

ASSESSMENT OF HISTONE TAIL MODIFICATIONS AND TRANSCRIPTIONAL
PROFILING DURING COLON CANCER PROGRESSION: EFFECT OF
CHEMOPROTECTIVE NATURAL COMPOUNDS

A Dissertation

by

KAREN TRIFF

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Robert S. Chapkin
Co-Chair of Committee	Mark Zoran
Committee Members,	Richard Gomer
	Joanne Lupton
Head of Department,	Tom McKnight

May 2016

Major Subject: Microbiology

Copyright 2016 Karen Triff

ABSTRACT

During colon cancer development, epigenetic alterations contribute to the dysregulation of major cellular functions and signaling pathways. Recent evidence suggests that nutritionally chemoprotective components that influence cellular dynamics in the colonic epithelium can also directly affect their epigenetic landscape. We hypothesize that the chemoprotective nutritional bioactives found in fish oil and fermentable fiber can act as epigenetic modifiers and mechanistically counteract epigenetic distortions associated with colonic tumorigenesis. Fermentable fiber generates short-chain fatty acids (SCFA), e.g., butyrate, in the lumen of the colon that can serve as a chemoprotective histone deacetylase inhibitor and/or as an acetylation substrate for histone acetylases. In addition, n-3 polyunsaturated fatty acids (n-3 PUFAs) in fish oil can affect the chromatin landscape by acting as ligands for tumor suppressive nuclear receptors.

In an effort to gain insight into the extensive dimension of post-translational modifications in histones (including H3K4me3 and H3K9ac) and elucidate the chemoprotective impact of dietary bioactive compounds on transcriptional control in a colon cancer preclinical model, we generated high-resolution genome-wide RNA (RNA-Seq) and “chromatin-state” (H3K4me3-seq and H3K9ac-seq) maps for intestinal (epithelial colonocytes) crypts in rats treated with a colon carcinogen and fed bioactive (i) fish oil (ii) butyrate (in the form of a fermentable fiber a rich source of SCFA), (iii) a combination of fish oil plus butyrate or (iv) control diets.

Poor correlation was observed between differentially transcribed (DE) and enriched genes (DERs) at multiple epigenetic levels in fat x fiber dietary combinations and in the presence/absence of carcinogen. The genome-wide carcinogen (AOM) effects were most prevalent at the RNA (116 DE genes) and K4me3 (7678 DERs including 3792 genes) levels. Pathway analysis of the differentially transcribed genes after AOM exposure indicated a strong link to interferon-associated innate immune responses often associated with anti-microbial activity, while K4me3 DERs were strongly linked to

colon tumorigenesis. However, these changes in K4me3 enrichment were not reflected at the transcriptional level during the early stages of cancer progression. Therefore, we propose that carcinogen-induced changes in genes with K4me3 DERs are harbingers of future transcriptional events, which drive malignant transformation of the colon cells.

We also demonstrated that the combinatorial diet (fish oil + pectin) was synergistically chemoprotective, and uniquely affected epigenetic profiles in the intestinal epithelium, e.g., upregulating lipid catabolism and beta-oxidation associated genes. These genes were linked to activated ligand-dependent nuclear receptors associated with n-3 PUFAs and were also correlated with the inhibition of lipogenesis and a decreased concentration of cholesterol. Interestingly, only a subset of these genes was affected in animals fed fish oil without pectin, and there was a markedly enhanced effect of biological mechanisms associated with n-3 PUFAs in the combinatorial diet.

In conclusion, we propose that the chemoprotective effects of the bioactive mediators of the combination fish oil + pectin diet during colon cancer progression are multifaceted and generate specific epigenetic modifications and transcriptional profiles. Our data contribute to the understanding of the regulatory action of chemoprotective bioactive compounds found in fish oil and readily fermentable fiber (n-3 PUFAs and SCFAs) in colonic crypts and provide mechanistic insight into current clinical and epidemiological findings.

DEDICATION

To my family and friends for their encouragement and love

ACKNOWLEDGEMENTS

Many thanks to Kranti Konganti, Matthew McLean, Jennifer Goldsby, and Jason Knight for assistance with data processing and bioinformatics. Also thanks to Evelyn Callaway for animal handling and diet preparation proficiency and Dr Ivan Ivanov for bioinformatics and statistics expertise. I am also thankful for the aid of all of the members of the Chapkin lab, including Jiahui (Echo) Pang, Dr Manasvi Shah and Dr Laurie Davidson.

I would like to especially thank my mentor, Dr Robert Chapkin for his guidance, patience and support. His insight and direction was indispensable in developing this body of work and I would also like to thank my committee members: Mark Zoran, Joanne Lupton, and Richard Gomer.

Lastly, I appreciate the love and support of my family and friends. My father and brother have always inspired and encouraged me. I would also like to especially thank my mom Dania Triff, for her intellectual insight, love and understanding throughout this process. None of this would have been possible without them.

This work was supported by the National Institutes of Health.

TABLE OF CONTENTS

	Page
ABSTRACT	ii
DEDICATION	iv
ACKNOWLEDGEMENTS	v
TABLE OF CONTENTS	vi
LIST OF FIGURES	x
LIST OF TABLES	xvii
CHAPTER I INTRODUCTION AND LITERATURE REVIEW	1
1.1 Colonic crypt physiology	1
1.2 Colorectal cancer	3
1.2.1 Histological progression of colon cancer	3
1.2.2 Molecular basis of colon cancer	4
1.2.3 Location of colon cancer	5
1.3 Epigenetic effects of n-3 polyunsaturated fatty acids (PUFA) and fiber in colon ...	6
1.3.1 Direct n-3 PUFA interaction with nuclear receptors during colon cancer	7
1.3.2 Indirect DHA regulation of transcription factors	9
1.3.3 DHA-mediated modulation of apoptosis regulatory pathways	13
1.3.4 Modulation of cytokines and growth factors	14
1.3.5 Effects of DHA on the cell cycle	15
1.3.6 Optimal chemoprevention: Interaction of DHA with butyrate	16
1.4 Summary and purpose of the study	17
1.5 Hypotheses and specific aims	18
CHAPTER II GENOME-WIDE ANALYSIS OF THE RAT COLON REVEALS PROXIMAL-DISTAL DIFFERENCES IN HISTONE MODIFICATIONS AND PROTO-ONCOGENE EXPRESSION	20
2.1 Introduction	20
2.2 Methods	22
2.2.1 Animals	22
2.2.2 Isolation of colonic crypts	22
2.2.3 Western blot	23
2.2.4 Chromatin immunoprecipitation and sequencing	24

2.2.5 Bioinformatic analyses	25
2.2.6 RNA isolation	25
2.2.7 RT-qPCR and ChIP-qPCR	25
2.2.8 RNA-Seq.....	27
2.2.9 Functional analyses	28
2.3 Results	30
2.3.1 Constructing an atlas of the distal and proximal colon epigenomes.....	30
2.3.2 Integration of the transcriptome and H3K4 trimethylation along the colonic longitudinal axis.....	38
2.3.3 Cataloging putative middle and long non-coding RNAs (ncRNAs)	41
2.3.4 Anatomically-linked signaling pathways.....	45
2.4 Discussion	50
2.4.1 Constructing an atlas of the distal and proximal colon epigenomes.....	51
2.4.2 Proto-oncogene networks influenced by location.....	52

**CHAPTER III ASSESSMENT OF HISTONE TAIL MODIFICATIONS AND
TRANSCRIPTIONAL PROFILING DURING COLON CANCER PROGRESSION
REVEALS A GLOBAL DECREASE IN H3K4ME3 ACTIVITY** 56

3.1 Introduction	56
3.2 Methods	57
3.2.1 Animals.....	57
3.2.2 Diets	58
3.2.3 Carcinogen treatment	58
3.2.4 Aberrant crypt foci scoring	59
3.2.5 Isolation of colonic crypts.....	59
3.2.6 Western blot.....	60
3.2.7 Chromatin immunoprecipitation.....	61
3.2.8 ChIP sequencing	61
3.2.9 RNA isolation	62
3.2.10 RNA sequencing.....	63
3.2.11 Non-coding RNA protocol.....	63
3.2.12 Ingenuity pathway analyses	66
3.3 Results	66
3.3.1 Global effects of AOM on transcription and histone tail modifications.	70
3.3.2 Poor correlation between AOM affected genes at the transcription and histone tail modification levels.....	71
3.3.3 Identification of upstream regulators perturbed by AOM.	74
3.3.4 Identification of pathways and networks perturbed by AOM.	77
3.3.5 Cataloging annotated middle and long ncRNAs and un-annotated H3K4me3 and H3K9ac enriched regions.	82
3.4 Discussion	83
3.4.1 AOM selectively modulates H3K4me3 genome association.	88

3.4.2 AOM induced transcriptional changes are strongly correlated with antimicrobial responses.....	89
3.4.3 Carcinogen induced changes in K4me3 DERs may predict for future transcriptional events	90
CHAPTER IV DIETARY FAT AND FIBER COMPOSITION ALTER GLOBAL HISTONE POST-TRANSLATIONAL EPIGENETIC PROGRAMMING IN A RAT COLON CANCER PROGRESSION MODEL	92
4.1 Introduction	92
4.2 Methods	94
4.3 Results	95
4.3.1 Enumeration of aberrant crypts.....	95
4.3.2 Overall effects of fish oil on transcription and histone tail modifications	96
4.3.3 Overall effects of pectin on transcription and histone tail modifications	102
4.3.4 Context specific epigenetic effects of fish oil and pectin	103
4.3.5 Identification of upstream regulators perturbed by dietary fat and fiber interaction.....	107
4.4 Discussion	110
4.4.1 Pectin preferentially affects H3K9 acetylation of saline over AOM injected rats	112
4.4.2 Fish oil triggers transcription of lipid metabolism associated genes	113
4.4.3 Fish oil plus pectin synergistically modulate transcription	115
CHAPTER V SUMMARY AND FUTURE DIRECTIONS	118
5.1 Aim 1 (Chapter II).....	119
5.2 Aim 2 (Chapter III)	121
5.3 Aim 3 (Chapter IV)	122
5.4 Future studies.....	125
5.4.1 ChIP-seq and sub-cellular localization of upstream regulators	125
5.4.2 Mitochondrial bioenergetic analysis.....	126
REFERENCES	128
APPENDIX A EXPERIMENTAL APPROACH	159
APPENDIX B ANIMAL WEIGHT LOG	161
APPENDIX C AZOXYMETHANE (AOM) INJECTION OF RATS	163
APPENDIX D LARR ANIMAL HANDLING	164

APPENDIX E LIPID ANALYSIS OF DIETS USING GC	165
APPENDIX F ACF BY RAT ID	167
APPENDIX G NUCLEI ISOLATION FROM COLON CRYPTS	171
APPENDIX H REAL TIME QPCR PROTOCOL FOR CHIP-QPCR.....	174
APPENDIX I BCA PROTEIN QUANTIFICATION ASSAY.....	179
APPENDIX J QUBIT PROTOCOL FOR MEASURING DNA CONCENTRATION	182
APPENDIX K WESTERN BLOT PROTOCOL.....	186
APPENDIX L CHIP PROTOCOL	192
APPENDIX M CHIP REAGENTS AND EQUIPMENT.....	197
APPENDIX N ISOLATION OF COLONIC CRYPTS	204
APPENDIX O QUALITY CONTROL REQUIREMENTS FOR CHIP-SEQ COMPARED TO ENCODE GUIDELINES	206
APPENDIX P SAMPLE PROCESSING; TYPES OF TISSUES COLLECTED	207
APPENDIX Q	208
APPENDIX R SAMPLE POOLING OF ANIMALS FOR SEQUENCING	209
APPENDIX S NUCLEAR PROTEIN LEVELS OF H3K4ME3 AND H3K9AC ARE NOT ALTERED BY DIET IN AOM INJECTED ANIMALS	211
APPENDIX T NUCLEAR LEVELS OF PAN-ACETYLATED H3 ARE NOT ALTERED BY A FISH OIL + PECTIN DIET.....	212
APPENDIX U SUMMARY OF DIFFERENTIAL EXPRESSION AND DIFFERENTIAL ENRICHMENT OUTPUT.....	213
APPENDIX V ANOVA TEST FOR RAT HM-ACF.....	214
APPENDIX W INFORMATION ON THE DATA IN EACH OF THE SUPPLEMENTARY FILES (IDENTIFIED BY BULLET AND BOLDED TEXT) IS PROVIDED BELOW	216

LIST OF FIGURES

	Page
Figure 1. Chemoprotective effects of n-3 PUFAs on the epigenetics of colon cancer cells. Intestinal genes that are up or down regulated by n-3 PUFAs at the mRNA and protein level including potential interaction patterns and segregated by cellular location. Red font represents gene up-regulation and blue font indicates gene down-regulation. Epigenetic levels of regulation in the nucleus are underlined. Nuclear genes are grouped by classification.	8
Figure 2. A. ChIP-Seq bioinformatics analysis workflow. B. Histone PTM binding at the transcription start site of representative active and silent genes.	26
Figure 3. General schematic of the experiments and analyses performed	31
Figure 4. Genome wide binding of histone modifications and RNA Pol II relative to the transcription start site (TSS) of annotated genes	32
Figure 5. Levels of trimethylated H3K4 are higher in the distal versus proximal colon. A. Representative immunoblots of nuclear protein from proximal (P) and distal (D) colon. B. The blue bars represent the proximal colon and red bars the distal colon. Quantification of band volumes using Quantity One software. Values are means \pm SE (n = 7). At least 2 independent assays were conducted.	34
Figure 6. Global RNA-Seq analysis. A. Detection of all annotated rat genes (y axis) and the intensity of their expression are presented in the x axis as $\log_{10}(\text{FPKM})$ values. Blue line represents proximal colon, orange line represents distal colon and the dotted line represents the minimum gene expression cutoff. B. Volcano plot representing differentially ($p < 0.05$) expressed genes, blue dots characterize significant differentially expressed genes; red dots denote genes that with no differential expression between the distal and proximal colon.....	36
Figure 7. Assessment of FABP1 ChIP (H3K4me3) and mRNA (qPCR and RNA-Seq) gene expression analyses in individual versus pooled samples. Patterns of gene expression in individual rats by qPCR analysis (using the $2^{-\Delta\Delta Ct}$ method) were consistent with sequencing analyses from pooled samples. ChIP and RNA sequencing was performed on a pooled sample containing the same amount of genetic material from 6 rats (A-F). A. H3K4me3 enrichment at the FABP1 transcription start site (pooled rats - ChIP-Seq) is compared with multiple (individual rats, A-F) ChIP-qPCR. B. mRNA expression of FABP1 (pooled rats - RNA-Seq) is compared with multiple (individual, A-F) rats using RT-qPCR. C. H3K4me3 enrichment at the NFIB and EPHB3 transcription start site (pooled rats - ChIP-Seq)	40

Figure 8. Occupancy of histone PTMs in highly differentially H3K4me3 regulated genes. Data from the distal colon are shown as blue peaks and proximal as black peaks. A. Genomic occupancy of differentially ($p < 0.05$) over-expressed Intestinal Immunity related genes, FOX1A, FOX2A, and TXN1, in the distal colon. B. Genomic occupancy of differentially ($p < 0.05$) under-expressed development regulating HOXB and OSR2 genes. The genome is scaled to 10kB. C. Genomic occupancy of differentially over-expressed long non-coding RNA genes; GAS5, NEAT1, and H19. The distal colon RNA transcripts are represented with orange bars and the proximal colon transcripts with red bars.42

Figure 9. Canonical Pathways. A. The PI3K signaling pathway was the most significantly affected canonical pathway in terms of H3K4me3 occupancy and RNA transcription. PI3K signaling activates cell survival, cell growth, cell cycle and contributes to colon carcinogenesis. Orange highlights indicate the differential expression of genes enriched at the H3K4 trimethylation and mRNA levels. Red highlighted genes represent over expression (distal > proximal) and green represents under expression in the distal colon at the RNA level. PI3K related genes differentially regulated solely at the H3K4me3 level are highlighted in pink and yellow, representing up and down regulated genes, respectively. B. Venn diagram showing the overlap between canonical pathways linked to changes in H3K4me3 occupancy and RNA transcription.46

Figure 10. Upstream analyses of differentially expressed genes are linked to key proto-oncogenic transcription factors (TFs). Validated gene targets of each transcriptional regulator and their direction of change were compared in the proximal and distal colon. The overlap p-values represent the statistical significance. Fisher’s Exact Test was used to determine overlap between the differentially expressed genes and those genes regulated by the TF. Activation z-scores indicate whether an upstream TF has significantly more “activated” than “inhibited” or vice versa.48

Figure 11. PPAR signaling in the proximal and distal colon is linked to disease-associated genes. Validated targets of PPAR genes and their direction of change were compared in the proximal and distal colon. Orange highlights indicate the differential expression of genes enriched at the H3K4 trimethylation and mRNA levels. Red highlighted genes represent over expression (distal > proximal) and green represents under expression in the distal colon at the RNA level.55

Figure 12. Bioinformatic workflow of sequencing data.67

Figure 13. AOM induced pathophysiology and epigenetic changes during cancer progression. A, High multiplicity aberrant crypt foci (HM ACF) number, a precursor to tumorigenesis is shown. Data represent 43 rats at 10 weeks post AOM injection. No HM-ACF were observed in saline-injected animals. ***, p-val < 0.001 (one-way ANOVA). B, Levels of trimethylated H3K4 and acetylated H3K9 levels remain the same in AOM vs saline exposed colonic epithelial cells. The results of Western blot analysis of nuclear protein extracts from AOM and saline injected rat colon. Blue bars represent AOM treated and red bars saline injected colon. Quantification of band volumes using Quantity One software. Values are means \pm SE ($n = 7$). At least 2 independent assays were conducted. n.s., p-val > 0.05. C, H3K4me3 is highly sensitive to AOM exposure. MAplots indicate the differential expression of all transcribed genes or histone tail enriched regions (y-axis, log-ratio of difference in intensity of histone tail modifications enriched regions) vs their overall intensity of expression (x-axis, log-average of read counts) following AOM vs saline treatment. Pink represents differentially expressed (DE) transcripts and differentially enriched regions (DERs) with a p-value < 0.05, and genes FDR < 0.1 in red, all other detected genes are highlighted in blue.....69

Figure 14. Correlation between transcriptional DE and histone tail DERs. A, Poor correlation is observed between DE transcripts and histone tail modifications with DERs. Scatterplots reveal low correlation between DE transcripts and histone tail modifications with DERs by comparing the log₂(fold changes) from AOM vs saline treatments. Specific contrasts include RNA vs. K4me3, RNA vs. K9ac, and K4me3 vs. K9ac. B, Venn diagram illustrating the number of genes modulated by AOM in common between the various epigenetic stages (H3K4me3, H3K9ac and RNA). OASL2, RTP4, and TPM2 were affected across all measured epigenetic states. C, Chromatin signature of H3K9ac and H3K4me3 occupancy following AOM exposure is associated with upregulation of OASL2 and RTP4 and downregulation of TMP2. Representative UCSC genome browser images of all DNA reads mapped and grouped by genomic location in highly differentially enriched H3K4me3-regulated genes. H3K4me3 data from AOM treated animals are shown as red peaks, and saline control as blue peaks. H3K9ac data from AOM treated animals are shown as purple peaks, and saline control as brown peaks.72

Figure 15. Top Upstream Regulators (URs) affected by AOM across multiple epigenetic states. A, Venn diagram illustrating the total number of URs in common across epigenetic states (H3K4me3, H3K9ac and RNA). The 5 URs predicted to be affected by AOM are listed. P-values and activity (z-scores) of URs were determined by Ingenuity Pathway Analysis of genes with an FDR<0.1. B-D, Representative networks of genes regulated by top

ranked upstream regulators (URs). Yellow fill represents the projected increase in UR activity, green fill represents projected decrease in UR activity. Blue fill indicates decreased gene activity (DE and/or DERs) and orange fill indicates increased gene activity (DE and/or DERs), deeper color hue indicates genes with greater $|\log_2(\text{fold change})|$. Solid lines represent direct and dashed lines represent indirect gene interaction. B, K4me3 DERs associated with activation of SOX1/3 development associated transcriptional regulators and inhibition of ANXA7, an ion channel with tumor suppressor function. C, Concurrent increase in mRNA DEs and K4me3 DERs predicted inhibition status of TRIM24 (Tif1 α), a tumor suppressor transcriptional inhibitor of proto-oncogenes that activate transcription, immune response and cell proliferation. D, Decrease in K4me3 DERs of key genes (blue) indicating AOM induced activation of KDM5B (histone K4me3 demethylase) associated with inhibition of KMT2D (histone K4 methylase). E, Heatmap of top URs in common between multiple epigenetic stages (H3K4me3, H3K9ac and RNA). Data represent $-\log(\text{Benjamini-Hochberg-adjusted p-value})$ with a score >1.29 indicating a B-H $p\text{-value} < 0.05$. URs are grouped according to families, color scale ranges from yellow to brown, with decreasing p-value corresponding to deeper shades of brown.76

Figure 16. A, Venn diagram depicts the poor correlation observed between canonical pathways affected by AOM across epigenetic states (H3K4me3, H3K9ac and RNA). B, Top canonical pathway significantly enriched for both DE genes and URs of DE genes. Purple outlined molecules indicate DE linked URs, yellow to orange color indicates DE genes.79

