annotated probe lists and deduplicated. ***(d)*** *Overlap of gene expression between Kcr, H3K18cr and H3K18ac.* Genes with no expression were removed and Venn diagram repeated as in (c) with the remaining genes.

Changes in crotonylation is greatly associated with gene expression (Fig. 4.2.2d), to focus on the link between crotonylation and gene expression the transcriptional start sites were looked at. Only the marks that were over TSS were compared with gene expression and I found that there is an overlap between increased crotonylation at TSS and increase in gene expression that is highly significant (Fig. 4.2.5) and likewise when there is a reduction in crotonylation at TSS the gene expression levels drop very significantly.

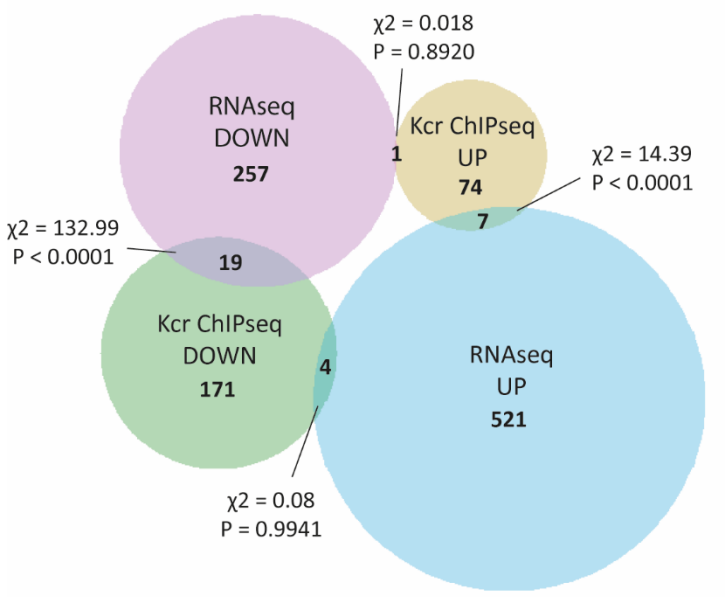


Figure 4.2.5. Transcription start sites of Kcr overlapped with gene expression

Four-way Venn diagram showing the overlap between up-crotonylated TSS, up-regulated gene expression, down-crotonylated TSS and down-regulated genes expression on antibiotics treatment. The significance of overlap was tested using a 2x2 two-tailed Chi-square test with Yates correction and the total number of expressed genes of 25896, as determined from the RNA-seq data set excluding unobserved probes.

4.2.6 Transcription factor binding site motif analysis of crotonylation and acetylation peaks

ChIPseq data can provide a rich source of information about the genome, including transcription regulation. The ChIPseq 'peaks', sites relatively enriched over a specific feature, e.g. histone modification, can be mapped by genomic regions and their corresponding transcription factors (TF)[143] binding sites. TFs are sequence-specific DNA binding proteins that modulate the transcription of sets of genes, these can then regulate the gene expression within the cell. Predicting the TFs are usually achieved by analysing sequence homology with previously characterised DNA-binding domains of TFs[214].

The MEME-suite web service can perform a number of analyses on genomic regions identified by ChIPseq; (i) *Ab initio* motif discovery which identifies novel sequence patterns (motifs) in peak regions identified by ChIPseq that may be due to TF binding sites. (ii) Motif enrichment analysis which looks for enrichment of known TF DNA-binding motifs in the data. (iii) Motif visualization displays the relative locations and binding strengths of TF binding sites in the input regions. (iv) Motif binding strength analysis computes an estimate of the total DNA-binding affinity of each input region for the TF corresponding to each discovered motif. Finally (v) Motif identification compares the *ab initio* motifs to known TF DNA-binding motifs[143]. These analyses give a prediction only of the TF binding at the enriched regions in the ChIP data. This could provide an insight into the potential link between these predicted TFs and the ChIPseq data.

The ChIPseq peaks from H3K18ac, H3K18cr and Kcr were inputted into the MEME-suite program. Analysis of the histone modification H3K18ac identified the most significant TF binding motif (Fig. 4.2.6a) as binding to the Ras-responsive element

binding protein 1 (Rreb1). This TF is involved in a number of biological processes including cell proliferation, transcriptional regulation and DNA damage repair[215]. Recent studies have found that it can function as both a transcriptional repressor and activator for transcriptional regulation of target genes[215] in cancer and metabolic diseases. Using GOMo, the gene ontology package within MEME-suite.org, the predicted biological processes from the same motif gave the terms, 'defence response', 'activation of immune response' and 'cell communication'.

Another predicted TF motif identified with H3K18ac is the nuclear receptor LRH-1 (Nr5a2), predominantly expressed in the ovaries and the liver but also to a lesser extent expressed in intestinal epithelium. This TF maintains intestinal epithelial health, protects against inflammatory damage[216,217] and plays a critical role in T cell maturation[217]. It is a gene expression activating TF driving the transcription of its target genes linked to colorectal cancer and IBD[218]. The gene ontology predicted biological processes linked to this motif produced the terms, 'translation', 'positive regulation of the cAMP' and 'mRNA processing'. Cyclic adenosine 3', 5' –monophosphate (cAMP) plays fundamental roles in cellular responses to many hormones and neurotransmitters[219]. One of the main effectors of cAMP are PKA which phosphorylates many metabolic enzymes[220].

Generally, the predicted motifs returned for the H3K18ac enriched ChIP, highlight the roles in intestinal maintenance and protection that this histone modification plays within a normal gut microbiota. When the microbiota is depleted by antibiotic treatment the enriched H3K18ac ChIP shows a different set of TF motifs. The most significant predicted TF motif is the *forkhead box J3 (Foxj3)* which is associated with tumour suppression of colorectal cancer[221]. Spi1 and SpiB transcription factors are also significantly enriched when the gut microbiota has been depleted. Most commonly

found in the bone marrow, brain and the respiratory system, this transcription factor activates the myeloid and B-lymphoid cell development[222]. Spi1 and its paralog SpiB regulates and activates the differentiation of intestinal microfold (M) cells initiating a mucosal immune response by uptake and transcytosis of luminal antigens[223]. Overall biological processes predicted by the gene ontology on H3K9ac TF gave terms that are associated more with an immune response (Fig. 4.2.6a) than with cell maintenance.

The predicted transcription factor binding motifs associated with H3K18cr provided almost the same motifs as H3K18ac within a normal microbiota. However, H3K18cr focuses slightly more on the side of immunity, tissue homeostasis and differentiation than H3K18ac (Fig. 4.2.6b). The most significant predicted TF motifs returned for H3K81cr are the *IFN regulatory factor 3 (IRF3)*, a transcription regulator of cellular responses and essential for innate immunity[224]. Another transcription factor identified from this motif is the Ets2, which can function as both a tumour repressor and promoter for different cancer types. Within the colon, Ets2 contributes to the repression of intestinal tumour formation[225] and maintains the intestinal epithelial cells homeostasis.

Moving the focus from site-specific (H3K18) histone modification to the global effect of crotonylation within the gut, the most significant predicted binding of TF for pan-crotonylation (Kcr) in both normal microbiota and depleted microbiota was FOXJ3, a transcription factor associated with tumour suppression in colorectal cancer[221] and the far upstream element binding protein 1 (FUBP1) which plays a role in apoptosis and cell proliferation[226] (Fig. 4.2.6C). Within the normal microbiota for Kcr the TFs ZN148 and SP1 are shown to be the next significantly identified binding motifs. Both are members of the Kruppel-like factor family. The ZN148 TF can activate intestinal apoptosis and can be used as a target to inhibit colon cancer[227]. The SP1 TF has

multiple functions associated with tumorigenesis, it can interact with oncogenes and tumour suppressors altering their expression and playing critical roles in the growth of gastrointestinal cancers[228] and is also recognised as an essential drug target. The gene ontology predicted biological processes linked to Kcr within a normal gut microbiota have been 'transcription and translation', 'GPCR signally pathway' and 'mRNA processing'. The general functions of these TF binding motifs play a role as both activator and suppresser of genes related to colorectal cancer.

Kcr peaks that were specifically enriched with microbiota depletion predicted some TF that differed from those predicted in the normal microbiota. Some differences were identified that were associated with negative regulation of transcription, a term that did not show within the normal microbiota – identified by GOMo. The TF Sp5, identified in the depleted microbiota of Kcr (Fig. 4.2.6c), is a direct down-stream target of the Wnt signalling pathway[229] and its expression is upregulated in many colon cancer cell lines[230]. EHF a member of the ETS transcription factor family and is highly expressed in a normal colon, and within the M cells, contributing to immune surveillance[231]. It has been suggested in a study by Taniue et al. (2011), that EHF is required for the survival colorectal cancer. This is due to EHF directly activating *RUVBL1,* an ATPase associated with chromatin-remodelling complexes. *RUVBL1* blocks p53-mediated apoptosis by repressing the expression of *p53* and its target genes[232]. Therefore EHF-mediated RUVBL1 expression allows colon tumour cells to avoid p53-mediated apoptosis.

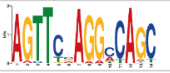
The peaks that showed a loss and those that showed an increase in crotonylation on antibiotics treatment were analysed for predicted conserved TF binding motifs. *Signal transducer and activator of transcription 1 (STAT1)* was identified in the increased crotonylation group. This TF is activated in association with inflammation, IBD and

colitis[233–235]. SP1 was also present, the TF associated with both oncogenes and tumour suppression. Within the decreased crotonylation group, the predicted binding motifs for *Early growth response 2 and 3 (EGR2 and EGR3)* were identified. These TFs have been shown to be associated with regulating autoimmunity and suppressing excessive immune responses[236,237] in T cells. Rreb1 was also identified within this group of predicted TF binding motifs and this TF can also function as both a transcriptional repressor and activator for regulation of target genes[215] in cancer and metabolic diseases.

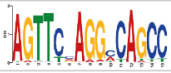
Despite Kcr being enriched or reduced within the depleted microbiota group, the predicted TFs that have been identified seem to be linked to colorectal cancer development and immune response. This suggests that if these TFs are present where they have been predicted to be, the influence of the microbiota linked to Kcr could directly affect the maintenance and homeostasis on a global level within the gut.

a H3K18ac

Normal microbiota

TF	Motif	E-value	Biological Processes
RREB1 BRAC		1.0e-036	Defense response Activation of immune response Cell communication
SALL1 ASCL1 SP4		8.3e-024	Transcription RNA splicing Postive regulation of RNA poly II
NR5A2 ERR2 STF1		9.0e-020	Translation Positive reg of cAMP mRNA processing
PITX2		1.7e-013	Translation Lipid transport Reg of gene silencing by RNA
SP1 ZBT17 EGR1		2.8e-013	Transcription Cell fate determination

Depleted microbiota

TF	Motif	E-value	Biological Processes
FOXJ3 SRY IRF3		4.3e -063	Transcription initiation from RNA polymerase II promoter
SPI1 SPIB ETV4		1.3e -023	B cell receptor signaling pathway Regulation of gene silencing by RNA
ERR2 NR5A2 STF1		9.2e -022	RNA splicing Immune response Lipid phosphorylation
PITX2 MSGN1 PDX1		4.5e -015	Lipid transport Translation RNA processing
FOXA3 ZN322		3.9e -014	Negative regulation of T cell mediated immunity

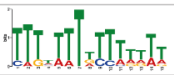
b H3K18cr

Normal microbiota

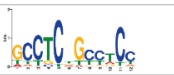
TF	Motif	E-value	Biological Processes
IRF3 ELF5 ETS2		5.2e-075	Postive regulation of transcription from RNA poly II Immune response
SP3 SP2 EGR1		4.5e-045	Transcription Negative regulation of transcription from RNA poly II
RREB1 EGR3 EGR2		1.3e-023	GPCR signalling pathway Signal transduction
NR5A2 STF1 ERR1		8.0e-009	Immune response RNA splicing mRNA processing
PITX2		2.2e-006	Reg of small RNA involved in gene silencing mRNA processing

c Kcr

Normal microbiota

TF	Motif	E-value	Biological Processes
FOXJ3 FUBP1		7.6e-031	GPCR signalling pathway mRNA processing Transcription initiatioin
ZN281 ZN148 SP1		5.7e-022	Transcription
NR5A2 STF1 ERR2		1.7e-008	Translation mRNA processing RNA splicing
PITX2		1.2e-007	Translation GCPR signalling pathway Reg of immune response
ZBT7A		1.2e-007	Transcription represssor

Depleted microbiota

TF	Motif	E-value	Biological Processes
FOXJ3 FUP1		5.9e -044	GPCR signalling pathway mRNA processing Transcription initiatioin
SP5 ZN281		7.1e -021	Negative reg of transcription from RNA polymerase II promotor
ETS1 ERG EHF		5.1e -008	mRNA processing RNA splicing Immune response
ZBT7A		8.5e -007	Transcription repressor
SMAD3 SALL1 ELF5		1.0e -004	Postive reg of transcription from RNA polymerase II promotor

Kcr Increased crotonylation

TF	Motif	E-value	Biological Processes
SP3 SP2 SP1		1.9e-031	Negative reg of transcription from RNA polymerase II promotor
STAT1 STAT2 IRF7		8.0e-015	GPCR signalling pathway Signal transduction Response to virus

Reduced crotonylation

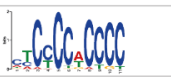
TF	Motif	E-value	Biological Processes
RREB1 EGR3		2.7e -018	GPCR signalling pathway Defense response to bacterium Positive reg of immune system
ZBT17 EGR2 MAZ		3.2e -008	Negative reg of transcription from RNA polymerase II promotor

Figure 4.2.6. Transcription factors identified with changes to histone acetylation and crotonylation antibiotic treatment.

Motif analysis was performed on the MACS peaks of colon epithelial cells treated with antibiotics using meme-suite.org. E-value is the enrichment p-value times the number of candidate motifs tested, calculated using Fisher's Exact Test for enrichment of the motif in the positive sequences. Biological processes where identified using 'GOMo', a gene ontology program within Meme Suite. The three most significant processes were noted in order of p-value, and specificity on the GO term hierarchy. ***(a)*** *Motif analysis on H3K18 acetylation in normal microbiota compared against a depleted microbiota after antibiotic treatment. Motif logo was identified with the corresponding transcription factors associated with it.* ***(b)*** *Motif analysis on H3K18 crotonylation in colon epithelial cells with a normal gut microbiota.* ***(c)*** *Motif analysis performed on pan crotonylated chromatin in colon epithelial cells in normal microbiota and antibiotics-depleted microbiota. Of those MACS peaks that showed changes when antibiotic treatment was given, the peaks were split according to whether the change was an increase in crotonylation or a decrease. These two groups were analysed through meme-suite to identify the changes occurring with the transcription factors due to microbiota depletion.*

4.3 Discussion

4.3.1 Crotonylation over regulatory elements connects the microbiota to gene expression

Antibiotic treatment is used frequently to deplete the gut microbiome for the purpose of combating bacterial infections and to understand the role it plays in gut homeostasis. The role the microbiota plays on colon epithelial cells in association with gene regulation was investigated and it was demonstrated that antibiotic treatment results in the dysregulation of hundreds of genes. The relationship between the microbiome, gene expression and crotonylation was looked at and found to have a direct correlation. Changes to gene-associated crotonylation on antibiotic treatment are positively correlated with changes in gene expression. However, this was not the case for histone acetylation, where links to changes in gene expression was very low if any at all[169]. The different responses of the two histone acylations suggests crotonylation directly responds to the changes in the colon environment.

Gene ontology analysis revealed many terms associated with biosynthetic regulation and developmental processes. Our findings support the evidence that histone crotonylation plays a general role in transcription and gene expression while simultaneously maintaining a healthy epithelial environment. However, despite the large number of genes that did change in crotonylation with antibiotic treatment, only a subset of those genes showed changes in gene expression. Pan anti-lysine-crotonylation antibody was used instead of a histone specific crotonyl-antibody such as anti-histone-3-lysine-18-acetylation, the antibody used for acetylation, this was because the histone specific antibody at the same residue location performed poorly in previous studies. Due to this, it cannot be excluded that crotonylation was detected on non-histone protein during analysis explaining the greater levels of change. Although, previous western blotting work done by colleagues (not shown in chapter) indicates the most abundant crotonylated proteins detected by the anti-Kcr antibody migrate in a way that is consistent with them being histones H3 and H4. Mass spec studies have also identified H3K18cr as the most abundant[170] Ongoing experiments and analysis with a different histone modification, H3K9cr and H3K9ac will give vital information on other site-specific crotonylation and acetylation and the corresponding gene expression.

4.3.2 Similarities in crotonylation and acetylation enrichment sites genome-wide

Enrichment sites (peaks) genome-wide of Kcr, H3K18ac and H3K18cr are very similar and greatly overlap. Both crotonylation and acetylation occur at the same location on the genome and it has been suggested that acetylation and crotonylation are in competition with each other in terms of occurrence (a single lysine residue cannot be acetylated and crotonylated at the same time). As they both require p300/CBP as a

transcriptional co-activator, whether the histone lysines are crotonylated or acetylated depends on the concentration levels of crotonyl-CoA and acetyl-CoA, linking metabolism to gene expression[111]. That being said, by looking at the gene expression percentage levels, crotonylation appears to show stronger link to gene expression than acetylation, which corresponds with the literature that histone crotonylation can directly stimulate transcription to a greater degree than histone acetylation[169]. It is possible that levels of cellular concentration of crotonyl-CoA, which can be altered through genetic and environmental disruptions such as microbiota depletion could lead to a reduction in acetyl-CoA and allow histones to be crotonylated more.

4.3.3 Microbiome depletion influences transcription factor expression

The number of different transcription factors that have been predicted to be enriched within the normal microbiota compared to the depleted microbiota suggests an argument that the epigenetic effect on gene expression and histone modification responses could be linked to microbiota population. The acetylated H3K18 residue within the normal gut microbiota showed predicted links to enriched TFs that would be involved in intestinal maintenance and protection, whereas when the microbiota was depleted at the same residue the predicted TFs enriched, were linked to immune responses. If these TF are proven to be at the binding sites they have been allocated this could suggest the feedback response by H3K18ac is to prepare to protect the gut from opportunistic organisms from colonising and causing inflammation and IBD. The crotonylated H3K18 residue within the normal gut microbiota, has been predicted to be linked with TFs that show similar homeostatic functions as the acetylated H3K18, but seems to be more strongly associated with immune responses and tumour suppression. Perhaps the crotonylated residue acts more in a surveillance and

protection role than the acetylated residue. Crotonylated histones on the global scale with a normal gut microbiota has shown TFs predicted to be enriched over sites that play a role as both activator and suppressor of genes related to colorectal cancer. With the depleted microbiota Kcr has shown predicted TFs that are linked to increases in potential colorectal cancer development and a reduction in immune responses. The next step in this investigation could be to validate the predicted TF binding motifs that the analysis has given with the ChIPseq data, through deep tools analysis, and to identify the location of these TF along the sequences giving an idea of the potential influence they would have on the gene expression with each histone residue.

This vast and varied response to the depletion of the microbiota suggests the complexity of the role that crotonylation plays within the gut. It seems that there is no one answer that can be proposed to describe and understand the full function of crotonylation within the gut. Instead, the best way to understand this modification is to look at histone residues individually and to focus on significant genes of interest associated with crotonylation.

Chapter 5. Effects of colon epithelium-specific *Smarcad1* gene knockout on histone modifications

Epigenetic modifications of chromatin and the manner in which chromatin is organized in the nucleus are important elements that can influence responses induced[1] by the gut microbiota. One of the previously mentioned epigenetic mechanisms involved in chromatin dynamics is the ATP-dependent chromatin remodelling complex. There are different types of ATP-dependant chromatin remodellers consisting of enzymes that determine nucleosome density and location, these can contribute to gene repression[129]. Others work with site-specific transcription factors and histone-modification enzymes to rearrange histones, allowing transcription factors to bind to the DNA[238] and some remodellers can replace major histones with histone variants[132]. The ATP-dependent chromatin remodelling enzymes are important in almost every chromosomal process, and without them a variety of diseases including cancer has the potential to develop[239].

As mentioned in the introduction, ATP-dependant chromatin remodellers have been classified into sub-families of chromatin remodelling enzymes; imitation switch (ISWI), chromodomain helicase DNA-binding (CHD), switch/sucrose non-fermentable (SWI/SNF) and Inositol auxotrophy 80 complex (INO80)[130]. Smarcad1 (SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily-A containing DEAD/H box-1) is an ATP-dependent chromatin remodeller in the INO80 family and has appeared to be involved in histone eviction[134]. It plays a key role in gene regulation, heterochromatin maintenance[135] and DNA repair[133]. One of the

suggested functions of the remodeller is facilitating deacetylation of histones H3 and H4, and the promotion of methylation. Mutations or deletions of Smarcad1 can cause a number of diseases, such as a variety of carcinomas and autosomal-dominant inherited diseases including Adermatoglyphia[139] and Basan Syndrome[140] (ectodermal dysplasia and absence of fingerprints).

Smarcad1 mediates gene silencing through heterochromatin markers and by recruiting methyltransferases G9a/GLP to modify the histones such as H3K9me3[137,240,241]. It can modulate the interaction of the histone deacetylase, HDAC1 and the transcriptional repressor KAP1 with heterochromatin[135]. At the same time, Smarcad1 has also been reported to have the dual function in regulating chromatin through the promotion of euchromatin[242].

Modifications of the lysine residues of the N-terminal tails of histone H3 and H4 control the accessibility of chromatin, regulating gene transcription. Modifications at H3K9 have a number of roles, when the histone is acetylated (H3K9ac) genes are activated and when methylated (H3K9me) they are silenced. Since H3K9 is a residue that can be acetylated or methylated, many possibilities have been explored in terms of the progression of modification changes at H3K9, including the proteins that can coordinate modifications at H3K9 in various model systems such as embryonic development or gene expression models [127]. Chromosome condensation is a complex mechanism that ensures genome integrity during mitosis. When H3K9 is acetylated, histone deacetylation is required prior to H3K9 methylation[127]. It has been suggested that the deacetylation of H3K9 is a prerequisite of, and H3K9me3 is also required for chromosome condensation.

Previous studies within the group explored the role of Smarcad1 by generating a new conditional deletion model of *Smarcad1* using C57BL/6 background male mice, kept in specific opportunistic pathogen free (SOPF) conditions at the Babraham Institute transgenic facility and fed CRM (P) VP diet (Special Diet Services) ad libitum[141]. Published work by Kazakyvech *et al.* (2020)[141] using this conditional deletion model includes data I generated, H3K9me3 ChIPseq data on colon epithelial cells, and is mentioned within this chapter. The focus of the paper was on the role of this remodelling factor within the intestinal epithelium. Analysis showed that a deletion of Smarcad1 can affect histone modifications linked to chromatin assembly in this tissue which, in turn, can modify gene expression and intestinal epithelium-microbiome interactions. This modification affects the colitis response in a DSS-induced mouse model. Without Smarcad1 the mice did not develop a colitis response to DSS suggesting that the changes occurring within the gut may provide a protective affect by inhibiting the innate immune response and preparing the intestinal epithelium to deal with a microbial challenge during the DSS treatment.

To follow on with this line of research, identifying the genes that are affected and the level of expression modified through the influence of Smarcad1 would provide a deeper insight into impact of chromatin remodelling. To achieve this, I needed to explore the effect Smarcad1 has on each of the histone modifications, H3K9me3, H3K9ac and H3K9cr within the colon. The conditional Smarcad1 deletion mouse model was used and ChIPseq analysis was performed on both the wild type (WT) and knockout (KO) colon samples. Dramatic effects were seen with the deletion of Smarcad1 in both the hetrochromatin-linked H3K9me3, and the euchromatin promoters H3K9ac and H3K9cr. H3K9me3 showed a striking increase over the entire gene body in Smarcad1 KO samples in several down regulated genes.