Figure 17. Top canonical pathways and networks affected by AOM. A, List of the top 3 enriched pathways, followed by top activated and inhibited pathways grouped by category. The RNA-Seq category includes the only 2 pathways that were activated/inactivated based on mRNA DE data analysis. All other canonical pathways listed were extrapolated from K4me3 DERs. Activation z-scores indicate whether an upstream transcription factor is significantly more “activated” or “inhibited” based on mRNA DE or K4me3 DER data. P-values are $-\log(\text{Benjamini-Hochberg-adjusted p-value})$ with any score >1.29 indicating a B-H $p\text{-value} < 0.05$. B-C, Blue fill indicates decreased gene activity (DE and/or DERs) and orange fill indicates increased gene activity (DE and/or DERs), deeper color hue indicates genes with greater fold change. Dashed lines indicate indirect interactions, solid lines indicate direct interactions. The arrow style indicates specific molecular relationships and the directionality of the interaction. B, Top network annotation of transcriptionally DE genes. C, Top network annotation of K4me3 DER genes.80

- Figure 18. H3K4me3 and H3K9ac enriched regions co-localize with ncRNAs. A, Putative ncRNA loci with K4me3 and K9ac enriched regions. (B-C) UCSC genome browser snapshots of ncRNA, as well as their conserved, thermodynamically stable secondary structures predicted by RNAfold (below); putative ncRNAs also include predicted annotation. The entropy color scale represents the values at the weakest spots of the structure where 0 entropy means no deviations, and entropy >0 indicates some deviations; the higher the entropy, the more likely for the folding structure to deviate. B, Genomic occupancy and structure of annotated and putative ncRNAs with unregulated K4me3 DERs (black box) and putative ncRNA homologous to a mouse annotated lncRNA.87
- Figure 19. Fish Oil + Pectin synergistically suppress malignant transformation of the colon. High Multiplicity Aberrant Crypt Foci (HM ACF) incidence per rat colon is shown. Bars not sharing the same superscript letters are significantly different. $p < 0.05$96
- Figure 20. Differential expression (DE) of all transcribed genes or histone tail differentially enriched regions (DERs). MAplots indicate the differential expression of all transcribed genes or histone tail enriched regions (y-axis, log-ratio of difference in intensity of histone tail modifications enriched regions) vs their overall intensity of expression (x-axis, log-average of read counts) following AOM vs saline treatment. Pink represents differentially expressed (DE) transcripts and differentially enriched regions (DERs) with a p-value < 0.05. Genes with an FDR < 0.1 are highlighted in red, and all other detected genes are highlighted in blue.....98
- Figure 21. Poor correlation between DE transcripts and histone tail modifications with DERs. Scatterplots reveal low correlation between DE transcripts and histone tail modifications with DERs by comparing the $\log_2(\text{Fold Changes})$ from AOM vs saline treatments. Specific contrasts include RNA vs. K4me3, RNA vs. K9ac, and K4me3 vs. K9ac.....99
- Figure 22. Poor correlation between DE transcripts and histone tail modifications with DERs modulated by fish oil. Overlap of differentially expressed (DE) or differentially enriched (DER) genes across multiple epigenetic levels (FDR < 0.1 and $p < 0.01$) is shown.100
- Figure 23. Top networks affected by fish oil. Network analysis of differentially expressed genes ($p < 0.01$) identified with Ingenuity Pathway Analysis (IPA) software in response to fish oil feeding (fcs+fps+fca+fpa versus ccs+cps+cca+cpa). Blue fill indicates decreased gene activity (DE and/or DERs) and orange fill indicates increased gene activity (DE and/or DERs). Deeper color hue indicates genes with greater fold change. Dashed lines

indicate indirect interactions and solid lines indicate direct interactions. The arrow style indicates specific molecular relationships and the directionality of the interaction. 102

Figure 24. Poor correlation between DE transcripts and histone tail modifications with DERs dysregulated by pectin. Overlap of differentially expressed (DE) or differentially enriched (DER) genes across multiple epigenetic levels (FDR<0.1 and p<0.01). 104

Figure 25. Top networks affected by pectin feeding. Network analysis of differentially expressed genes ($p < 0.01$) identified with Ingenuity Pathway Analysis (IPA) software in response to pectin. Blue fill indicates decreased gene activity (DE and/or DERs) and orange fill indicates increased gene activity (DE and/or DERs). Deeper color hue indicates genes with greater fold change. Dashed lines indicate indirect interactions, and solid lines indicate direct interactions. The arrow style indicates specific molecular relationships and the directionality of the interaction. 105

Figure 26. Fish oil + pectin transcriptionally enhance β -oxidation and decrease lipid concentration associated genes. Analysis of differentially expressed genes associated with lipid metabolism. Lines indicate the association between a biological process and each gene. Colors of biological functions indicate whether the differential expression of genes were associated with activation or inhibition of that biological process ($p < 0.001$). Differentially expressed genes with and FDR<0.1 are bolded. Ovals and star indicate genes also dysregulated in diets/treatments different from fpa by any one variable including fat, fiber or carcinogen. 109

Figure 27. The chemoprotective effects of Pectin + Fish Oil during cancer progression included predicted activation of nuclear receptors. Heatmap of A. ligand-dependent nuclear receptors and B. chromatin modifiers (HDACs) and interferon associated URs, indicates the top Upstream Regulators in fish oil + pectin fed rats during cancer progression. Data represent $-\log(\text{Benjamini-Hochberg-adjusted } p\text{-value})$ with a score > 1.29 indicating a B-H $p\text{-value} < 0.05$. URs are grouped according to families, color scale ranges from yellow to brown, with decreasing p-value corresponding to deeper shades of brown. 111

Figure 28. Fish oil + pectin feeding enhances activation of ligand dependent nuclear receptors (NRs) associated with n-3 PUFAs (fpa v cca). Yellow fill represents the projected increase in NR activity. Blue fill indicates decreased gene activity (DE) and orange fill indicates increased gene activity (DE). Deeper color hue indicates genes with greater

<p>$\log_2(\text{FoldChange})$. Solid lines represent direct and dashed lines represent indirect gene interaction.</p>	114
Figure 29. Summary of epigenetic findings from Aim 2 and 3.	124
Figure 30. Future Studies Methodology and Protein of Interest. Red color indicates predicted activation and blue color predicted inhibition of transcription factors and co-factors, nuclear receptors, and histone modifying enzymes in response to chemoprotective treatment.	125
Figure 31. We propose a putative mechanism for future investigation in which 1. decreased RXR/RAR activity at specific sites (in a context specific manner) during colon cancer (Tang and Gudas 2011) leads to a 2. decrease in KDM5B binding (Zhang, Liang et al. 2014) therefore 3. TRIM24 cannot bind to trimethylated H3K4 (Tsai, Wang et al. 2010). Thus an increase in K4me3 subsequently leads to a decrease in TRIM24 (Tisserand, Khetchoumian et al. 2011) binding, inducing IFN/STAT1 associated colon cancer genes (Kikuchi, Okumura et al. 2009).	126

LIST OF TABLES

	Page
Table 1. The experimental details associated with n-3 PUFA differentially expressed target genes.	11
Table 2. The experimental details associated with n-3 PUFA plus fermentable fiber differentially expressed target genes	12
Table 3. Annotation of histone ChIP-Seq island peaks using SICER software.....	33
Table 4. Relationship between levels of H3K4 trimethylation, H3K9 trimethylation and mRNA expression.	34
Table 5. Top 20 most up-regulated and down-regulated genes in the distal versus proximal colon. RNA expression and H3K4me3 enrichment are represented as fold change DvP, distal divided by proximal colon.....	37
Table 6. Intestinal genes display similar patterns across analytical platforms.....	38
Table 7. Transcribed genes correlate with histone PTMs associated with active gene expression.	39
Table 8. Candidate middle and long non-coding RNAs. Transcripts >100 bp mapped to an intronic or intergenic region of the genome, considered candidates for middle and long non-coding RNAs, also exhibited signs of epigenetic regulation at the histone level.	43
Table 9. Annotated middle and long non-coding RNAs. Transcripts identified as non-coding RNAs by sequence using the NONCODE and fRNA databases are listed.....	44
Table 10. Pathway analysis was used to identify the top biological functions associated with differentially expressed genes. RNA and H3K4me3 up and down regulated genes that exhibited significant differential expression between the colonic tissue types were associated with a biological function (p-value) and a statistical score (Activation z-score). Right-tailed Fisher’s exact test was used to calculate p-values determining the probability that each biological function was due to chance alone. The activation state (“Increased” or “Decreased”) in the distal over proximal colon was inferred by the activation z-score. These z-scores define relationships in the molecular networks that represent experimentally observed causal associations between genes and their functions.....	49

Table 11. Summary total of all detected genes at the H3K4me3 and H3K9ac and mRNA level.	68
Table 12. List of AOM induced differentially expressed (DEs) and differentially enriched regions (DERs) detected across multiple epigenetic states.....	73
Table 13. Description of lncRNA genes modulated by AOM at the H3K4me3, H3K9ac or transcription level.....	84
Table 14. Summary of diet effects on differentially expressed (DE) and differentially enriched (DERs) genes.	100
Table 15. Summary of differentially expressed (DEs) and differentially enriched (DERs) genes in common between saline and AOM injected rats fed the same diet.	107
Table 16. Top categories of biological functions affected by fish oil + pectin under carcinogenic conditions at each epigenetic level (parenthesis), for each treatment the listed category was predicted by pathway analysis with a certainty of $p < 0.05$	108

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW*

1.1 Colonic crypt physiology

The epithelial layer, at the luminal surface of the colon, consists of a single sheet of columnar epithelial cells which form finger-like invaginations into the underlying connective tissue of the lamina propria to form the basic functional unit of the intestine, the crypt. Within the colon are millions of crypts and all the evidence points to the location of the stem-cell population at the base of the crypt. Upon asymmetrical divisions, the daughter cells of stem cells in the bottom of the crypt undergo differentiation and migrate upward to give rise in turns to transit-amplifying (TA) precursors and terminally differentiated cells (1,2).

A crypt contains approximately 2000 cells, around 30-50 cells high, and is supported by the cells of the underlying lamina propria. The rate of cell turnover in the colon is high, epithelial cell lineages have complete self-renewal every 2-7 days under normal circumstances. Crypts themselves also reproduce by a process called crypt fission beginning with basal bifurcation and followed by longitudinal division (3).

There are three major terminally differentiated epithelial lineages in the colon, the colonocytes, also termed absorptive enterocytes, the mucus-secreting goblet cells, and the less abundant enteroendocrine cells. Paneth cells, which secrete growth factors, digestive enzymes and antimicrobial peptides such as cryptidins are also sometimes present in the proximal colon and in certain disease states.

* Portions of this chapter reprinted with permission from Karen Triff, Eunjoo Kim and Robert S. Chapkin (2015) Chemoprotective epigenetic mechanisms in a colorectal cancer model: Modulation by n-3 PUFA in combination with fermentable fiber. *Curr Pharmacol Rep* 1, 11-20 Copyright 2015 by Springer International Publishing

Colonocytes are simple columnar cells important in absorption and movement of mucus in the colon. They contain apical microvilli, increasing the luminal surface area of the cell by 14- to 40 fold, and the nucleus is closer to the basal aspect of the cell (1).

Enteroendocrine cells are specialized endocrine cells of the gastrointestinal tract that constitute the diffuse neuroendocrine system of the gut and secrete a variety of products. Although the principal function of enteroendocrine cells is classically to synthesize and secrete gut hormones, enteroendocrine cells may also express TLRs which are important for innate immunity in response to environmental challenge. These substances act via neurocrine, endocrine, and paracrine mechanisms and maintain homeostasis by allowing communication and reciprocal regulation between the nervous, endocrine, and immune systems. At least ten different types of enteroendocrine cells have been described. Generally, the colon is known to contain secretin cells, glucagon secreting L cells, and gastric inhibitory peptide (GIP) synthesizing K cells. Major hormones in the colon include glucagon, peptide-YY, neurotensin and guanylyl cyclase-activating protein 1 and 2 (GCAP2) (4).

Goblet cells are unicellular glands that function to secrete mucin which dissolves in water to form mucus. The majority of the cell's cytoplasm is occupied by mucinogen granules, except at the bottom. The nucleus, and other organelles are concentrated in the basal portion and the apical plasma membrane projects microvilli to increase surface area for secretion of mucus is by exocytosis of secretory granules. Goblet cells have two pathways for secretion, constitutive or basal secretion, low level, unregulated and essentially continuous secretion. This pathway is dependent on cytoskeletal movement of secretory granules. The other is stimulated secretion, regulated exocytosis of granules in response to extracellular stimuli. This pathway provides an ability to dramatically increase mucus secretion (5).

In experimental models of colonic carcinogenesis, aberrant crypt foci (ACF), collections of abnormal appearing colonic crypts (2-3X larger), are the earliest detectable abnormality and precede adenomas. A number of mutations within important

tumor related genes have also been identified within these lesions, including K-ras, B-catenin and APC (2).

1.2 Colorectal cancer

Colon cancer is a major public health concern due to the high prevalence of the disease both globally and in the United States (6). Colon cancer is third in cancer incidence in both men and women and the second leading overall cause of cancer mortality (7), with incidence and death rates of colon cancer being higher in men than in women(7). In the United States alone, approximately 140,000 new cases occur each year, with colon cancer causing the death of approximately 50,000 people every year (6). Therefore, developing strategies for prevention and treatment of this pernicious disease is of the utmost importance. Both genetic and environmental factors have been indicated as mediators of colon cancer risk (8). Additionally, many modifiable factors, including diet, exercise, and smoking are suggested to play a role in more than 50% of colon cancer cases (9,10).

1.2.1 Histological progression of colon cancer

Formation of colon cancer is hypothesized to be a multi-step process that results from a systematic accumulation of both genetic and epigenetic perturbations that cause normal colonic epithelial cells to transform and progress into cancer. Colonic transformation develops through multiple distinct histologically abnormal stages.

Aberrant crypts foci (ACF), indicated as putative biomarkers of colonic carcinogenesis, are very early, microscopic lesions composed of hyperplastic epithelium (11). It currently remains unclear whether ACF are precursors for colon cancer, but many studies have shown that ACF incidence increases with increased risk factors for colon cancer (12).

1.2.2 Molecular basis of colon cancer

The multi-step process of carcinogenesis is hypothesized to include 3 major steps: initiation, promotion, and progression. Each of these stages arises due to specific molecular and genetic changes that occur. The molecular basis of colon cancer has been extensively researched. Three major molecular contributors to colon cancer are chromosomal instability, inactivation of tumor suppressor genes, and activation of oncogenes.

Genomic instability: Genomic instability has been found to be a major contributor to colon cancer. Genetic instability in colon cancer is often divided by researchers into two classes: microsatellite instability or chromosomal instability (13) (14). Chromosomal instability is most often the source of genomic instability in colon cancer (15). These two classes of genomic instability induce distinct perturbations to the genome. Microsatellite instability displays a high rate of alterations in short, tandemly repeated nucleotide sequences, whereas chromosomal instability manifests as major abnormalities in chromosome structure and number (16). Numerous genes which function to maintain genetic stability become inactivated in colonic epithelium to result in instability that leads to cancer (17). Specifically, many colon cancer patients have inactivated mismatch repair genes, including MLH1 and MSH2 (18), and mismatch repair deficiency contributes to microsatellite instability (15). In contrast, chromosomal instability is often linked to loss of genes that function as critical mitotic checkpoints (16). Gene mutation and aberrant DNA methylation commonly drive the inactivation of mismatch repair genes and mitotic checkpoint genes (19). In fact, aberrant methylation has been clearly demonstrated to be a major mechanism of inactivation of MLH1 (20). Additionally, recent work assessing the colon cancer epigenome has shown that virtually all colon cancers have hundreds to thousands of abnormally methylated genes (21).

Hypermethylation of promoter-associated CpG islands of genes leads to transcriptional silencing of genes. Widespread CpG island promoter methylation, known as the CpG island methylator phenotype (CIMP), is frequently present in colon cancer

(22). CIMP is most commonly observed in colon cancers with microsatellite instability (22,23).

Inactivation of tumor suppressor genes: Genes that restrict the growth of cells are referred to as tumor suppressor genes. Inactivation of tumor suppressor genes unfetters cells from the constraints implemented by these genes, allowing for uncontrolled cell growth. Similarly to mismatch repair genes, many tumor suppressor genes are inactivated in colon cancer through genetic mutation or aberrant methylation. Major tumor suppressor genes known to be inactivated in colon cancer include APC, TP53, and TGFBR2. APC inhibits the nuclear localization of the oncoprotein β -catenin and targets it for proteolysis. A deficiency in APC results in constitutive activation of the Wnt signaling cascade, which is an initiating event in colon cancer (18). Germ-line APC mutations lead to a predisposition for colon cancer, and somatic mutations and deletions of APC have been found to occur in most sporadic colon cancers (24). Mutation of TP53 is also key in colon cancer. TP53 encodes p53, which is central in controlling the cell cycle (25). TP53 inactivation most often occurs at an intermediate stage in cancer development during the transition from adenoma to carcinoma (18).

1.2.3 Location of colon cancer

Colon cancer can arise in either the right side (proximal) or left side (distal) of the colon. Interestingly, research findings have implicated differing genetic alterations as the cause of colonic transformation in left-sided compared to right-sided colorectal cancers (14,26,27). Microsatellite instability has been shown to be primarily located in right-sided cancers (Cancer Genome Atlas, 2012), whereas chromosomal instability is more common in left-sided colon cancer (28). CIMP incidence most often occurs in right-sided colon cancers (23). Furthermore, mutations in TP53 and KRAS genes occur predominantly in left-sided carcinomas (14,29). The differences between left- and right-sided colon cancers provide evidence of different mechanisms contributing to colonic transformation.

1.3 Epigenetic effects of n-3 polyunsaturated fatty acids (PUFA) and fiber in colon

Over the past 25 years, hundreds of published papers have described the effects of polyunsaturated fatty acids (PUFA) on normal and cancer cell types, including differences between n-6 and n-3 PUFA with respect to their mechanisms of action (30-33). From this body of work, there is now mounting evidence that n-3 PUFA, namely, docosahexaenoic acid (DHA, 22:6n-3) and eicosapentaenoic acid (EPA, 20:5n-3) found in fish and algal oils, exert anti-inflammatory properties in the colon, enhance the efficacy of chemotherapeutic drugs, suppress chronic inflammatory biomarkers associated with obesity/diabetes, and reduce colon cancer risk (34-39). The actions of n-3 PUFA appear to involve multiple mechanisms that link the cell membrane, cytosol, and the nucleus (33)*, (40). For example, n-3 PUFA modulate membrane and nuclear receptors, and sensors/ion channels, thereby regulating signaling processes that influence patterns of gene expression. These effects appear to be mediated, in part, via the incorporation of n-3 PUFA into cell membranes (33)*, (41). Moreover, these changes in membrane composition can affect membrane order, the formation of lipid rafts, and intracellular signaling processes (31).

With respect to the cell nucleus, nutri-epigenomics is an emerging field of research that is focused on the interaction between nutrition and the epigenome. Epigenetics refers to a group of heterogeneous processes that regulate transcription without changing the DNA coding sequence. These changes include covalent histone modifications, principally acetylation and methylation of lysine residues but also phosphorylation and ubiquitination, DNA methylation, transcriptional machinery and noncoding RNA activities (42-44). Epigenetic marks can exhibit plasticity throughout the life course, albeit to varying degrees, and can be modified by environmental factors including diet (45). One implication of the interaction between the diet and the epigenome is that it may be possible to reprogram epigenetic marks that are associated with increased disease risk by nutritional or lifestyle interventions. This review will focus on the nutri-epigenomic role of n-3 PUFA, particularly DHA, in relation to colon

cancer. To summarize the effects of fish oil as a chemoprotective bioactive we consider published data on fish oil along with its chemoprotective components including; long chain n-3 PUFAS, and also more specifically DHA/EPA.

1.3.1 Direct n-3 PUFA interaction with nuclear receptors during colon cancer

DHA and EPA and their oxidative metabolites have been shown to interact with specific ligand dependent nuclear receptors including CAR, HNF4A, PPARG, PXR and RXRA (**Figure 1**) (46). In this fashion, n-3 PUFA regulate the function of nuclear receptors and their impact on transcriptional processes. For example, DHA bound PPARG can be transported to the nucleus where it controls energy balance by regulating fatty acid homeostasis in part via enhancing the expression of genes associated with membrane-bound fatty acid transporting proteins, and β -oxidation of fatty acids in peroxisomes and mitochondria (47). Interestingly, impaired expression and function of PPARG is associated with inflammatory bowel diseases (IBD) and colon cancer (48,49). RXRA, which is implicated in cancer chemoprevention, also preferentially binds to n-3 PUFA in colonocytes (50). Activation of PPARG as well as heterodimers formed with RXR play an important role in the antitumor effects of n-3 PUFAs (48).

LXRs are transcriptional regulators of cholesterol metabolism that control cholesterol uptake into cells, catabolism, and efflux (51). This is noteworthy, because cholesterol can control cell proliferation; and disruptions in cholesterol metabolism have been associated with the development of colon cancer (52-54). LXRs also function by heterodimerizing with RXRA and binding to direct repeats with four nucleotide spacers (DR4 elements), termed LXR response elements (LXREs), in the promoter regions of target genes (51). Interestingly, n-3 PUFA activated LXR α blocks proliferation of human colorectal cancer cells and slows the growth of xenograft tumors in mice (55).