H3K9ac however showed a marked decrease in acetylated genes with a significant number of differentially expressed genes associated with this reduction in acetylated genes. H3K9cr also showed a major decrease in Smarcad1 KO and a significant number of down regulated expressed genes linked to them. H3K9cr was also found to be down crotonylated at the YEATS4 gene in Smarcad1 KO, while the acetylation of the same residue (H3K9ac) was not affected. Kcr is recognised by proteins containing the YEATS domain such as YEATS4 – a protein involved in chromatin remodelling. This finding could suggest a specific role for crotonylation over this gene and a potential feedforward loop between histone crotonylation and its reader.

A significant number of genes were both down acetylated and down crotonylated, mainly linked to metabolic processes. While another set of genes were significantly both up acetylated and crotonylated and linked to developmental processes. This would indicate that Smarcad1 deletion could disrupt the normal functioning of the intestinal epithelium.

Overall, my analysis confirms that Smarcad1 deletion can noticeably affect histone modifications in the gut - specifically chromatin activators, modifying gene expression linked to many functional biological processes.

5.1 Optimisation of chromatin immunoprecipitation (ChIP) conditions using pilot ChIP-qPCR

5.1.1 Identifying appropriate genomic targets for each ChIP

Pilot experiment with 9 female mice (colons of 3 animals pooled per ChIP) was performed to ensure the oligonucleotide primers worked and their genomic targets showed sufficient enrichment by chromatin immunoprecipitation to enable detectable differences between the groups.

All ChIPs provided some chromatin recovery (Fig. 5.1). As expected, there was anti-correlation between H3K9-methylation and both acetylation and crotonylation.

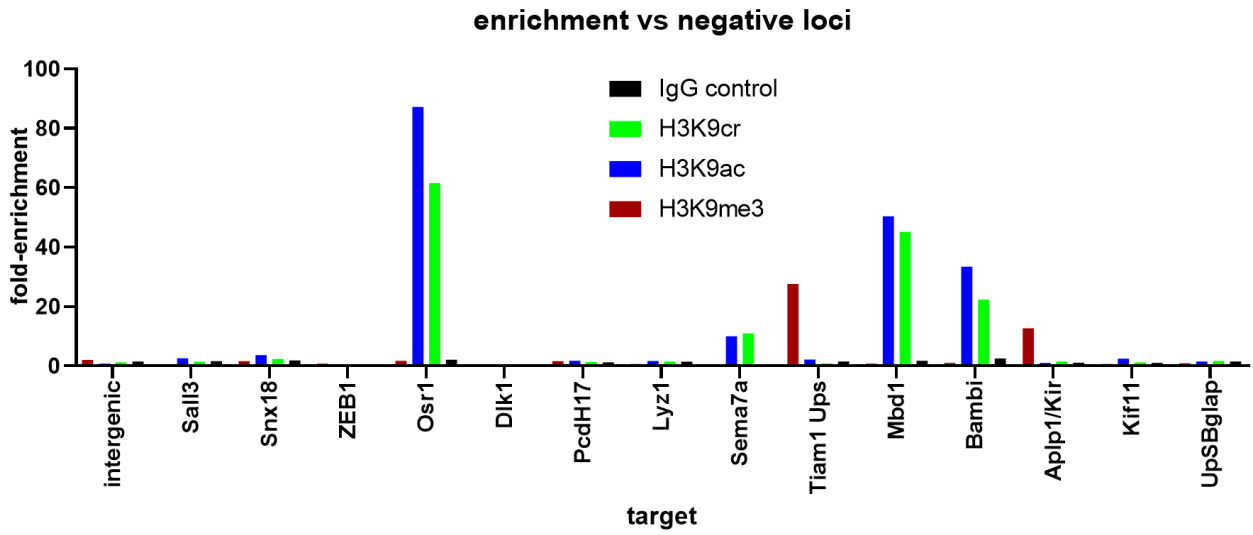


Figure 5.1.1 Enrichment of H3K9me3, H3K9ac, H3K9cr (and IgG control) in mice colon epithelial cells over several candidate loci.

ChIP enrichment can be seen between the positive targets, Ors1, Sema7a, Mbd1 and Bambi for both acetylation and crotonylation and two different positive targets for H3K9-trimethylation, Tiam1 and Aplp1/Kirrel2. Negative targets for acetylation and crotonylation appeared to be Sall3, Klf11 and 'intergenic 93/94' and for H3K9-trimethylation, Bambi, Klf11 and 'intergenic 93/94'. All targets were negative for the IgG control.

In this initial test, four targets were enriched for H3K9 acetylation and crotonylation, and only two targets enriched with H3K9me3 so more primers were designed to

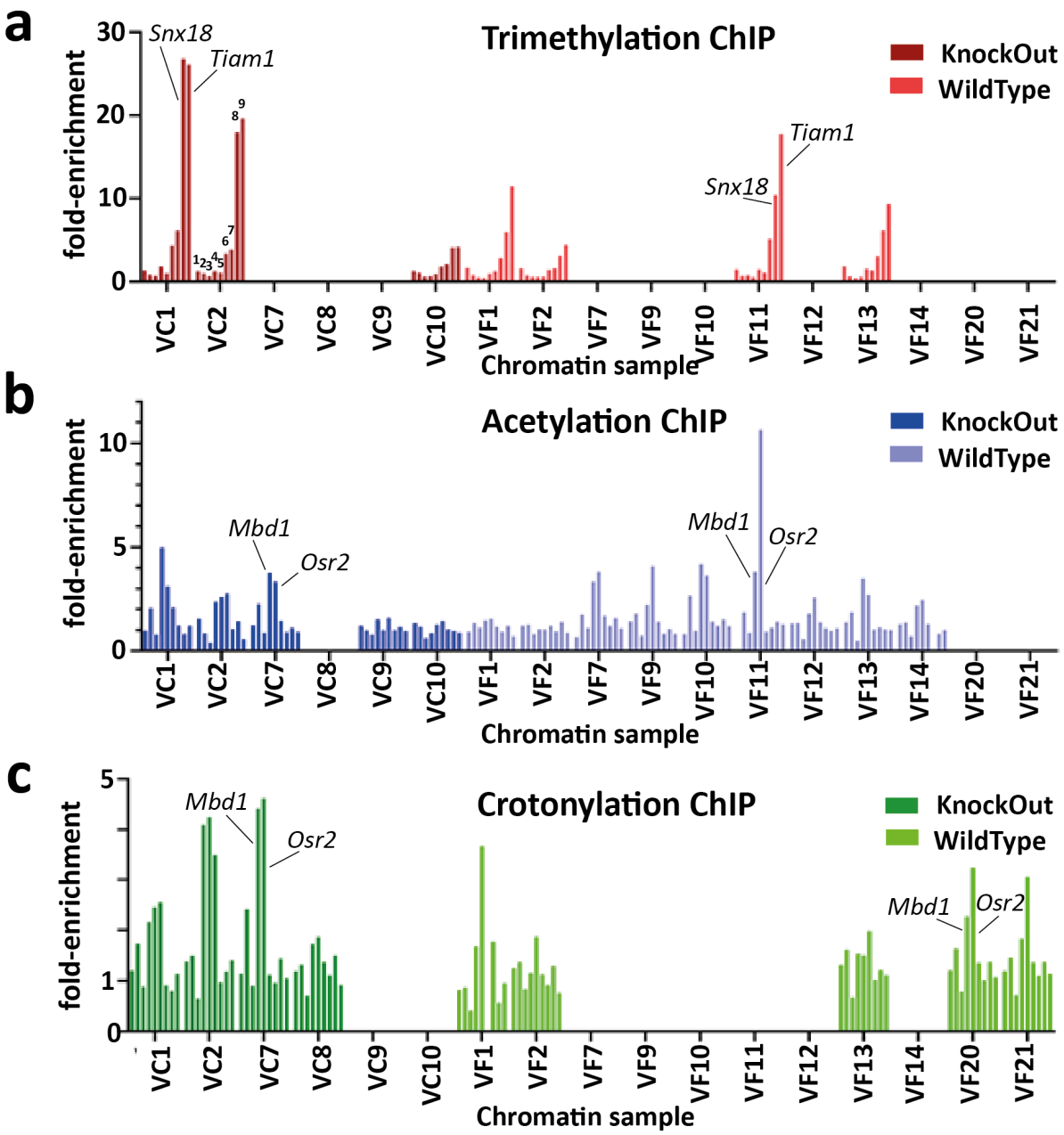
increase the number of positive targets for H3K9me3. Overall, the recovery of acetylation was relatively low and the next step was to optimize recovery. This was done by removing a high salt wash after the antibodies bound to the chromatin. I also attempted to optimize the crotonylation and H3K9-trimethylation ChIP by reducing the number of cells to enable the use of a single mouse per ChIP. The cells were sorted using flow cytometry, but the failed at the sonication step due to the dramatically reduced number of cells. This flow cytometry step was removed from the protocol.

5.1.2 ChIP-qPCR results

For the optimisation steps, the final ChIP enrichment of all the samples have been complied together in Figure 5.1.2. The H3K9-trimethylation ChIP (Fig. 5.1.2a) was enriched using 8ng of chromatin and provided a high of 27 and 26 fold enrichment (in reference to the negative controls) in *Snx18* and *Tiam1* upstream loci respectively in the KO phenotype and in the WT a fold enrichment of 10 and 18 in *Snx18* and *Tiam1* upstream respectively. The acetylation ChIPs (Fig. 5.1.2b) were more successful in terms of number of ChIPs that worked in comparison but with an overall lower enrichment rate. The KO samples gave a maximum of a 5-fold enrichment for *Mbd1* and 3-fold enrichment for *Osr2*. The WT samples all have a 3-4-fold enrichment in 4 of the ChIP-PCRs performed, with one giving an 11-fold enrichment in *Osr2*. The crotonylation ChIP (Fig. 5.1.2c) is within similar the range of the two other ChIPs, H3K9ac and H3K9me3. Four KO ChIPs were successful with a maximum fold enrichment between 4 and 4.6 for *Mbd1* and *Osr2* respectively. The WT is very similar to the KO with a 2-4-fold enrichment in *Mbd1* and *Osr2*.

Three ChIP samples for each group with the highest enrichment levels across the targets were chosen and sent off for sequencing (Fig. 5.2d). The samples coloured the same are from the same chromatin sample for each ChIP. There were a number

of ChIP samples that failed qPCR which could have been due to human error when loading the 384-well PCR plates, a loss of DNA material during extraction or inefficient binding of the antibodies during the immunoprecipitation step. However, all samples were plotted including the failed samples with just the label on the graph for clarity to visually compare the same chromatin sample across all three histone modifications.



d

	Me3	Ac	Cr
Wild Type	VF1 (WT1)	VF7 (WT1)	VF1 (WT1)
	VF11 (WT2)	VF11 (WT2)	VF20 (WT2)
	VF13 (WT3)	VF10 (WT3)	VF21 (WT3)
Knock Out	VC1 (KO1)	VC1 (KO1)	VC1 (KO1)
	VC2 (KO2)	VC2 (KO2)	VC2 (KO2)
	VC10 (KO3)	VC7 (KO3)	VC7 (KO3)

Figure 5.1.2. Total enrichment of methylation, acetylation and crotonylation ChIP from all colon samples.

ChIP enrichment plotted by loci target and grouped by chromatin sample, example of target order are numbered in (a) Trimethylation ChIP VC2. From left to right the targets were; (1)94/94 Intergenic, (2)Bambi, (3)Klf11, (4)Mbd1, (5)Osr2, (6)PcdH17, (7)Sall3, (8)Snx18 and (9)Tiam1 Upstream. The labelled targets on each graph were the most enriched with that particular modification. Samples with no results were failed ChIPs, but were still labelled in the graph to compare between each histone modification. ***a)*** *H3K9me3 successful ChIPs are plotted by fold enrichment. Smarcad-1 KO samples are coloured in dark red and WT are bright red. Three KO and four WT samples worked.* ***b)*** *H3K9ac successful ChIPs are plotted by fold enrichment. Smarcad-1 KO samples are coloured in dark blue and WT are pale blue. Five KO and nine WT samples worked.* ***c)*** *H3K9cr successful ChIPs are plotted by fold enrichment. Smarcad-1 KO samples are coloured in dark green and WT are light green. Four KO and five WT samples worked.* ***d)*** *Table showing the chosen ChIP samples sent off for sequencing. Across the groups, where the same colon sample was used was highlighted in the same colour. The sample names from this point onwards have been changed to WT1 for wild type 1 and KO1 for knock out 1, as indicated in* ***d)****. Each sample will be specified when referring to the corresponding histone modification, e.g H3K9me3_WT1.*

5.1.3 Quality control analysis of sequenced samples

The ChIPseq data was analysed using the Seqmonk v47.2 Software. To analyse the data, a command line tool known as MACS (Model-based Analysis of ChIPseq) was used within the Seqmonk software. The MACS can be used to identify transcription factor binding sites and histone modification enriched regions[243]. The uploaded parameters can be found in Materials and Methods chapter 2.3.

5.1.3.1 Histone 3 lysine 9 acetylation quality control analysis

After running the MACS peaks analysis of the grouped data sets, WT and KO versus input, a number of quality control tests were performed to check the condition of the sequenced samples. A data store tree was used to view the degree of relatedness of all the data stores. This uses a Pearson correlation to calculate a distance matrix

between all of the currently visible data stores. The datasets for H3K9ac revealed one KO dataset, H3K9ac_KO3 to be closely related to the WT datasets compared to the other two KO counterparts (Fig. 5.1.3a). This sample was flagged.

A cumulative distribution plot was performed in Seqmonk to compare the distribution of probe values for the same probe list across several data stores. Samples that are well normalised with each other will have very similar profiles on this plot, and any differences seen here indicate a systematic difference between the samples, which might have either a technical or a biological source. The H3K9ac_KO3 dataset appears to have grouped with the WT replicates (Fig. 5.1.3b). Suggesting this sample could possibly be more like a WT sample than a KO.

The t-SNE plot is a dimensionality reduction technique used to graphically show the similarity between large datasets. Within Seqmonk it can be used to cluster data stores on the basis of the current quantitation across a large number of probes. Using this technique to plot the similarities between my datasets reveal that H3K9ac_KO3 appears to be more closely related to the WT replicates than the KO (Fig. 5.1.3c).

Each replicate sample in both groups were individually analysed against the input dataset to check for their quality. MACS peaks analysis was performed on each sample, quantitated by read counts and corrected for total read count then log transformed. After plotting each sample against input on a scatterplot (Fig. 5.1.3d), the distribution of MACS peaks probes highlighted the stark difference between H3K9ac_KO3 and the other two KO samples. The browser screen shot of the samples (Fig. 5.1.3e) also confirmed this by showing the peak patterns of KO3 is appearing to respond more like the WT samples than the KO samples. To confirm whether this sample was a true KO, the Smarcard1 gene was looked at for read counts and for all

three KO samples the absent reads were just as expected, there was however a small area that did appear to have a small number of reads with KO3. This suggests that the KO may not have occurred as cleanly resulting in an intermediate KO mouse.

It was decided that the H3K9ac_KO3 sample was to be removed entirely for the rest of the analysis. Figure 5.1.3f, g and h show the resulting relationship and distributions between the remaining samples.

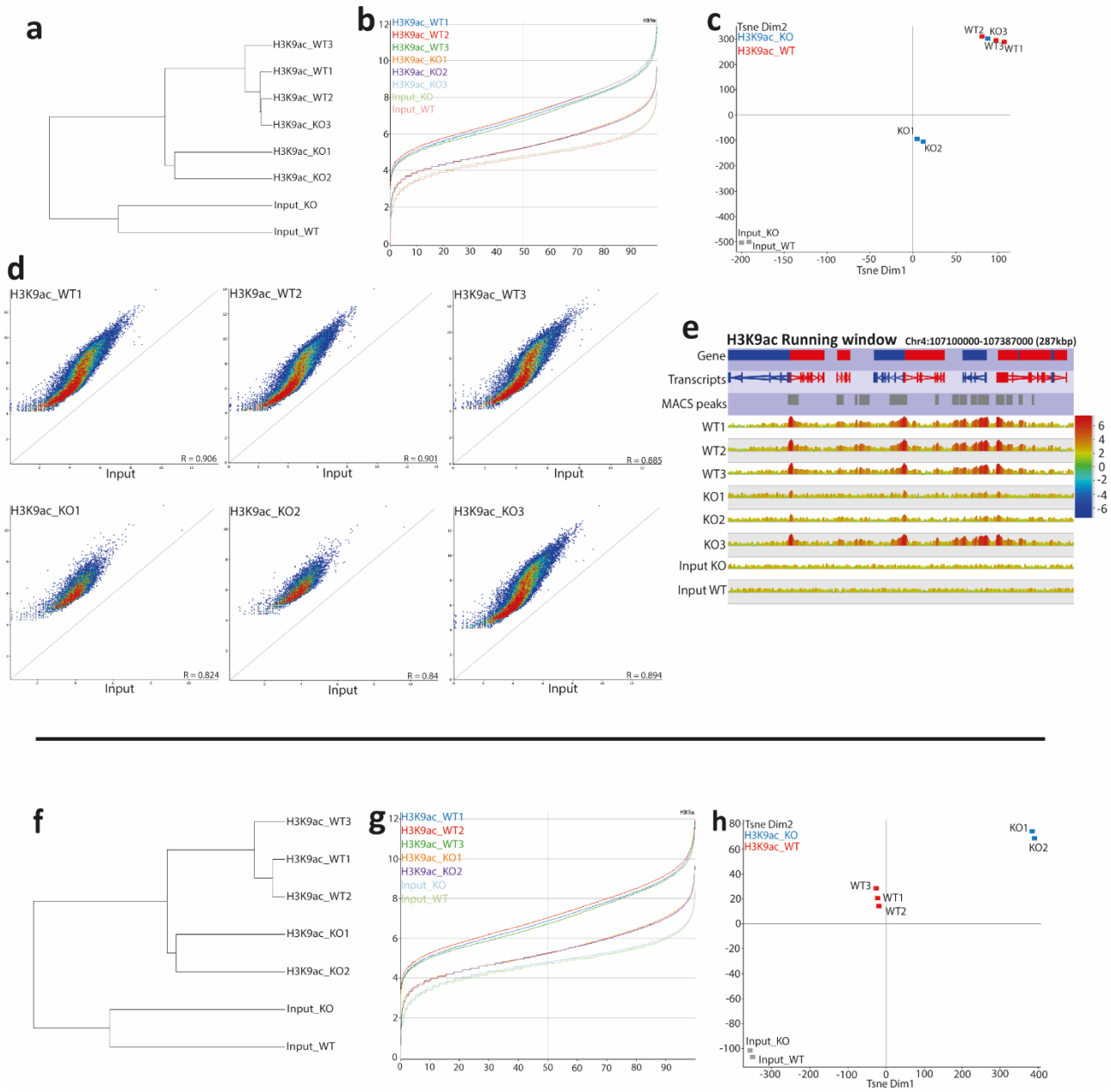


Figure 5.1.3 Quality control analysis of H3K9ac ChIP datasets

a) *Data store tree mapped the likeness of all replicate datasets. H3K9ac WT vs H3K9ac KO vs Input.* ***b)*** *Cumulative distribution plot of all replicate datasets, H3K9ac WT, H3K9ac KO and Input. Each ChIPseq dataset has been coloured individually* ***c)*** *TSNE plot measure the degree of similarities against each dataset****,*** *WT coloured red, KO coloured blue, Input coloured grey.* ***d)*** *MACS peaks analysis of each dataset replicate vs Input, plotted on a scatterplot (log2) Input vs biological replicate. e****)*** *Browser screen shot of running window, H3K9ac WT, H3K9ac KO and Input WT and KO. The top annotation tracks indicate the gene, exons, and H3K9ac MACS peaks. Read counts are shown in linear scale, color coded. Chromosomal position and length in kbp are indicated.* ***f)*** *Data store tree of selected replicate datasets.* ***g)*** *Cumulative distribution plot of selected replicate datasets.* ***h)*** *TSNE plot of selected replicate datasets. WT coloured red, KO coloured blue, Input coloured grey.*

5.1.3.2 Histone 3 lysine 9 crotonylation quality control analysis

The same quality control analysis was performed on the H3K9 crotonylation sequenced ChIP samples. H3K9cr_WT1 and H3K9cr_KO3 both appeared to not correlate with the other two replicates in their corresponding groups. Each analysis in figure 5.1.3a-e, suggests that the samples may have been mixed up or swapped during the experiments as the WT1 has grouped very closely with the two KO samples and vice versa, the KO3 has grouped closely with the two WT. However, looking at read counts at the Smarcad1 location on the sequences, they have been correctly labelled, albeit not as clear cut as one would like. This discrepancy has forced the conclusion that the two samples would be removed to avoid skewing of the analyses. Figure 5.1.4 f-h show the remaining selected samples going forward.

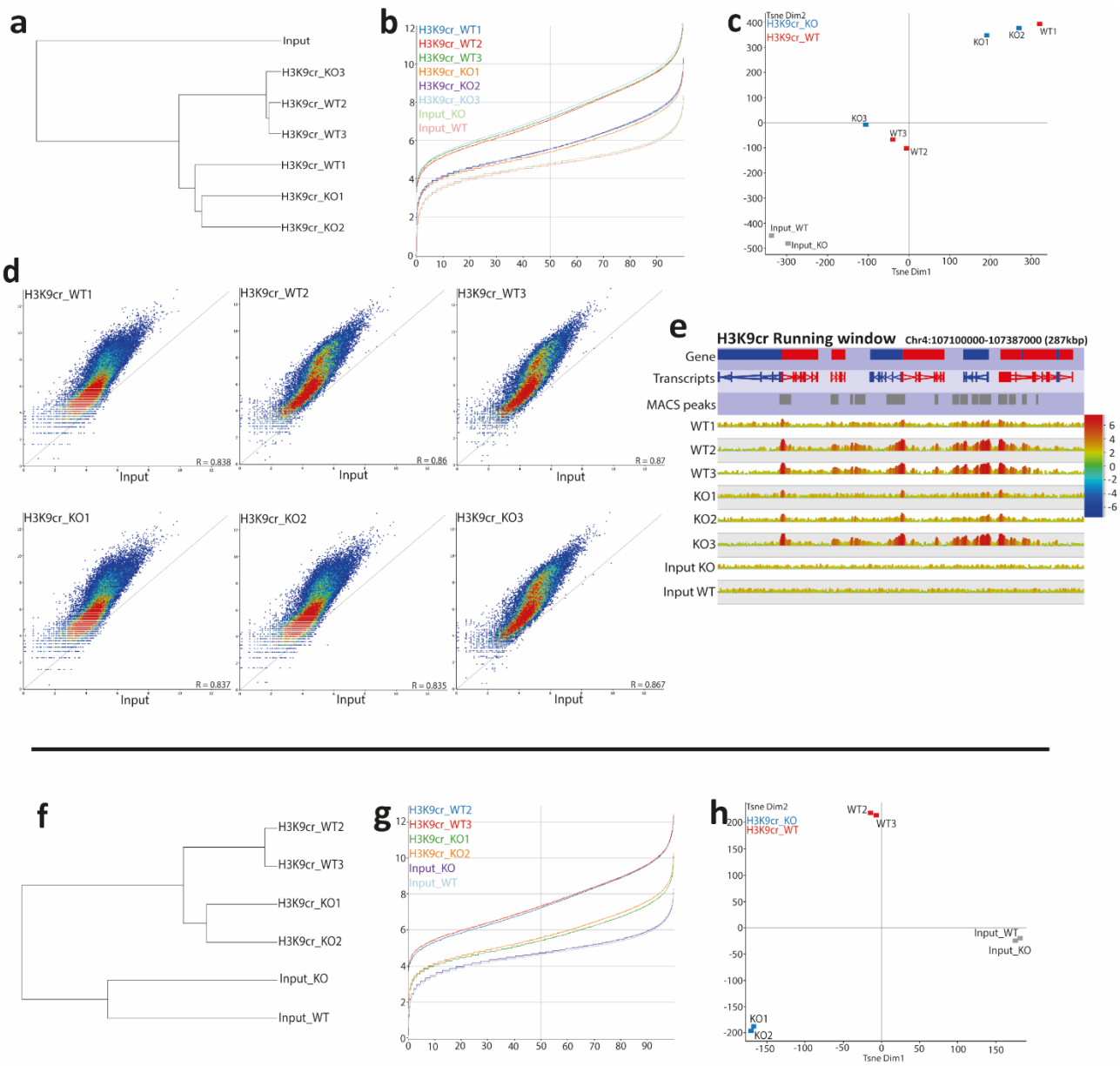


Figure 5.1.4 Quality control analysis of H3K9cr ChIP datasets

a) *Data store tree mapped the likeness of all replicate datasets. H3K9cr WT vs H3K9cr KO vs Input.* ***b)*** *Cumulative distribution plot of all replicate datasets, H3K9cr WT, H3K9cr KO and Input. Each ChIPseq dataset has been coloured individually.* ***c)*** *TSNE plot measure the degree of similarities against each dataset****,*** *WT coloured red, KO coloured blue, Input coloured grey.* ***d)*** *MACS peaks analysis of each dataset replicate vs Input, plotted on a scatterplot (log2) Input vs biological replicate.* ***e)*** *Browser screen shot of running window, H3K9cr WT, H3K9cr KO and Input WT and KO. The top annotation tracks indicate the gene, exons, and H3K9cr MACS peaks. Read counts are shown in linear scale, color coded. Chromosomal position and length in kbp are indicated.* ***f)*** *Data store tree of selected replicate datasets.* ***g)*** *Cumulative*

distribution plot of selected replicate datasets. ***h)*** *TSNE plot of selected replicate datasets. WT coloured red, KO coloured blue, Input coloured grey.*

5.1.3.3 Histone 3 lysine 9 trimethylation quality control analysis

The quality of the H3K9me3 sequenced ChIP samples were analysed. The data store tree grouped the samples with the WT replicate sets together and the KOs together (Fig. 5.1.5a). The cumulative distribution graph (Fig. 5.1.5b) and the tSNE plot (Fig. 5.1.5c) show that samples within the replicate groups also cluster together, however the separation between the two groups is quite low. This could just indicate that the overall effects of Smarcad1 KO on H3K9me3 was not large. A browser shot of a running window-based quantification (Fig. 5.1.5e) illustrates a clear difference between WT and KO over a select region. Overall the QC results are satisfactory to go ahead with the analysis.

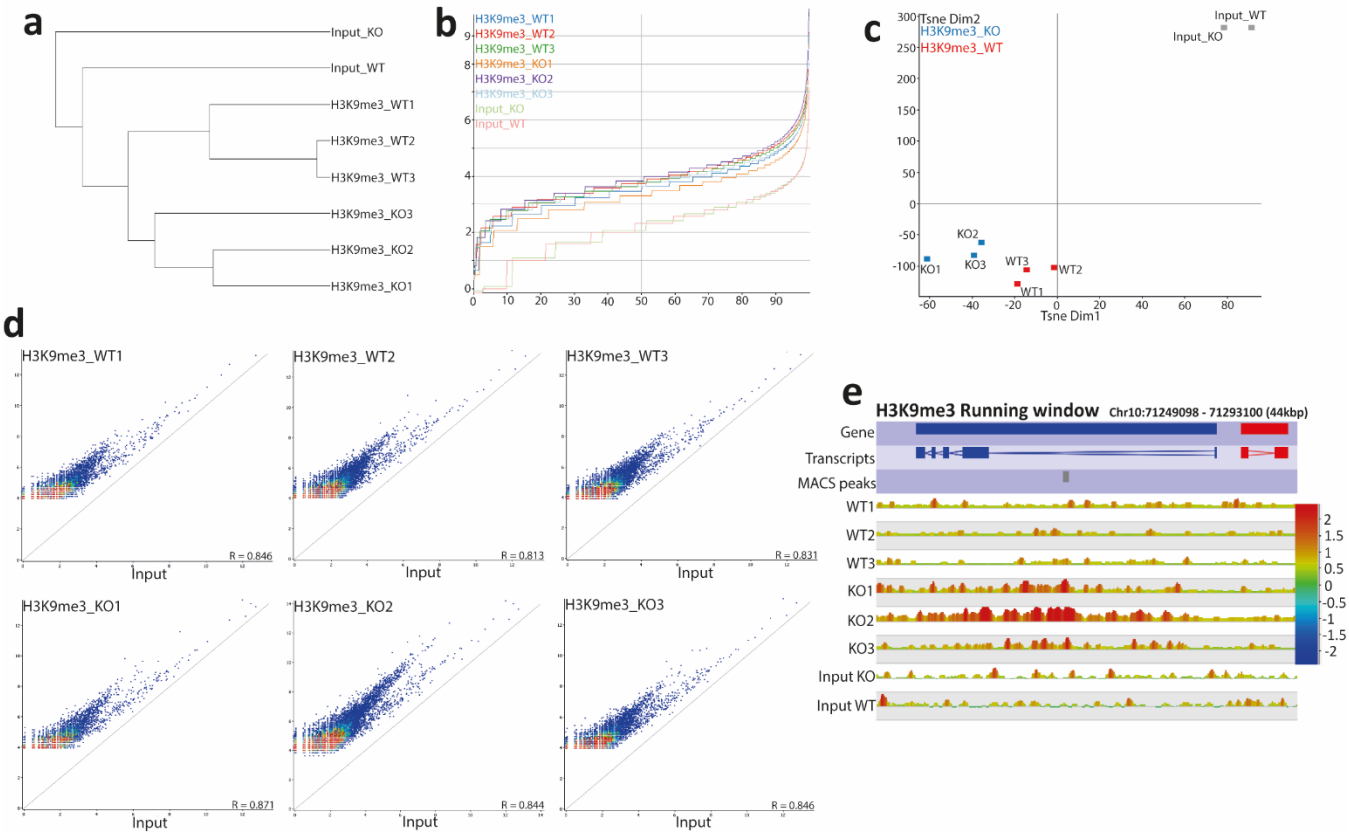


Figure 5.1.5 Quality control analysis of H3K9me3 ChIP datasets

a) *Data store tree mapped the similarity of all replicate datasets. H3K9me3 WT, H3K9me3 KO, Input.* ***b)*** *Cumulative distribution plot of all replicate datasets, H3K9me3 WT, H3K9me3 KO and Input. Each ChIPseq dataset has been coloured individually.* ***c)*** *TSNE plot measure the degree of similarities against each dataset****,*** *WT coloured red, KO coloured blue, Input coloured grey.* ***d)*** *MACS peaks analysis of each dataset replicate vs Input, plotted on a scatterplot (log2) Input vs biological replicate. e****)*** *Browser screen shot of running window, H3K9me3 WT, H3K9me3 KO and Input WT and KO. The top annotation tracks indicate the gene, exons, and H3K9me3 MACS peaks. Read counts are shown in linear scale, color coded. Chromosomal position and length in kbp are indicated.*

5.2 The link between Smarcad1 and H3K9 modifications

5.2.1 Effects of Smarcad1 KO on H3K9me3

The influence of Smarcad1 on modifications at H3K9 was determined by developing a Smarcad1 deletion mouse model. The Smarcad1 knockout (KO) colon samples had ChIPseq analysis performed on them to explore the effects of this chromatin remodeller on the histone modification H3K9me3.

MACS peaks analysis was performed on each of the ChIPseq data sets, the Smarcad1 KO ChIP sequences were analysed against WT to identify any changes. Analysis of the ChIP data shows that H3K9me3 MACS peaks gained trimethylation in n = 573 (right of scatterplot) and lost n = 209 (left of scatterplot) probes upon Smarcad1 deletion (Figure 5.2.1a). From the list of genes that have gained trimethylation due to Smarcad1 KO, 3 genes were shown to have a significant down-regulation in gene expression (Fig 5.2.1a & b); *Cmah, Snx18* and *Zfp503*. *Cmah* (*Cytidine monophosphate –N)* catalyses the sialic acid N-acetylneuraminic acid (Neu5Ac) into N-glycolylneuraminic acid (Neu5Gc)[244,245], these are components of carbohydrate chains of glycoconjugates. These components are involved in cell-cell recognition and cell-pathogen interaction playing an important function in the innate immune system[245,246]. As a consequence of reduced expression in *Cmah,* the depletion of Neu5Gc could affect responses against pathogens and inflammation[244] and possibly boost tumour development[247]. The second gene that has reduced expression is *Snx18,* (Sorting Nexins 18) this gene promotes autophagy[248,249] and is involved in endocytosis and intracellular vesicle trafficking[250]. It has been found to be strongly associated with immune cells that supress tumour growth[251,252]. The reduced gene expression of both of these genes, could indicate that an inhibition of the mechanisms

supporting an immune response and tumour growth regulation could be attributed to Smarcad1-KO. The third gene to have been down regulated with an increase in H3K9me3 when Smarcad1 has been depleted was the *Zfp503* (zinc finger protein 503), a gene linked to tissue development in the mammary gland and possibly plays a crucial role in tissue homeostasis in the intestines, lung and kidneys[253]. The corresponding transcription factor has also been found to be strongly associated with colon cancer cell proliferation and tumour progression[254].

Only one gene was up-regulated with an increase in trimethylation, *Mt1* (methallothionein 1), a gene part of a mechanism within the immune system that reduces inflammation[255]. The *Mt1* gene – methallothionein 1 is a copper and zinc ion binding protein. It is transcriptionally regulated by both heavy metals and glucocorticoids – a class of steroid hormones that are part of a feedback mechanism in the immune system which reduces inflammation[255]. Copper and zinc are important minerals that contribute to energy production[256], immune function and inflammatory response[257,258]. Where H3K9me3 has been reduced one gene was highlighted to be down regulated in gene expression, the *Mapre3* (Microtubule-associated protein RP/EB family member 3). This is a microtubule associated protein localised in the cytoplasm[259]. It is a promotor gene that supports and gives shape to the cell. Within the colon, the Mapre3 protein localises to the cytoplasmic microtubule network and binds to the adenomatous polyposis coil, a tumour suppressor gene[260].

When looking at the total MACS peaks (8046), 88 genes are linked to a down-regulation of H3K9me3 and 98 genes linked to an up regulation of H3K9me3 (Fig. 5.2.1b). Of the total number of expressed genes (25896), 110 are differentially expressed (DESEQ2) with Smarcad1 deletion, 64 genes have reduced expression

and 45 show an increase in gene expression. Of those differentially expressed genes overlapped with the differentially trimethylated MACS peaks of H3K9me3, 1 gene (*Mt1*) as highlighted in Fig. 5.2.1a, has shown to be increased in expression when there is an increase in H3K9me3. This increase in expression when H3K9me3 is increased in KO could suggest that *Mt1* contributes towards the reduced immune response when Smarcad1 is missing. However, the Chi^2 analysis was not significant, so the expression of *Mt1* may not have been a direct result of Smarcad1-KO or an artefact of Smarcad1 deletion.

Smarcad1 KO did result in a significant reduced expression of 3 genes previously mentioned, *Cmah, Snx18* and *Zfp503.* The inhibited expression of these three genes with functions linked to immune response and inflammation, suppression to CRC and development and homeostasis, respectively gives an insight into the overall effects of Smarcad1 deletion and the roles this remodeler in protection and maintenance of the gut.

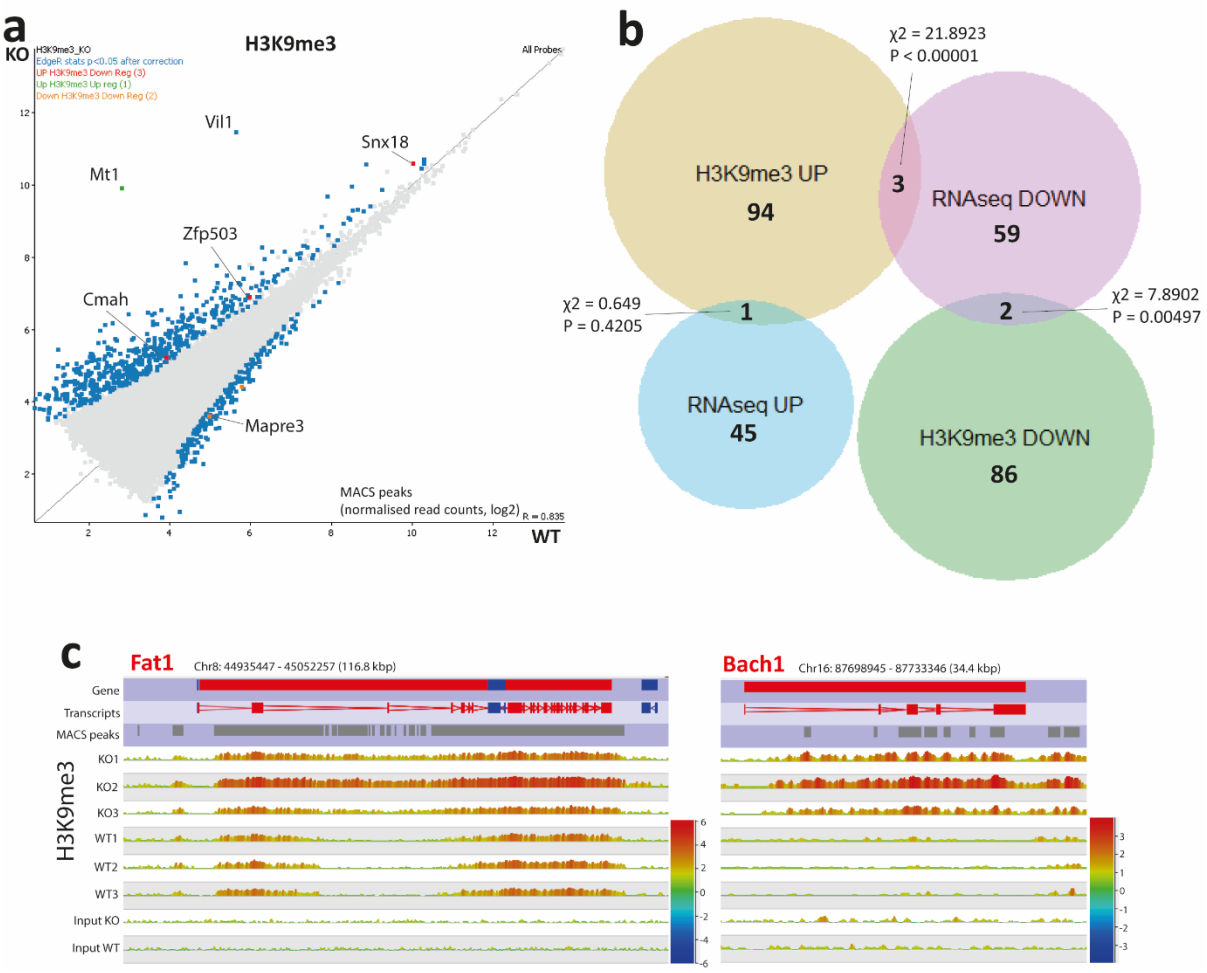


Figure 5.2.1 H3K9me3 changes linked to *Smarcad1* knockout in the colon.

(a) *EdgeR analysis (p<0.05) of read counts over H3K9me3 peaks, coloured probes to the right of the scatterplot have been reduced in trimethylation and those left of the scatterplot have an increase in trimethylation. Genes that have been identified by DESeq2 analysis of the RNAseq data set (p<0.05) are coloured according to an increase or decrease in gene expression. Red coloured genes have been increased in trimethylation and reduced in gene expression, green coloured genes show an increase in trimethylation and an increase in gene expression and the orange coloured genes showed a reduction in trimethylation and gene expression.* ***(b)*** *Four-way venn diagram showing the overlap between up-trimethylated, up-regulated, down-trimethylated and down-regulated genes with Smarcad1-KO, changes assigned to genes closest to peaks within 2kbp. The significance of overlap was tested using a 2x2 two-tailed Chi-square test with Yates correction and the total number of expressed genes of 25896, as determined from the RNA-seq data set excluding unobserved probes.* ***(c)*** *H3K9me3 changes over selected regions. The top annotation tracks indicate the gene, exons, and H3K9me3*

MACS peaks. Read counts are shown in linear scale, color coded. Chromosomal position and length in kbp are indicated.

5.2.2 Effects of Smarcad1 KO on H3K9ac

ChIPseq analysis of H3K9ac in colon epithelial cells when Smarcad1 has been knocked out revealed many H3K9ac changes linked to genes. The MACS peaks analysis highlighted 6524 differentially regulated probes using the EdgeR statistical analysis ($p < 0.05$) of read counts over H3K9ac peaks (Fig. 5.2.2a). Probes highlighted in red (n = 2637) showed an increase in acetylation when Smarcad1 was knocked out and probes highlighted in blue (n = 3887) showed a decrease in acetylation. A volcano plot of the differentially acetylated probes was graphed with the differentially expressed genes overlaid on top (Fig. 5.2.2b). Smarcad1 gene (labelled blue point) has been highlighted to have reduced expression when knocked out as expected. The gene Mapre3, has been highlighted to show a change in enrichment of the histone lysine residue H3K9ac as it did with H3K9me3. However, the reduction in acetylation of Mapre3 was linked to a reduction in gene expression (Fig. 5.2.2b – blue point). The effects of Smarcad1-KO at H3K9 reveals genes that have a decrease in the acetylation modification link to a significant change in gene expression in both directions. There were 25 genes that have had a reduction in acetylation and down-regulated gene expression ($\chi^2 = 46.036$, $p < 0.00001$) and 15 genes that were up-regulated in gene expression where acetylation was reduced ($\chi^2 = 18.537$, $p < 0.00001$) (Fig 5.2.2c). Some of the genes that are both down acetylated and down-regulated in gene expression (Fig. 5.2.2c) have been labelled in Fig. 5.2.2.b. One of these genes, *Lrp1* (low density lipoprotein receptor-related protein 1), has been found to be involved in tumorigenesis, and acts as a tumour suppressor[261] by mediating the clearance of

enzymes involved in cancer spread. Some studies have identified that a reduction in *Lrp1* expression can lead to a worse colorectal cancer outcome[262]. *Tnip1* (TNFAIP3 interacting protein 1) also known as *ABIN-1* is a protein that plays a critical role in inhibiting necroptosis, apoptosis and NF-κB activation[263]. A study by Cai *et al.* (2021)[263] discovered that a down-regulation of this gene promotes CRC suppression. Another gene that has found to have been down-regulated and reduced H3K9ac was the gene *Galnt6* (polypeptide N- acetylgalactosaminyltransferase 6) known to suppress the progression of CRC[205].

The set of genes that were up-regulated in gene expression and significantly linked to the reduction in H3K9ac revealed genes that function in suppressing CRC or have been found to be down-regulated or absent in CRC cells. *Rnf152* (ring finger protein 152) has been found to suppress CRC, and with high expression levels resulting in better prognosis for patients[264]. This gene has been identified to be up-regulated when Smarcad1 was knocked out, with a decrease in H3K9ac. Another gene, *MCT1/SLC16A1* has been described in other studies to be downregulated in colorectal carcinoma[265] possibly linked to switching from butyrate as an energy source to glucose[265]. The up-regulation of this gene when Smarcad1 has been knocked out could be a counteracting response to the disruption within the colon epithelial cells. Other genes that have been up-regulated, function to maintain the homeostatic environment both with normal cells and during aberration. These include *SULT1B1* (sulfotransferase family 1B, member 1), a gene that is found to be expressed at highest levels throughout the human colon and is vital in the regulation of messenger molecules and removal of xenobiotics[266].

The effects of Smarcad1 depletion at H3K9ac are varied and contradictory. It is difficult to assign a particular function linked to H3K9ac, as a number of affected genes are

known to suppress the development of CRC and another set have been seen to promote tumorigenesis. What can be established is that Smarcad1 is part of a sophisticated mechanism and it is possible that with H3K9ac, it functions to ensure an overall balance of gene functions and regulated gene expression to maintain a healthy colon and counteract the potential cancer development.

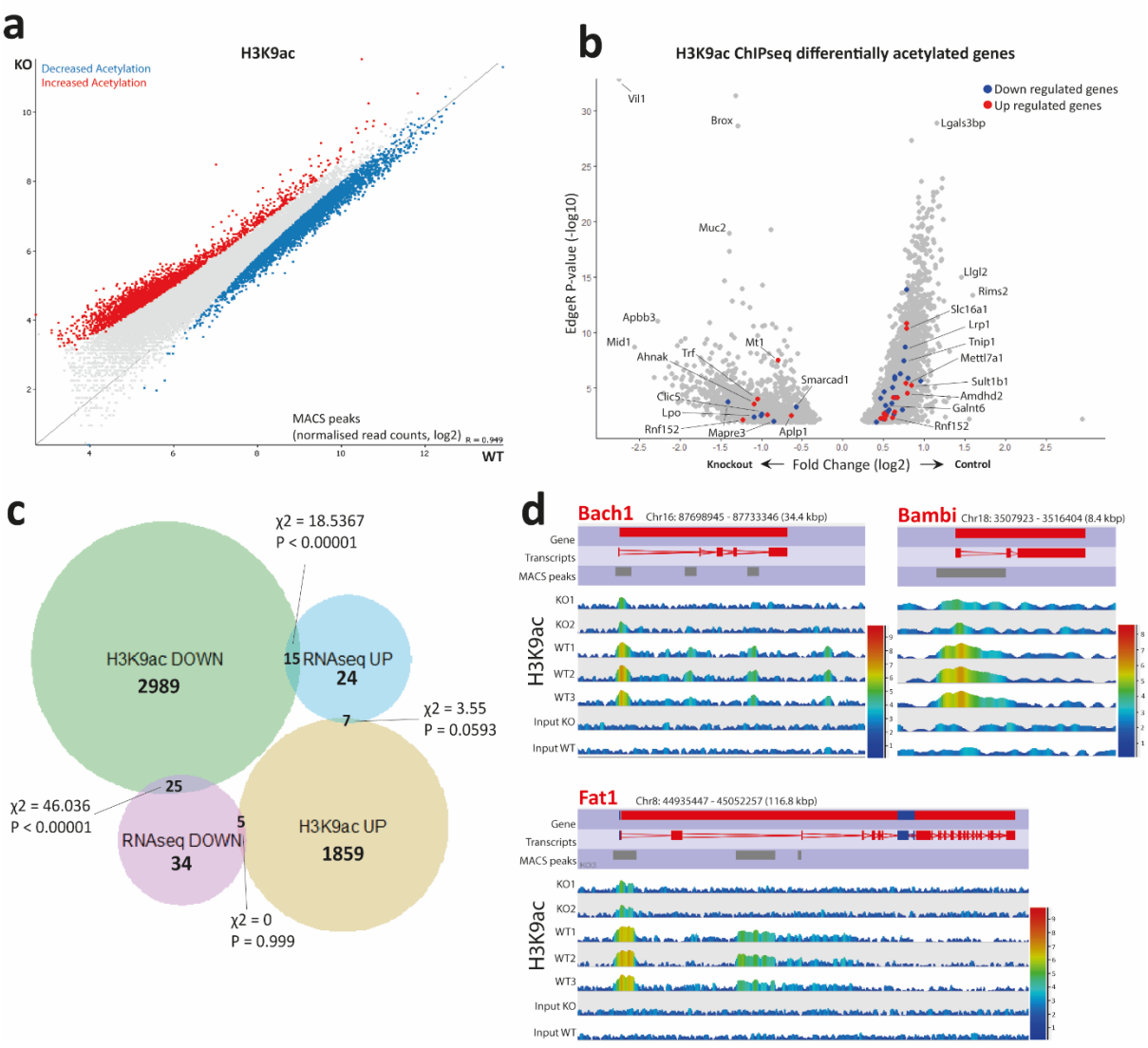


Figure 5.2.2 H3K9ac changes linked to *Smarcad1* knockout in the colon.