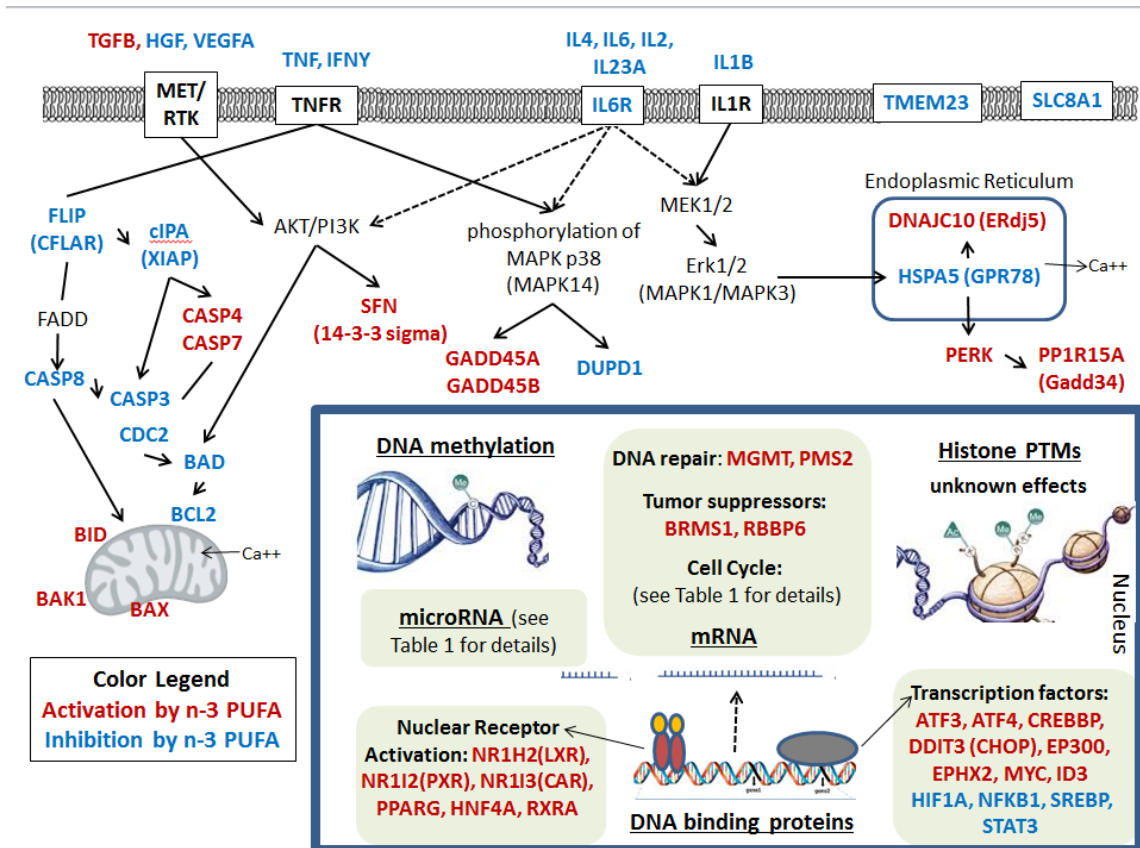


Figure 1. Chemoprotective effects of n-3 PUFAs on the epigenetics of colon cancer cells. Intestinal genes that are up or down regulated by n-3 PUFAs at the mRNA and protein level including potential interaction patterns and segregated by cellular location. Red font represents gene up-regulation and blue font indicates gene down-regulation. Epigenetic levels of regulation in the nucleus are underlined. Nuclear genes are grouped by classification.

PXR (NR1I2) has been shown to regulate the expression of genes involved in the oxidation, conjugation, and the transport of xenobiotics, and promotes the metabolism, elimination and detoxification of chemotherapeutic agents (56). The transcription of PXR increases in the presence of n-3 PUFA (57). This is noteworthy, because PXR can suppress the proliferation and tumorigenicity of colon cancer cells (58). CAR (NR1I3) is likewise transcriptionally increased by n-3 PUFAs in epithelial colorectal adenocarcinoma cells and similarly regulates genes involved in xenobiotic detoxification and energy homeostasis (57).

HNF4A maintains epithelial cell function and normal colon physiology via regulation of the balance between proliferation and differentiation, immune function, ion transport, epithelial barrier function and oxidative stress (59,60). P1-, but not P2-HNF4A, expression is lost in colorectal carcinomas in humans and it is predicted that treatments that increase nuclear P1-HNF4 α protein levels, such as n-3 PUFA, could help slow colon cancer progression (61,62).

1.3.2 Indirect DHA regulation of transcription factors

Since the original description of dietary fat as a regulator of gene expression over a decade ago, many transcription factors have been identified as prospective indirect targets for n-3 PUFA regulation. For example, DHA can increase the activity of CREBBP, EP300, and MYC and decrease activity of NF-kB (NFKB1), and STAT3 (46). However, DHA does not directly bind to this class of transcription factors. With respect to colon cancer, DHA exhibits a protective suppressive effect against hyperactivated STAT3 and may reestablish the equilibrium between STAT3 and PPARG (63). The ability to decrease STAT3 activity may be associated with the ability of n-3 PUFA ligands to trigger PPARG-RXR heterodimers to localize at their cognate PPAR response elements (PPREs) and exchange corepressors for coactivators such as cyclic AMP response element binding protein (CREB) and p300 (49).

The cytotoxic effects of DHA are also associated with signaling pathways involving lipid metabolism and endoplasmic reticulum (ER) stress. DHA induced depletion of free cholesterol in the ER can lead to ER stress, resulting in the growth arrest/apoptosis of metastatic tumor cells (64). It has been suggested that these alterations in the sterol content of the ER by DHA mediate growth reduction partly by down regulating nuclear SREBP, an important manager of lipid homeostasis and cell growth regulation (65). Induction of ER stress mediators by DHA also promotes expression of the kinase PERK, which in turn promotes translation of transcription factors ATF3, 4 and 6 (64). Furthermore, an elevation in PERK activity can increase

levels of ER protein GADD34 (PPP1R15A) and the pro-apoptotic transcription factor DDIT3 (CHOP) along with its downstream target TRIB3 (66). The experimental details associated with differentially expressed target genes are described in **Table 1** and **Table 2**.

Colon adenocarcinomas exhibit defective expression of the adenomatous polyposis coli (APC) gene, which is a critical regulator of the Wnt signaling pathway. This and other developmental pathways play an important role in both genetic (familial) and sporadic epithelial cancers (67). From a chemoprevention perspective, *in vivo* studies demonstrate that fish oil-derived n-3 PUFA suppress the formation of intestinal tumors in mice and humans with a defective APC allele (68,69). The downstream APC signaling oncogene, MYC, is an important regulator of cell proliferation, and the lack of MYC expression is associated with a reduced number of intestinal adenomas (70). Interestingly, patients with an amplified MYC gene and wild type p53 have a greater response to anticancer therapies (71). In colon cancer cells, DHA increases the level of MYC, which is believed to induce a chemoprotective, pro-apoptotic phenotype (72).

Table 1. The experimental details associated with n-3 PUFA differentially expressed target genes.

n-3 PUFA Treated								
Up regulated								
Gene symbols	Reference	Assay type	Carcinogenesis method	Chemoprotective treatment method	Dose	Organism	Level	Cancer stage
FOXO3A,FAF1,BCL10,DIFFA,TNFRSF1A,GADD45A,CASP7,MOAP1,DAP3,TNFRSF10B,GADD45B,CASP4,CIP1/P21/CDKN1A,CENG2,SFN/14-3-3,PPP1R15A/GADD34,SFN/14-3-3,TRIB3	Slagsvold et al, 2010	Genome Arrays	colon adenocarcinoma cell line, SW620	Solution in vitro	DHA 70 mmol/l	Human	RNA	adenocarcinoma
ATF6,GCLC,OSBP,CAPN7,NPC2,VCP,SOD1,VLDLR,NRF2,XBP1,HSP47,CAPN2,CAMLG,ATF4,PERK,CASP7,PSMD1/RPN2,IP3R1,LDLR,TXNRD1,CASP4,GCLM,GADD34,ATF3,DNAJB1,NPC1,BAG3,HSPA1B,TRIB3,SQSTM1,SQSTM1,HSPA1A/B,HMOX1,HMOX1	Jakobsen et al., 2008	Genome Arrays	colon adenocarcinoma cell line, SW620	Solution in vitro	DHA 70 mmol/l	Human	RNA	adenocarcinoma
ERDJ5,PERK	Fasano et al., 2012	Western Blot	colon adenocarcinoma cell line, SW480	Solution in vitro	DHA 30 mmol/l	Human	protein	adenocarcinoma
BID,BAK,BAX	Giros et al., 2010	Western Blot	colon adenocarcinoma cell line, HT-29 and Caco-2	Solution in vitro	DHA 60 mmol/l	Human	protein	adenocarcinoma
miR-18a, miR-27b, miR-93, miR-200c, miR-497	Shah et al., 2011	Low-density array	Azoxymethane (AOM) injection (15 mg/kg bw)	Diet	11.5% fish oil (ad-libitum)	Rat distal colon	RNA	cancer progression
miR-30c, miR-141	Gil-Zamorano et al., 2014	q-PCR	colon adenocarcinoma cell line, Caco-2	in vitro	DHA 200 mmol/l	Human	RNA	adenocarcinoma
Down regulated								
Gene symbols	Reference	Assay type	Carcinogenesis method	Treatment method	Dose	Organism	Level	Cancer stage
CCND1,CCND3,CDK2,CDC42,CDC25C,CDC45L,CDC20,CDK4,E2F1,CENPE,AKT1/PKB,BAD,CCNA2,CCNF,DC25B,TNFRSF1B,CDK2AP,AURKA,BUB1,CDC7,PCNA,BIK,BIRC5,BIRC5,UNG,CCNA2,CCNB2,STMN1,CDC2/CDK1,AURKB,PLK1	Slagsvold et al, 2010	Genome Arrays	colon adenocarcinoma cell line, SW620	Solution in vitro	70 mmol/l	Human	RNA	adenocarcinoma
FDPS,CAT,CAV1,DHCR7,DHCR24,PMVK,TM7SF2,CCND1,HMGR,SREBP2	Jakobsen et al., 2008	Genome Arrays	colon adenocarcinoma cell line, SW621	Solution in vitro	DHA 70 mmol/l	Human	RNA	adenocarcinoma
GRP78	Fasano et al., 2012	Western Blot	colon adenocarcinoma cell line, SW480	Solution in vitro	DHA 30 mmol/l	Human	protein	adenocarcinoma
IL2,IL4,IFNG,TNF,IL6,IL1B	Purasiri et al., 1994	ELISA	CRC patient serum	supplement	50% n-3 PUFA supplement	Human	protein	adenocarcinoma
XIAP,FLIP,BAD,BCL2,COX2	Giros et al., 2009	Western Blot	colon adenocarcinoma cell line, HT-29 and Caco-2	Solution in vitro	DHA 60 mmol/l	Human	protein	adenocarcinoma
BID,BAK,BAX	Giros et al., 2010	Western Blot	colon adenocarcinoma cell line, HT-29 and Caco-3	Solution in vitro	DHA 60 mmol/l	Human	protein	adenocarcinoma
VEGF	Calviello et al., 2009	Western Blot	colon adenocarcinoma cell line, SW480	Solution in vitro	DHA 30 mmol/l	Human	protein	adenocarcinoma
miR-21	Shah et al., 2011	Low-density array	Azoxymethane (AOM) injection (15 mg/kg bw)	Dietary	11.5% fish oil (ad-libitum)	Rat distal colon	RNA	cancer progression

Table 2. The experimental details associated with n-3 PUFA plus fermentable fiber differentially expressed target genes

n-3 PUFA plus fermentable fiber								
Up regulated								
Gene symbols	Reference	Assay type	Carcinogenesis method	Treatment method	Dose	Organism	Level	Cancer_stage/classification
RBBP6, ID3, BRMS1, MTMR4, MGMT, PMS2	Cho et al., 2011	CodeLink Rat Genome Arrays	Azoxymethane (AOM) injection (15mg/kg body weight)	Dietary	11.5% fish oil (ad-libitum)	Rat distal colon	RNA	tumor
Down regulated								
Gene symbols	Reference	Assay type	Carcinogenesis method	Treatment method	Dose	Organism	Level	Cancer_stage/classification
HIPK2, FEM1B, SLC8A1, PTHR2, DUPD1, IL6R, MFN1, SMOC1, TMEM23, HGF, IL23A, STX1A, ADAM3, PPP1R7, CYP2S1, NRN1, MMP2, SNIP	Cho et al., 2011	CodeLink Rat Genome Arrays	Azoxymethane (AOM) injection (15mg/kg body weight)	Dietary	11.5% fish oil (ad-libitum)	Rat distal colon	RNA	tumor

NF- κ B activity can be inhibited by DHA (73). This is relevant because NF- κ B mediates signaling pathways that control the transcriptional activation of genes important for the regulation of many cellular processes and is aberrantly activated in many types of cancer (74,75). n-3 PUFA treatment inhibits the expression and activity of NF- κ B in many cell types, however, the exact mechanism is not fully understood (66). This has implications in chronic disease management because the DHA - mediated decrease in NF- κ B activity has been shown to sensitize tumor cells to gamma-irradiation and promote the induction of apoptosis (48).

1.3.3 DHA-mediated modulation of apoptosis regulatory pathways

It has been demonstrated that DHA contributes to the down-regulation of BCL2, a well-known anti-apoptotic molecule (76), which can block lipid peroxidation and thus apoptosis induction. Additionally, DHA induces caspase-dependent apoptosis in colon adenocarcinoma cells and adenoma cells (77). There is also evidence of upregulation of CASP4 and CASP7 (64) along with increased activation of the intrinsic apoptotic pathway as demonstrated by CASP9 and Bid cleavage (77). CASP4 activation may also be linked to augmented expression of ER resident factor ERdj5 and downregulation of anti-apoptotic GRP78 (78). The major involvement of the intrinsic apoptotic pathway following DHA treatment is through increased expression and activation of BAX and BAK (66), depolarization of the mitochondrial membrane, and the subsequent release of cytochrome c and Smac/Diablo into the cytosol (79). Once these factors are released from mitochondria, apoptosis is accelerated (80). These findings have been confirmed both in vitro (81) as well as in vivo (82).

n-3 PUFA can also act as efficient modulators of both the level and activity of endogenous caspase inhibitors. For example, DHA and EPA decrease XIAP (an X-linked inhibitor of apoptosis protein) at both the protein and mRNA levels, which may in part explain their antineoplastic effects (79). High XIAP expression correlates with poor clinical outcome, resistance to chemotherapy and radiotherapy in different colon cancer

cell lines (79). DHA also down-regulates mRNA and protein levels of two other inhibitors of apoptosis, survivin (BIRC5) and livin (BIRC7) in cancer cells (66). Furthermore, the immediate and dramatic down-regulation of FLIP (CFLAR), a potent inhibitor of caspase-8 (CASP8) activation, appears to be linked to the induction of apoptosis in colon cancer cells following DHA and EPA supplementation (79).

DHA can inhibit the expression of antioxidant enzymes or deplete cells of antioxidants (83). It has also been suggested that DHA may have anti-inflammatory/pro-apoptotic effects in colon cancer cell lines by inhibiting the expression and activity of a key rate-limiting cyclooxygenase enzyme, COX-2 (84). This is noteworthy, because COX-2 is often overexpressed in colon tumors and is able to confer a pro-inflammatory niche, which contributes to epithelial cell resistance to apoptosis (85,86). Activation of NF- κ B and the PPAR-BCL2 feedback loop may control the life-death continuum in colon cells and has been associated with the expression of COX-2 (85). Chemoprotective suppression of the activation of NF- κ B by DHA reduces the production of pro-proliferative eicosanoids produced by COX-2 (87). Moreover, DHA may suppress tumor cell growth directly by inhibition of the COX-2 derived metabolite, PGE₂, which stimulates cell proliferation and suppresses apoptosis (86). However, it is possible that DHA may also act via mechanisms independent of COX-2 inhibition (88), because suppression of tumor growth also occurs in cell lines that do not express COX at the protein level. Moreover, the growth of these cells in culture and in nude mice is not affected by overexpression of COX-1 or COX-2 (89). Additional DHA-dependent pro-apoptotic mechanisms impacting colon adenocarcinomas include the upregulation of several growth arrest DNA-damage-inducible proteins such as GADD445A and GADD45B, likely through the stimulation of p38 MAPK phosphorylation (66).

1.3.4 Modulation of cytokines and growth factors

Cytokines, including IL1 β , IL2, IL4, IFN γ , and TNF α increase in the early stages of carcinogenesis. n-3 PUFA suppression of NF- κ B activity is at least partly responsible

for the reduction in cytokine levels, including IL2, IL4, IFN γ , TNF α , IL6, and IL1 β (90,91). These cytokines (IL1 β , TGF β , TNF α and IL6) further regulate transcription factor, e.g., HNF4A, function through modulation of proteosomal degradation, DNA binding affinity, transcriptional activity and cofactor interaction (92)*. Thus, n-3 PUFA cytokine regulatory control can extend to ion transport, epithelial barrier function and oxidative stress via effects on this transcription factor.

The protective role of n-3 PUFA can also be attributed to an increase in the expression of TGF β through inhibition of the Akt pathway in intestinal epithelial cells (93) and fat-1 transgenic mice (94). This is noteworthy, because the reduction of TGF β expression increases chemical-induced colon carcinogenesis (95). Furthermore, both EPA and DHA decrease the growth of colon tumors by reducing VEGF and TNF α expression through inhibition of ERK1/2 phosphorylation and hypoxia-induced factor HIF1 α protein expression (96).

1.3.5 Effects of DHA on the cell cycle

There is some evidence that DHA has a selective dose dependent growth inhibitory effect on colon cancer but not normal colonic cells (97). Several key genes involved in the regulation of both the G1 and G2 phases of the cell cycle are affected by DHA treatment in colon cancer. Generally, molecules involved in cell cycle progression, such as Cdc25c, Cdc25b, Cdc20, CDK1, CDK2, and cyclin D, A, and B, are down-regulated (66) by DHA incubation as compared to control. In comparison, genes involved in cell cycle arrest such as cyclin-dependent kinase inhibitors (CDKN1A, CDKN1B, CDKN1C, CDKN2A) and stratifin are up-regulated by DHA (98). Some studies additionally show that activated PXR inhibits the proliferation and tumorigenicity of colon cancer cells by targeting the cell cycle at the G(0)/G(1) cell phase via modulation of the p21(WAF1/CIP1) and E2F/Rb signaling pathways (58). In addition, in some cell contexts, DHA induces cell cycle arrest and down-regulates the

nuclear form of sterol regulatory element-binding proteins (SREBP1 and 2) in colon cancer cell lines, indicating a possible relationship between disturbances in lipid homeostasis and cell cycle arrest (64,65,72). While a large number of mechanisms are linked to DHA anti-proliferative effects in cancer, several reports have focused on whether p53 protein plays a role in DHA- induced growth inhibition. DHA inhibits the growth of p53-wildtype colon cell lines as well as of those with inactivating p53 mutations; thus its action does not seem to be dependent on p53 status (72).

1.3.6 Optimal chemoprevention: Interaction of DHA with butyrate

It has been proposed by us and others that n-3 PUFA and butyrate (fiber fermentation product) interact in the colon to profoundly suppress colon cancer (99-101). Interaction of dietary fiber-derived compounds in the colonic lumen can have a substantial impact on the metabolism and kinetics of the colon epithelial cell population and suppress inflammation and neoplasia (102-104). For example, butyrate, a four-carbon short-chain fatty acid, is produced during anaerobic fermentation of dietary fiber by endogenous bacteria present in the colon. This agent has pleiotropic effects in the colon (105,106). It acts as a principal energy source and a survival factor for normal colon cells, whereas it exerts anti-proliferative, differentiation- and apoptosis-inducing effects in cancer cells (107). In addition to the regulation of basic cytokinetic processes, butyrate has also been shown to affect cell adhesion, morphology, invasiveness, metastasis, oxidative metabolism, angiogenesis, and the activity of different enzymes and transcription factors. These effects are linked in part to butyrate's function as a histone deacetylase inhibitor, which mechanistically links it to gene expression (108).

Studies published by our group describe the protective effects of fish oil containing DHA, compared to corn oil and its interaction with fiber using rat and mouse model colon carcinogenesis models (31,35). These data demonstrate that the combination of n-3 PUFA and butyrate (fermentable fiber) treatment maximally enhances cell cycle arrest, by inhibiting expression of cell cycle genes (**Table 2**), shifting

the balance between differentiation and apoptosis depending on the cell transformation status of the model (104,109,110). These findings demonstrate that dietary n-3 PUFA and fermentable fiber can act synergistically to protect against colon carcinogenesis primarily by enhancing the deletion of DNA damaged cells (86,100,101,111,112).

Temporal gene expression profiles from exfoliated rat colonocytes have revealed at the cancer initiation stage that fish oil plus fermentable fiber (FO/F) downregulates the expression of genes involved with cell adhesion and enhances apoptosis compared to the non-chemoprotective control of corn oil plus cellulose (CO/C) (109). In addition, at the cancer progression stage, the expression of genes involved in cell cycle promotion are downregulated while DNA mismatch repair genes, MGMT and PMS2, are upregulated. FO/F also increases apoptosis and the expression of genes that promote apoptosis at the tumor stage (109). The chemoprotective gene profiles at the tumor stage include the up-regulation of the pro-apoptotic inhibitor of DNA binding ID3 and tumor suppressors BRMS1 and RBBP6, and also downregulation of anti-apoptotic genes HGF and TMMEM23, and down-regulation of cytokine signaling, IL23A and receptor IL6RA (109). Signal transduction related genes such as MAPK, DUPD1 and PPP1R7, and calcium signaling receptor SLC8A1 were also downregulated (109). In addition, the chemotherapeutic effect of the FO/F dietary extends to translational activation of the xenobiotic metabolizing phase I enzyme EPHX2 and tumor suppressor retinoblastoma-associated protein RB1. These novel findings demonstrate that the effects of the chemotherapeutic (FO/F) diet on epithelial cell gene expression can be monitored noninvasively throughout the tumorigenic process by analysis of exfoliated colonocytes.