(a) *EdgeR analysis (p<0.05) of read counts over H3K9ac peaks, probes coloured blue have been reduced in acetylation and those coloured red have an increased in acetylation.* ***(b)*** *Volcano plot of genes containing MACS peaks closest to 2kbp that changed in acetylation, as identified using EdgeR statistical testing of the H3K9ac ChIPseq data set, using a threshold p-value of 0.05 and adjusting for multiple testing. The most significant genes and those that*

changed the most on treatment are labelled. Genes that increase or decrease in gene expression, as identified by RNA-seq, are coloured red or blue respectively. ***(c)*** *Four-way venn diagram showing the overlap between up-acetylated, up-regulated, down-acetylated and down-regulated genes with Smarcad1-KO. The significance of overlap was tested using a 2x2 two-tailed Chi-square test with Yates correction and the total number of expressed genes of 25896, as determined from the RNA-seq data set excluding unobserved probes.* ***(d)*** *H3K9ac changes over selected regions. The top annotation tracks indicate the gene, exons, and H3K9ac MACS peaks. Read counts are shown in linear scale, color coded. Chromosomal position and length in kbp are indicated.*

5.2.3 Effects of Smarcad1 KO on H3K9cr

MACS peaks analysis showed 10069 probes to have been differentially crotonylated when Smarcad1 has been knocked out. Of those, 4336 probes have been up-crotonylated, and 5733 probes have been down-crotonylated (Fig. 5.2.3a) identified by EdgeR analysis ($p<0.05$). Differentially expressed genes were overlaid on top of the differentially crotonylated probes (Fig. 5.2.3b). *Mt1* is up-regulated and *Mapre3* is down-regulated the same as H3K9ac. An interesting finding is that four of the down regulated MACS peaks are over the YEATS4 gene when the Smarcad1 gene is knocked out (Fig. 5.2.3b), MACS peaks over YEATS4 gene demonstrated in Figure 5.2.3d. YEATS4 has been identified as a candidate crotonylation reader by the Varga-Weisz group (Stoyanova et al. Manuscript in preparation), and with this reduction in H3K9cr enrichment with Smarcad1-KO could suggest a possible feedback loop. The Smarcad1 gene could be associated with the YEATS4 crotonylation reader.

The overlap between differentially crotonylated genes and differentially expressed genes was plotted on a Venn diagram and statistically tested with a Chi^2 and corrected with Yates correction (Fig. 5.2.3c). The only significant overlap was the down crotonylated genes against down-regulated expressed genes (n = 30 genes, χ^2 =

23.885, p <0.00001). Some of these genes can be seen labelled in Fig. 5.2.3b (blue points). One of the genes that has been found to have a great reduction in H3K9cr is the *Sectm1b* (Secreted and transmembrane1B). This gene functions to modulate the activation of regulatory T cells[267] and the down regulation of this gene would support the reduction in immune responses that are seen with Smarcad1 reduction. Another gene to have been greatly reduced in crotonylation and expression is the *RhoC* (Ras homolog gene family member C) gene. This gene has reportedly been shown to control actin organisation and has been demonstrated to affect cancer cells' motility, which in turn affects invasion and metastasis and speeds up the course of carcinoma[268]. The reduction in *RhoC* as seen on Smarcad1 knockout has been linked in other studies to the inhibition of CRC[269]. An activator of the Rho family of GTPases, is the *Tiam1* (T cell lymphoma invasion and metastasis 1) gene, also found to be reduced in gene expression and crotonylation. It has been reported to be crucial in the invasion and metastasis of colorectal cancer[270]. Like *RhoC*, a reduced expression of *Tiam1* has resulted in a good prognosis of patients with CRC[271].

The link between Smarcad1 and crotonylation can be strongly suggested by the major reduction in crotonylation across many genes at H3K9cr, including the crotonylation reader and remodelling protein YEATS4.

Smarcad1- KO appears to have had a far greater effect on H3K9cr than H3K9ac. The differentially crotonylated genes are almost 2-fold greater than differentially acetylated genes. By comparison between H3K9ac and H3K9cr, it could be suggested that Smarcad1-mediated responses with H3K9ac act in more of a maintenance and homeostasis role, whereas H3K9cr acts in a protective role. To explore the similarities and differences between the two modifications their differentially acylated genes were compared.

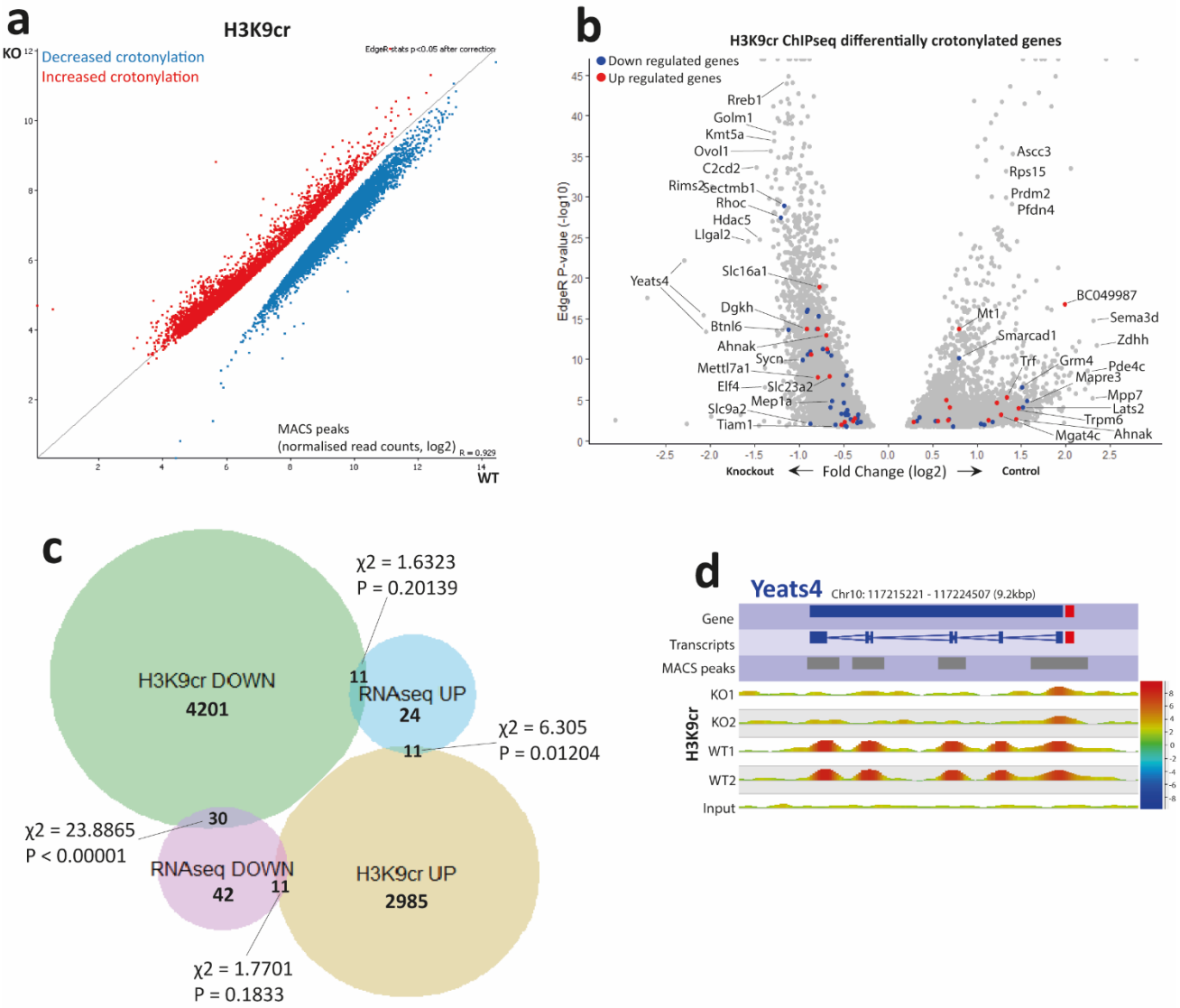


Figure 5.2.3 - H3K9cr changes linked to Smarcad1 knockout in the colon.

(a) *EdgeR analysis (p<0.05) of read counts over H3K9cr peaks, probes coloured blue have been reduced in acetylation and those coloured red have an increased in acetylation.* ***(b)*** *Volcano plot of genes containing MACS peaks that changed in crotonylation, as identified using EdgeR statistical testing of the H3K9cr ChIPseq data set, using a threshold p-value of 0.05 and adjusting for multiple testing. The most significant genes and those that changed the most on treatment are labelled. Genes that increase or decrease in gene expression, as identified by RNA-seq, are coloured red or blue respectively.* ***(c)*** *Four-way venn diagram showing the overlap between up-crotonylated , up-regulated, down-crotonylated and down-regulated genes with Smarcad1-KO, changes assigned to genes closest to peaks within 2kbp. The significance of overlap was tested using a 2x2 two-tailed Chi-square test with Yates correction and the total number of expressed genes of 34245, as determined from the RNA-*

seq data set excluding unobserved probes. ***(d)*** *H3K9cr changes over Yeats4. The top annotation tracks indicate the gene, exons, and H3K9cr MACS peaks. Read counts are shown in linear scale, color coded. Chromosomal position and length in kbp are indicated.*

5.3 Comparison between H3K9ac and H3K9cr with Smarcad1 KO

5.3.1 H3K9ac and H3K9cr have many overlapping genes that are differently modified with Smarcad1 KO

With the Smarcad1 KO, many genes were differentially modified at H3K9. ChIPseq analysis showed that with both H3K9ac and H3K9cr, there were large numbers of genes that had been enriched or reduced with these residues. The question whether H3K9ac and H3K9cr effected the same genes was investigated, and the gene lists generated by EdgeR analysis were compared. From a total of 4241 genes that had been reduced in crotonylation, 2665 of them also had a reduction in acetylation (χ^2 = 16941.54, p <0.00001) (Fig. 5.3.1a 1-3). Conversely, of the total genes that had been enriched with crotonylation (n = 2973), 982 were also enriched with acetylation (total n = 1871, χ^2 = 4609.4, p <0.00001) (Fig. 5.3.1a 4-6).

Gene ontology analysis was performed on these generated lists to identify patterns in biological processes associated with either or both histone modifications. Figure 5.3.1b 1, shows the biological processes of the genes that had reduced enrichment in both H3K9ac and H3K9cr. The most common process was metabolic, with the greatest number of genes giving terms such as “phosphorus metabolic process” and “phosphate-containing compound metabolic process”. These are crucial components of nucleic acids and the cell membrane, acting as key intracellular signalling mediators, and regulates protein activity[272]. This complements the second largest group of genes linked to signal transduction. This could suggest that both histone modifications have common functions regulating the cell's metabolism, shape, and gene expression.

Gene ontology of the 1576 genes that were down crotonylated only, showed similar results with “metabolic processes” being the most common occurring and the greatest number genes gave terms such as ““phosphorus metabolic process” and “phosphate-containing compound metabolic process”. These reoccurring terms present in the down crotonylated only gene list could suggest that the effect of Smarcad1 deletion affects crotonylation (Fig. 5.3.1b 2) far greater than it does with acetylation (Fig. 5.3.1b 3) in genes related to metabolic processes.

One of the most enriched terms under the signalling transduction group is the “Apoptotic signalling pathway”, suggesting that a large number of genes are that are linked to cell death and regulation of carcinomas may be linked to down-crotonylation (Fig. 5.3.1b 2).

Gene ontology analysis of the list of genes that were increased in both H3K9ac and H3K9cr on Smarcad1 KO (Fig. 5.3.1b 4) resulted in a smaller range of processes occurring in both up acylated groups. However, the most common process was linked to development. The most enriched term within this process is “establishment of cell polarity” and the highest number of genes present within this group gave terms such as “nervous system development” and “neuron differentiation”.

The genes that were up-crotonylated only (Fig. 5.3.1b 5), resulted in a smaller variety of biological processes, the most common being cellular processes. The most enriched gene ontology term within this group is “plasma membrane organisation” and the majority of the genes present in this group gave the term “cellular component biogenesis”.

Gene ontology analysis of up acetylated genes after applying quality control filters, (see methods, 2.4.3) gave no terms.

This analysis has revealed that when Smarcad1 has been depleted, whether H3K9ac or H3K9cr has been up or down acylated the response affects the development and maintenance of the cells function. In all the analysis, a large number of terms have been linked to immune response, however the number of genes and enrichment levels of these are too low to be seen in the graphs. It could possibly suggest that the Smarcad1 influence on the immune response is present but not greatly affected by or linked to these histone modifications. There may be another mechanism that triggers a down regulation in immunity when Smarcad1 is reduced.

The interesting finding with the reduced groups, is that with H3K9ac only, H3K9cr only and both together, all three give processes that are linked to metabolism (Fig. 5.3.1b 1,2,3), which suggests that both histone modifications have similar, if not the same functions but not necessarily with all the exact same genes. This is not seen with the increased H3K9ac and H3K9cr group but that could be because there are less genes that have been enriched, so the other common processes are not recognised. It could also be that the increase in both histone modifications main focus is on developmental processes.

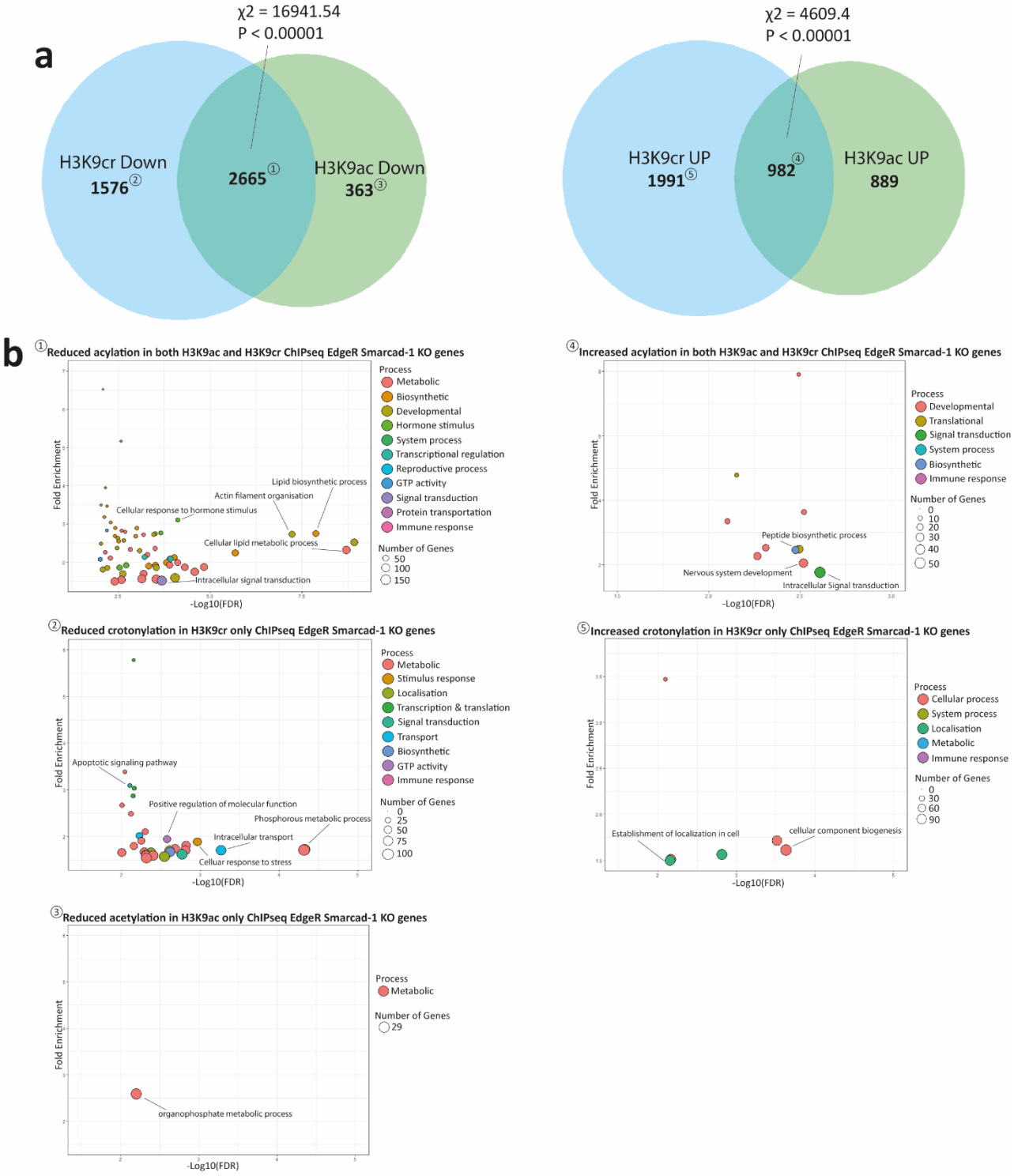


Figure 5.3.1 - The similarities between H3K9ac and H3K9cr

(a) *Venn diagram of all the genes with reduced crotonylation peaks of H3K9cr overlapped with those showing reduced H3K9ac (left), Venn diagram of all the genes with increased crotonylation overlapped with those with increased H3K9ac (right). Chi^2 test was performed with Yates correction.* ***(b)*** *Gene ontology analysis, plotted by FDR against fold enrichment, (1) genes that have been reduced in both crotonylation and acetylation. (2) genes that have*

been reduced with crotonylation only. (3) genes that have been reduced with acetylation only. (4) genes that have been enriched with both crotonylation and acetylation, and (5) genes that have been encriched with crotonylation only.

5.3.2 Smarcad1 KO can affect H3K9ac and H3K9cr in different ways

The vast number of genes that exhibited in an up/down crotonylation or and up/down acetylation when Smarcad1 was deleted showed the processes that may have been affected by either histone modification or both. What was left to discover, was whether there were any genes that were differentially acylated with the opposing modification. The gene lists of differentially acylated genes were cross referenced with the opposing histone medication residue, i.e., the genes that were up crotonylated with Smarcad1-KO against genes that were down acetylated, and vice versa.

The results showed that where there was up crotonylation on H3K9, 430 genes were also down acetylated at this residue (χ^2 = 112.32, p <0.00001) (Fig. 5.3.2.a). These genes were analysed through the gene ontology program Panther, and the most common processes returned were “regulation of cellular components”. Terms such as “actin filament organisation” and “positive regulation of cytoskeleton organisation” are the dominating terms, suggesting that an increase in crotonylation and a reduction in acetylation could encourage a positive regulation of genes linked to muscle structure and mechanical support of the gut. 573 genes were up-acetylated and down-crotonylated, (χ^2 = 566.84, p <0.00001) (Fig. 5.3.2b). Gene ontology analysis showed that these genes gave a greater variety of biological processes including a “negative regulation of actin filament polymerization” and “negative regulation of protein-containing complex assembly” the opposite processes present in the up crotonylated/down acetylated group and the most enriched of all the terms. The large

diversity of biological processes with up acetylation/down crotonylation when Smarcad1 has been deleted could suggest that in a normal gut, the function of H3K9cr would be responsible for more biological processes than H3K9ac. As they are both chromatin activators, the reduction in the modification when Smarcad1 has been depleted supports the idea that both crotonylation and Smarcad1 are involved with a variety of functions and processes that acetylation are not linked to.

In conclusion despite the similarities in functions and processes with a Smarcad1 and these histone modifications, it seems that they affect both identical and different locations along the genome and enrich different genes with the influence of Smarcad1.

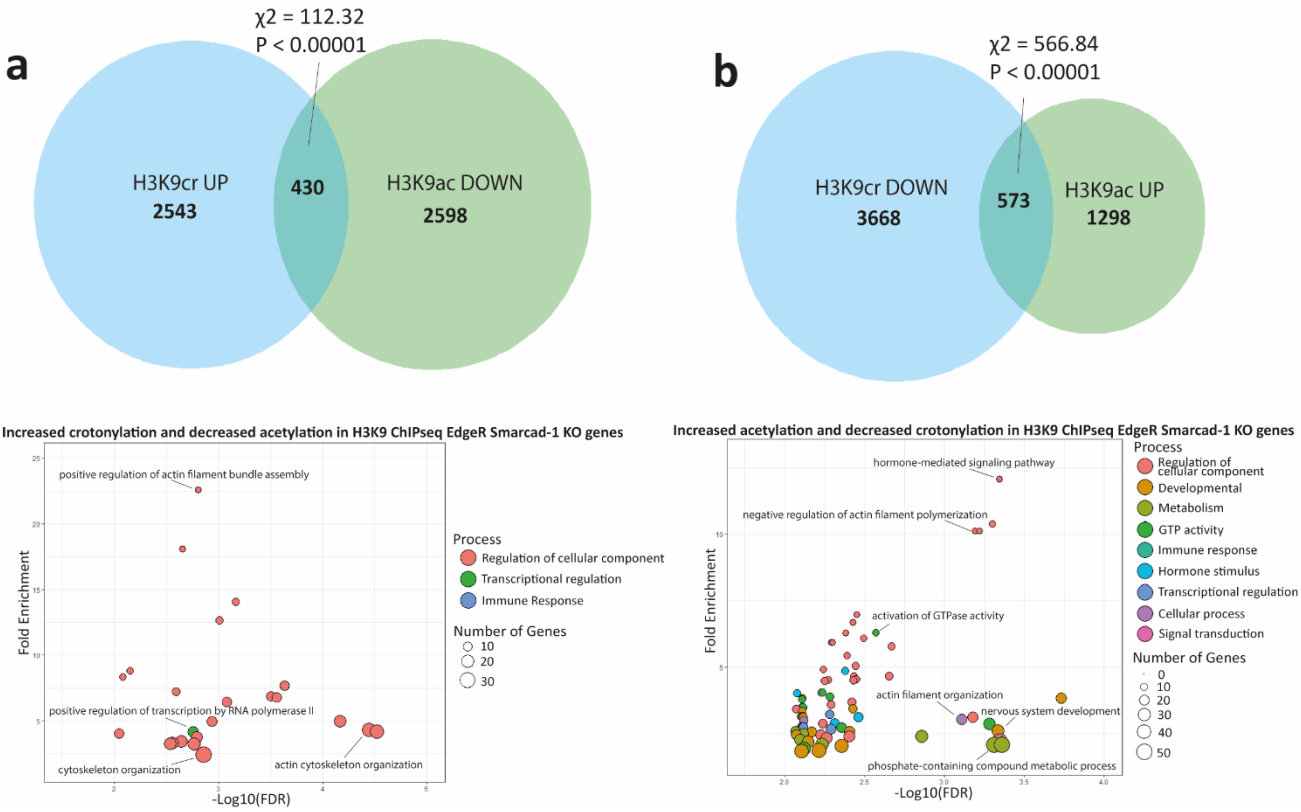


Figure 5.3.2 – Opposing response to Smarcad1-KO of H3K9ac and H3K9cr

(a) *Venn diagram of all the genes with opposing responses to Smarcad1-KO, with increased H3K9cr peaks overlapped with the reduced H3K9ac peaks (top), Gene ontology analysis, plotted by FDR against fold enrichment, genes that have been enriched with crotonylation and*

reduced in acetylation (bottom). ***(b)*** *Venn diagram of all the reduced H3K9cr peaks overlapped with the enriched H3K9ac peaks (top). Chi^2 test was performed with Yates correction. Gene ontology analysis, plotted by FDR against fold enrichment, genes that have been enriched in acetylation and reduced in crotonylation (bottom).*

5.4 Discussion

Smarcad1 deletion leads to changes in gene expression in the colon epithelium and colitis resistance. Absence of *Smarcad1* can result in significant changes in all three histones H3K9me3, H3K9ac and H3K9cr over many sites, including genes that are differentially regulated upon *Smarcad1* deletion.

Although it has been well reported that Smarcad1 is a heterochromatin promoter, my research has shown that it also has a strong role in euchromatin promotion. The largest effect of the three histone modifications was H3K9cr with a significant number of genes down crotonylated with *Smarcad1* deletion. These findings suggest that Smarcad1 may have a larger effect on crotonylation at H3K9 over the other two modifications. It cannot be ruled out that this result may be due to a technical discrepancy where H3K9cr ChIP just worked better and revealed more sites. Despite this, Smarcad1 plays significant roles with H3K9me3 with gene expressions linked to immune response and CRC suppression. The roles Smarcad1 plays with H3K9ac was also significant in that the effects of the deletion of the remodeller resulted in a dramatic reduction of H3K9ac and a variety of differentially expressed genes as a consequence.

It is well recognised that several genes, distinct signal transduction pathways, and the involvement in numerous significant biological processes all have a role in the initiation, development, invasion, and metastasis of colorectal cancer. The genes that

have shown a reduction in H3K9ac seems to both possess these roles and block them simultaneously. Which suggests that H3K9ac may not be necessarily labelled as a CRC promotor or suppressor, but more as a regulator of changes that may occur within the gut, to act in a way that would maintain homeostasis.

The immense changes with H3K9cr over genes that occurred with Smarcad1-KO, gives an indication of the cross-talk that is happening between this particular post-translational modification and Smarcad1. This possible link between Smarcad1 and crotonylation could infer some interesting theories. With the Smarcad1 deletion, reduced crotonylated genes are mainly linked to a reduction in the immune response, and to supporting the suppression of CRC. This could suggest that crotonylation with Smarcad1 acts a protective mechanism against potential carcinomas and reduced level of inflammatory responses.

The interesting revelation between H3K9ac and H3K9cr is that despite their similarities in many functions, they also show their independence from each other. Through the deletion of *Smarcad1*, many more genes had shown a reduction in crotonylation than acetylation, confirming the strong link between this chromatin remodeler and histone modification. It was also noted that without Smarcad1, a significant number of genes were both reduced in crotonylation and enriched with acetylation. These particular genes gave a wide variety of biological processes, and considering that these histone modifications are both chromatin activators, it could be that the acetylated genes may counteract response to the absence of crotonylation via the Smarcad1-mediated regulation of histone modifications.

Smarcad1 plays a very complex role when regulating the histone modifications, both coordinated and antagonistic functional relationships between nucleosome

remodelling and modifying machineries. Understanding the combined roles of the complexes that change nucleosome interactions, location, and stability is crucial, especially considering the increasing recognition of their importance in human health and disease.

Chapter 6. Discussion

My thesis' purpose was to gain better knowledge of the gut microbiome's effect on epigenetic relationships, functions and processes of host colon cells. The intention was to provide to a better understanding of the development of IBD and the genomic factors that would influence the immune response and increase the risk of CRC. This information could ultimately facilitate and develop preventative and treatment measures.

The interactions between the chromatin dynamics and the gut microbiome are very sophisticated. There are many pathways that facilitate the smooth running of the gut and protect it from disruptions to its environment. I have identified that the SCFA crotonate, possesses a protective effect on DSS-induced colitis, demonstrating the inhibition of symptomatic colitis and weight loss compared to DSS-only control mice. This suggests a potential pre-treatment for IBD patients could be developed through the use of crotonic acid supplemented into their diets.

I was able to characterise the overall impacts of the microbiome on histone modifications, especially histone crotonylation, providing insight into the role crotonylation may play in tumour development and suppression, as well as immune response modulation.

Finally, I have identified some of the functions that the chromatin remodeller Smarcad1 plays in histone modifications and a possible connection with the microbiome. Most notably, Smarcad1 modulates the activity of H3K9ac and H3K9cr, emphasising linkages to gut maintenance and possible regulation of tumour formation and suppression. With the major impact Smarcad1 has on H3K9cr and the significant rise

of this histone modification in recolonised germ-free mice, I was able to identify a probable relationship between Smarcad1 and the microbiome.

Overall, establishing the roles of these host-microbiome interactions has enabled potentially new approaches of exploration for developing IBD and CRC preventative strategies and prospective treatment targets.

6.1 The SCFA crotonate provides a significant level of protection against DSS-induced colitis

Crotonate has been shown to provide a significant level of protection within the gut when colitis has been induced by DSS. This SCFA has unexpectedly shown better results in protecting the gut than butyrate, a well-studied SCFA that plays a critical role in cellular metabolism and intestinal homeostasis[74,150]. Butyrate is the preferred energy source for the gut epithelial cells and is a key nutrient in modulating biological responses such as inflammation regulation and promoting genomic stability[74]. The unexpected result that butyrate did not provide the level of protection that was anticipated could be due to a number of reasons. This includes butyrate's high affinity with GPR109A – a GPCR that activates the inflammasome pathways in colonic macrophages and results in T-cell production[54]. This increased influx of immune cells would result in a greater level of inflammation facilitating the colitis response.

However, the significant level of protection that crotonate has provided could give potential therapeutic routes in preventing colitis. Crotonate is already used as a herbal remedy in India to treat people with digestive problems, constipation and diarrhoea. But understanding how crotonate affects the gut needs further investigation. The link between crotonate and histone crotonylation could give a clue, as histone crotonylation has been reported to be involved in immune response regulation. There

could also be a link between crotonate and the gut microbiome. Perhaps the result of DSS-induced colitis, causing the microbiota to be depleted it would not have much of a downstream effect on the crotonate concentration or ability to protect because crotonate is naturally present in the gut at low levels whereas butyrate relies heavily on the microbiota to be produced.

Studies by Tsukuda *et al.* (2021)[273] have shown that different composition of microbiota results in different gut SCFAs profiles. Although their study was based around early-stage development of the microbiota, they have been able to demonstrate that the composition and diversity of the microbiota can be a limiting factor in determining the concentrations of each of the SCFAs. This supports the idea that the disruptions to the microbiota from DSS-induced colitis may have noticeably affected butyrate production and therefore inhibited its protective abilities.

The idea that the mice in DSS experiments would have had differing base lines of butyrate and crotonate in the gut should be considered. To take this factor in to consideration in future studies, the concentration levels of both butyrate and crotonate should be measured in blood and stool prior to treatment, then again before DSS is given and then at the end of the experiment. This would also confirm the efficiency and/or loss of both SCFAs during their migration through the digestive tract towards the colon.

6.2 The gut microbiome greatly impacts histone crotonylation within the colon epithelium

The gut microbiome has been shown to greatly influence the function of HPTMs mediated by SCFAs that induce HATs and HDAC responses[55]. I was able to show the profound impact the microbiota has on histone crotonylation. With the depletion of the

microbiota, significant changes with the histone crotonylation, revealed a direct correlation with differential gene expression. The dysregulation of hundreds of expressed genes was linked to functions associated with both tumour suppression and CRC promotion. This link is present in both up-regulated and down-regulated genes suggesting the global effect of microbiota depletion over crotonylated genes could be acting in way to maintain balance and counteract the effect of a depleted microbiome on some genes. This was also seen with the predicted TFs linked to Kcr; irrespective of being enriched or reduced within the depleted microbiota group, the TFs identified seem to be linked to colorectal cancer development and immune response. The gene ontology analysis of genes linked to reduced Kcr with a depleted microbiota returned terms linked to “regulation of transcription”, “regulation of biological quality” and “negative regulation of cellular processes” including “negative regulation of transcription”, suggesting that Kcr linked to microbiota supports the development and regulation of cell production and processes. Gene groups that were linked to an increase in crotonylation returned terms associated with an increased immune response, supporting the idea that crotonylation directly responds to changes within the colon environment in way to protect the gut.

This response to the depletion of the microbiota was not seen with histone acetylation where changes were very low. This suggests that crotonylation is affected far greater than acetylation by the presence or absence of the microbiota. It has been reported that acyl-CoA forms such as crotonyl-CoA, the co-factor of histone crotonylation, is produced during periods of low energy availability and low oxygen consumption[73], whereas acetyl-CoA is mainly generated during high levels of energy and is used during oxygen consumption[274], albeit studies were done in *S. cerevisiae*[275]. It is possible that increase of cellular concentration of crotonyl-CoA, which can be altered

through genetic and environmental disruptions such as microbiota could lead to a reduction in acetyl-CoA and allow histones to be crotonylated more. This also supports the idea that histone acetylation and crotonylation respond differently to the depletion of the microbiota and thus the molecular pathways and cellular processes linked to these modifications differ.

It is difficult to determine how other site-specific histone modifications of acetylation and crotonylation would respond to the depletion of the microbiota. Especially considering the effect of germfree mice after recolonization with a microbiota, showed a significant increase in H3K9cr and H3K9ac, but also H3K18ac – which was not as significant but still increased. Further research needs to be done using ChIPseq with H3K9cr and H3K9ac to compare these two modifications locally. Kac and H3K18cr would also be insightful to investigate to confirm whether the global effects of the microbiome do indeed effect crotonylation more than acetylation.

6.3 Smarcad1 linked to histone modifications H3K9ac and H3K9cr

Smarcad1 is a complex and sophisticated chromatin remodeller, and has been found to be involved in many mechanisms, and regulates hetrochromatin and euchromatin in parallel. The deletion of this gene revealed its interactions with H3K9me3, H3K9ac and H3K9cr at varied degrees. Notably, H3K9ac and H3K9cr were greatly affected by the deletion of Smarcad1 and resulted in a significant change in regulation of gene expression. The down regulation of H3K9ac was connected to a significant number of differentially expressed, up-regulated and down-regulated genes with links to tumour regulation. H3K9cr however showed the greatest change when Smarcad1 was deleted, with the most significant effect on the down regulation of expressed genes

linked to the reduction of H3K9cr. These genes were associated with immune response and CRC suppression.

Crotonylation has been reported to be abundant over promoters and possibly enhancers of active genes, boosting transcription to a greater extent than acetylation[108]. Crotonyl lysine and acetyl lysine chromosomal distribution studies indicate functional distinctions between the two, as they compete for sites along the histone N-terminal tails. It has been suggested that these modifications are not functionally redundant and have distinct roles in regulating gene function resulting in diverse biological outcomes[109]. I have been able to support this suggestion in regards to Smarcad1 as I identified a significant number of genes that were up crotonylated and down acetylated when Smarcad1 was deleted, and vice-versa. The remaining genes that did not overlap showed a range of diverse biological processes also, supporting the fact that the two acyl-modifications possess unique functions from each other.

However, what is yet to discover, is the direct pathway between Smarcad1 and histone modification changes; specifically H3K9cr. Perhaps, further research could be carried out with the YEATS4 crotonylation reader, a gene that was found to be notably down crotonylated with Smarcad1-KO. It may give us an insight into Smarcad1 – crotonylation interactions.

We cannot rule out that the different effect of Smarcad1 on all three histone modifications is in part due to a technical issue where different strengths of antibodies are seen against the modifications, this should be tested further, e.g. using different antibodies from different companies.

Overall, my analysis shows that Smarcad1 deletion can noticeably affect histone modifications in the gut - specifically those linked to gene expression.

6.4 Conclusions

This study has probed a few questions about the role of histone modifications, in particular crotonylation in the host-microbiome interactions, and in turn has opened up many avenues to research further. Evidently, there is a connection between histone crotonylation and immune response, gene activation and maintaining homeostasis within the gut, all of which plays role in supporting the suppression of colorectal cancer.

There seems to be at this stage a poorly defined link between crotonate, histone crotonylation, microbiota and Smarcad1. I have seen a strong association between crotonate and the reduced response in DSS-induced colitis. Smarcad1-KO also inhibits DSS-colitis, seen in a study by Kazakevych *et al.* (2020)[141]. Smarcad1 is also strongly linked to histone crotonylation, specifically H3K9cr. The microbiota also greatly impacts the presence of histone crotonylation and significantly increased H3K9cr when germfree mice were recolonised. All these observations could be interlinked at some point along the pathways towards cell and gut function, but they could also be working alongside each other or in a synergistic manner.

6.5 Future studies

To take this research further, it would be interesting to explore the effects of crotonate and butyrate on the gut at a genomic level. A way to approach this would be to analyse extracts from the colons from these mice by western blot to identify HPTM changes with the presence of crotonate and butyrate. It would be informative to also focus on the genomic effects linked to HPTMs, looking specifically at H3K9cr, H3K9ac, Kcr and

Kac by running a ChIPseq on the intestinal epithelial colon cells. This would shine a light on the extent of influencing factors that could modulate these histone modifications. Luminal content should be measured before, during (just before DSS treatment) and after the experiment for the SCFAs levels and microbiota diversity, ELISA to measure the cytokines and finally histology sections taken to visualise the extent of the damage DSS would have on each test group. Mice that are mono-colonised with bacteria that specifically produce butyrate such as *Faecalibacterium prausnitzii,* could be used alongside this experiment to measure the effects of the microbiota population with SCFA supplementation.

To fully understand whether crotonylation is greatly impacted by the loss of the microbiota over acetylation, ChIPseq experiments need to be performed on H3K9ac, H3K9cr and Kac upon antibiotics treatment and controls to complete the dataset and perhaps identify a site-specific response to the microbiota.

An interesting experiment with Smarcad1, may be to investigate the effects of supplemented SCFAs to Smacard1-KO mice. They both play a role in regulating the histone modifications, and there could be a link between them.

There are still many questions that need to be answered but what has been revealed are the potential routes of investigation in understand better the interactions between the microbiome and the host, leading to treatment and prevention strategies for IBD patients.

7. References

1. Ray, G. & Longworth, M. S. Epigenetics, DNA Organization, and Inflammatory Bowel Disease. *Inflamm Bowel Dis* **25**, 235–247 (2019).
2. Roggeveen, M. J., Tismenetsky, M. & Shapiro, R. Ulcerative Colitis. *RadioGraphics* **26**, 947–951 (2006).
3. Xu, H.-M. *et al.* Selection strategy of dextran sulfate sodium-induced acute or chronic colitis mouse models based on gut microbial profile. *BMC Microbiol* **21**, 279 (2021).
4. van der Flier, L. G. & Clevers, H. Stem Cells, Self-Renewal, and Differentiation in the Intestinal Epithelium. *Annu. Rev. Physiol.* **71**, 241–260 (2009).
5. Levine, J. & Burakoff, R. Chapter 3: Inflammatory bowel disease: medical considerations. in *Current Diagnosis & Treatment: Gastroenterology, Hepatology, & Endoscopy* (McGraw-Hill, 2012).
6. Arias Vallejo, E. [Ulcerative colitis. Introduction]. *Rev Esp Enferm Apar Dig* **38**, 623 (1972).
7. Chang, M. D. & Liu, X. Chapter 5 - Overview of Histopathology of Ulcerative Colitis and Crohn's Disease. in *Interventional Inflammatory Bowel Disease: Endoscopic Management and Treatment of Complications* (ed. Shen, B.) 49–68 (Academic Press, 2018). doi:10.1016/B978-0-12-811388-2.00005-1.
8. Tata, L. News - Rates of Crohn's and Colitis have been vastly underestimated for decades, says new study - University of Nottingham. https://www.nottingham.ac.uk/news/rates-of-crohns-and-colitis-have-been-vastly-underestimated-for-decades-says-new-study (2022).
9. Freeman, K. *et al.* The incidence and prevalence of inflammatory bowel disease in UK primary care: a retrospective cohort study of the IQVIA Medical Research Database. *BMC Gastroenterology* **21**, 139 (2021).

10. Queensland, C. H. Medication for Inflammatory Bowel Disease fact sheet. *Children's Health Queensland* https://www.childrens.health.qld.gov.au/fact-sheet-medication-for-inflammatory-bowel-disease-ibd/ (2016).
11. Archer, R. *et al.* Infliximab, adalimumab and golimumab for treating moderately to severely active ulcerative colitis after the failure of conventional therapy (including a review of TA140 and TA262): clinical effectiveness systematic review and economic model. *Health Technol Assess* **20**, 1–326 (2016).
12. Sparrow, M. P. Adalimumab in ulcerative colitis - efficacy, safety and optimization in the era of treat-to target. *Expert Opin Biol Ther* **17**, 613–621 (2017).
13. Román, A. L. S. & Muñoz, F. Comorbidity in inflammatory bowel disease. *World J Gastroenterol* **17**, 2723–2733 (2011).
14. Skomorochow, E. & Pico, J. Toxic Megacolon. in *StatPearls* (StatPearls Publishing, 2022).
15. Truelove, S. C. & Witts, L. J. Cortisone in Ulcerative Colitis. *Br Med J* **2**, 1041–1048 (1955).
16. Pabla, B. S. & Schwartz, D. A. Assessing Severity of Disease in Patients with Ulcerative Colitis. *Gastroenterol Clin North Am* **49**, 671–688 (2020).
17. Stidham, R. W. & Higgins, P. D. R. Colorectal Cancer in Inflammatory Bowel Disease. *Clin Colon Rectal Surg* **31**, 168–178 (2018).
18. Rezapour, M. *et al.* Reclassifying Pseudopolyps in Inflammatory Bowel Disease: Histologic and Endoscopic Description in the New Era of Mucosal Healing. *Crohn's & Colitis 360* **1**, otz033 (2019).
19. Monteserín, L. *et al.* Colonic pseudopolyposis in inflammatory bowel disease. *Gastroenterol Hepatol* **40**, 375–376 (2017).
20. Politis, D. S., Katsanos, K. H., Tsianos, E. V. & Christodoulou, D. K. Pseudopolyps in inflammatory bowel diseases: Have we learned enough? *World J Gastroenterol* **23**, 1541–1551 (2017).

21. Sharma, R. *et al.* Global, regional, and national burden of colorectal cancer and its risk factors, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet Gastroenterology & Hepatology* **7**, 627–647 (2022).

22. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians* **68**, 394–424 (2018).

23. Ewing, I., Hurley, J. J., Josephides, E. & Millar, A. The molecular genetics of colorectal cancer. *Frontline Gastroenterol* **5**, 26–30 (2014).

24. Rawla, P., Sunkara, T. & Barsouk, A. Epidemiology of colorectal cancer: incidence, mortality, survival, and risk factors. *Prz Gastroenterol* **14**, 89–103 (2019).

25. Arnold, M. *et al.* Global patterns and trends in colorectal cancer incidence and mortality. *Gut* **66**, 683–691 (2017).

26. Yamauchi, M. *et al.* Colorectal cancer: a tale of two sides, or a continuum? *Gut* **61**, 794–797 (2012).

27. Sawicki, T. *et al.* A Review of Colorectal Cancer in Terms of Epidemiology, Risk Factors, Development, Symptoms and Diagnosis. *Cancers* **13**, 2025 (2021).

28. Thomas, J. S. & Shi, C. Chapter 24 - Molecular Testing in Colorectal Cancer. in *Diagnostic Molecular Pathology* (eds. Coleman, W. B. & Tsongalis, G. J.) 305–320 (Academic Press, 2017). doi:10.1016/B978-0-12-800886-7.00024-8.

29. Galandiuk, S. *et al.* Colon and Rectum. in *Sabiston Textbook of Surgery The Biological Basis of Modern Surgical Practice E-Book* 1320–1400 (Elsevier Inc., 2022).

30. Pritchard, C. C. & Grady, W. M. Colorectal Cancer Molecular Biology Moves Into Clinical Practice. *Gut* **60**, 116–129 (2011).

31. Markowitz, S. D. & Bertagnolli, M. M. Molecular Origins of Cancer. *N Engl J Med* **361**, 2449–2460 (2009).

32. Guinane, C. M. & Cotter, P. D. Role of the gut microbiota in health and chronic gastrointestinal disease: understanding a hidden metabolic organ. *Therap Adv Gastroenterol* **6**, 295–308 (2013).

33. Koh, A., Vadder, F. D., Kovatcheva-Datchary, P. & Bäckhed, F. From Dietary Fiber to Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites. *Cell* **165**, 1332–1345 (2016).

34. Frank, D. N. *et al.* Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *PNAS* **104**, 13780–13785 (2007).

35. Ley, R. E., Turnbaugh, P. J., Klein, S. & Gordon, J. I. Microbial ecology: Human gut microbes associated with obesity. *Nature* **444**, 1022–1023 (2006).

36. Wang, Z. *et al.* Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* **472**, 57–63 (2011).

37. Karlsson, F. H. *et al.* Symptomatic atherosclerosis is associated with an altered gut metagenome. *Nature Communications* **3**, 1245 (2012).

38. Tan, J. *et al.* The role of short-chain fatty acids in health and disease. *Adv. Immunol.* **121**, 91–119 (2014).

39. Qin, Y. & Wade, P. A. Crosstalk between the microbiome and epigenome: messages from bugs. *J Biochem* **163**, 105–112 (2018).

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

41. Flint, H. J., Scott, K. P., Louis, P. & Duncan, S. H. The role of the gut microbiota in nutrition and health. *Nat Rev Gastroenterol Hepatol* **9**, 577–589 (2012).

42. Flint, H. J., Scott, K. P., Duncan, S. H., Louis, P. & Forano, E. Microbial degradation of complex carbohydrates in the gut. *Gut Microbes* **3**, 289–306 (2012).

43. Walker, A. W. *et al.* Dominant and diet-responsive groups of bacteria within the human colonic microbiota. *ISME J* **5**, 220–230 (2011).

44. Salonen, A. *et al.* Impact of diet and individual variation on intestinal microbiota composition and fermentation products in obese men. *ISME J* **8**, 2218–2230 (2014).

45. Duncan, S. H. *et al.* Reduced dietary intake of carbohydrates by obese subjects results in decreased concentrations of butyrate and butyrate-producing bacteria in feces. *Appl Environ Microbiol* **73**, 1073–1078 (2007).

46. David, L. A. *et al.* Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505**, 559–563 (2014).

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

48. Ramirez, J. *et al.* Antibiotics as Major Disruptors of Gut Microbiota. *Front Cell Infect Microbiol* **10**, 572912 (2020).

49. Wu, Y., Wang, C.-Z., Wan, J.-Y., Yao, H. & Yuan, C.-S. Dissecting the Interplay Mechanism between Epigenetics and Gut Microbiota: Health Maintenance and Disease Prevention. *Int J Mol Sci* **22**, 6933 (2021).

50. Kennedy, E. A., King, K. Y. & Baldridge, M. T. Mouse Microbiota Models: Comparing Germ-Free Mice and Antibiotics Treatment as Tools for Modifying Gut Bacteria. *Frontiers in Physiology* **9**, (2018).

51. Fukushima, K. *et al.* Non-pathogenic bacteria modulate colonic epithelial gene expression in germ-free mice. *Scand J Gastroenterol* **38**, 626–634 (2003).

52. Sharma, M., Li, Y., Stoll, M. L. & Tollefsbol, T. O. The Epigenetic Connection Between the Gut Microbiome in Obesity and Diabetes. *Frontiers in Genetics* **10**, (2020).

53. Spalinger, M. R. *et al.* Administration of the Hyper-immune Bovine Colostrum Extract IMM-124E Ameliorates Experimental Murine Colitis. *Journal of Crohn's and Colitis* **13**, 785–797 (2019).