1.4 Summary and purpose of the study

n-3 PUFA from fish oil and fermentable fiber are ideal colon cancer chemotherapeutics because (1) they are toxicologically innocuous and free of safety problems intrinsic to drugs administered over long periods of time, (2) are relatively inexpensive, and (3) provide additional health benefits, such as reduction in mortality

(69,113,114). In addition, the simultaneous ingestion of fish oil and fermentable fiber can improve their efficacy in colon cancer prevention/therapy (111,112,115-117).

From an epigenetic perspective, there is still much to be discovered in terms of the effects of n-3 PUFA and fermentable fiber in the colon at the transcriptional and chromatin state level. In this study, we employ novel technologies and bioinformatics algorithms, such as next-generation sequencing, in order to explore 'nutri-epigenomics' at a genome-wide level by using an *in-vivo*, pre-clinical model to better comprehend the importance of epigenetic mechanisms related to chemoprevention. Our objective is to contribute to the understanding of the regulatory action of chemoprotective bioactive compounds found in fish oil and readily fermentable fiber (n-3 PUFAs and SCFAs) in terms of malignant transformation of the colonic crypt. Mechanistic insights gained in this study are needed to interpret human clinical and epidemiological findings (118,119).

1.5 Hypotheses and specific aims

Aim 1: Determine chromatin structure associated with differential expression and phenotypic responses in the distal versus proximal colon epithelial crypts by correlating ChIP-Seq to mRNA transcription data.

Hypothesis 1.1: Gene expression at the mRNA level in the distal versus proximal colon will differ in a manner correlated with histone tail modifications.

Hypothesis 1.2: Genes known to be hypermethylated (ESR1, MYOD1) at the DNA level in the proximal colon will also exhibit higher levels of H3K9me3. Genes known to be reciprocally expressed in the proximal vs distal colon, e.g., PPAR-gamma, L-FABP, will contain higher levels of H3K4me3 at the promoter region (a histone tail modification associated with active transcription) and/or lower levels of H3K9me3 a histone tail modification often found in heterochromatin.

Aim 2: Monitor early pre-tumorigenic epigenetic molecular events driving colon cancer by generating high-resolution genome-wide transcriptional and "chromatin-state"

maps for intestinal epithelial cell crypts in rats injected with a colon-specific carcinogen or saline.

Hypothesis 2.1: Carcinogen (AOM) will perturb transcription (RNA) and chromatin state in a proto-oncogenic manner. Tumor suppressing genes, including clinical biomarkers of colon cancer such as FOXO3, H2AFX, MSH2, NR3C1, and PDCD4 will be down regulated.

Aim 3: Gain epigenetic insight into the mechanisms of nutritional chemoprevention by generating high-resolution transcriptional genome-wide “chromatin-state” maps for intestinal epithelial cell crypts in rats (i) injected with a colon-specific carcinogen or saline, and (ii) fed a diet containing fish oil (n-3 PUFA), (iii) pectin (butyrate) or (iv) a combination of both fish oil and pectin.

Hypothesis 3.1: Fish oil (n-3 PUFA) will modulate intestinal genome-wide chromatin state maps, in part, by suppressing pro-inflammatory gene pathways and activating tumor suppressing ligand induced nuclear receptors (e.g., PPARs, FXR, RXR).

Hypothesis 3.2: Fermentable fiber (pectin) will suppress colon cancer risk by altering acetylated histone modifications (H3K9ac) at the promoter and throughout the body of carcinogen-associated genes. Butyrate will act as a Histone Deacetylase Inhibitor and/or as an acetylation substrate and induce chromatin hyperacetylation at the promoter region of genes associated with apoptosis and tumor suppression (e.g., TP53, CEBPA, RB1).

Hypothesis 3.3: The interaction of DHA + But (from fish oil + pectin) in the diet will modulate intestinal genome-wide transcriptional and chromatin state maps to synergistically suppress tumorigenesis. This will be accomplished, in part, by altering gene expression and chromatin acetylation and methylation of genes associated with apoptosis and lipid metabolism (e.g., FABP1, AQP8).

CHAPTER II
GENOME-WIDE ANALYSIS OF THE RAT COLON REVEALS PROXIMAL-
DISTAL DIFFERENCES IN HISTONE MODIFICATIONS AND PROTO-
ONCOGENE EXPRESSION*

2.1 Introduction

The complexity of cellular chromatin structure is in part dictated by the Histone Code, consisting of a variety of histone tail post-translational modifications (PTMs) that can be found in each nucleosome. For example, actively expressed genes are associated with decreased gene-suppressive histone H3 lysine 9 trimethylation (H3K9me3) modifications (120). Conversely, gene-activating H3K4 trimethylation (H3K4me3) promotes mRNA expression. This pattern is observed in nucleosomes near transcription start sites (TSSs) of active and stalled protein-coding and non-coding RNAs that are enriched for H3K4me3. Collectively, these and other modifications are now considered indicators of nascent gene and non-coding RNA transcripts (121).

Many diseases of the colon exhibit an anatomical bias. For example, in the distal colon these maladies include ulcerative colitis, chromosomal instability (CIN) induced cancer, Crohn's disease, and diverticulitis. Contrastingly, in the proximal colon, ischemic colitis, collagenous colitis and microsatellite instability (MIS) induced cancer, are highly prevalent (122-124). Furthermore, proximal and distal colorectal cancers have been reported to exhibit distinct gene-specific methylation profiles, transcriptional profiles, and molecular and clinical characteristics (125-127). .

* Portions of this chapter reprinted with permission from Karen Triff, Kranti Konganti, Sally Gaddis, Beiyan Zhou, Ivan Ivanov, and Robert S. Chapkin (2013) Genome-wide analysis of the rat colon reveals proximal-distal differences in histone modifications and proto-oncogene expression. *Physiological genomics* 45, 1229-1243 Copyright 2013 by APS

Since there are no cell lines that uniquely represent the proximal or distal colon, the need to identify regional epigenetic differences at the chromatin level using an in vivo model remains unfulfilled.

The rodent colon tumor system is a valuable model for studying dietary and environmental modifiers of human colon cancer risk (99,128). For example, the rat azoxymethane (AOM) carcinogenesis induced model mimics the neoplastic processes observed in humans during both the initiation and promotion/progression stages of carcinogenesis, including similarities in (i) light and electron microscopic morphology, (ii) histochemical properties, and (iii) biological behavior of colon carcinomas (129,130). However, despite the advantages of animal models, genome-wide assessment of chromatin state and transcriptional regulation has been hampered because of the challenges associated with chromatin immunoprecipitation of organ tissues for high throughput sequencing.

To date, a genome-wide histone PTM occupancy analysis has not been performed using an in-vivo model of intestinal development. In vitro studies from cell lines in which chromatin immunoprecipitation, combined with high throughput sequencing (ChIP-Seq), provide access to only the subset of chromatin PTMs that are active in a given cell type under culture standardized conditions. This limits insight into their in vivo function. Therefore, we determined the chromatin structure associated with gene expression profiles in the rat proximal and distal colon by globally correlating chromatin immunoprecipitation next generation sequencing analysis (ChIP-Seq) with mRNA transcription (RNA-Seq) data. This characterization provides the first available in vivo genome-wide map of histone PTMs in the rat colon with respect to the relationship between chromatin alterations and gene transcription. For this purpose, we globally identified the transcriptome and sites of H3K4 trimethylation (H3K4me3) in colonic crypt epithelial cells. By assessing the correlation between histone PTMs and transcriptional data, we identified canonical pathways influenced by differential H3K4me3 occupancy and RNA expression. We were also able to catalogue the middle and long non-coding RNAs transcribed in the colon, including select long non-coding

RNAs (lncRNAs) formerly only detected in the rat nervous system. In addition, in silico analysis of transcription factor (TF) activity revealed tumor suppressors and oncogenes unique to the distal and proximal colon. Knowledge of the proximal versus distal colon epigenetic landscape will assist in the improved detection, therapy and prognoses of colonic disease.

2.2 Methods

2.2.1 Animals

Seven male Sprague–Dawley rats, 14 weeks of age, were individually housed in the same room and maintained in a temperature and humidity-controlled animal facility with a daily 15 h light/9 h dark photoperiod. The animal use protocol was approved by the University Animal Care Committee of Texas A&M University and conformed to NIH guidelines. Rats were fed a semi-purified diet containing (grams/100 gram diet): dextrose, 51.00; casein, 22.40; D,L-methionine, 0.34; AIN-76 salt mix, 3.91; AIN-76 vitamin mix, 1.12; choline chloride, 0.13, pectin, 6.00, and 15.0 g corn oil/100 g diet.

2.2.2 Isolation of colonic crypts

Rats were terminated by CO₂ asphyxiation, followed by cervical dislocation. The large intestine was resected from the junction between the cecum and the rectum, and was opened longitudinally and washed in 1× PBS. Subsequently, the visible “herring bone” folds were used to identify the proximal colon. The region distal to this point was referred to as the distal colon. The colonic sections were incubated with HBSS (Hank’s Balanced Salt Solution, without Ca⁺⁺ and Mg⁺⁺) containing 1 mM glutamine, 0.1% BSA, 30 mM EDTA, and 5 mM DTT, adjusted to pH 7.5, for 15 min in a 37°C

shaking incubator as we have previously described (131). Following incubation, tissue sections were placed in a petri dish on ice, and the colonic crypts isolated by scraping with a rubber policeman. Isolation of crypts was verified by histological examination to ensure that epithelial cells were removed and the lamina propria and muscle layers remained intact. Cells were washed with HBSS and centrifuged at $100 \times g$, for 15 min. The pellet was resuspended in HBSS and an aliquot of the crypts was subsequently used to create mRNA expression profile libraries. The remaining crypts were immediately crosslinked for CHIP analysis.

2.2.3 Western blot

Colonic crypts were rocked in 50 mM HEPES-KOH, pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40, 0.25% Triton X-100 with protease inhibitors for 10 min at 4°C, followed by centrifugation at 1,350g, 4°C for 7 min. The crypt-containing pellet was subsequently resuspended in 10 mM Tris-HCl, pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA with protease inhibitors and incubated by gently rocking at room temperature for 10min. Nuclei were pelleted by centrifugation at 1,350g, 4°C for 7min and resuspended in RIPA buffer (50 mM HEPES-KOH, pKa 7.55, 500 mM LiCl, 1 mM EDTA, 1.0% NP-40, 0.7% Na-deoxycholate with protease inhibitors). Nuclear lysates (2 ug protein) were treated with 1× pyronin sample buffer and subjected to SDS polyacrylamide gel electrophoresis (PAGE) in precast 4–20% Tris-glycine mini gels (Invitrogen). After electrophoresis, proteins were electroblotted onto a polyvinylidene fluoride membrane with the use of a Hoefer Mighty Small Transphor unit at 400 mA for 90 min. Following transfer, the membrane was incubated in 5% milk and 0.1% Tween 20 in TBS (TBST) at room temperature for 3 h with shaking, followed by incubation with shaking overnight at 4°C with primary antibody diluted in 5% milk in TBST. Membranes were washed with TBST and incubated with secondary peroxidase conjugated secondary antibody as per manufacturer's instructions. Bands were developed using Pierce SuperSignal West Femto™ maximum sensitivity

substrate. Blots were scanned using a Fluor-S Max MultiImager system (Bio-Rad, Hercules, CA). Quantification of bands was performed using Quantity One software (Bio-Rad). Primary antibodies were used against histone H3 tri methyl K4 (ab1012) and histone H3 (ab1791). Peroxidase conjugated goat anti-rabbit IgG was purchased from Kirkegaard and Perry Laboratories (Gaithersburg, MD).

2.2.4 Chromatin immunoprecipitation and sequencing

ChIP-Seq analyses were performed in order to determine global histone mapping in crypts isolated from the proximal versus distal colon (**Figure 2A**). A modified version of the ChIP protocol described by Lee et al (132) was utilized. Specifically, cells were cross-linked by adding freshly prepared formaldehyde at a 1% concentration for 25 min at room temperature followed by quenching with glycine. Cells were then lysed and sheared in 3 mL tubes at 4°C using a Covaris S2 sonicator to obtain DNA distributions of ~300 bp (range of 200–400 bp): duty cycle –20%, intensity –8, cycles/burst –200 and time 40 min. ChIP antibodies included: ChIP Grade (Abcam, ab8898) anti-histone H3 (tri methyl K9) antibody, ChIP Grade (ab1012) anti-histone H3 (tri methyl K4) antibody, ChIP Grade (ab5408) RNA Polymerase II CTD repeat YSTSPS phospho Ser-5 antibody. The specificities of all antibodies were tested by Western blot and ChIP-qPCR. Antibody–chromatin complexes were captured using Dynabeads G Protein coupled (Dyna) and eluted with 1% SDS in 50 mM Tris-HCl pH8.0 10, and 10 mM EDTA, after extensive washing. Cross-linking between DNA and chromatin proteins was reversed by incubation at 65°C overnight. DNA was purified by QIAquick PCR Purification Kit (Qiagen 28004) and dissolved into 50 µL EB (10 mM Tris pH8.0) buffer per immunoprecipitation. Equal amounts of ChIPed DNA from each of the 7 rats were pooled and used for high throughput sequencing.

2.2.5 Bioinformatic analyses

Sequence reads were aligned to the rat genome (rn4) using standard Illumina Pipeline Analysis software and only non-identical uniquely mapped reads were retained. Sequence read numbers were summed into 200 bp non-overlapping windows with read position shifted 75 bp to represent the DNA fragment center. Output data from the pipeline were converted to browser extensible data (BED) files for viewing in the UCSC genome browser. Genomic regions with histone modification marks were identified using SICER with an E-value of 100 (133). ChIP-Seq data are presented from both BED files and SICER-processed BED files. A submodule of the SICER program was employed to select regions that exhibited significant differences in the tag counts between the two colonic regions (P -value < 0.00001 ; $FC > 2$). The nearest gene to each island, i.e., within 2 kb of the island, was identified using closestBed from the BEDTools software suite (134) and the refGene table downloaded from the UCSC Genome Browser for the Baylor 3.4/rn4 assembly (135).

2.2.6 RNA isolation

For total RNA isolation, colonic crypts were homogenized on ice in lysis buffer (RNAqueous Isolation kit, Ambion) and frozen at -80°C until RNA was isolated. Subsequently, total RNA was isolated using the RNAqueous kit, followed by DNase treatment. RNA integrity was analyzed on an Agilent Bioanalyzer to assess RNA integrity.

2.2.7 RT-qPCR and ChIP-qPCR

In order to assay enrichment of select regions in ChIP DNA and to confirm H3K4me3 and RPOLII binding at transcription start site (TSS) regions of interest,

quantitative PCR (qPCR) was carried out in triplicate with primers specific for these regions using SYBR Green Mastermix (Applied Biosystems). Data are presented as fold-enrichment using the ddCT method comparing the ChIP antibody signal to a negative control IgG for the distal and the proximal colon regions. Fold enrichment = $2^{(Ct_{IgG}-Ct_{IP})_{distal\ colon}/2^{(Ct_{IgG}-Ct_{IP})_{proximal\ colon}}$ (136).

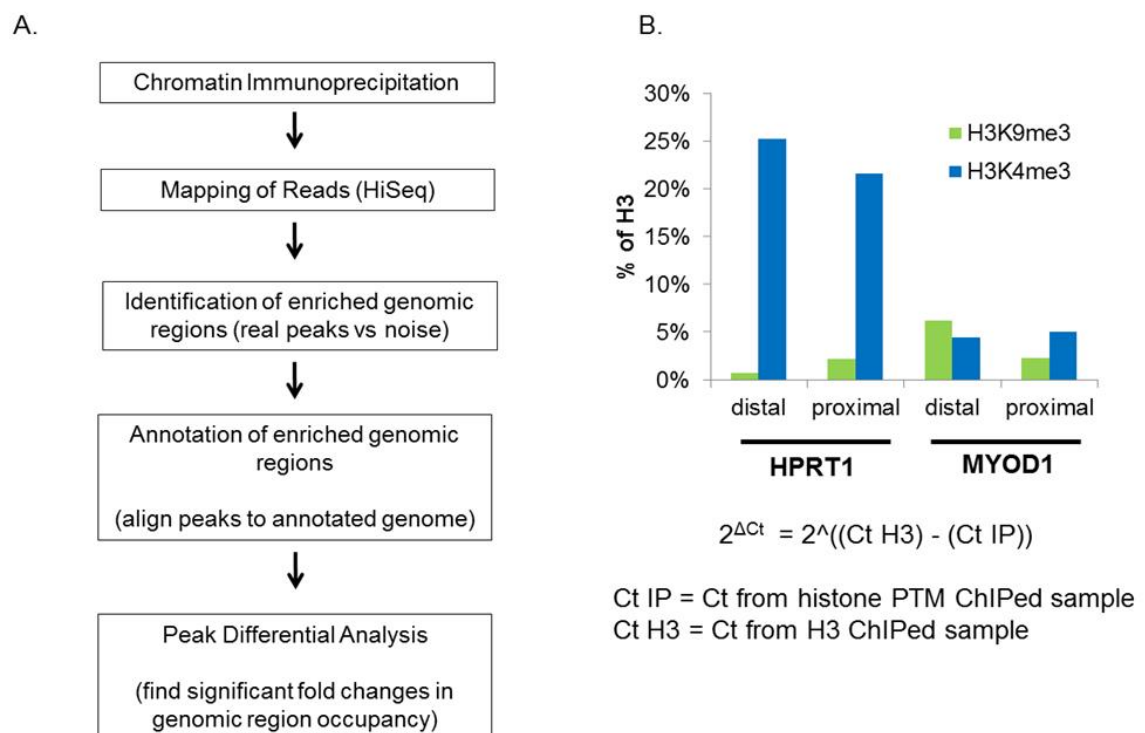


Figure 2. A. ChIP-Seq bioinformatics analysis workflow. B. Histone PTM binding at the transcription start site of representative active and silent genes.

Primers used for ChIP-qPCR include; EPHB3 near TSS (+15) FP 5'-CGAGGACCAGCAGAAGTGAG-3' and, RP 5'-GGGGTAGGAGCCGGTATCA-3', EPHB3 downstream (+18605) FP 5'-CCAGAGACTGACTCAGAGAGC-3' and RP 5'-AGCTTGTACACACTCATGCT-3'. FABP1 near TSS (+419) FP 5'-

CATTATTCAAGGCCAACCATATCTC-3' and RP 5'-
 CCCCATCTGTCTCCCTGGTA-3', FABP1 downstream (+2541) FP 5'-
 CATGAATGCTGGAAAACCTGAA-3' and RP 5'-TGGTGCCGATTCTAAGCTGAA-
 3'. HPRT1 near TSS (+579) FP 5'-GCCTGGGCTTTGCCTCTAAT-3' and RP 5'-
 GGAGCAATTGCGCATTCC-3', HPRT1 downstream (+5284) FP 5'-
 GGCGGCTATTCAGGAACTCTCT-3' and RP 5'-
 GATGGTATGGTGGATGGTTTGG-3'. MYOD1 near TSS (+224) FP 5'-
 TTGGGTTGAGCGAGAAGCA-3' and RP 5'-CGGAGTGGCGGCGATA-3', MYOD1
 downstream (+2031) FP 5'-TGCCCCTGGGTCCTTCAT-3' and RP 5'-
 GGACAATTGGGAGGAGTGTCA-3'. NFIB near TSS (+152) FP 5'-
 GAAGTAGTGAGGAGTTGCGG-3' and RP 5'-CTGGGATGGGCGTATAAGGT-3',
 and NFIB downstream (+18239) FP 5'-CTGTGGTGTGCCTTTTCTCT-3' and RP 5'-
 TATCGTCTGAACGGCAACTG-3'.

To assay for gene expression, cDNA was synthesized from 100 ng of total RNA using random hexamers and oligo dT primers with Superscript II reverse transcriptase (Invitrogen, Carlsbad, CA). PCR was performed (primer sequences available online) using Taqman rat assays and Taqman master mix (Applied Biosystems, Foster City, CA) and normalized to HPRT1. Data are presented as fold enrichment for the gene of interest versus HPRT1 using the ddCT method comparing the distal to proximal colon (137).