54. Liu, H. *et al.* Butyrate: A Double-Edged Sword for Health? *Adv Nutr* **9**, 21–29 (2018).

55. Fellows, R. *et al.* Microbiota derived short chain fatty acids promote histone crotonylation in the colon through histone deacetylases. *Nature Communications* **9**, 105 (2018).

56. Donohoe, D. R. & Bultman, S. J. Metaboloepigenetics: Interrelationships between energy metabolism and epigenetic control of gene expression. *J Cell Physiol* **227**, 3169–3177 (2012).

57. Martin-Gallausiaux, C., Marinelli, L., Blottière, H. M., Larraufie, P. & Lapaque, N. SCFA: mechanisms and functional importance in the gut. *Proceedings of the Nutrition Society* **80**, 37–49 (2021).

58. Cummings, J. H., Pomare, E. W., Branch, W. J., Naylor, C. P. & Macfarlane, G. T. Short chain fatty acids in human large intestine, portal, hepatic and venous blood. *Gut* **28**, 1221–1227 (1987).

59. Tazoe, H. *et al.* Roles of short-chain fatty acids receptors, GPR41 and GPR43 on colonic functions. *J. Physiol. Pharmacol.* **59 Suppl 2**, 251–262 (2008).

60. Miller, T. L. & Wolin, M. J. Pathways of acetate, propionate, and butyrate formation by the human fecal microbial flora. *Appl Environ Microbiol* **62**, 1589–1592 (1996).

61. Reichardt, N. *et al.* Phylogenetic distribution of three pathways for propionate production within the human gut microbiota. *ISME J* **8**, 1323–1335 (2014).

62. Louis, P. & Flint, H. J. Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine. *FEMS Microbiol Lett* **294**, 1–8 (2009).

63. Sleeth, M. L., Thompson, E. L., Ford, H. E., Zac-Varghese, S. E. K. & Frost, G. Free fatty acid receptor 2 and nutrient sensing: a proposed role for fibre, fermentable carbohydrates and short-chain fatty acids in appetite regulation. *Nutrition Research Reviews* **23**, 135–145 (2010).

64. Roediger, W. E. Role of anaerobic bacteria in the metabolic welfare of the colonic mucosa in man. *Gut* **21**, 793–798 (1980).

65. Hong, Y.-H. *et al.* Acetate and propionate short chain fatty acids stimulate adipogenesis via GPCR43. *Endocrinology* **146**, 5092–5099 (2005).

66. Chang, P. V., Hao, L., Offermanns, S. & Medzhitov, R. The microbial metabolite butyrate regulates intestinal macrophage function via histone deacetylase inhibition. *Proceedings of the National Academy of Sciences* **111**, 2247–2252 (2014).

67. Furusawa, Y. *et al.* Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nature* **504**, 446–450 (2013).

68. Smith, P. M. *et al.* The microbial metabolites, short chain fatty acids, regulate colonic Treg cell homeostasis. *Science* **341**, 10.1126/science.1241165 (2013).

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

70. Bardhan, K. *et al.* IFNγ Induces DNA Methylation-Silenced GPR109A Expression via pSTAT1/p300 and H3K18 Acetylation in Colon Cancer. *Cancer Immunol Res* **3**, 795–805 (2015).

71. Sabari, B. R., Zhang, D., Allis, C. D. & Zhao, Y. Metabolic regulation of gene expression through histone acylations. *Nat Rev Mol Cell Biol* **18**, 90–101 (2017).

72. Sabari, B. R. *et al.* Intracellular Crotonyl-CoA Stimulates Transcription Through p300-Catalyzed Histone Crotonylation. *Mol Cell* **58**, 203–215 (2015).

73. Hiltunen, J. K. *et al.* The biochemistry of peroxisomal β-oxidation in the yeast Saccharomyces cerevisiae. *FEMS Microbiology Reviews* **27**, 35–64 (2003).

74. Conlon, M. A. & Bird, A. R. The Impact of Diet and Lifestyle on Gut Microbiota and Human Health. *Nutrients* **7**, 17–44 (2015).

75. Gujral, P., Mahajan, V., Lissaman, A. C. & Ponnampalam, A. P. Histone acetylation and the role of histone deacetylases in normal cyclic endometrium. *Reproductive Biology and Endocrinology* **18**, 84 (2020).

76. Huang, H. *et al.* Fine-mapping inflammatory bowel disease loci to single-variant resolution. *Nature* **547**, 173–178 (2017).

77. Mokry, M. *et al.* Many inflammatory bowel disease risk loci include regions that regulate gene expression in immune cells and the intestinal epithelium. *Gastroenterology* **146**, 1040–1047 (2014).

78. Gangaraju, V. K. & Bartholomew, B. Mechanisms of ATP Dependent Chromatin Remodeling. *Mutat Res* **618**, 3–17 (2007).

79. Vignali, M., Hassan, A. H., Neely, K. E. & Workman, J. L. ATP-Dependent Chromatin-Remodeling Complexes. *Mol Cell Biol* **20**, 1899–1910 (2000).

80. Kaikkonen, M. U., Lam, M. T. Y. & Glass, C. K. Non-coding RNAs as regulators of gene expression and epigenetics. *Cardiovasc Res* **90**, 430–440 (2011).

81. Wei, W. *et al.* Class I histone deacetylases are major histone decrotonylases: evidence for critical and broad function of histone crotonylation in transcription. *Cell Research* **27**, 898–915 (2017).

82. Jostins, L. *et al.* Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491**, 119–124 (2012).

83. Liu, J. Z. *et al.* Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet* **47**, 979–986 (2015).

84. Scarpa, M. & Stylianou, E. Epigenetics: Concepts and Relevance to IBD Pathogenesis. *Inflammatory Bowel Diseases* **18**, 1982–1996 (2012).

85. Allfrey, V. G., Faulkner, R. & Mirsky, A. E. ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS*. *Proc Natl Acad Sci U S A* **51**, 786–794 (1964).

86. Clayton, A. L., Hazzalin, C. A. & Mahadevan, L. C. Enhanced Histone Acetylation and Transcription: A Dynamic Perspective. *Molecular Cell* **23**, 289–296 (2006).

87. Barnes, C. E., English, D. M. & Cowley, S. M. Acetylation & Co: an expanding repertoire of histone acylations regulates chromatin and transcription. *Essays Biochem* **63**, 97–107 (2019).

88. Hałasa, M. *et al.* H3K18Ac as a Marker of Cancer Progression and Potential Target of Anti-Cancer Therapy. *Cells* **8**, 485 (2019).

89. Richman, R., Chicoine, L. G., Collini, M. P., Cook, R. G. & Allis, C. D. Micronuclei and the cytoplasm of growing Tetrahymena contain a histone acetylase activity which is highly specific for free histone H4. *J Cell Biol* **106**, 1017–1026 (1988).

90. Roth, S. Y., Denu, J. M. & Allis, C. D. Histone acetyltransferases. *Annual Review of Biochemistry* **70**, 81–120 (2001).

91. Lee, K. K. & Workman, J. L. Histone acetyltransferase complexes: one size doesn't fit all. *Nat Rev Mol Cell Biol* **8**, 284–295 (2007).

92. Spencer, T. E. *et al.* Steroid receptor coactivator-1 is a histone acetyltransferase. *Nature* **389**, 194–198 (1997).

93. Chen, H. *et al.* Nuclear receptor coactivator ACTR is a novel histone acetyltransferase and forms a multimeric activation complex with P/CAF and CBP/p300. *Cell* **90**, 569–580 (1997).

94. Voss, A. K. & Thomas, T. MYST family histone acetyltransferases take center stage in stem cells and development. *Bioessays* **31**, 1050–1061 (2009).

95. Dekker, F. J. & Haisma, H. J. Histone acetyl transferases as emerging drug targets. *Drug Discovery Today* **14**, 942–948 (2009).

96. Yan, G. *et al.* Selective Inhibition of p300 HAT Blocks Cell Cycle Progression, Induces Cellular Senescence and Inhibits the DNA Damage Response in Melanoma Cells. *J Invest Dermatol* **133**, 2444–2452 (2013).

97. Ait-Si-Ali, S. *et al.* CBP/p300 histone acetyl-transferase activity is important for the G1/S transition. *Oncogene* **19**, 2430–2437 (2000).

98. Seto, E. & Yoshida, M. Erasers of Histone Acetylation: The Histone Deacetylase Enzymes. *Cold Spring Harb Perspect Biol* **6**, a018713 (2014).

99. Lee, J. & R, S. H. Cancer Epigenetics: Mechanisms and Crosstalk of a HDAC Inhibitor, Vorinostat. *Chemotherapy (Los Angel)* **2**, 14934 (2013).

100. Reichert, N., Choukrallah, M.-A. & Matthias, P. Multiple roles of class I HDACs in proliferation, differentiation, and development. *Cell. Mol. Life Sci.* **69**, 2173–2187 (2012).

101. Bosch-Presegué, L. & Vaquero, A. The dual role of sirtuins in cancer. *Genes Cancer* **2**, 648–662 (2011).

102. Villagra, A. *et al.* The histone deacetylase HDAC11 regulates the expression of interleukin 10 and immune tolerance. *Nat Immunol* **10**, 92–100 (2009).

103. Chen, I.-C., Sethy, B. & Liou, J.-P. Recent Update of HDAC Inhibitors in Lymphoma. *Frontiers in Cell and Developmental Biology* **8**, (2020).

104. Xiong, X. *et al.* Selective recognition of histone crotonylation by double PHD fingers of MOZ and DPF2. *Nature Chemical Biology* **12**, 1111–1118 (2016).

105. Tan, M. *et al.* Identification of 67 Histone Marks and Histone Lysine Crotonylation as a New Type of Histone Modification. *Cell* **146**, 1016–1028 (2011).

106. Liu, X. *et al.* MOF as an evolutionarily conserved histone crotonyltransferase and transcriptional activation by histone acetyltransferase-deficient and crotonyltransferase-competent CBP/p300. *Cell Discovery* **3**, (2017).

107. Li, Y. *et al.* Molecular Coupling of Histone Crotonylation and Active Transcription by AF9 YEATS Domain. *Molecular Cell* **62**, 181–193 (2016).

108. Crespo, M. *et al.* Multi-omic analysis of gametogenesis reveals a novel signature at the promoters and distal enhancers of active genes. *Nucleic Acids Res* **48**, 4115–4138 (2020).

109. Ntorla, A. & Burgoyne, J. R. The Regulation and Function of Histone Crotonylation. *Front Cell Dev Biol* **9**, 624914 (2021).

110. Ortega, E. *et al.* Transcription factor dimerization activates the p300 acetyltransferase. *Nature* **562**, 538–544 (2018).

111. Baumann, K. Crotonylation versus acetylation. *Nat Rev Mol Cell Biol* **16**, 265–265 (2015).

112. Kaczmarska, Z. *et al.* Structure of p300 in complex with acyl-CoA variants. *Nat Chem Biol* **13**, 21–29 (2017).

113. Feldman, J. L., Baeza, J. & Denu, J. M. Activation of the Protein Deacetylase SIRT6 by Long-chain Fatty Acids and Widespread Deacylation by Mammalian Sirtuins ♦. *J Biol Chem* **288**, 31350–31356 (2013).

114. Kelly, R. D. W. *et al.* Histone deacetylase (HDAC) 1 and 2 complexes regulate both histone acetylation and crotonylation in vivo. *Sci Rep* **8**, 14690 (2018).

115. Klein, B. J. *et al.* Structural insights into the π-π-π stacking mechanism and DNA-binding activity of the YEATS domain. *Nat Commun* **9**, 4574 (2018).

116. Zhao, D. *et al.* YEATS2 is a selective histone crotonylation reader. *Cell Research* **26**, 629–632 (2016).

117. Feoli, A. *et al.* Lysine methyltransferase inhibitors: where we are now. *RSC Chemical Biology* **3**, 359–406 (2022).

118. Fallah, M. S., Szarics, D., Robson, C. M. & Eubanks, J. H. Impaired Regulation of Histone Methylation and Acetylation Underlies Specific Neurodevelopmental Disorders. *Frontiers in Genetics* **11**, (2021).

119. Klose, R. J. & Zhang, Y. Regulation of histone methylation by demethylimination and demethylation. *Nat Rev Mol Cell Biol* **8**, 307–318 (2007).

120. Jambhekar, A., Dhall, A. & Shi, Y. Roles and regulation of histone methylation in animal development. *Nat Rev Mol Cell Biol* **20**, 625–641 (2019).

121. Chen, Y. *et al.* The role of histone methylation in the development of digestive cancers: a potential direction for cancer management. *Sig Transduct Target Ther* **5**, 1–13 (2020).

122. Davies, R. J., Miller, R. & Coleman, N. Colorectal cancer screening: prospects for molecular stool analysis. *Nat Rev Cancer* **5**, 199–209 (2005).

123. McCleland, M. L. *et al.* Cdk8 deletion in the ApcMin murine tumour model represses EZH2 activity and accelerates tumourigenesis. *The Journal of Pathology* **237**, 508–519 (2015).

124. Yokoyama, Y. *et al.* Cancer-associated upregulation of histone H3 lysine 9 trimethylation promotes cell motility in vitro and drives tumor formation in vivo. *Cancer Sci* **104**, 889–895 (2013).

125. Huang, T. *et al.* Targeting histone methylation for colorectal cancer. *Therap Adv Gastroenterol* **10**, 114–131 (2017).

126. Olcina, M. M. *et al.* H3K9me3 facilitates hypoxia-induced p53-dependent apoptosis through repression of APAK. *Oncogene* **35**, 793–799 (2016).

127. Park, J.-A. *et al.* Deacetylation and Methylation at Histone H3 Lysine 9 (H3K9) Coordinate Chromosome Condensation during Cell Cycle Progression. *Mol Cells* **31**, 343–349 (2011).

128. Lutz, L. *et al.* Histone modifiers and marks define heterogeneous groups of colorectal carcinomas and affect responses to HDAC inhibitors in vitro. *Am J Cancer Res* **6**, 664–676 (2016).

129. Clapier, C. R., Iwasa, J., Cairns, B. R. & Peterson, C. L. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat Rev Mol Cell Biol* **18**, 407–422 (2017).

130. Petty, E. & Pillus, L. Balancing chromatin remodeling and histone modifications in transcription. *Trends Genet* **29**, 621–629 (2013).

131. Winston, F. & Carlson, M. Yeast SNF/SWI transcriptional activators and the SPT/SIN chromatin connection. *Trends in Genetics* **8**, 387–391 (1992).

132. Singh, R. K. *et al.* Transient Kinetic Analysis of SWR1C-Catalyzed H2A.Z Deposition Unravels the Impact of Nucleosome Dynamics and the Asymmetry of Histone Exchange. *Cell Rep* **27**, 374-386.e4 (2019).

133. Costelloe, T. *et al.* The yeast Fun30 and human SMARCAD1 chromatin remodellers promote DNA end resection. *Nature* **489**, 581–584 (2012).

134. Markert, J., Zhou, K. & Luger, K. SMARCAD1 is an ATP-dependent histone octamer exchange factor with de novo nucleosome assembly activity. *Sci Adv* **7**, eabk2380 (2021).

135. Rowbotham, S. P. *et al.* Maintenance of silent chromatin through replication requires SWI/SNF-like chromatin remodeler SMARCAD1. *Mol. Cell* **42**, 285–296 (2011).

136. Tong, Z.-B., Ai, H.-S. & Li, J.-B. The Mechanism of Chromatin Remodeler SMARCAD1/Fun30 in Response to DNA Damage. *Front Cell Dev Biol* **8**, 560098 (2020).

137. Taneja, N. *et al.* SNF2 Family Protein Fft3 Suppresses Nucleosome Turnover to Promote Epigenetic Inheritance and Proper Replication. *Mol Cell* **66**, 50-62.e6 (2017).

138. Doiguchi, M. *et al.* SMARCAD1 is an ATP-dependent stimulator of nucleosomal H2A acetylation via CBP, resulting in transcriptional regulation. *Sci Rep* **6**, (2016).

139. Nousbeck, J. *et al.* A Mutation in a Skin-Specific Isoform of SMARCAD1 Causes Autosomal-Dominant Adermatoglyphia. *Am J Hum Genet* **89**, 302–307 (2011).

140. Marks, K. C., Banks III, W. R., Cunningham, D., Witman, P. M. & Herman, G. E. Analysis of two candidate genes for Basan syndrome. *American Journal of Medical Genetics Part A* **164**, 1188–1191 (2014).

141. Kazakevych, J. *et al.* Smarcad1 mediates microbiota-induced inflammation in mouse and coordinates gene expression in the intestinal epithelium. *Genome Biology* **21**, 64 (2020).

142. Pellizzon, M. Choice of laboratory animal diet influences intestinal health. *Lab Animal* **45**, 238–239 (2016).

143. Machanick, P. & Bailey, T. L. MEME-ChIP: motif analysis of large DNA datasets. *Bioinformatics* **27**, 1696–1697 (2011).

144. Kulakovskiy, I. V. *et al.* HOCOMOCO: a comprehensive collection of human transcription factor binding sites models. *Nucleic Acids Res* **41**, D195–D202 (2013).

145. Kolmykov, S. *et al.* GTRD: an integrated view of transcription regulation. *Nucleic Acids Research* **49**, D104–D111 (2021).

146. Kim, N.-K., Tharakaraman, K. & Spouge, J. L. Adding sequence context to a Markov background model improves the identification of regulatory elements. *Bioinformatics* **22**, 2870–2875 (2006).

147. Kiesler, P., Fuss, I. J. & Strober, W. Experimental Models of Inflammatory Bowel Diseases. *Cellular and Molecular Gastroenterology and Hepatology* **1**, 154–170 (2015).

148. González Hernández, M. A., Canfora, E. E., Jocken, J. W. E. & Blaak, E. E. The Short-Chain Fatty Acid Acetate in Body Weight Control and Insulin Sensitivity. *Nutrients* **11**, 1943 (2019).

149. Banasiewicz, T., Domagalska, D., Borycka-Kiciak, K. & Rydzewska, G. Determination of butyric acid dosage based on clinical and experimental studies – a literature review. *Prz Gastroenterol* **15**, 119–125 (2020).

150. Guilloteau, P. *et al.* From the gut to the peripheral tissues: the multiple effects of butyrate. *Nutrition Research Reviews* **23**, 366–384 (2010).

151. Slavin, J. Fiber and Prebiotics: Mechanisms and Health Benefits. *Nutrients* **5**, 1417–1435 (2013).

152. Fung, K. Y. C., Cosgrove, L., Lockett, T., Head, R. & Topping, D. L. A review of the potential mechanisms for the lowering of colorectal oncogenesis by butyrate. *British Journal of Nutrition* **108**, 820–831 (2012).

153. Eichele, D. D. & Kharbanda, K. K. Dextran sodium sulfate colitis murine model: An indispensable tool for advancing our understanding of inflammatory bowel diseases pathogenesis. *World J Gastroenterol* **23**, 6016–6029 (2017).

154. Cui, J. *et al.* Improvement of magnesium isoglycyrrhizinate on DSS-induced acute and chronic colitis. *International Immunopharmacology* **90**, 107194 (2021).

155. Kim, J. J., Shajib, M. S., Manocha, M. M. & Khan, W. I. Investigating Intestinal Inflammation in DSS-induced Model of IBD. *JoVE (Journal of Visualized Experiments)* e3678 (2012) doi:10.3791/3678.

156. Gao, Z. *et al.* Butyrate Improves Insulin Sensitivity and Increases Energy Expenditure in Mice. *Diabetes* **58**, 1509–1517 (2009).

157. de Clercq, N. C., Groen, A. K., Romijn, J. A. & Nieuwdorp, M. Gut Microbiota in Obesity and Undernutrition. *Advances in Nutrition* **7**, 1080–1089 (2016).

158. den Besten, G. *et al.* Short-Chain Fatty Acids Protect Against High-Fat Diet–Induced Obesity via a PPARγ-Dependent Switch From Lipogenesis to Fat Oxidation. *Diabetes* **64**, 2398–2408 (2015).

159. Zeng, H. & Briske-Anderson, M. Prolonged Butyrate Treatment Inhibits the Migration and Invasion Potential of HT1080 Tumor Cells. *The Journal of Nutrition* **135**, 291–295 (2005).

160. Jiminez, J. A., Uwiera, T. C., Abbott, D. W., Uwiera, R. R. E. & Inglis, G. D. Butyrate Supplementation at High Concentrations Alters Enteric Bacterial Communities and Reduces Intestinal Inflammation in Mice Infected with Citrobacter rodentium. *mSphere* **2**, e00243-17.

161. Lynch, S. V. & Pedersen, O. The human intestinal microbiome in health and disease. *New England Journal of Medicine* **375**, 2369–2379 (2016).

162. Besten, G. den *et al.* The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. *Journal of Lipid Research* **54**, 2325–2340 (2013).

163. Takahashi, H., McCaffery, J. M., Irizarry, R. A. & Boeke, J. D. Nucleocytosolic acetyl-coenzyme A synthetase is required for histone acetylation and global transcription. *Molecular Cell* **23**, 207–217 (2006).

164. Friis, R. M. N. *et al.* A glycolytic burst drives glucose induction of global histone acetylation by picNuA4 and SAGA. *Nucleic Acids Research* **37**, 3969–3980 (2009).

165. Wellen, K. E. *et al.* ATP-citrate lyase links cellular metabolism to histone acetylation. *Science* **324**, 1076–1080 (2009).

166. Cai, L., Sutter, B. M., Li, B. & Tu, B. P. Acetyl-CoA induces cell growth and proliferation by promoting the acetylation of histones at growth genes. *Molecular Cell* **42**, 426–437 (2011).

167. Donohoe, D. R. *et al.* The warburg effect dictates the mechanism of butyrate-mediated histone acetylation and cell proliferation. *Molecular Cell* **48**, 612–626 (2012).

168. Bulusu, V. *et al.* Acetate recapturing by nuclear acetyl-CoA synthetase 2 prevents loss of histone acetylation during oxygen and serum limitation. *Cell Reports* **18**, 647–658 (2017).

169. Sabari, B. R. *et al.* Intracellular Crotonyl-CoA Stimulates Transcription through p300-Catalyzed Histone Crotonylation. *Molecular Cell* **58**, 203–215 (2015).

170. Fellows, R. *et al.* Microbiota derived short chain fatty acids promote histone crotonylation in the colon through histone deacetylases. *Nature Communications* **9**, 1–15 (2018).

171. González-Bosch, C., Boorman, E., Zunszain, P. A. & Mann, G. E. Short-chain fatty acids as modulators of redox signaling in health and disease. *Redox Biology* **47**, 102165 (2021).

172. Vinolo, M. A. R. *et al.* SCFAs Induce Mouse Neutrophil Chemotaxis through the GPR43 Receptor. *PLOS ONE* **6**, e21205 (2011).

173. Shaidullov, I. F. *et al.* Short chain fatty acids and colon motility in a mouse model of irritable bowel syndrome. *BMC Gastroenterology* **21**, 37 (2021).

174. Nakkarach, A. *et al.* Anti-cancer and anti-inflammatory effects elicited by short chain fatty acids produced by Escherichia coli isolated from healthy human gut microbiota. *Microbial Cell Factories* **20**, 36 (2021).

175. Koh, A., De Vadder, F., Kovatcheva-Datchary, P. & Bäckhed, F. From dietary fiber to host physiology: Short-chain fatty acids as key bacterial metabolites. *Cell* **165**, 1332–1345 (2016).

176. Corrêa-Oliveira, R., Fachi, J. L., Vieira, A., Sato, F. T. & Vinolo, M. A. R. Regulation of immune cell function by short-chain fatty acids. *Clinical & Translational Immunology* **5**, e73 (2016).

177. Wei, W. *et al.* Class I histone deacetylases are major histone decrotonylases: evidence for critical and broad function of histone crotonylation in transcription. *Cell Research* **27**, 898–915 (2017).

178. Kelly, R. D. W. *et al.* Histone deacetylase (HDAC) 1 and 2 complexes regulate both histone acetylation and crotonylation in vivo. *Scientific Reports* **8**, 1–10 (2018).

179. Madsen, A. S. & Olsen, C. A. Profiling of Substrates for Zinc-dependent Lysine Deacylase Enzymes : HDAC3 Exhibits Decrotonylase Activity In Vitro. *Angewandte Chemie - International Edition* **51**, 9083–9087 (2012).

180. Chen, Y. *et al.* Lysine propionylation and butyrylation are novel post-translational modifications in histones. *Molecular & Cellular Proteomics* **6**, 812–819 (2007).

181. Montgomery, D. C., Sorum, A. W. & Meier, J. L. Chemoproteomic profiling of lysine acetyltransferases highlights an expanded landscape of catalytic acetylation. *Journal of the American Chemical Society* **136**, 8669–8676 (2014).

182. Li, Y. *et al.* Molecular Coupling of Histone Crotonylation and Active Transcription by AF9 YEATS Domain. *Molecular Cell* **62**, 181–193 (2016).

183. Liu, X. *et al.* MOF as an evolutionarily conserved histone crotonyltransferase and transcriptional activation by histone acetyltransferase-deficient and crotonyltransferase-competent CBP/p300. *Cell Discovery* **3**, (2017).

184. Tan, M. *et al.* Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* **146**, 1016–1028 (2011).

185. Sabari, B. R. *et al.* Intracellular Crotonyl-CoA Stimulates Transcription through p300-Catalyzed Histone Crotonylation. *Molecular Cell* **58**, 203–215 (2015).

186. Flynn, E. M. *et al.* A subset of human bromodomains recognizes butyryllysine and crotonyllysine histone peptide modifications. *Structure* **23**, 1801–1814 (2015).

187. Andrews, F. H. *et al.* The Taf14 YEATS domain is a reader of histone crotonylation. *Nature chemical biology* 1–4 (2016) doi:10.1038/pj.2016.37.

188. Zhang, Q. *et al.* Structural Insights into Histone Crotonyl-Lysine Recognition by the AF9 YEATS Domain. *Structure* **24**, 1606–1612 (2016).

189. Zhao, D. *et al.* YEATS2 is a selective histone crotonylation reader. *Cell Research* **26**, 629–632 (2016).

190. Xiong, X. *et al.* Selective recognition of histone crotonylation by double PHD fingers of MOZ and DPF2. *Nature chemical biology* **12**, 1111–1118 (2016).

191. Wan, J., Liu, H. & Ming, L. Lysine crotonylation is involved in hepatocellular carcinoma progression. *Biomedicine and Pharmacotherapy* **111**, 976–982 (2019).

192. Abu-Zhayia, E. R., Machour, F. E. & Ayoub, N. HDAC-dependent decrease in histone crotonylation during DNA damage. *Journal of Molecular Cell Biology* **00**, 1–3 (2019).

193. Wang, Z. *et al.* NEAT1 regulates neuroglial cell mediating Aβ clearance via the epigenetic regulation of endocytosis-related genes expression. *Cellular and Molecular Life Sciences* (2019) doi:10.1007/s00018-019-03074-9.

194. Jiang, G. *et al.* HIV latency is reversed by ACSS2-driven histone crotonylation. *Journal of Clinical Investigation* **128**, 1190–1198 (2018).

195. Ruiz-Andres, O. *et al.* Histone lysine crotonylation during acute kidney injury in mice. *Disease Models & Mechanisms* **9**, 633–645 (2016).

196. Boffa, L. C. *et al.* Modulation of colonic epithelial cell proliferation, histone acetylation, and luminal short chain fatty acids by variation of dietary fiber (wheat bran) in rats. *Cancer research* **52**, 5906–12 (1992).

197. Krautkramer, K. A. *et al.* Diet-microbiota interactions mediate global epigenetic programming in multiple host tissues. *Molecular Cell* **64**, 982–992 (2016).

198. Dalile, B., Van Oudenhove, L., Vervliet, B. & Verbeke, K. The role of short-chain fatty acids in microbiota–gut–brain communication. *Nat Rev Gastroenterol Hepatol* **16**, 461–478 (2019).

199. Paschos, G. K. & FitzGerald, G. A. Circadian clocks and metabolism: implications for microbiome and aging. *Trends in Genetics* **33**, 760–769 (2017).

200. Claesson-Welsh, L. How the matrix metalloproteinase MMP14 contributes to the progression of colorectal cancer. *J Clin Invest* **130**, 1093–1095 (2020).

201. Li, M. *et al.* Immune Infiltration of MMP14 in Pan Cancer and Its Prognostic Effect on Tumors. *Frontiers in Oncology* **11**, (2021).

202. Bao, X. *et al.* Identification of 'erasers' for lysine crotonylated histone marks using a chemical proteomics approach. *eLIFE* 1–18 (2014) doi:10.7554/eLife.02999.

203. Schwarz, D. A., Katayama, C. D. & Hedrick, S. M. Schlafen, a new family of growth regulatory genes that affect thymocyte development. *Immunity* **9**, 657–668 (1998).

204. Glowacka, W. K., Alberts, P., Ouchida, R., Wang, J.-Y. & Rotin, D. LAPTM5 Protein Is a Positive Regulator of Proinflammatory Signaling Pathways in Macrophages. *J Biol Chem* **287**, 27691–27702 (2012).

205. Duan, J. *et al.* GALNT6 suppresses progression of colorectal cancer. *Am J Cancer Res* **8**, 2419–2435 (2018).

206. Albert, M. & Bennett, A. PWE-089 The roles of CYP2C40 and CYP2C55 in preventing colon cancer. *Gut* **61**, A333–A333 (2012).

207. Choi, J. H. *et al.* Essential requirement for nicastrin in marginal zone and B-1 B cell development. *Proc Natl Acad Sci U S A* **117**, 4894–4901 (2020).

208. Volodko, N., Gordon, M., Salla, M., Ghazaleh, H. A. & Baksh, S. RASSF tumor suppressor gene family: Biological functions and regulation. *FEBS Letters* **588**, 2671–2684 (2014).

209. Stemmler, S., Parwez, Q., Petrasch-Parwez, E., Epplen, J. T. & Hoffjan, S. Association of variation in the LAMA3 gene, encoding the alpha-chain of laminin 5, with atopic dermatitis in a German case–control cohort. *BMC Dermatology* **14**, 17 (2014).

210. Ryan, M. C., Lee, K., Miyashita, Y. & Carter, W. G. Targeted Disruption of the LAMA3 Gene in Mice Reveals Abnormalities in Survival and Late Stage Differentiation of Epithelial Cells. *J Cell Biol* **145**, 1309–1324 (1999).

211. Melendez, J. *et al.* Cdc42 Coordinates Proliferation, Polarity, Migration, and Differentiation of Small Intestinal Epithelial Cells in Mice. *Gastroenterology* **145**, 808–819 (2013).

212. Li, H., Zhao, L., Lau, Y. S., Zhang, C. & Han, R. Genome-wide CRISPR screen identifies LGALS2 as an oxidative stress-responsive gene with an inhibitory function on colon tumor growth. *Oncogene* **40**, 177–188 (2021).

213. Musa, J., Aynaud, M.-M., Mirabeau, O., Delattre, O. & Grünewald, T. G. MYBL2 (B-Myb): a central regulator of cell proliferation, cell survival and differentiation involved in tumorigenesis. *Cell Death Dis* **8**, e2895–e2895 (2017).

214. Kim, G. B., Gao, Y., Palsson, B. O. & Lee, S. Y. DeepTFactor: A deep learning-based tool for the prediction of transcription factors. *Proceedings of the National Academy of Sciences* **118**, e2021171118 (2021).

215. Deng, Y.-N., Xia, Z., Zhang, P., Ejaz, S. & Liang, S. Transcription Factor RREB1: from Target Genes towards Biological Functions. *Int J Biol Sci* **16**, 1463–1473 (2020).

216. Bayrer, J. R. *et al.* LRH-1 mitigates intestinal inflammatory disease by maintaining epithelial homeostasis and cell survival. *Nat Commun* **9**, 4055 (2018).

217. Seitz, C. *et al.* The orphan nuclear receptor LRH-1/NR5a2 critically regulates T cell functions. *Science Advances* **5**, eaav9732 (2019).

218. Fernandez-Marcos, P. J., Auwerx, J. & Schoonjans, K. Emerging actions of the nuclear receptor LRH-1 in the gut. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **1812**, 947–955 (2011).

219. Sutherland, E. W. & Rall, T. W. Fractionation and characterization of a cyclic adenine ribonucleotide formed by tissue particles. *Journal of Biological Chemistry* **232**, 1077–1091 (1958).

220. Sassone-Corsi, P. The Cyclic AMP Pathway. *Cold Spring Harbor Perspectives in Biology* **4**, a011148–a011148 (2012).

221. Laissue, P. The forkhead-box family of transcription factors: key molecular players in colorectal cancer pathogenesis. *Mol Cancer* **18**, (2019).

222. Zakrzewska, A. *et al.* Macrophage-specific gene functions in Spi1-directed innate immunity. *Blood* **116**, e1–e11 (2010).

223. Kanaya, T. *et al.* The Ets Transcription Factor Spi-B Is Essential for the Differentiation of Intestinal Microfold (M) Cells. *Nat Immunol* **13**, 729–736 (2012).

224. Yanai, H. *et al.* Revisiting the role of IRF3 in inflammation and immunity by conditional and specifically targeted gene ablation in mice. *PNAS* **115**, 5253–5258 (2018).

225. Múnera, J., Ceceña, G., Jedlicka, P., Wankell, M. & Oshima, R. G. Ets2 regulates colonic stem cells and sensitivity to tumorigenesis. *Stem Cells* **29**, 430–439 (2011).

226. Tang, Q. *et al.* Role of far upstream element binding protein 1 in colonic epithelial disruption during dextran sulphate sodium-induced murine colitis. *Int J Clin Exp Pathol* **7**, 2019–2031 (2014).

227. Law, D. J., Labut, E. M. & Merchant, J. L. Intestinal overexpression of ZNF148 suppresses ApcMin/+ neoplasia. *Mamm Genome* **17**, 999–1004 (2006).

228. Prathyusha, A. M. V. N., Nawadkar, R. & Bramhachari, P. V. Role of Sp1 Transcriptional Factor in Gastrointestinal Carcinogenesis. in *Role of Transcription Factors in Gastrointestinal Malignancies* (eds. Nagaraju, G. P. & Bramhachari, P. V.) 193–204 (Springer, 2017). doi:10.1007/978-981-10-6728-0_13.

229. Takahashi, M., Nakamura, Y., Obama, K. & Furukawa, Y. Identification of SP5 as a downstream gene of the beta-catenin/Tcf pathway and its enhanced expression in human colon cancer. *Int J Oncol* **27**, 1483–1487 (2005).

230. Miyamoto, M., Hayashi, T., Kawasaki, Y. & Akiyama, T. Sp5 negatively regulates the proliferation of HCT116 cells by upregulating the transcription of p27. *Oncol Lett* **15**, 4005–4009 (2018).

231. Luk, I. Y., Reehorst, C. M. & Mariadason, J. M. ELF3, ELF5, EHF and SPDEF Transcription Factors in Tissue Homeostasis and Cancer. *Molecules* **23**, 2191 (2018).

232. Taniue, K., Oda, T., Hayashi, T., Okuno, M. & Akiyama, T. A member of the ETS family, EHF, and the ATPase RUVBL1 inhibit p53-mediated apoptosis. *EMBO Rep* **12**, 682–689 (2011).

233. Krishnan, M. & McCole, D. F. T cell protein tyrosine phosphatase prevents STAT1 induction of claudin-2 expression in intestinal epithelial cells. *Ann N Y Acad Sci* **1405**, 116–130 (2017).

234. Chiriac, M. T. *et al.* Activation of Epithelial Signal Transducer and Activator of Transcription 1 by Interleukin 28 Controls Mucosal Healing in Mice With Colitis and Is Increased in Mucosa of Patients With Inflammatory Bowel Disease. *Gastroenterology* **153**, 123-138.e8 (2017).

235. Giles, E. M. *et al.* Regulation of human intestinal T-cell responses by type 1 interferon-STAT1 signaling is disrupted in inflammatory bowel disease. *Mucosal Immunol* **10**, 184–193 (2017).

236. Morita, K. *et al.* Emerging roles of Egr2 and Egr3 in the control of systemic autoimmunity. *Rheumatology* **55**, ii76–ii81 (2016).

237. Okamura, T., Yamamoto, K. & Fujio, K. Early Growth Response Gene 2-Expressing CD4+LAG3+ Regulatory T Cells: The Therapeutic Potential for Treating Autoimmune Diseases. *Front. Immunol.* **9**, (2018).

238. Li, M. *et al.* Dynamic regulation of transcription factors by nucleosome remodeling. *eLife* **4**, e06249 (2015).

239. Ho, L. & Crabtree, G. R. Chromatin remodelling during development. *Nature* **463**, 474–484 (2010).

240. Sachs, P. *et al.* SMARCAD1 ATPase activity is required to silence endogenous retroviruses in embryonic stem cells. *Nat Commun* **10**, 1335 (2019).

241. Yu, Q., Zhang, X. & Bi, X. Roles of Chromatin Remodeling Factors in the Formation and Maintenance of Heterochromatin Structure. *J Biol Chem* **286**, 14659–14669 (2011).

242. Xiao, S. *et al.* SMARCAD1 contributes to regulation of naïve pluripotency by interacting with histone citrullination. *Cell Rep* **18**, 3117–3128 (2017).

243. Feng, J., Liu, T. & Zhang, Y. Using MACS to Identify Peaks from ChIP-Seq Data. *Curr Protoc Bioinformatics* **CHAPTER**, Unit2.14 (2011).

244. Uno, Y., Kawakami, S., Ochiai, K. & Omi, T. Molecular characterization of cytidine monophospho-N-acetylneuraminic acid hydroxylase (CMAH) associated with the erythrocyte antigens in dogs. *Canine Genetics and Epidemiology* **6**, 9 (2019).

245. Tangvoranuntakul, P. *et al.* Human uptake and incorporation of an immunogenic nonhuman dietary sialic acid. *Proceedings of the National Academy of Sciences* **100**, 12045–12050 (2003).

246. Rutishauser, U., Acheson, A., Hall, A. K., Mann, D. M. & Sunshine, J. The Neural Cell Adhesion Molecule (NCAM) as a Regulator of Cell-Cell Interactions. *Science* **240**, 53–57 (1988).

247. Okerblom, J. & Varki, A. Biochemical, Cellular, Physiological, and Pathological Consequences of Human Loss of N-Glycolylneuraminic Acid. *ChemBioChem* **18**, 1155–1171 (2017).

248. Raudenska, M., Balvan, J. & Masarik, M. Crosstalk between autophagy inhibitors and endosome-related secretory pathways: a challenge for autophagy-based treatment of solid cancers. *Molecular Cancer* **20**, 140 (2021).

249. Knævelsrud, H. *et al.* Membrane remodeling by the PX-BAR protein SNX18 promotes autophagosome formation. *Journal of Cell Biology* **202**, 331–349 (2013).

250. Søreng, K. *et al.* SNX18 regulates ATG9A trafficking from recycling endosomes by recruiting Dynamin-2. *EMBO reports* **19**, e44837 (2018).

251. Zhang, P., Holowatyj, A. N., Ulrich, C. M. & Edgar, B. A. Tumor suppressive autophagy in intestinal stem cells controls gut homeostasis. *Autophagy* **15**, 1668–1670 (2019).

252. Hu, B., Yin, G. & Sun, X. Identification of specific role of SNX family in gastric cancer prognosis evaluation. *Sci Rep* **12**, 10231 (2022).

253. Shahi, P. *et al.* The Transcriptional Repressor ZNF503/Zeppo2 Promotes Mammary Epithelial Cell Proliferation and Enhances Cell Invasion. *J Biol Chem* **290**, 3803–3813 (2015).

254. Zheng, Y. *et al.* ZNF503 promotes colon cancer cell proliferation and mediates tumor progression via activating Myc. *International Journal of Clinical and Experimental Pathology* **10**, 5306–5315 (2017).

255. van der Velden, V. H. Glucocorticoids: mechanisms of action and anti-inflammatory potential in asthma. *Mediators Inflamm* **7**, 229–237 (1998).

256. Water, N. R. C. (US) C. on C. in D. *Physiological Role of Copper*. (National Academies Press (US), 2000).

257. Prasad, A. S. Zinc in Human Health: Effect of Zinc on Immune Cells. *Mol Med* **14**, 353–357 (2008).

258. Roohani, N., Hurrell, R., Kelishadi, R. & Schulin, R. Zinc and its importance for human health: An integrative review. *J Res Med Sci* **18**, 144–157 (2013).

259. Kim, Y. R. *et al.* Frameshift mutation of MAPRE3, a microtubule-related gene, in gastric and colorectal cancers with microsatellite instability. *Pathology* **42**, 493–496 (2010).

260. Lesko, A. C. *et al.* The APC tumor suppressor is required for epithelial cell polarization and three-dimensional morphogenesis. *Biochim Biophys Acta* **1854**, 711–723 (2015).

261. Le, C. C. *et al.* LRP-1 Promotes Colon Cancer Cell Proliferation in 3D Collagen Matrices by Mediating DDR1 Endocytosis. *Frontiers in Cell and Developmental Biology* **8**, (2020).

262. Boulagnon-Rombi, C. *et al.* LRP1 expression in colon cancer predicts clinical outcome. *Oncotarget* **9**, 8849–8869 (2018).

263. Cai, J. *et al.* ABIN-1 is a key regulator in RIPK1-dependent apoptosis (RDA) and necroptosis, and ABIN-1 deficiency potentiates necroptosis-based cancer therapy in colorectal cancer. *Cell Death Dis* **12**, 1–15 (2021).

264. Cui, X. *et al.* Ring finger protein 152 inhibits colorectal cancer cell growth and is a novel prognostic biomarker. *Am J Transl Res* **10**, 3701–3712 (2018).

265. Lambert, D. W., Wood, I. S., Ellis, A. & Shirazi-Beechey, S. P. Molecular changes in the expression of human colonic nutrient transporters during the transition from normality to malignancy. *Br J Cancer* **86**, 1262–1269 (2002).

266. Tibbs, Z. E., Guidry, A. L., Falany, J. L., Kadlubar, S. A. & Falany, C. N. A High Frequency Missense SULT1B1 Allelic Variant (L145V) Selectively Expressed in African Descendants Exhibits Altered Kinetic Properties. *Xenobiotica* **48**, 79–88 (2018).

267. Howie, D., Garcia Rueda, H., Brown, M. H. & Waldmann, H. Secreted and Transmembrane 1A Is a Novel Co-Stimulatory Ligand. *PLoS One* **8**, e73610 (2013).

268. Thomas, P., Pranatharthi, A., Ross, C. & Srivastava, S. RhoC: a fascinating journey from a cytoskeletal organizer to a Cancer stem cell therapeutic target. *Journal of Experimental & Clinical Cancer Research* **38**, 328 (2019).

269. Wang, H. *et al.* Silencing of RhoA and RhoC expression by RNA interference suppresses human colorectal carcinoma growth in vivo. *J Exp Clin Cancer Res* **29**, 123 (2010).

270. Liu, L., Wu, D.-H. & Ding, Y.-Q. Tiam1 gene expression and its significance in colorectal carcinoma. *World J Gastroenterol* **11**, 705–707 (2005).

271. Ding, J., Yang, F. & Wu, W. Tiam1 high expression is associated with poor prognosis in solid cancers: A meta-analysis. *Medicine* **98**, e17529 (2019).

272. Choi, N.-W. Kidney and Phosphate Metabolism. *Electrolyte Blood Press* **6**, 77–85 (2008).

273. Tsukuda, N. *et al.* Key bacterial taxa and metabolic pathways affecting gut short-chain fatty acid profiles in early life. *ISME J* **15**, 2574–2590 (2021).

274. Cai, L., Sutter, B. M., Li, B. & Tu, B. P. Acetyl-CoA induces cell growth and proliferation by promoting the acetylation of histones at growth genes. *Mol Cell* **42**, 426–437 (2011).

275. Gowans, G. J. *et al.* Recognition of Histone Crotonylation by Taf14 Links Metabolic State to Gene Expression. *Mol Cell* **76**, 909-921.e3 (2019).

8. Appendices

8.1 H3 Normalisation calculation of germfree western blot

Table 8.1 – H3 Normalisation calculation of germfree western blot

To calculate the volume to load 4µg of sample into the gel.

	sample	pixels	sqrt	average	%		µl
H3 Normalisation	4000	5717.761	75.62		120.74	79.26	3.17
	4001	1341.012	36.62		58.47	141.53	5.66
	4002	1592.962	39.91		63.73	136.27	5.45
	4003	3042.447	55.16		88.08	111.92	4.48
	4004	437.92	20.93		33.41	166.59	6.66
	4005	636.991	25.24	62.63	40.30	159.70	6.39
	2431	5648.225	75.15		120.01	79.99	3.20
	2432	11470.49	107.10		171.02	28.98	1.16
	2441	11677.49	108.06		172.55	27.45	1.10
	2442	9623.66	98.10		156.64	43.36	1.73
	2443	2208.912	47.00		75.05	124.95	5.00

8.2 Step wise protocol for harvesting colon samples and ChIP

2018-12-13 harvesting Co for H3K9me3 H3K9ac H3K9cr ChIP test w/o FACS
VF(WT) and VC(KO)

input adult 6x **Bl6 male** (VF 2x2, VC 2x2)

O **prep DNAse aliquots (Qiagen)**

O **add NaBut to buffers (fresh for the day)**

work in RT PBS if not otherwise indicated, use DNA-low bind tubes

O switch on 37°C water bath
O open mouse, extract all small intestine and colon (cut off fat, caecum)
O cut open longitudinally (works best dry)
O wash in petri dish 1-2 times
O cut ~5mm pieces to a 50 ml falcon with 25 ml PBS, **pool 3 colons per tube**
OOO wash 3x in 25 ml PBS-EDTA (5mM), 10x very gentle inversions in between
O incubate in PBS-EDTA 30' rocking at 20 rpm **add Butyrate to PBS-EDTA (everything must be done under 1 hour)**

O wash tissue 1x
O add fresh25 ml PBS + EDTA
O shake vigorously 5' -> Co-crypts
O filter 1st fraction Co-crypts SN through cell strainer 70 µm

O centrifuge 400 g 5'
O collect to 2-ml eppis **(2 ml PBS per 3 colons, 1ml for 1 colon)**
O wash 1x in PBS-EDTA to get rid of enzymes, remove SN

Fixation:

O resuspend in 737.5 ml PBS-EDTA (5mM), add 62.5 µl 16% FA (1%final) (double for 3 colons)
O Fix 10' RT with low agitation
O quench with 1/10 vol. 1.25 M glycine 5' RT with low agitation
OOO wash 3x with 1 ml PBS at 4°C, spin 2400g, 1min
O spin down at 5000g 2', remove SN
O resuspend in ~0.05 ml PBS for following sonication, 9x0.05 ml

No sorting this time

(O) saving point, freezable at -80 °C/-20 °C short term (proceed with 4xCo directly)

Prepare antibody-bead complexes:

Prepare this at the same time as preparing the chromatin to allow for 2-4 h incubation so that bead-antibody mix is ready when chromatin has been sonicated, alternatively in parallel with sonication check

OOO Wash 80 µl of beads thrice (50/50 A+G Dynabeads – 40 µl of each) with 3000 µl of RIPA buffer

Pipette to mix, place on magnetic rack until solution becomes clear and then remove supernatant

O Divide beads into 4x 1.5 ml low bind tubes, add antibodies:

- 20 µl AG beads + 5 µl (5 µg) Ab H3K9me3 ab8898 1mg/ml M
- 20 µl AG beads + 5 µl (5 µg) Ab H3K9ac Millipore 06-942 1mg/ml? A
- 20 µl AG beads + 6 µl (5 µg) Ab H3K9cr PTM516 C
- 20 µl AG beads + 0.5 µl (5 µg) rabbit IgG Sigma 1-5006 10 mg/ml (normal IgG box +4°C) I

O Incubate for 2-4 h at 4°C on an rotating wheel in 500 µl RIPA buffer

OOO Wash beads-antibody mix thrice with 1 ml of RIPA buffer

As above, pipette to mix, place on magnetic rack until solution is clear and then remove supernatant

Sonicate chromatin: 4x Co

O Add **360 µl** room temperature lysis buffer + PIC/PMSF to pellet, mix by pipette or vortex 2x5''

O Leave on ice for 5 minutes and resuspend cells by vortexing

O Sonicate 45x 10''/80'' ON/OFF high

O Take 30 µl to check sonication and chromatin amount (2% of single ChIP material or 20k cells)

O run sonication check before proceeding to have input amounts (kit isolation, Nanodrop, gel)

O Add **1200** µl **RIPA ChIP** buffer (to approx. 210 µl in tube) and mix by vortexing

O Centrifuge at 16,200 g for 10' at 4 °C, then carefully aspirate 700 µl SN and transfer it into a clean 1.5 ml tube chilled on ice (leave approximately 100 µl supernatant in the tube).