2.2.8 RNA-Seq

Equal amounts of RNA from the distal or proximal colon from individual rats were pooled and the 2 samples were used for high throughput sequencing. A total of 150 million (74 million proximal and 76 million distal colon) 40 bp single-end Illumina reads were obtained from a multiplexing run on a single lane. Reads were filtered for adapter sequences and trimmed based on sequence quality (threshold of Q20) using ea-utils toolkit (<https://code.google.com/p/ea-utils/>). Filtered reads less than 20 bp in length

were discarded. Spliced alignment was performed against the rat genome (rn4) using tophat2 (138) without allowing any mismatches to reduce the number of false positives and the resultant alignments were further processed using cufflinks2 (139) to perform reference annotation based transcript assembly with bias and multi-read correction. Differential expression analyses were performed using cuffdiff2 and cummeRbund (140). The cuffcompare program of cufflinks2 pipeline (reference annotation based method) was used to classify the assembled transfrags into different classes. The transfrags annotated with class codes **u** (unknown intergenic transcript) and **i** (a transfrag falling entirely within a reference intron) were selected to annotate against known lncRNA's. A custom perl script was written to fetch sequences for the coordinate information produced by cuffcompare. [NONCODE v3](#), UCSC, and [fRNAdb v3.0](#) databases were downloaded and only non-coding RNA (ncRNA) sequences belonging to the rat genome whose length was greater than 100 bp were used to create a unified database (141-143). The fetched sequences from UCSC were then queried for ncRNA sequence similarity with blastn by thresholding at 1e-10 (144,145). High confidence hits, % with high-scoring segment pairs (HSPs) length greater than 100 and whose identity was greater than 95% were retained. Whenever possible, BLAST results were further filtered to remove hits that did not correspond to known ncRNA coordinate information. Due to the consistent lack of coordinate information for the known ncRNA sequences, the results include matches to multiple genomic locations.

2.2.9 Functional analyses

Ingenuity Pathway Analyses. “Functional enrichment” analysis was performed using Ingenuity Pathway Analysis (IPA) version 2.0 software (Ingenuity Systems Inc., Redwood City, CA). To perform IPA analysis, all differentially expressed genes (adjusted $P < 0.005$) in the distal or proximal colon were uploaded into three columns for the purpose of generating Illumina probe ID, t-value (fold change) and adjusted P-value (FDR) data. By convention, genes that were up-regulated in the distal colon are

shown in red and genes that were down-regulated are shown in green. By default, during IPA analysis, only molecules from the data set associated with the Ingenuity Knowledge Base repository (Ingenuity Systems Inc.) were considered. Functional Analysis identified the biological functions and/or diseases that were most significant to the data set. The significance of the association between the data set and the specific pathways of interest was determined in three ways: (1) ratio of the number of molecules from the data set that mapped to the pathway divided by the total number of molecules that mapped to the Ingenuity Knowledge Base pathway, (2) Fisher's exact test was used to calculate a *P* value determining the probability that the association between the genes in the data set and the pathway of interest could be explained by chance alone, and (3) activation state (“Increased” or “Decreased”) was inferred by the activation z-score. The derivations of the z-scores are based on relationships in the molecular network that represent experimentally observed causal associations between genes and those functions.

“Canonical pathway” analysis was used to identify networks from the IPA library that were most significantly modulated across anatomical sites. Significance of the association between each data set and the canonical pathway was measured in 2 ways: (1) A ratio of the number of molecules from the data set that mapped to the pathway divided by the total number of molecules that mapped to the canonical pathway, and (2) Fisher’s exact test was used to calculate p-values determining the probability that the association between genes in the dataset and each canonical pathway was explained by chance alone.

“Upstream regulator” analysis was based on prior knowledge of expected associations between transcriptional regulators and their target genes stored in the Ingenuity® Knowledge database, and significance for each TF was measured in 2 ways: (1) Fisher’s Exact Test (p-value) was used to identify differentially expressed genes from the RNA-Seq data set which overlapped with genes known to be regulated by a TF. Since the regulation direction (“activating” or “inhibiting”) of an edge is not taken into account for the computation of overlap p-values, the underlying network also included

findings without associated directional attribute, such as protein-DNA (promoter) binding. In addition, the activation score (z-score) was used to infer the status of predicted transcriptional regulators by comparing the observed differential regulation of genes (“up” or “down”) in the dataset relative to the literature-derived regulation direction which can be either “activating” or “inhibiting”.

2.3 Results

2.3.1 Constructing an atlas of the distal and proximal colon epigenomes

To characterize genome-wide changes in histone trimethylation and gene transcription, we compared crypts isolated from the rat proximal and distal colon, as separated by the hepatic flexure. The overall experimental strategy is shown in **Figure 3**. Chromatin immunoprecipitation (ChIP) was performed using antibodies recognizing trimethylated histone H3 at lysine residues 9 or 4 (H3K9me3 and H3K4me3) and active RNA Polymerase II (RPOLII). The immunopurification of soluble chromatin was assessed for enrichment using qPCR at genomic regions consistent with active and inactive chromatin states using *HPRT1* and *MYOD1* as positive and negative markers, respectively. As expected, enrichment of H3K4 trimethylation at the TSS was associated with the active expression of the *HPRT1* gene (**Figure 2B**). Next, we analyzed hypertrimethylated sequences of immunopurified chromatin using the Illumina HiSeq platform. Proximal and distal colon libraries generated from H3K9me3, H3K4me3, and RPOLII ChIP were sequenced on 3 lanes, generating >450 million 40-bp tags. Following sequencing, the reads were mapped and annotated as outlined in **Figure 2B**.

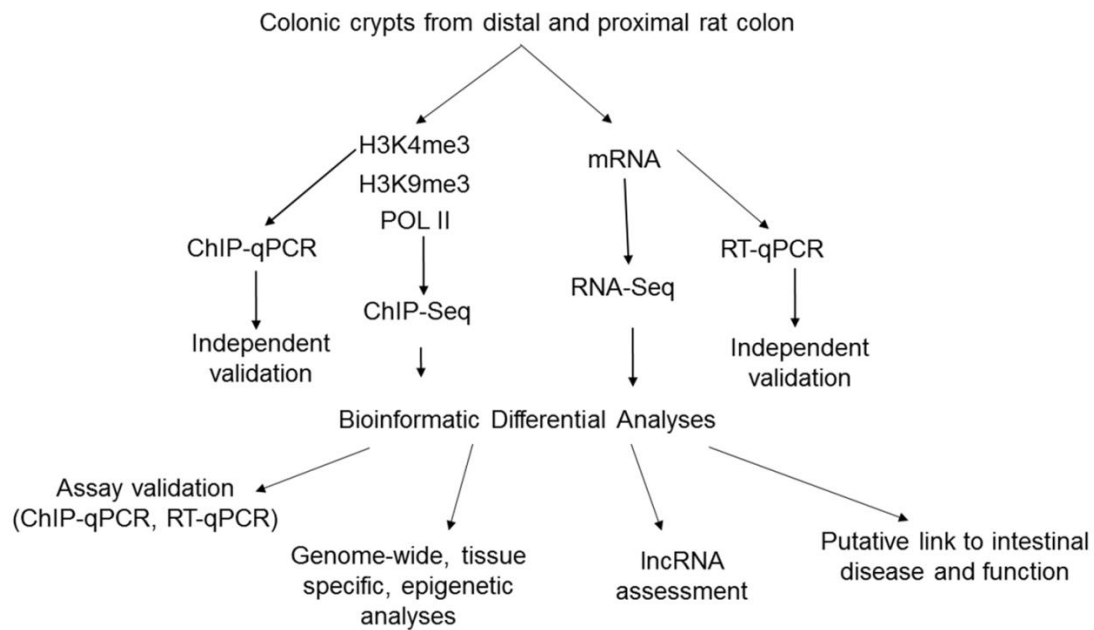


Figure 3. General schematic of the experiments and analyses performed

The trimethylation of H3K4 was regionalized close to the TSS and up to 5 kb downstream from actively transcribed genes (ENCODE Project Consortium 2007), which typically extends into the gene body and is associated with RPOLII enrichment (146,147). Plotting the density of all generated sequence tags versus the entire set of RefSeq TSS positions affirmed that H3K4 trimethylation was predominantly regionalized (5 kbp upstream/downstream) in these genes (**Figure 4**).

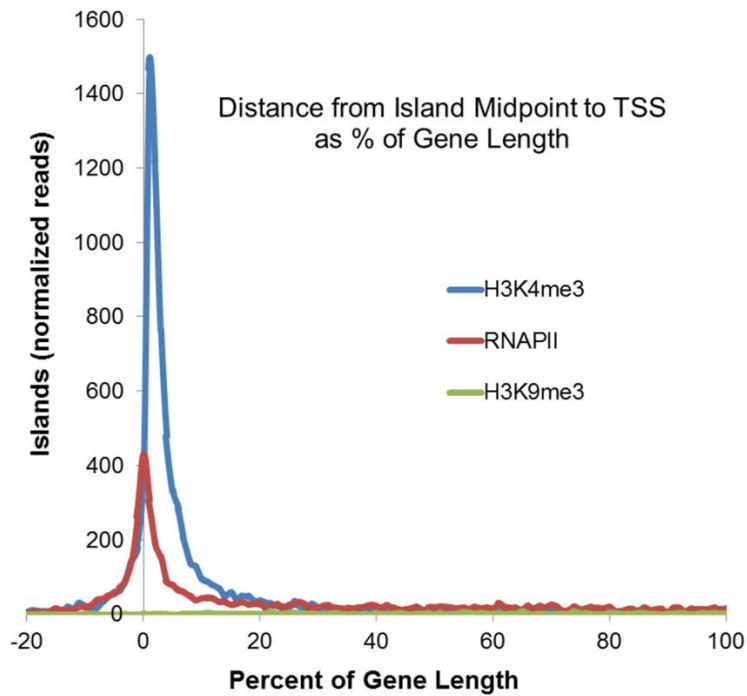


Figure 4. Genome wide binding of histone modifications and RNA Pol II relative to the transcription start site (TSS) of annotated genes

As anticipated, H3K9me3 peaks were not prevalent at the TSS of genes and instead were primarily associated with heterochromatic (71%) and intragenic (14%) regions of the genome. Greater than 18,000 regions across the genome were enriched for H3K4me3 and over 8,000 regions for H3K9me3 (**Table 3**). Sequence tags generated from proximal and distal colonic crypt ChIP-Seq were analyzed in 150-bp bins. ChIP-Seq generated tags were subsequently aligned to the rat genome reference m4.

Table 3. Annotation of histone ChIP-Seq island peaks using SICER software..

Annotation Summary
Proximal H3K4me3: 92.8% of islands with a score ≥ 100 are located within 2 kb of a gene Islands: 18,797 Islands with threshold score ≥ 100 : 11,171 Islands with threshold score ≥ 100 within 2 kb of gene: 10,368
Distal H3K4me3: 92.2% of islands with a score ≥ 100 are located within 2 kb of a gene Islands: 20,938 Islands with threshold score ≥ 100 : 11,707 Islands with threshold score ≥ 100 within 2 kb of gene: 10,794
Proximal H3K9me3: 61.3% of islands with a score ≥ 100 are located within 2 kb of a gene Islands: 8,826 Islands with threshold score ≥ 100 : 1,410 Islands with threshold score ≥ 100 within 2 kb of gene: 865
Distal H3K9me3: 60.0% of islands with a score ≥ 100 are located within 2 kb of a gene Islands: 11,745 Islands with threshold score ≥ 100 : 1,502 Islands with threshold score ≥ 100 within 2 kb of gene: 907
Distal RNA Polymerase II: 81.3% of islands with a score ≥ 100 are located within 2 kb of a gene Islands: 35940 Islands with threshold score ≥ 100 : 8200 Islands with threshold score ≥ 100 within 2 kb of gene: 6667

The total number of islands, number of islands with a minimal intensity score ≥ 100 , and number of islands within 2 kb of an annotated gene are described. H3K4me3, histone H3 lysine 4 trimethylation; H3K9me3, histone H3 lysine 9 trimethylation.

In total, 10,898 regions in the genome with H3K4 trimethylation peaks were compared with the input data set. Based on anatomical location, 857 regions exhibited significant hypertrimethylation, while 98 regions were classified as hypotrimethylated (**Table 4**). Surprisingly, assessment of H3K4me3 nuclear levels revealed that the distal

colon contained significantly higher (p-value 0.045) amounts of trimethylated H3K4me3 as compared to the proximal colon (**Figure 5**)

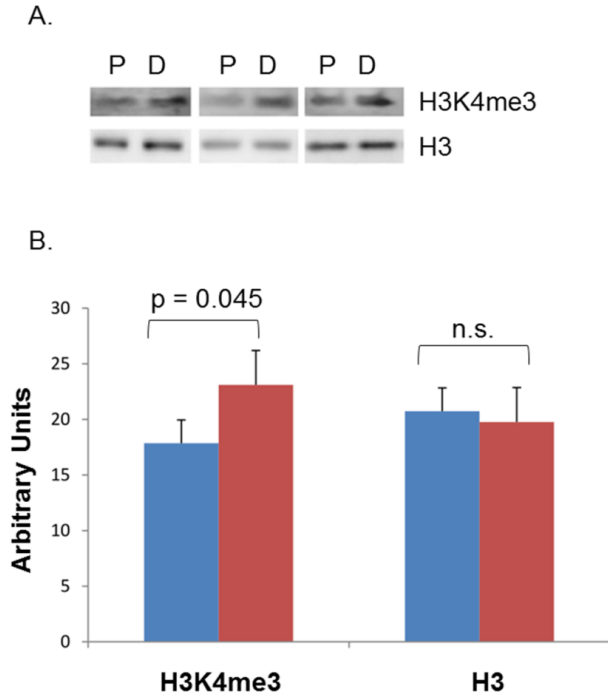


Figure 5. Levels of trimethylated H3K4 are higher in the distal versus proximal colon. A. Representative immunoblots of nuclear protein from proximal (P) and distal (D) colon. B. The blue bars represent the proximal colon and red bars the distal colon. Quantification of band volumes using Quantity One software. Values are means \pm SE (n = 7). At least 2 independent assays were conducted.

Table 4. Relationship between levels of H3K4 trimethylation, H3K9 trimethylation and mRNA expression.

	# genes detected by region		# genes differentially expressed Distal over Proximal colon	H3K4 trimethylation changes			
				upregulated	downregulated	no change	
H3K4me3	Distal colon	9,542	increase	857	0	0	
	Proximal colon	9,248	decrease	98	0	98	
RNA	Distal colon	9,706	increase	265	121	0	144
	Proximal colon	9,591	decrease	282	5	31	246
H3K9me3	Distal colon	907	increase	119	9	2	108
	Proximal colon	865	decrease	37	2	1	34

In order to assess patterns of gene expression throughout the longitudinal axis, pooled RNA was sequenced to obtain proximal (74,013,548) and distal (76,948,093) colon reads which are represented by the area under the curve in **Figure 6A**. Regional gene expression was subsequently calculated using FPKM, which measures the molar concentration of a transcript by normalizing read counts to the respective mRNA length and the total number of reads in each sample. We based our gene expression pattern analysis on moderately abundant transcripts, defined as FPKM > 1 in at least 1 sample, with the cutoff represented as the dotted line in **Figure 6A**. This corresponds approximately to 1 copy of RNA per intestinal cell (148). Using these criteria, the number of expressed genes totaled 9,867 in the distal colon and 9,591 in the proximal colon (**Table 4**). A total of 547 differentially expressed genes were detected with a q-value cutoff < 0.05 (**Supplemental Data 1A**) as represented by the blue dots in **Figure 6B**. This group included tumor and inflammation markers *CD44*, *PIK3CD*, *ALDH1A1*, *FOXP1* and *MS4A1* (149,150). Other cancer progression associated genes also enclosed in our dataset were *MMP9*, *TFF3*, *TNFRSF25*, *CCND2*, *CDKN1A*, *TCF3* and *MUC1* (127).

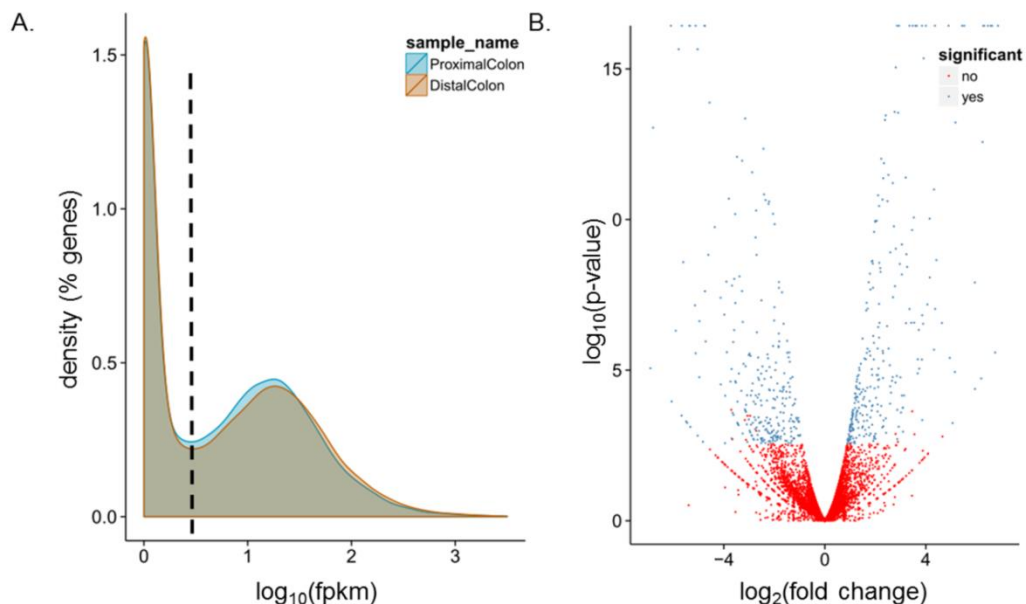


Figure 6. Global RNA-Seq analysis. A. Detection of all annotated rat genes (y axis) and the intensity of their expression are presented in the x axis as $\log_{10}(\text{FPKM})$ values. Blue line represents proximal colon, orange line represents distal colon and the dotted line represents the minimum gene expression cutoff. B. Volcano plot representing differentially ($p < 0.05$) expressed genes, blue dots characterize significant differentially expressed genes; red dots denote genes that with no differential expression between the distal and proximal colon.

Forty microRNAs were also detected, from which 13 were overexpressed in the proximal, and 14 in the distal colon (**Supplemental Data 1A**). The data in **Table 5** show the top 20 differentially expressed genes in the distal and proximal colon, of which a subgroup of 7 were jointly found in both the H3K4me3 and RNA datasets. Some of these genes, including ATP12A, FABP1 have been previously shown to be differentially expressed in the mouse colon (151). Overall, 282 genes were over expressed and 265 genes under expressed in the distal colon, with all exhibiting H3K4me3 binding near the TSS (**Table 4**). The computed correlation coefficient between the differentially transcribed genes and the differentially regulated H3K4me3 regions was 0.72 ($p\text{-value} < 0.001$). For validation purposes, a subset of genes involved in gastrointestinal

homeostasis were further examined by RT-qPCR (**Table 6**). Patterns of high, moderate and low expressed genes were consistent across analytical platforms.

Table 5. Top 20 most up-regulated and down-regulated genes in the distal versus proximal colon. RNA expression and H3K4me3 enrichment are represented as fold change DvP, distal divided by proximal colon.

RNA	D/P		D/P	H3K4me3	D/P		D/P
<u>upregulated</u>	<u>Fold</u>	<u>downregulated</u>	<u>Fold</u>	<u>upregulated</u>	<u>Fold</u>	<u>downregulated</u>	<u>Fold</u>
HOXD13	188.3	REG3B	-715.4	HOXB13	6.7	RGD1561678	-8.0
SLC37A2	116.0	OSR2	-507.4	B3GNT7	6.5	HOXB6	-5.2
HOXD10	107.4	REG3G	-459.4	VSIG1	6.2	PAX8	-4.9
HOXB13	98.5	BMPER	-169.2	RGD1306811	5.1	LOC302473	-4.9
SVAL1	94.9	HOXB3	-141.9	IGF2BP3	5.0	LOC681383	-4.8
CHST5	87.8	PRL2A1	-120.2	EVX2	4.9	OSR2	-4.3
FXYD4	79.3	HOXB4	-111.5	SSTR1	4.8	LOC691277	-4.3
B3GNT7	77.6	COLEC10	-68.5	ADH1	4.6	HOXB5	-3.8
SCGB1A1	75.9	HOXB5	-55.4	LOC100165759	4.6	LOC100359646	-3.7
TGM3	46.0	HOXB8	-50.8	TOPP8	4.5	OTUD6B	-3.5
ATP12A	43.4	LGALS2	-42.1	HOXD13	4.5	HOXB4	-3.2
LBP	36.1	NAPSA	-41.4	AKR1CL1	4.4	MUM1	-3.1
BEST2	31.0	LOC691277	-40.7	LOC100807064	4.3	P2RX7	-3.1
PLA2G2F	30.9	SLC38A4	-35.5	BC006965	4.3	SMOK2B	-3.0
SFRP2	29.9	HOXB7	-34.8	ANK1	4.1	ARHGDI B	-3.0
CEACAM10	24.9	ARHGAP20	-32.9	SOX5	4.1	ZFAT	-3.0
KLK1C9	21.4	MS4A1	-31.3	RGD1306282	4.1	HOXD4	-2.9
RGD1309350	20.7	FABP1	-27.3	TMEM232	4.1	RGD1562551	-2.9
VSIG1	20.5	SLC13A1	-26.9	MPP7	4.0	LOC100359646	-2.9
CCDC153	17.8	POU2AF1	-24.5	ARHGEF9	4.0	SNURF	-2.8

Table 6. Intestinal genes display similar patterns across analytical platforms.

		RT-qPCR	RNA-Seq	RNA-Seq		
		dist/prox	dist/prox	prox	dist	
Official Symbol	Official Full Name	(fold change)	(fold change)	Signal (FPKM)	Signal (FPKM)	
high	FABP1	fatty acid binding protein 1, liver	-6.78	-4.77	1,152.68	42.18
	ATP12A	ATPase, H+/K+ transporting, nongastric, alpha polypeptide	-1.00	-0.92	61.32	32.38
	SLC5A8	solute carrier family 5 (iodide transporter), member 8	1.32	1.25	66.28	157.49
moderate	Pparg	peroxisome proliferator-activated receptor gamma	7.16	5.44	0.47	20.42
	Hprt1	hypoxanthine phosphoribosyltransferase 1	0.00	-0.01	18.62	18.51
	Socs3	suppressor of cytokine signaling 3	-0.94	-1.14	8.43	3.82
low	GAPDH	glyceraldehyde-3-phosphate dehydrogenase	0.00	0.38	2.74	3.56
	TNFa	tumor necrosis factor	-2.00	-2.07	1.45	0.35
	COX2	prostaglandin-endoperoxide synthase 2	0.38	0.61	0.50	0.76
	ESR1	estrogen receptor 1	-2.64	Not detected	0	0
	ESR2	estrogen receptor 2	-2.00	-1.83	0.21	0.06
	iNOS	nitric oxide synthase 2	-1.00	-0.23	0.02	0.02

2.3.2 Integration of the transcriptome and H3K4 trimethylation along the colonic longitudinal axis

Genome-wide integration of histone PTM position with RNA-Seq data revealed that a similar number of regions in the distal and proximal colon were enriched for H3K4me3 and H3K9me3. Comparison between genome wide RNA expression and occupancy of each histone modification by colonic region (total number of matching genes detected within each ChIP-Seq and RNA-Seq dataset) is shown in **Table 7**. Overall, the occupancy of H3K4me3 in the distal colon was up regulated >2 fold in 857 genes and down regulated >2 fold in 98 genes (**Table 4 and Supplemental Data 1B**). H3K9me3 occupancy in the distal colon was up regulated in 119 genes and down regulated in 37 genes (**Table 4 and Supplemental Data 1C**). Of the genes upregulated at the RNA level, 45% also exhibited an elevation in H3K4 trimethylation. In comparison, 11% of the genes downregulated at the RNA level were also down modulated with respect to H3K4 trimethylation. None of the genes transcriptionally upregulated exhibited downregulation at the H3K4me3 level and 5 of the genes transcriptionally downregulated in the distal colon exhibited upregulation in H3K4me3 levels. The correlation between changes in H3K4 trimethylation and H3K9

trimethylation was less distinct, with less than 10% of all upregulated and downregulated genes at the H3K9me3 level exhibiting any changes at the H3K4me3 level (**Table 4-2**). Location analysis of histone modifications revealed that H3K4me3 was present at the transcription start site of all transcribed genes, even when expressed at relatively low levels. With respect to gene regulation, approximately 23% of the genes in which H3K9 trimethylation was detected also produced RNA transcripts (**Table 7**).

Table 7. Transcribed genes correlate with histone PTMs associated with active gene expression.

	H3K4me3	H3K9me3	RPOLII
D colon	(9,542)	(510)	(4,737)
mRNA (9,706)	6,748	123	3,249
% detected	70.7	23.8	68.6
P colon	(9,248)	(458)	
mRNA (9,591)	6,632	107	
% detected	71.7	23.4	

Comparison of genome-wide RNA expression with occupancy of each histone modification by colonic region (total number of genes detected within each ChIP-Seq and RNA-Seq dataset are listed within parentheses). For each colonic region, the annotated genes (RNA-Seq) that overlap with ChIP-Seq data are presented. The percentage of genes in each ChIP-Seq dataset that were also transcribed are noted as “% detected”. PTM, posttranslation modification.

To verify that ChIP-Seq data from pooled samples were consistent with genomic features obtained from individual rats, we performed ChIP-qPCR analysis targeting the transcription start site of the fatty acid binding protein 1 gene (*FABP1*), which regulates fatty acid trafficking in the colon. H3K4 trimethylation differences between the proximal and distal colon ranged from 1.8-fold to 11.1-fold in individual samples, compared to the ChIP-Seq at 2.1 fold (**Figure 7A**). A similar analysis was performed on *FABP1* RNA expression to assess individual sample variation. Transcriptional differences between the proximal and distal colon ranged from 5.8-fold to 8.2-fold in individual samples compared to the RNA-Seq at 5.1 fold (**Figure 7B**). H3K4

trimethylation testing of chromatin state differences between the distal and proximal colon of genes that lacked significant differential expression at the RNA level was also performed. At the TSS of *NFIB* H3K4me3 ranged from 12-fold to 1.2-fold higher in the distal colon of individual samples, compared to the ChIP-Seq at 3.3, and in *EPHB3* ranged from 2.8-fold to 1.3-fold in individual samples, compared to the ChIP-Seq at 2.2 (**Figure 7C**). These findings indicate that reliable coverage of the minimal regional difference of H3K4 trimethylation using colonic crypt epithelial cells was obtained.

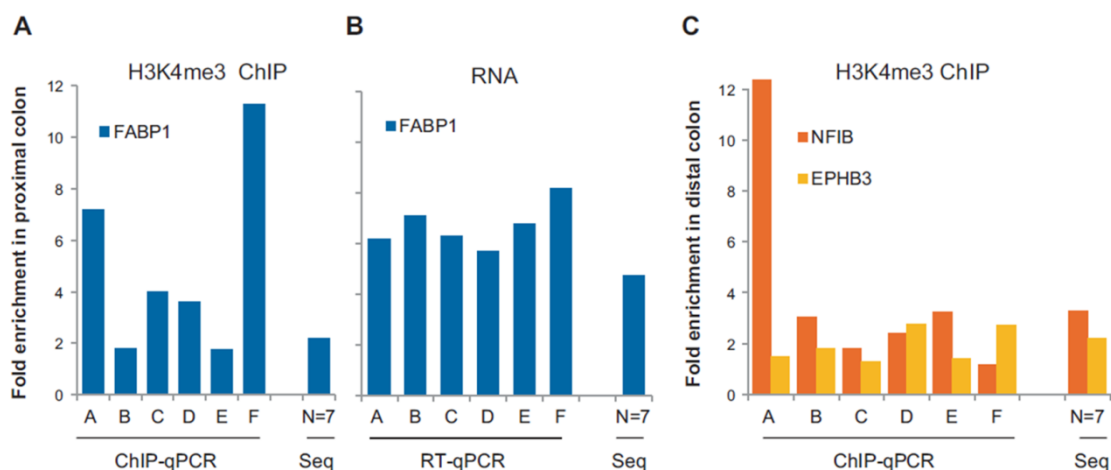


Figure 7. Assessment of FABP1 ChIP (H3K4me3) and mRNA (qPCR and RNA-Seq) gene expression analyses in individual versus pooled samples. Patterns of gene expression in individual rats by qPCR analysis (using the $2^{-\Delta\Delta Ct}$ method) were consistent with sequencing analyses from pooled samples. ChIP and RNA sequencing was performed on a pooled sample containing the same amount of genetic material from 6 rats (A-F). A. H3K4me3 enrichment at the FABP1 transcription start site (pooled rats - ChIP-Seq) is compared with multiple (individual rats, A-F) ChIP-qPCR. B. mRNA expression of FABP1 (pooled rats - RNA-Seq) is compared with multiple (individual, A-F) rats using RT-qPCR. C. H3K4me3 enrichment at the NFIB and EPHB3 transcription start site (pooled rats - ChIP-Seq)

With respect to the association between H3K4 trimethylation patterns and targeted mRNA expression, genes involved with cellular movement, including *FOX1A* and *FOX2A*, hepatocyte nuclear factors 3-alpha and 3-beta, and thioredoxin 1 (*TXN1*),

were characterized by higher levels of H3K4me3 and gene expression in the distal colon (**Figure 8A and Supplemental Data 1A and 1B**). In comparison, genes with increased H3K4 trimethylation in the proximal colon included a cluster of gut development regulators, the homeobox B genes (*HOXB8*, *HOXB7*, *HOXB5*, and *HOXB4*) and zinc finger transcription factor *OSR2* (**Figure 8B, Supplemental Data 1A and 1B**)

2.3.3 Cataloging putative middle and long non-coding RNAs (ncRNAs)

For the purpose of assessing candidate middle and lncRNAs, sequences over 100 nucleotides in length were extracted from proximal and distal colon RNA-Seq data by isolating intragenic (located in the intron of genes) and intergenic reads. Of the 135,670 regions detected in the distal colon, 65,180 (48%) were enriched with H3K4me3 (**Table 8 and Supplemental Data 2A**). In comparison, in the proximal colon, 60,744 intragenic and intergenic regions were detected, of which 31,025 (51%) were enriched with H3K4me3 (**Supplemental Data 2B**). A higher proportion of candidate ncRNA regions in the distal colon, 38,825 (29%) were enriched with H3K9me3 as compared to the proximal colon, where 11,195 regions (18%) were enriched with H3K9me3 (**Table 8 and Supplemental Data 2C**). Approximately, 25% of the candidate ncRNAs in the distal and 36% of the intergenic region candidate ncRNAs in the proximal colon were enriched in co-trimethylated H3K4 and H3K9 (**Supplemental Data 2D**).

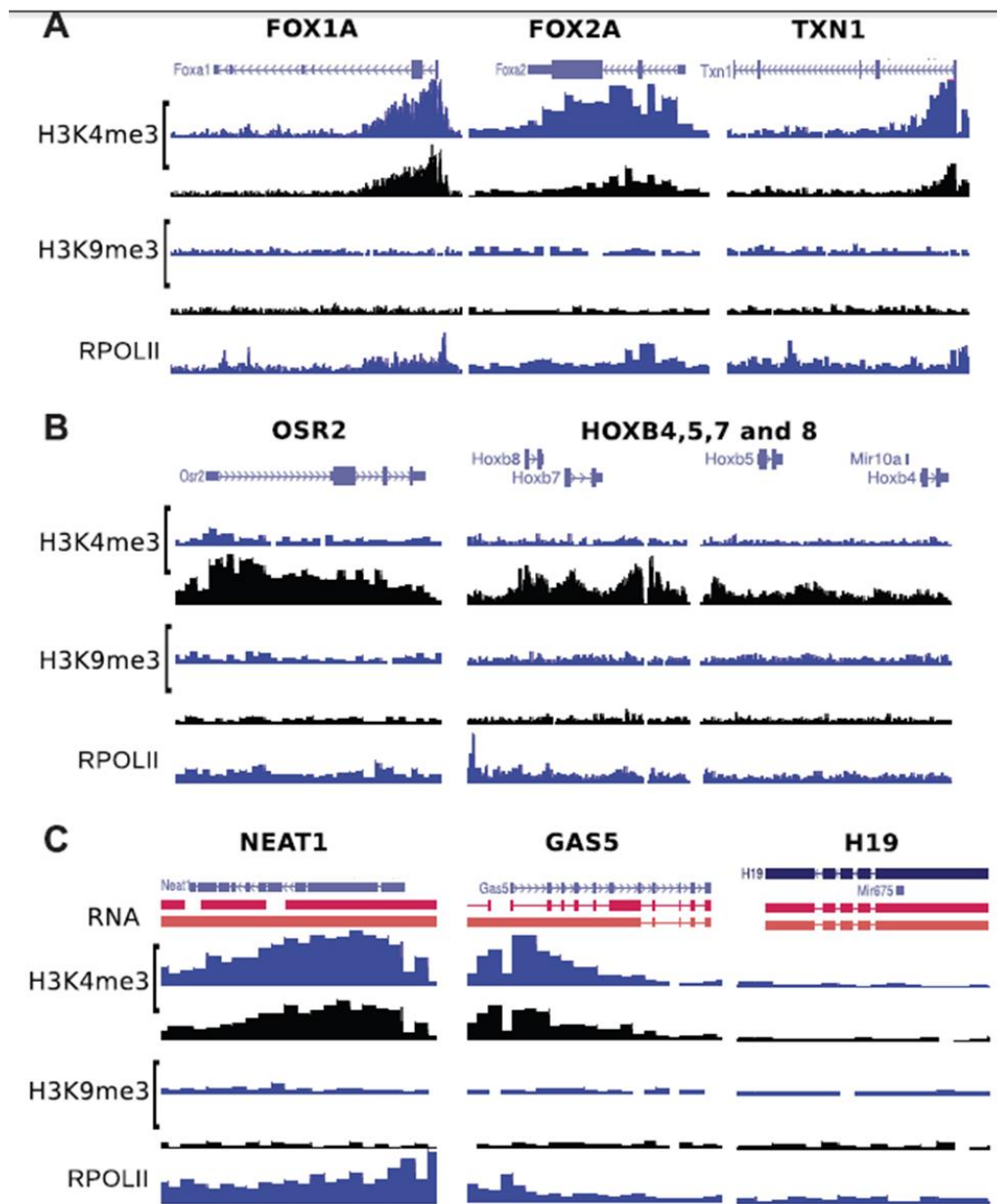


Figure 8. Occupancy of histone PTMs in highly differentially H3K4me3 regulated genes. Data from the distal colon are shown as blue peaks and proximal as black peaks. A. Genomic occupancy of differentially ($p < 0.05$) over-expressed Intestinal Immunity related genes, FOX1A, FOX2A, and TXN1, in the distal colon. B. Genomic occupancy of differentially ($p < 0.05$) under-expressed development regulating HOXB and OSR2 genes. The genome is scaled to 10kB. C. Genomic occupancy of differentially over-expressed long non-coding RNA genes; GAS5, NEAT1, and H19. The distal colon RNA transcripts are represented with orange bars and the proximal colon transcripts with red bars.

Table 8. Candidate middle and long non-coding RNAs. Transcripts >100 bp mapped to an intronic or intergenic region of the genome, considered candidates for middle and long non-coding RNAs, also exhibited signs of epigenetic regulation at the histone level.

	total # of transcripts	genomic location	colocalization with histone modifications		
			H3K4me3	H3K9me3	H3K4me3 and H3K9me3
Distal colon	135,670	intron	24,777	15,275	6,146
		intergenic region	40,403	23,554	10,079
Proximal colon	60,704	intron	11,323	3,919	2,393
		intergenic region	19,702	7,276	7,184

Since a genome wide catalog of long and middle ncRNAs has not been previously generated in the colon, we compiled annotated middle and noncoding RNAs using rat ncRNA databases, including NONCODE, UCSC, and fRNAdb to compose an established ncRNA database. Initial "potential ncRNAs" included protein coding gene transcripts yet to be annotated that were aligned against the established annotated ncRNA database while allowing for zero mismatches (see Methods) in order to identify ncRNAs in the epithelia of the large intestine. A total of 600 and 108 genomic regions were identified as unique middle and long ncRNA transcripts in the distal and proximal colon, respectively (**Table 9 and Supplemental Data 2E**). The majority of middle and long non-coding RNAs identified, belonged to the small cytoplasmic (scRNA) class of housekeeping ncRNAs. We also identified small nucleolar RNA (snoRNA), vault ribonucleoprotein (vault RNA), and small nuclear RNA (snRNA) transcripts. As expected, housekeeping non-coding RNAs *H19*, endoribonuclease RNase *MRP*, and *7SK* were detected. Collectively, these findings are the first evidence of the presence of *7H4*, *EVF1*, *GAS5*, *NEAT1*, *NTAB*, *BSR* RNAs, and *BC1 RNA*, in the colon (**Table 9**).

Table 9. Annotated middle and long non-coding RNAs. Transcripts identified as non-coding RNAs by sequence using the NONCODE and fRNA databases are listed.

<u>Distal colon</u>		<u>Proximal colon</u>	
<u>Name</u>	<u>Type of ncRNA</u>	<u>Name</u>	<u>Type of ncRNA</u>
7SK RNA	7SK RNA	GAS5	lncRNA
EVF1 RNA	lncRNA	H19	lncRNA
GAS5	lncRNA	MRP RNA	RNase MRP RNA
H19	lncRNA	NEAT1	lncRNA
NEAT1	lncRNA	NTAB RNA	lncRNA
U1	snRNA	U1	snRNA
U2	snRNA	U3B	snoRNA
U3B	snoRNA	U4 SNRNA	snRNA
U4 SNRNA	snRNA	U6 SNRNA	snRNA
U5	snRNA	VAULT RNA	vault RNA
U6 SNRNA	snRNA	BSR RNA	lncRNA
VAULT RNA	vault RNA	BC1	lncRNA
Y1	Y RNA	7H4	7H4 RNA
Y3	Y RNA	SCRNA	scRNA
BSR	lncRNA		
BC1	lncRNA		
7H4	7H4 RNA		
SCRNA	scRNA		

Most of the lncRNAs detected were accompanied by H3K4me3 occupancy at the TSS (**Figure 8C**). LncRNAs that were not transcribed in the colon also lacked H3K4me3 occupancy at the corresponding genomic region. This is not consistent with protein coding gene regions, which generally exhibited H3K4me3 enrichment in the absence of mRNA transcription (**Table 7**).

2.3.4 Anatomically-linked signaling pathways

Gene networks were assessed using the IPA library of canonical pathways across colonic regions at the H3K4 trimethylation and the transcriptional levels. Specifically, for transcriptional data, genes that were differentially expressed by at least 1.9-fold between the distal and proximal colon, with a q-value cutoff < 0.05 , were associated with canonical pathways using the Ingenuity Knowledge Base. The top 100 genes with at least 2-fold enriched occupancy of H3K4me3 and e-value < 0.01 were also analyzed (**Supplemental Data 3A**). Significance of the association between each data set and specific canonical pathways was measured in 2 ways: 1) The ratio of the number of genes from the data set was mapped to annotated pathways divided by the total number of genes known to comprise that canonical pathway, and 2) Fisher's exact test was used to determine the probability that the association between the genes in the dataset and the canonical pathway were explained by chance alone. Using this approach, the phosphatidylinositol-3 kinase (PI3K) signaling cascade was the most significantly impacted canonical pathway in terms of differential H3K4me3 occupancy and RNA transcription (**Figure 9A**). This is noteworthy, because PI3K signaling is involved in all aspects of epithelial cell biology, including the development and progression of colon carcinogenesis (152,153). Interestingly, over 97% of the canonical pathways associated with differential H3K4 trimethylation were also affected at the transcriptional level. Specifically, 169 canonical pathways exhibited differential H3K4 trimethylation and 264 pathways differential expression between the distal and proximal colon (**Figure 9B and Supplemental Data 3A**).

In complementary analyses, we applied IPA Upstream Regulator analysis software to identify the cascade of transcription factors linked to the differentially expressed genes. This in silico analysis is based on prior knowledge of expected effects between transcriptional regulators and their target genes stored in the Ingenuity® Knowledge Base.