O Add **500** µl **RIPA ChIP** buffer to the remaining pellet, vortex, centrifuge 16,200g for 5' at 4°C

O Remove 200 µl SN, leaving approximately 60 µl with the (invisible) pellet, discard the pellet.

(O) Pool all six supernatants together i.e. those from the first and second spin from the three tubes

O wait for sonication check here, evtl. Adjust input/ChIP quantities

O Take 1/64 of single ChIP vol (_ µl) chromatin and put in a 1.5 ml tube on ice->INPUT (1:64)

(O) check residual volume is > 3*800 µl

O Remove buffer from antibody-bead complexes, resuspend with ~800 µl chromatin solution

Check sonication:

Optional: treat with 1uL RNase (bringing up to 50 µl with Elution buffer) for 30min at 37°C

O Bring **up to** 4-fold volume with Elution buffer - **120µl total**

O Add 1ul of Proteinase K (10mg/ml or 20 mg/ml)

O Incubate 1 h at 62-65°C with low agitation (300rpm)

O purifiy with a PCR-purification kit, e.g. Qiagen or NEB

25 µl eluted in 50 µl (__% of ChIP input/__ µl -> ___% per µl (1:___)

O Load at least 450-500 ng of sonicated material on a 1-2% agarose gel. (15 µl)

optimal size range is between 200 and 500 bp for ChIP-seq, 300-700 for ChIP-qPCR

(O) Nanodrop for approx. quantity to normalize samples for ChIP

Immunoprecipitation:

O Incubate antibody, beads and chromatin mix on a rotator at 4 °C overnight

Day 2:

OO.O Spin tubes briefly and place them on magnetic rack until solution becomes clear
OO.O Discard the supernatant and add 1 ml ice-cold **RIPA low salt** buffer.
OO.O Resuspend the bead complexes by gentle manual agitation and place on a rotator for 4' at 4°C
Repeat the previous three steps once more and then once again with **RIPA high salt** buffer **(avoid the High Salt step in future for the H3K9ac ChIP)**
O Remove supernatant and add 1 ml TE buffer
O Incubate on a rotator at 4 °C for 4 min, spin briefly and place on magnetic rack
O Remove the TE buffer

Isolation of ChIP DNA and Input:
O Add 200 µl elution buffer (without SDS)
O incubate for 20' at 65 °C on the thermomixer at 1300 rpm
O add 2 µl (6 µl for input) RNase solution (20mg/ml) and 20 µl of 10% SDS
O Incubate 30' 37°C on the thermomixer
O Add 2 µl (4 µl for input) of proteinase K (20 mg/ml)
O incubate 2 h 65 °C, 1300 rpm mixing
O Centrifuge 1 min at 11,600 g and save the SN (~200 µl)
O Add 150 µl elution buffer (without SDS) to the remaining ChIP material
O incubate on the thermomixer 5' 65 °C, 1300 rpm
O Centrifuge 2' 11,600 g and pool SN (~170 µl)
O Add __ µl elution buffer (without SDS) to a final volume of 500 µl

Chloroform/Phenol DNA extraction:
O In fume hood, add an equal volume of phenol-chloroform-isoamylalcohol (500 µl) Pipette up and down several times to allow the vapor to equalize
O Vortex well and centrifuge at 15,000 g for 2 minutes
O Transfer 490 µl of the aqueous (top) phase to a clean tube. Tilt the tube to remove easier, it is ok to leave a small amount in the tube, do not to transfer any of the lower phase into fresh tube
O Add an equal volume of chloroform (500 µl) Chloroform is stored with a layer of water on top, avoid taking it. Pipette up and down several times to allow the vapor to equalize.
O Vortex well and centrifuge at 15000 g for 1-2 minutes
O Transfer 480 µl of the aqueous (top) phase to a clean tube 2 ml tube

DNA Precipitation:
O Add 3 vol. 100% EtOH (1.5 ml), 1/10 volume of 3M KOAc pH 5.5 (50 µl) and 1 µl glycogen.
O Mix thoroughly and incubate for at least 1 hour at -80 °C or 90' at -20 °C
O Thaw the tubes, vortex briefly and centrifuge at maximum speed (20817g) for 30 min at 4 °C
Make sure that the content is liquid; it might need a minute to thaw
O Remove SN, add 1 ml of 70 % ethanol at -20 °C and vortex briefly to wash the DNA pellet
O Centrifuge max speed for 10 minutes at 4 °C
O Remove the supernatant and dry tubes open (Takes 5-20 minutes, 37°C block can be used)
O Dissolve the DNA in 22 µl H_2O
O quantitate input by Nanodrop, qubit and/or bioanalyzer to assure quantity and size
O Freeze at -20°C
O qPCR

Reagents and solutions:

Lysis-buffer (at RT):	**405 ml H_20+**	
50 mM Tris-HCl, pH 8.0	25 ml	1M
10 mM EDTA	10 ml	0.5 M
1 % SDS (wt/vol)	50 ml	10%
Fresh 1 mM PMSF (100x stock)		
1x PIC (1 tablet in 0.5ml = 100x)		
5 mM Na-butyrate	fresh	

RIPA-LOW SALT buffer (at 4 °C):	**405 ml H_20+**	
10 mM Tris-HCl, pH 7.5	5 ml	1M
140 mM NaCl	14 ml	5M
1 mM EDTA	1 ml	0.5 M
0.5 mM EGTA	0.5 ml	0.5 M
1 % (vol/vol) Triton X-100	25 ml	20%
0.1 % (wt/vol) SDS	5 ml	10%
0.1 % Na-deoxycholate	0.5 g	
5 mM Na-butyrate	fresh	

RIPA ChIP-buffer at 4 °C:

RIPA-buffer +
fresh 1 mM PMSF and 1x PIC

RIPA-HIGH SALT buffer (at 4 °C):

Add 14.4 ml 5M NaCl to 185.6 ml RIPA LOW SALT for 500 mM final
5 mM Na-butyrate fresh

TE-buffer at 4 °C:

10 mM Tris-HCl, pH 8.0
10 mM EDTA

Elution-buffer at 4 °C:	**94 ml H2O+**	
20 mM Tris-HCl, pH 7.5	2 ml	1 M
5 mM EDTA	1 ml	0.5 M
50 mM NaCl	1 ml	5 M

Complete elution buffer:

Elution-buffer +
1 % (wt/vol) SDS

8.3 ChIP PCR scheme

samples		VC4 A	VC5 A	VC6 A	VC4 M	VC5 M	VC6 M	VC4 C	VC5 C	VC6 C	VC4 Input	VC5 Input	VC6 Input	VF4 A	VF5 A	VF6 A	VF4 M	VF5 M	VF6 M	VF4 C	VF5 C	VF6 C	VF4 Input	VF5 Input	VF6 Input
amount		15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15
trip/dup		2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
volume		10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
Primervolume (3µM premixed)		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
security factor		1.10	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1
Mastermixvolume		297	297	297	297	297	297	297	297	297	297	297	297	297	297	297	297	297	297	297	297	297	297	297	297
each		9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9

		VC4 A	VC5 A	VC6 A	VC4 M	VC5 M	VC6 M	VC4 C	VC5 C	VC6 C	VC4 Input	VC5 Input	VC6 Input	VF4 A	VF5 A	VF6 A	VF4 M	VF5 M	VF6 M	VF4 C	VF5 C	VF6 C	VF4 Input	VF5 Input	VF6 Input	
(SYBR green)	Buffer 2x	165	165	165	165	165	165	165	165	165	165	165	165	165	165	165	165	165	165	165	165	165	165	165	165	
	H20	124	124	124	124	124	124	124	124	124	124	124	124	124	124	124	124	124	124	124	124	124	124	124	124	
	DNA	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8 (64 x less input then ChIP material)

pippetting scheme		VC4 A	VC4 A	VC5 A	VC5 A	VC6 A	VC6 A	VC4 M	VC4 M	VC5 M	VC5 M	VC6 M	VC6 M	VC4 C	VC4 C	VC5 C	VC5 C	VC6 C	VC6 C	VC4 Input	VC4 Input	VC5 Input	VC5 Input	VC6 Input	VC6 Input
129/30 UpS Bglap	A																								
131/2 Aplp1/Kirrel2	B																								
133/4 Bambi	C																								
135/6 Mbd1	D																								
137/8 Tiam1 UpS	E																								
139/40 Sema7a	F																								
141/2 Lyz1	G																								
143/4 PcdH17	H																								
145/6 Dlk1	I																								
147/8 Osr2	J																								
151/2 ZEB1	K																								
153/4 Snx18	L																								
155/6 Sall3	M																								
Klf11	N																								
93/94	O																								
	P																								
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24

Program:

95 °C	10'	
95 °C	15" 45x	
62 °C	1' 45x	detect
MELTCURVE		detect

use 386-well plate white

NTC control to determine primer dimers

pippetting scheme		VF4 A	VF4 A	VF5 A	VF5 A	VF6 A	VF6 A	VF4 M	VF4 M	VF5 M	VF5 M	VF6 M	VF6 M	VF4 C	VF4 C	VF5 C	VF5 C	VF6 C	VF6 C	VF4 Input	VF4 Input	VF5 Input	VF5 Input	VF6 Input	VF6 Input
129/30 UpS Bglap	A																								
131/2 Aplp1/Kirrel2	B																								
133/4 Bambi	C																								
135/6 Mbd1	D																								
137/8 Tiam1 UpS	E																								
139/40 Sema7a	F																								
141/2 Lyz1	G																								
143/4 PcdH17	H																								
145/6 Dlk1	I																								
147/8 Osr2	J																								
151/2 ZEB1	K																								
153/4 Snx18	L																								
155/6 Sall3	M																								
Klf11	N																								
93/94	O																								
	P																								
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24

8.4 Smarcad1 ChIPseq dataset

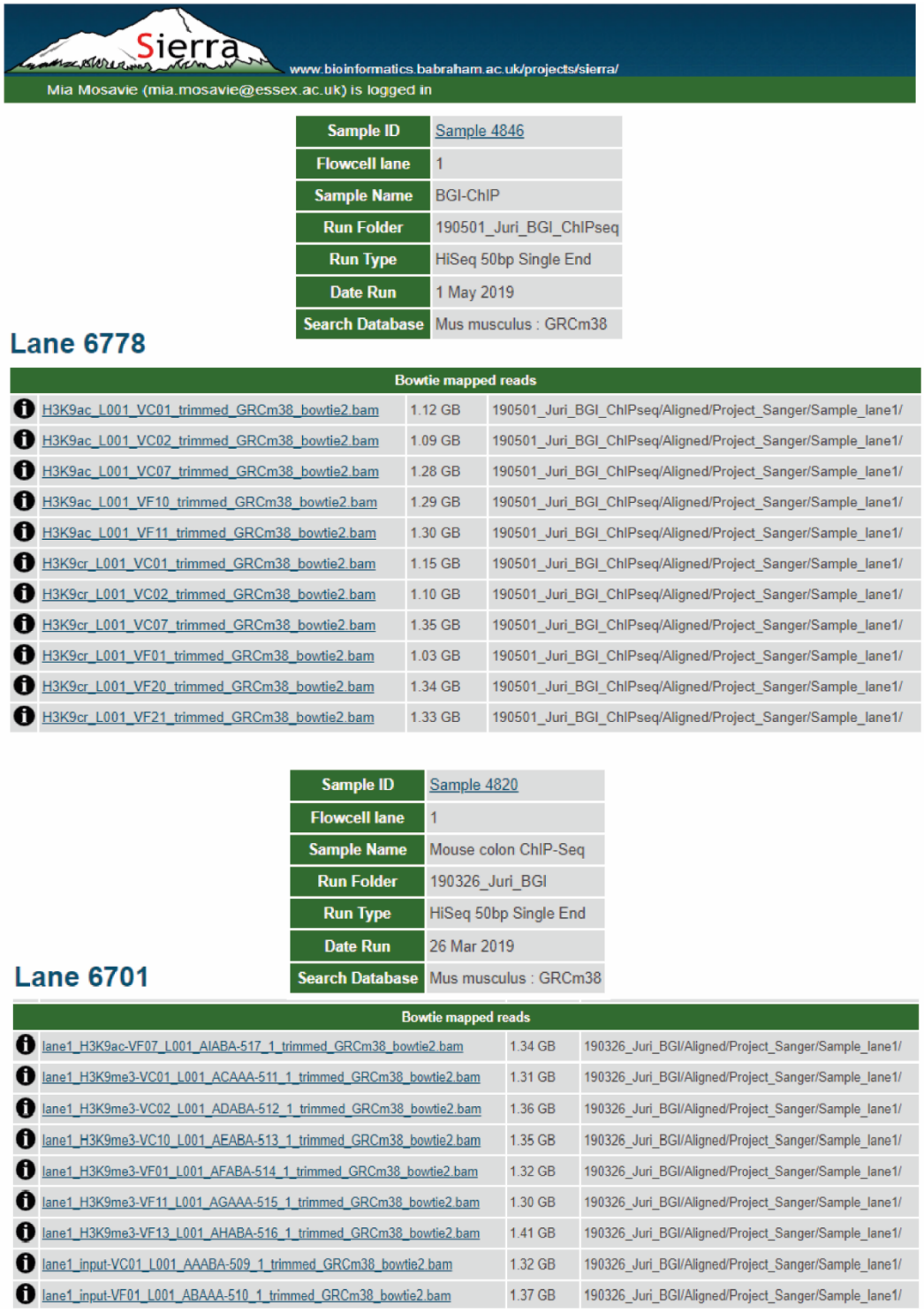


Figure 8.4.1 Smarcad1 ChIPseq dataset uploaded into Sierra. *Runs were split into two groups. H3K9ac and H3K9cr are in one group (Lane 6778). The remaining H3K9ac and H3K9me3 are in the second group (Lane 6701) Samples labelled with VF are WT samples, VC are KO samples.*

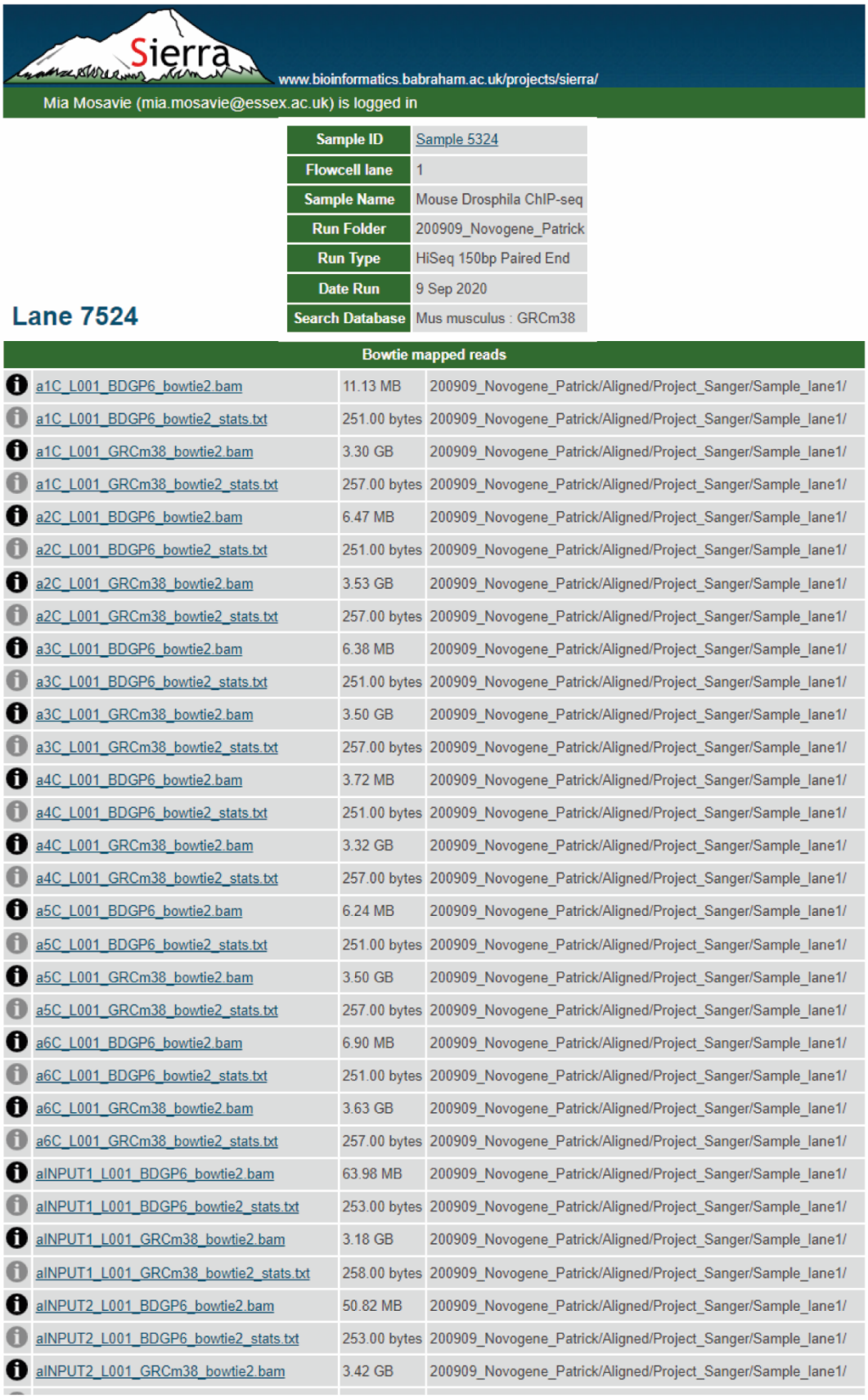


Figure 8.4.2 Antibiotic treated H3K18ac ChIPseq dataset uploaded into Sierra. *Both mouse and drosophila datasets uploaded. A1c, a2c and a3c are control samples CT4, CT5 and CT6 respectively. A4c, a5c and a6c are antibiotic treated samples AB4, AB5 and AB6 respectively.*

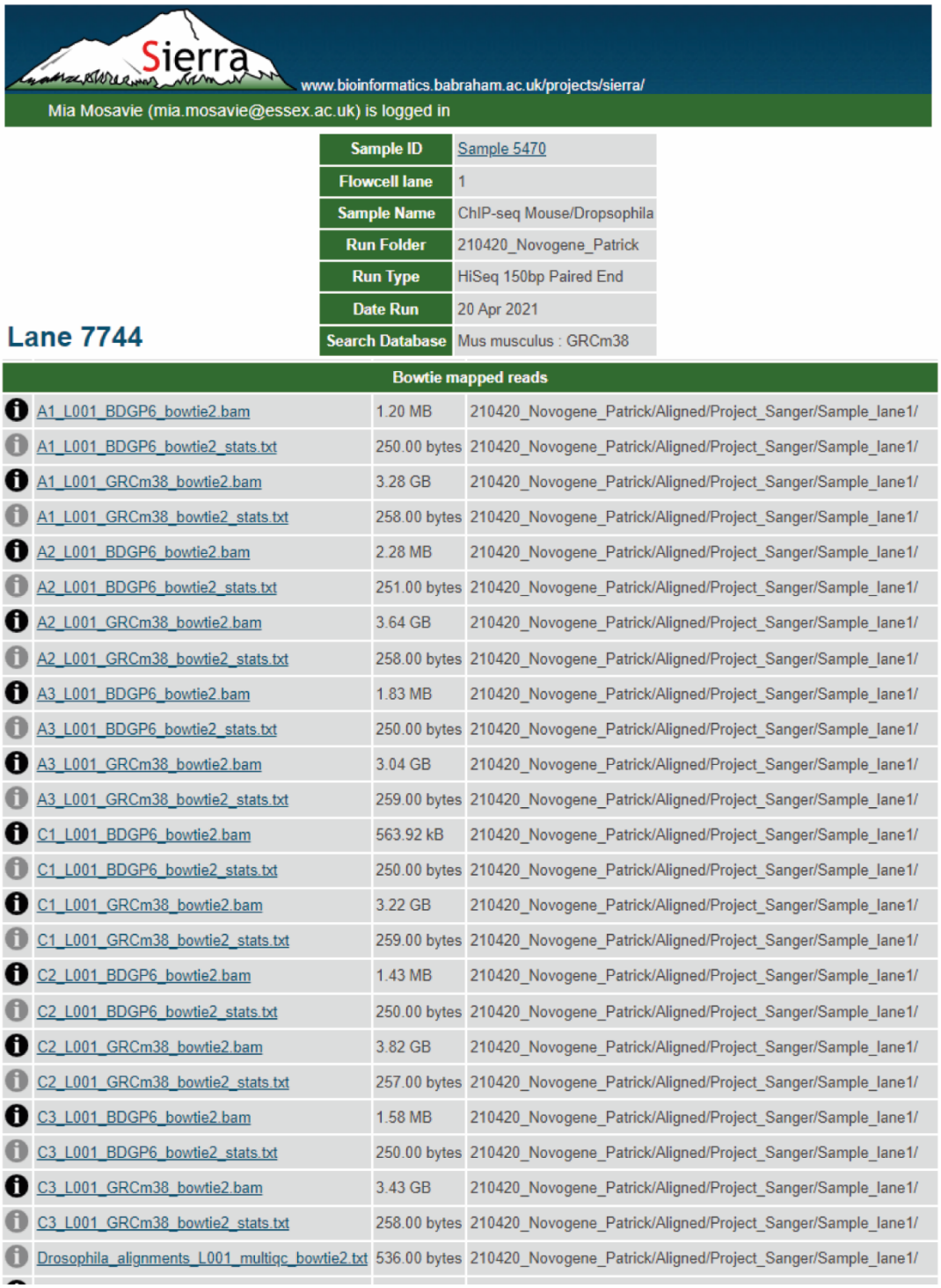


Figure 8.4.3 Antibiotic treated H3K18ac ChIPseq dataset uploaded into Sierra. *Both mouse and drosophila datasets uploaded. C1, C2 and C3 are control samples CT1, CT2 and CT3 respectively. A1, A2 and A3 are antibiotic treated samples AB1, AB2 and AB3 respectively.*