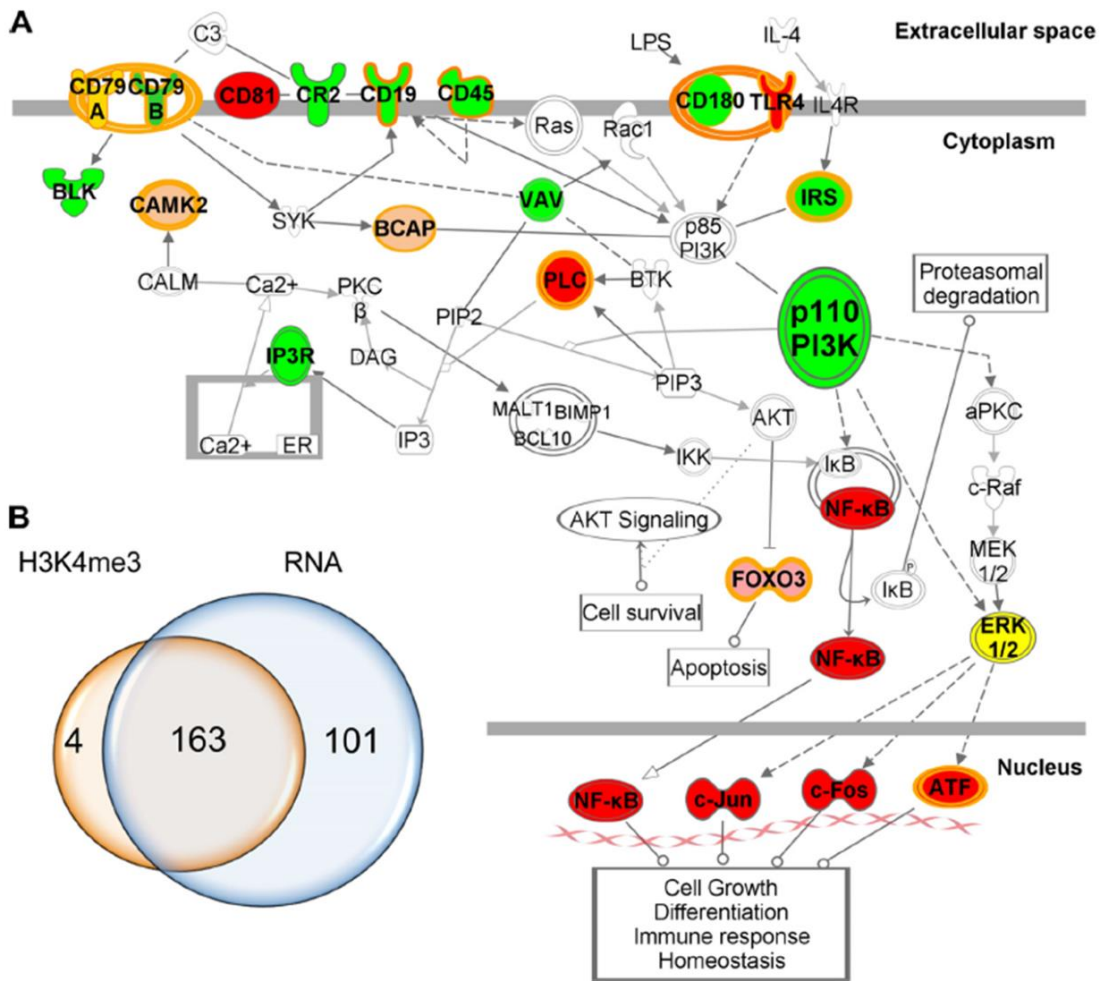
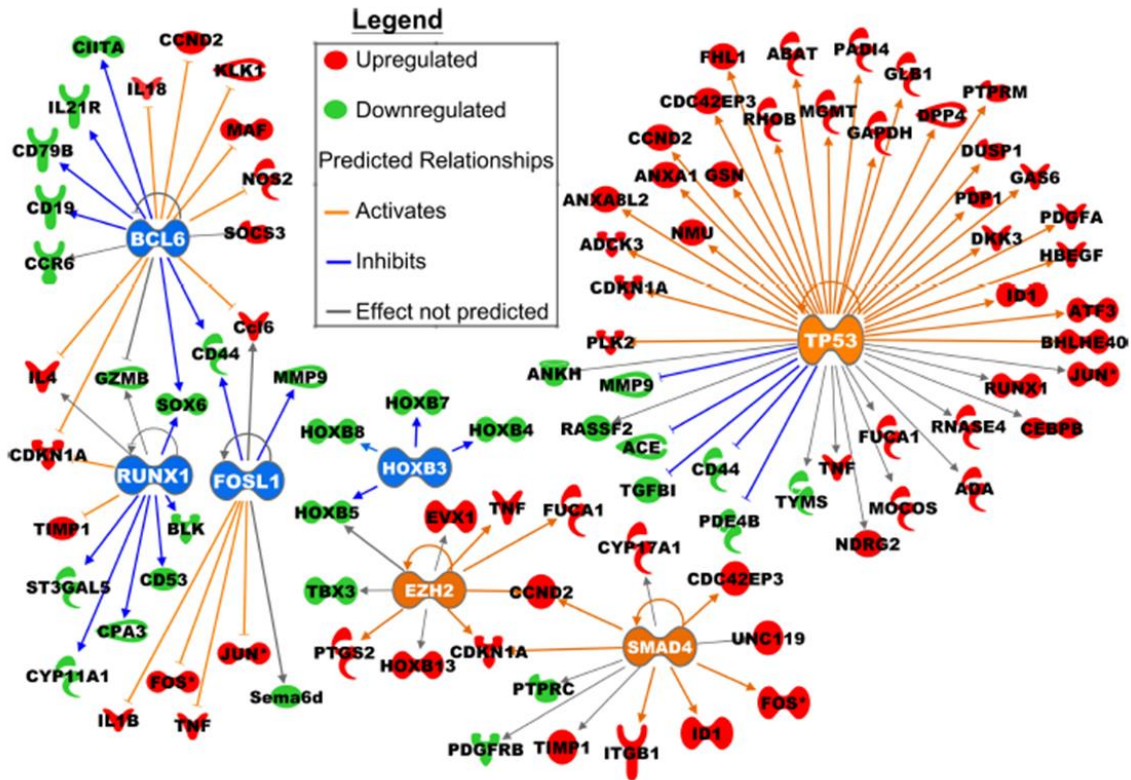


Figure 9. Canonical Pathways. A. The PI3K signaling pathway was the most significantly affected canonical pathway in terms of H3K4me3 occupancy and RNA transcription. PI3K signaling activates cell survival, cell growth, cell cycle and contributes to colon carcinogenesis. Orange highlights indicate the differential expression of genes enriched at the H3K4 trimethylation and mRNA levels. Red highlighted genes represent over expression (distal > proximal) and green represents under expression in the distal colon at the RNA level. PI3K related genes differentially regulated solely at the H3K4me3 level are highlighted in pink and yellow, representing up and down regulated genes, respectively. B. Venn diagram showing the overlap between canonical pathways linked to changes in H3K4me3 occupancy and RNA transcription.

Initially, we quantified known targets of transcriptional regulators present in our dataset and compared their direction of change (over or under expression) in order to predict likely relevant transcription factors (**Supplemental Data 3B**). For individual TFs, two statistical measures, an overlap p-value and an activation z-score, were computed. The overlap p-values were used to rank upstream regulators based on significant overlap between differentially expressed genes and known gene targets regulated by each TF. In addition, the activation z-score was used to infer likely activation states of upstream regulators based on comparison with a model that assigned random regulation directions. Using this approach, over 150 transcription regulators had a p-value < 0.05, of these 9 received an activation z score >2.0 (**Supplemental Data 3B**). The transcription factors with predicted inhibitory activity in the distal colon included *SPI1* (proviral integration oncogene), *BCL6* (B-cell CLL/lymphoma 6) and *RUNX1* (runt-related transcription factor 1). Transcription factors with projected increased activity in the distal colon included *TP53* (tumor protein p53), and *SMAD4* (SMAD family member 4) and histone trimethylator *EZH2* (enhancer of zeste homolog 2). Resultant values and differentially regulated genes are shown in **Figure 10**.

We next identified those biological functions most significantly associated with the proximal and distal RNA-Seq datasets using the IPA Functional Analysis feature. Genes from the transcriptional dataset that were associated with biological functions in the Ingenuity Knowledge Base were considered for the analysis, of which 501 unique biological functions were detected with p-values <0.001 (**Supplemental Data 3C**). Assessment of predicted activation states were based on prior knowledge of expected associations between genes and their biological functions. Therefore, we compared the direction of change (over or under expression) to expected values from the literature. A total of 34 biological functions were assigned a predicted activation state in the distal colon by generating an activation z-score >2.0 (**Supplemental Data 3C**). The same analysis was performed with the H3K4me3 dataset, of which 499 unique biological functions were assigned p-values < 0.001, and a total of 44 were assigned a predicted

activation state. The top 8 most differentially active biological functions are presented in Table 10.



Upstream regulator	Predicted Activation Region	# of target molecules in dataset	Predicted Activation State	Activation z-score	p-value of overlap
TP53	Distal	59	Active	3.15	3.78E-04
SMAD4	Distal	13	Active	2.376	1.83E-03
EZH2	Distal	8	Active	2.191	1.56E-01
BCL6	Proximal	12	Inhibited	-3.273	1.4E7-04
RUNX1	Proximal	10	Inhibited	-2.305	3.59E-04
FOSL1	Proximal	8	Inhibited	-1.725	6.59E-03

Figure 10. Upstream analyses of differentially expressed genes are linked to key proto-oncogenic transcription factors (TFs). Validated gene targets of each transcriptional regulator and their direction of change were compared in the proximal and distal colon. The overlap p-values represent the statistical significance. Fisher’s Exact Test was used to determine overlap between the differentially expressed genes and those genes regulated by the TF. Activation z-scores indicate whether an upstream TF has significantly more “activated” than “inhibited” or vice versa.

Table 10. Pathway analysis was used to identify the top biological functions associated with differentially expressed genes. RNA and H3K4me3 up and down regulated genes that exhibited significant differential expression between the colonic tissue types were associated with a biological function (p-value) and a statistical score (Activation z-score). Right-tailed Fisher’s exact test was used to calculate p-values determining the probability that each biological function was due to chance alone. The activation state (“Increased” or “Decreased”) in the distal over proximal colon was inferred by the activation z-score. These z-scores define relationships in the molecular networks that represent experimentally observed causal associations between genes and their functions.

Biological Function	Functional Annotation	P Value	Genes, <i>n</i>	Predicted Activation State (D/P)	Activation <i>z</i> Score
<i>RNA</i>					
Lipid metabolism	concentration of lipid	9.72E-11	74	increased	2.93
Lymphoid tissue structure and development	development of lymphatic system component	6.73E-12	45	increased	2.24
Infectious disease	infection of mammalia	1.41E-09	40	increased	2.05
Cellular movement	cell movement of cancer cells	2.92E-04	4	increased	2.00
Humoral immune response	quantity of IgG	1.26E-13	37	decreased	-3.23
Hematological system development & function	cell movement of phagocytes	3.22E-10	54	decreased	-3.14
Cell-to-cell signaling and interaction	recruitment of granulocytes	1.70E-08	26	decreased	-2.93
Cellular movement	leukocyte migration	3.27E-17	92	decreased	-2.60
<i>H3K4me3</i>					
Cellular function and maintenance	microtubule dynamics	2.26E-04	57	increased	4.61
Organismal development	size of body	8.16E-05	55	increased	4.52
Cellular assembly and organization	growth of plasma membrane projections	2.42E-03	26	increased	3.23
Cell death and survival	cell survival	2.62E-03	63	increased	3.14
Developmental disorder	agenesis	1.22E-03	6	decreased	-2.21
Metabolic disease	glucose metabolism disorder	8.92E-03	49		-1.84
Gastrointestinal disease	diabetes mellitus	1.11E-02	42		-1.72
Humoral immune response	quantity of IgG3	1.32E-02	7		-1.30

RNA and H3K4me3 up- and downregulated genes that exhibited significant differential expression between the colonic tissue types were associated with a biological function (*P* value) and a statistical score (activation *z* score). Right-tailed Fisher’s exact test was used to calculate *P* values determining the probability that each biological function was due to chance alone. The activation state (“increased” or “decreased”) in the distal over proximal colon was inferred by the activation *z* score. These *z* scores define relationships in the molecular networks that represent experimentally observed causal associations between genes and their functions.

2.4 Discussion

To our knowledge, this is the first in vivo study to globally assess the chromatin state and transcriptome profile of colonic crypt epithelial cells. Chromatin structure affects several processes important to transcription, including transcription factor and polymerase recruitment, and transcriptional initiation, elongation, and cessation (147,154). Traditionally, because of the many technical challenges associated with ChIP-Seq, e.g., the large number of cells required, most investigators have used tissue culture or blood derived primary cells. However, in order to study epigenetic changes in gastrointestinal diseases, it is imperative that physiologically relevant animal models be utilized. Only then can the complexity of the colonic crypt be examined. Typically, the colonic crypt is composed of (85%) enterocyte cells, a few stem cells, (2-5%) enteroendocrine cells and (10%) goblet cells. Therefore, the results from this study are primarily, but not solely, a reflection of the biological state of enterocyte cells in the colon.

Many gastrointestinal diseases are region-specific, e.g., cancer and inflammatory bowel disease (IBD) (122,149,155). The distal and proximal colon, offer a unique and relevant backdrop to study epigenetic regulation at the trimethylation level of histones. Differential levels of H3K4 and H3K9 trimethylation and DNA binding across the intestinal longitudinal axis, suggest that many of the pathway and gene activity related anomalies previously linked to intestinal disease states may be anatomically imprinted along the longitudinal axis of the colon (151,156,157). Consistent with the findings in this study, site-specific differences in global histone modification patterns and proto-oncogene expression were detected. Most notably, CD44, an independent and strong predictive factor of proximal tumors, was upregulated in the proximal colon (149). This emphasizes the need to assess the molecular profiles of the proximal and distal colon and treat them as separate entities. Along these lines, the chemoprotective non-steroidal anti-inflammatory drug (NSAID) sulindac is an effective chemopreventive agent in the distal colon of AOM treated mice. Conversely, in the proximal colon, it triggers a

distinct profile of inflammatory factors and induces lesions (30). Thus, this NSAID has both beneficial and harmful effects in vivo, which are associated with the distinct microenvironments within the colon of experimental mice (127). As expected, colonic crypts from our non-carcinogenic model exhibited only very low levels of iNOS and COX2. We propose that the distinct epigenetic profile of the colonic regional axis prior to the onset of disease may influence the ability of chemoprotective agents to favorably modulate cell transformation in the ascending (proximal) and descending (distal) colon.

2.4.1 Constructing an atlas of the distal and proximal colon epigenomes

It is known that dense H3K4me3 nucleosomes associated with gene activation can indicate the presence of non-coding RNA transcripts, and are disrupted during the onset of disease (158). In our study, the number of upregulated H3K4me3 nuclear binding sites across the genome was higher in the distal colon (**Table 4**). Assessment of H3K4me3 nuclear levels revealed that the distal colon contained 30% higher levels of trimethylated H3K4me3 as compared to the proximal colon (**Figure 5**). These findings suggest the existence of an H3K4me3 gradient along the colonic longitudinal axis. Interestingly, the levels of H3K4me3 binding in individual animals pooled for ChIP-Seq analysis, fluctuated to a greater extent relative to mRNA expression in the proximal colon (**Figure 7**). These data suggest that H3K4 trimethylation has to reach a certain threshold level of upregulation in the proximal colon to trigger increased transcription. After the threshold is reached, any greater increase in H3K4me3 binding may have no effect on transcription.

The accuracy and relevance of the colonic gene expression patterns are in agreement with previously reported protein profiles from the distal and proximal colon of mice and colorectal cancer (CRC) patients, including an increase in *CD44* in patients suffering from proximal colon cancer (127,149). Interestingly, it has been predicted that over a thousand genes have yet to be annotated in the murine genome (159). By correlating genome wide H3K4me3 binding with transcription data, we were able to

increase detection accuracy of unannotated genes in the colonic mucosa. As summarized in **Table 8**, many RNA transcripts colocalized with H3K4me3 binding regions can be used as a means to identify strong candidate genes for future coding and non-coding RNAs annotation (146,160). Consistent with previous findings in other cell types, several long non-coding RNA genes (*NEAT1*, *H19*, and *GAS5*) exhibited H3K4me3 binding at their TSS regions (**Figure 8C**). With the exception of *H19*, our findings demonstrate for the first time the presence of 7 lncRNAs (**Table 9**) in the colon. In accordance with messenger RNA annotation and classical methodology for RNA-Seq, we excluded all transcripts under 100 base pairs. Therefore, it is not surprising that only 44 out of 383 annotated microRNAs were detected in our RNA-Seq dataset (**Supplemental Data 1A**).

2.4.2 Proto-oncogene networks influenced by location

A common therapeutic strategy in treatment of disease includes targeting biological pathways as opposed to an individual gene. Pathway analyses, including biological function, canonical pathway, and upstream transcriptional regulator analyses, can expose the underlying vulnerabilities in the healthy colon to specific diseases. Our biological function analyses of the RNA-Seq data revealed a location bias associated with gastrointestinal diseases in the distal and proximal colon. These included the expression of 28 inflammatory bowel disease (IBD) associated genes, among which *PIK3CD*, *TLR4*, *CD44*, *B3GNT6*, *CASP1*, *MMP9*, *MUC1*, *STAT4*, *TFF3*, and *TLR5* have been linked to digestive tract tumorigenesis (**Supplemental Data 1A and 3B**) (161). H3K4me3 occupancy in these IBD associated genes was not differentially regulated. Surprisingly, our functional analysis of H3K4me3-Seq data indicated that genes associated with diabetes mellitus, and not IBD, were differentially regulated at the H3K4me3 occupancy level. However, few of these genes exhibited changes at the RNA level. Collectively, these data suggest that concurrent analysis of gene expression and histone PTM provides higher sensitivity for determining susceptibility to intestinal

disease onset.

The physiological differences between the proximal and distal colon can be further illuminated by examining the activity of upstream transcriptional regulators. Upstream Regulator IPA analysis provided insight into signaling directionality along the colonic axis. By linking gene expression data with known targets regulated by a TF, we were able to predict which of the observed changes may mediate disease progression. For example, the HOXB gene cluster associated with gastrointestinal development and regulation was prevalently active in the proximal colon, where it seems transcriptional upregulation of *HOXB3* activity leads to chromatin modification and RNA downregulation of *HOXB4*, 5, 7 and 8 (**Figures 8B and 10**). These increases in HOXB genes have also been previously documented in hypoplasia and Hirschsprung's, diseases prevalent in the proximal colon (162).

SMAD4, a biomarker for prognosis of colorectal cancer, was identified as a highly active transcription factor in the distal colon (**Figure 10**). Gene expression profiles in our study are consistent with *SMAD4* upregulation during the promotion of tumorigenesis in the distal colon (163). These observed patterns include increased activity of *SMAD4* and *TGFB* and increased transcription of *FOS* and *BMP2* (**Supplemental Data 1A and 3C**). Based on these findings, we propose that the projected increase in *EZH2* TF activity in the normal distal colon could predispose this region of the intestine to future *EZH2* mediated oncogenic activity (164). We suggest the pathway is only partially activated under normal physiological states. The promotion of cancer would require the aberrant expression of a priming signal, e.g., the expression of *IGFBP2*, which is known to appear solely in the distal colon during carcinogenesis (165). This would subsequently alter PI3K-Akt-induced activation of *PRC2*, thereby promoting tumorigenesis (165,166). As further proof of the relevance of our upstream regulator analyses, *TP53*, *EZH2*, and *RUNX1* mutations are currently used as biomarkers of poor prognosis in colon cancer (167-169) (**Figure 10 and Supplemental Data 3C**).

Peroxisome Proliferator Activator Receptor gamma (*PPARgamma*), which is more highly expressed in the distal colon, is the target of many antidiabetic and

chemotherapeutic drugs (170). PPAR signaling also plays a key regulatory role in lipid metabolism, inflammation, and cellular differentiation (171). In **Figure 11** we suggest the function of the PPAR regulatory pathway in our model. Our data including genes *ATP12A*, *FABP1*, *SOCS3*, *SLC5A8*, *APQ3*, *SLC37A2* and *PPARgamma*, are corroborated by previous findings indicating that *PPARgamma* activity in the colonic epithelium is complex and likely regulated by a number of factors that remain largely unknown (151,170). Many of the identified connections have only been detected following deregulation of cellular function, further supporting the concept that innate, epigenetic differences play an important role in the colon (172-174).

In summary, we have documented for the first time the chromatin structure associated with gene expression profiles in the rat proximal and distal colon by correlating ChIP-Seq with RNA-Seq data. Globally, approximately 500 genes were differentially expressed between the proximal and distal colon. With regard to differentially expressed genes, a high correlation was observed between H3K4me3 occupancy and RNA-Seq data. Gene ontology analysis indicated that colonic crypt location significantly impacted both chromatin and transcriptional regulation of genes involved in cell transformation, lipid metabolism, lymphatic development and immune cell trafficking. We were also able to detect a range of lncRNAs that have not been previously reported in the colon. In addition, gene function analysis indicated that the PI3-Kinase signaling pathway was regulated in a site-specific manner. In conclusion, distinct combinatorial patterns of histone modifications exist in the proximal versus distal colon. These site-specific differences may explain the differential effects of bioactive chemoprotective agents on cell transformation in the ascending (proximal) and descending (distal) colon.

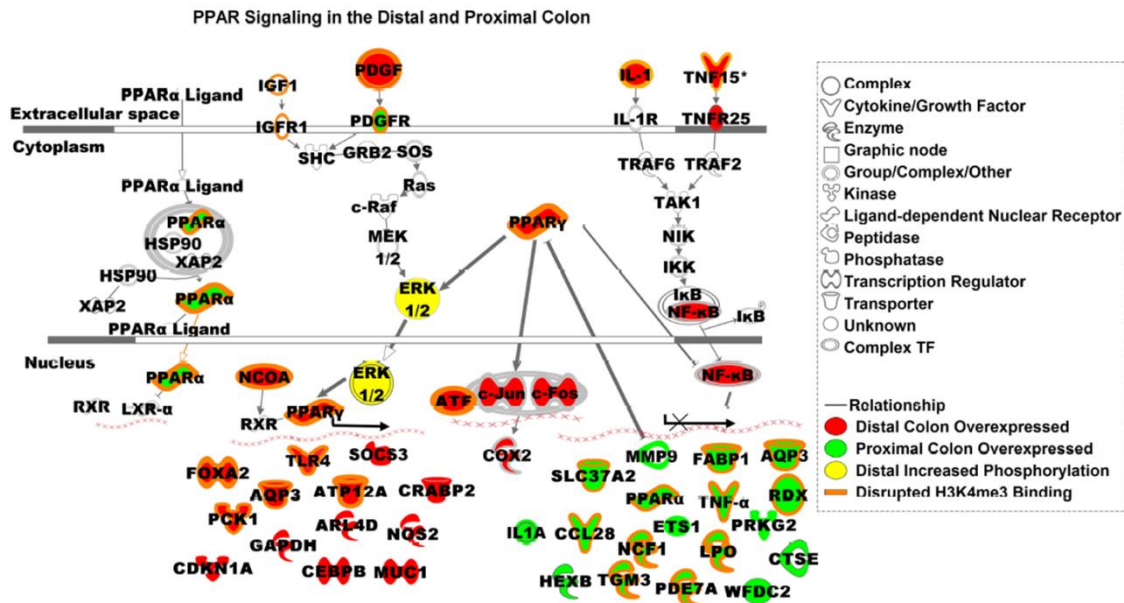


Figure 11. PPAR signaling in the proximal and distal colon is linked to disease-associated genes. Validated targets of PPAR genes and their direction of change were compared in the proximal and distal colon. Orange highlights indicate the differential expression of genes enriched at the H3K4 trimethylation and mRNA levels. Red highlighted genes represent over expression (distal > proximal) and green represents under expression in the distal colon at the RNA level.

CHAPTER III
ASSESSMENT OF HISTONE TAIL MODIFICATIONS AND TRANSCRIPTIONAL
PROFILING DURING COLON CANCER PROGRESSION REVEALS A GLOBAL
DECREASE IN H3K4ME3 ACTIVITY

3.1 Introduction

Colorectal cancer (CRC) ranks as the third leading cause of cancer death among adults. Every year in the United States, more than 150,000 cases of colorectal cancer are diagnosed and approximately 57,000 patients die of the disease (18). During carcinogenesis, major cellular functions and pathways, including drug metabolism, cell cycle regulation, potential to repair DNA damage or to induce apoptosis, response to inflammatory stimuli, cell signaling, and cell growth control and differentiation are dysregulated (18). Epigenetic alterations contribute to these cellular defects. For example, the epigenetic modulation of master transcription factors by promoter methylation and modification of histones and non-histone proteins lead to genomic instability and perturb the expression of gene sets associated with cell adhesion and apoptosis (175-178).

Chromatin signatures are tightly linked to epigenetic regulation. For instance, transcriptionally active genes are characterized by active chromatin marks, such as trimethylated histone H3 lysine 4 (K4me3) and acetylated histone H3 lysine 9 (K9ac) (175,178). Alterations in these histone modifications can drive oncogenic processes, such as proliferation, invasion, angiogenesis, and dedifferentiation, by perturbing normal gene expression patterns (178). This is particularly relevant to CRC, because altered K4me3 levels are associated with onset of colorectal cancer (179).

The azoxymethane (AOM) chemical carcinogenesis model serves as one of the most definitive means of assessing mechanisms related to human colon cancer risk (180). We have previously demonstrated that at 10 week post AOM injection, the colonic mucosa is precancerous, equivalent to the progression stage of colon cancer, e.g.,

high multiplicity aberrant crypt foci are apparent (181) . By ~34 weeks post AOM injection, macroscopic tumors become detectable (180).

Current evidence indicates that genetic mutations, epigenetic changes, and aberrant immunologic signaling pathways are major contributors to CRC (182). However, until recently, appreciation of epigenetic alterations has lagged behind genetic mutations with regard to their contributions to human cancer development. To date, preclinical animal models as well as human studies that elucidate the pre-tumorigenic epigenetic molecular events driving CRC are limited. Colon cancer profiles for H3K4me3 and H3K9ac enrichment and other histone modifications have been previously generated, however, in most cases cancer cell lines and/or genome-limited methodology were used (179,183). In addition, few studies to date have integrated diverse epigenetic inputs in an in-vivo model of CRC. This type of analysis would improve the statistical and interpretative power of the changes in transcription and chromatin state during cancer progression. Since distinct genomic and epigenetic events drive the initiation, promotion and progression of colon cancer (18), in this study we integrated global chromatin immunoprecipitation sequencing (ChIP-Seq) and RNA-Seq data in order to explore the progression of colon cancer at a genome-wide epigenetic level.

3.2 Methods

3.2.1 Animals

Sixty-seven weanling male Sprague Dawley rats (Harlan, Houston, TX) were individually housed and acclimated for 1 week in the same room, maintained in a temperature and humidity-controlled animal facility with a daily 15 h light/9 h dark photoperiod. The animal use protocol was approved by the University Animal Care Committee of Texas A&M University and conformed to NIH guidelines. The study

examined the effects of two treatments (AOM or saline control). Animals were stratified by body weight after the acclimation period so that mean initial body weights did not differ. Body weight and food intake were monitored during the study.

3.2.2 Diets

After 1-week acclimation on standard pelleted diet, rats were assigned to one of four diet groups which differed in type of fat plus fiber. The diets contained (g/100 g diet): dextrose, 51.06; casein, 22.35; D,L-methionine, 0.34; AIN-76 salt mix, 3.91; AIN-76 vitamin mix, 1.12; and choline chloride, 0.22. The total fat content of each diet was 15% by weight as follows: n-6 fat diet, 15.00 g corn oil/100 g diet; n-3 fat diet, 11.50 g fish oil/100 g diet (OmegaPure TE from Omega Protein Inc); and 3.50 g corn oil/100 g diet (Dyets Inc.). The total fiber content of each diet was 6% by weight of pectin (fermentable fiber from Gum Technology) or cellulose (non-fermentable fiber from Bio-Serv). To prevent formation of oxidized lipids, diets were stored at -20°C and freely provided to the animals fresh each day. To protect against lipid oxidation during storage, 0.025% tertiary butylhydroquinone and mixed tocopherols (MTS-50; ADM, Decatur, IL) were added to the oils.

3.2.3 Carcinogen treatment

After 2 weeks of feeding, 24 rats were injected with saline (control), and 43 rats were AOM (Sigma, St. Louis, MO) injected s.c. at 15 mg/kg body weight. Each rat subsequently received a second AOM or saline injection 1 week later and animals were terminated 10 weeks after the first AOM injection.

3.2.4 Aberrant crypt foci scoring

Immediately after removal, colons (11 per diet group with AOM injection and 2 per diet with saline injection) were flattened between Whatman 1 filter paper and fixed in 70% ethanol for 24 h. Subsequently, the whole mount colon was stained with 0.5% methylene blue in PBS for 30 sec, placed on a plastic sheet with a 5-mm grid, and examined under the microscope at $\times 400$. The number of aberrant crypts (putative colon cancer precursors), as singlets and multiples, was determined. Crypts were classified as aberrant using the morphologic characteristics described previously (184). The number of high multiplicity ACF (more than three aberrant crypts per foci) was scored.

3.2.5 Isolation of colonic crypts

The large intestine was resected from the junction between the cecum and the rectum, and was opened longitudinally and washed in $1\times$ PBS. Subsequently, the visible “herringbone” folds were used to identify the proximal colon. The region distal to this point was referred to as the distal colon. The distal colon was subsequently incubated with HBSS (Hank’s Balanced Salt Solution, without Ca^{++} and Mg^{++}) containing 1 mM glutamine, 0.1% BSA, 30 mM EDTA, and 5 mM DTT, adjusted to pH 7.5, for 15 min in a 37°C shaking incubator as we have previously described (131). Following incubation, tissue sections were placed in a petri dish on ice, and the colonic crypts isolated by scraping with a rubber policeman. Isolation of crypts was verified by histological examination to ensure that epithelial cells were removed and the lamina propria and muscle layers remained intact. Cells were washed with HBSS and centrifuged at $100\times$ g for 15 min. The pellet was resuspended in HBSS and an aliquot of the crypts was subsequently used to create mRNA expression profile libraries. The remaining crypts were immediately crosslinked for ChIP analysis.

3.2.6 Western blot

Colonic crypts were rocked in 50 mM HEPES-KOH, pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40, 0.25% Triton X-100 with protease inhibitors for 10 min at 4°C, followed by centrifugation at 1,350 x g, 4°C for 7 min. The crypt-containing pellet was subsequently resuspended in 10 mM Tris-HCl, pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA with protease inhibitors and incubated by gently rocking at room temperature for 10 min. Nuclei were pelleted by centrifugation at 1,350 x g, 4°C for 7 min and resuspended in RIPA buffer (50 mM HEPES-KOH, pKa 7.6, 500 mM LiCl, 1 mM EDTA, 1.0% NP-40, 0.7% Na-deoxycholate with protease inhibitors). Nuclear lysates (2 ug protein) were treated with 1× pyronin sample buffer and subjected to SDS polyacrylamide gel electrophoresis (PAGE) in precast 4–20% Tris-glycine mini gels (Invitrogen). After electrophoresis, proteins were electroblotted onto a polyvinylidene fluoride membrane with the use of a Hoefer Mighty Small Transphor unit at 400 mA for 90 min. Following transfer, the membrane was incubated in 5% milk and 0.1% Tween 20 in TBS (TBST) at room temperature for 3 h with shaking, followed by incubation with shaking overnight at 4°C with primary antibody diluted in 5% milk in TBST. Membranes were washed with TBST and incubated with secondary peroxidase conjugated secondary antibody as per manufacturer's instructions. Bands were developed using Pierce SuperSignal West Femto™ maximum sensitivity substrate. Blots were scanned using a Fluor-S Max MultiImager system (Bio-Rad, Hercules, CA). Quantification of bands was performed using Quantity One software (Bio-Rad). Primary antibodies were used to detect histone H3 tri methyl K4 (Active Motif 39160), H3 acetyl K9 (ab10812), and histone H3 (ab1791). Peroxidase conjugated goat anti-rabbit IgG was purchased from Kirkegaard and Perry Laboratories (Gaithersburg, MD).

3.2.7 Chromatin immunoprecipitation

ChIP-Seq analyses were performed in order to determine global histone mapping in crypts isolated from the distal colon. The ChIP protocol described by Lee et al (132) was utilized (185). Specifically, cells were cross-linked by adding freshly prepared formaldehyde at a 1% concentration for 15 min at room temperature followed by quenching with glycine. Cells were then lysed and sheared in 3 mL tubes at 4°C using a Covaris S2 sonicator to obtain DNA distributions of ~300 bp (range of 200–400 bp): duty cycle –20%, intensity –8, cycles/burst –200 and time 25 min. ChIP antibodies included: ChIP Grade (Active Motif 39160) anti-histone H3 (tri methyl K4) antibody, ChIP Grade anti-histone H3 acetyl K9 antibody (ab10812). The specificities of all antibodies were tested by Western blot and ChIP-qPCR. Antibody–chromatin complexes were captured using Dynabeads G Protein coupled (Dyna) and eluted with 1% SDS in 50 mM Tris-HCl pH 8.0, and 10 mM EDTA, after extensive washing. Cross-linking between DNA and chromatin proteins was reversed by incubation at 65°C overnight. DNA was purified by treatment with RNaseA, proteinase K and QIAquick PCR Purification Kit (Qiagen 28004) and dissolved into 50 µL EB (10 mM Tris pH8.0) buffer per sample for immunoprecipitation purposes. Equal amounts of (200–500 bp) ChIPed DNA from 2-3 AOM rats with the same dietary treatment (biological replicates) were pooled into 16 barcoded groups (representing 43 individual rats), and the saline biological replicates were similarly pooled into 12 barcoded groups (representing 24 individual rats) prior to high throughput sequencing.

3.2.8 ChIP sequencing

BioScientific NETflex (ChIPseq kit 5143-01, Barcodes kit 514120) multiplex libraries from ChIPed DNA (10 nM) were sequenced using a Illumina Hiseq 2000 DNA Sequencer. Sequence reads with poor quality bases or other contaminants were filtered.

The remaining reads (>290 million total per sample) were mapped to the reference rat genome (rn4) with commonly used Burrows-Wheeler Aligner (BWA) for Illumina (version 1.2.3) settings, only non-identical and uniquely mapped reads were retained. The peak caller programme MACS (version 1.4.1) (186) was used to identify peaks with the following parameter settings: --bw=300 --keep-dup=1. Islands were defined (using merge function of BEDTools (134)) as the genomic areas enriched with the ChIPed protein (peaks) in at least one sequenced sample, and reads were quantified using coverageBed function of BEDTools (134). The UCSC Genome Browser was used to visualize bigwig data tracks. The nearest gene to each island, i.e., within 5 kb of the island was identified using closestBed from the BEDTools software suite (134) and the refGene table downloaded from the UCSC Genome Browser for the Baylor 3.4/rn4 assembly files (135).

Regions showing differences in histone modification were identified using the edgeR package (187,188) for the R software environment (186,187). In order to increase the statistical power of our analysis (higher number of samples per treatment) and detect key AOM effects associated with cancer progression, rats were pooled across the various diet groups described under the methods section “Chromatin Immunoprecipitation”. Read counts per gene were normalized using the scaling factor method of Anders and Huber (189). Differential expression testing of genes was performed using likelihood ratio tests on the negative binomial GLMs estimated by edgeR (187,188). Regions with FDR < 0.1 and minimal threshold of one count per million mapped reads in at least four samples were selected as differentially enriched regions (DERs).

3.2.9 RNA isolation

For total RNA isolation, colonic crypts were homogenized on ice in lysis buffer (RNAqueous Isolation kit, Ambion) and frozen at -80°C until RNA was isolated. Subsequently, total RNA was isolated using the RNAqueous kit, followed by DNase

treatment. RNA integrity was analyzed on an Agilent Bioanalyzer to assess RNA integrity.

3.2.10 RNA sequencing

Total RNA (1000 ng) was used to generate multiplex libraries for whole-transcriptome analysis following Illumina's TruSeq RNA v2 sample preparation protocol. Libraries from 24 individual rats per treatment were sequenced on an Illumina HiSeq 2000. At least 151 million, 50 bp single-end reads per treatment were obtained for each sample. Reads were mapped with the STAR aligner using the default parameters and rn5 genome assembly (190). More than 85% of reads aligned uniquely to the rat genome. Genes that did not have at least one read count per million mapped reads in at least four samples were removed. Read counts per gene were normalized using the scaling factor method of Anders and Huber (189). The read counts were modeled directly using negative binomial distribution and GLM accounting for the differences in diet Bioconductor package edgeR (187,188). Differential expression was then tested using likelihood ratio tests involving the fitted models. Genes with adjusted p-value (FDR) <0.1 were selected as differentially expressed transcripts (DE).

3.2.11 Non-coding RNA protocol

Non-coding RNAs were identified using the lncRNApipe (<https://github.com/biocoder/Perl-for-Bioinformatics/releases>) software.

Pipeline steps with further technical details:

A total of approximately 1.25 billion (including all samples and replicates) 50bp single-end Illumina reads were obtained from a multiplexing run.

- A. Reads were trimmed for low quality bases (phred score $<Q20$) and any adapter sequences using Trimmomatic (191). Any reads less than 25bp in length were discarded after trimming.
- B. A total of approximately 1.2 billion reads survived trimming which were then aligned against rn4 rat assembly (from UCSC) using tophat (192) with multi-hit filtering for each replicate of a treatment.
- C. A total of approximately 590 million reads uniquely mapped to the transcriptome and resulting alignments for each treatment were merged using Sambamba (193).
- D. Merged read alignments for each treatment were processed with cufflinks (192) to perform reference annotation based transcript assembly with bias and multi-read correction.
- E. Cuffcompare (194) was run to compare assembled transcripts with reference annotation (rn4) and assign various class codes for transcript assembly of each treatment.
- F. Any transcripts that belonged to class codes “j”, “i”, “o”, “u” and “x” were considered and categorized into 5 categories: long intergenic lncRNAs (LincRNAs), intronic contained lncRNAs (Incs), partially overlapping lncRNAs (Poncs), completely overlapping lncRNAs (Concs) and exonic overlaps (LncRNAs with sense or antisense overlap with reference exon) (195) by comparison to known RefSeq (196) gene annotation (rn4). When calculating exonic overlaps with reference exon boundaries, only cufflinks assembled transcripts whose exon (at least one) overlaps with reference exon by at least 80% of the exon length was considered. Single exon transcripts were also retained.
- G. Since RefSeq gene annotation consists both protein-coding and non protein-coding genes, the resulting catalog of categorized ncRNAs were compared to identify novel ncRNAs in rn4 i.e., any assembled transcripts from cufflinks

which were not overlapping with RefSeq genes were marked as potential novel ncRNAs.

- H. FASTA sequences were created for the potential novel ncRNAs to run CPC (Coding Potential Calculator) (197) with UniRef90 protein database, and any transcripts which were flagged as “coding” by CPC were discarded.
- I. Cmscan from INFERNAL1.1 (198) was run with default E-value cutoff of 10 on the resulting transcripts using Rfam covariance models (CMs). Annotation from cmscan result was applied to putative non-coding RNA if at least 10% of query sequence was matched with the resulting best hit.
- J. Finally, RNAfold (199) was run to predict minimum free energy structure of putative novel ncRNAs.

After the pipeline was run on each treatment, the individual putative novel lncRNA assemblies for each treatment were merged with cuffmerge (192) to create unified ncRNA transcript catalog for this experiment. BLAST (200) homology search was performed against hg38, mm10 and rn6 RefSeq genes for the merged transcripts with e-value cutoff of $1e-5$, 95% sequence identity and 90% query sequence coverage thresholds to eliminate any similar or newly identified ncRNAs in closely related species to rat. A final list of potential novel long non-coding RNAs (transcript length ≥ 150) was created. Cuffdiff (192) was run using the final list of putative novel lncRNAs as reference between AOM and saline (control) conditions to identify differentially expressed putative novel lncRNAs. A total of 324 transcripts were identified as putative novel lncRNAs of which, 266 have been assigned RNA families by INFERNAL with at least 10% query sequence coverage and E-value cut off of 10. None of the putative novel lncRNAs were significantly differentially expressed.

3.2.12 Ingenuity pathway analyses

“Functional enrichment” analysis was performed using Ingenuity Pathway Analysis (IPA) version 2.0 software (Ingenuity Systems Inc., Redwood City, CA) as we have previously described (Triff, 2013).

To perform IPA analysis, all differentially expressed genes (adjusted FDR < 0.1) in the distal colon were uploaded into three columns for the purpose of generating Illumina probe ID, t-value (fold change) and adjusted p-value (FDR) data. By convention, genes that were up-regulated by AOM are shown in orange and genes that were down-regulated are shown in blue. Analysis was performed as previously described in CHAPTER II methods “Ingenuity pathway analyses”

3.3 Results

In an effort to identify carcinogen-induced genes that contribute to CRC progression, transcriptomic and histone tail H3K4me3 and H3K9ac alterations in early precancerous colonic epithelial cells were examined. Gene expression profiling by Next Generation Sequencing of RNA and ChIPed DNA (**Figure 12**) was performed. The majority of expressed genes (88%) contained H3K4me3 and H3K9ac enrichment within 5kb of the TSS (**Table 11**). Differentially expressed transcripts (DE) and differentially enriched chromatin regions (DERs) were determined by comparing rats 10 weeks post AOM injection (cancer progression stage) to saline (control).

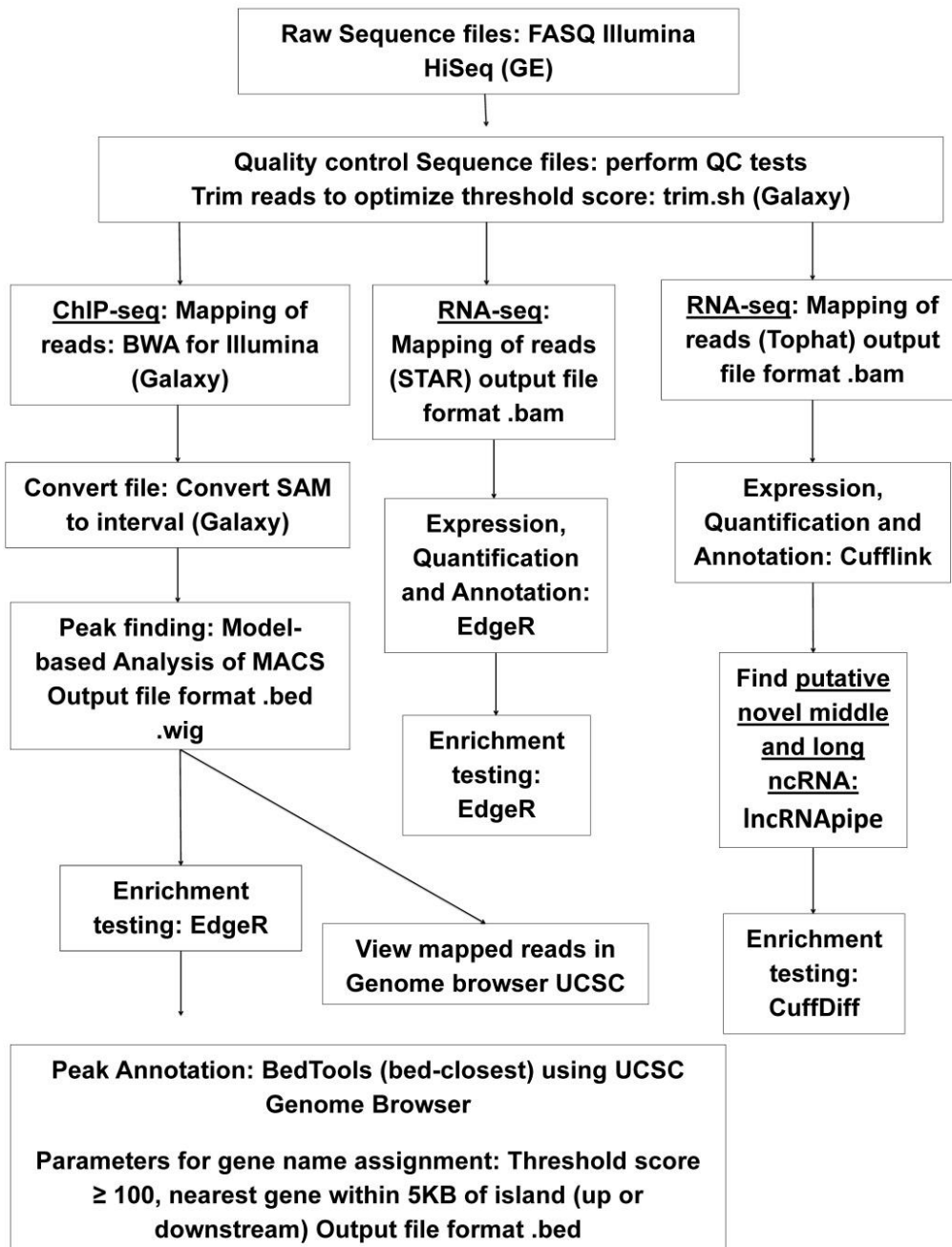


Figure 12. Bioinformatic workflow of sequencing data.

Table 11. Summary total of all detected genes at the H3K4me3 and H3K9ac and mRNA level.

Category of Annotated Genes	Description	# of genes
H3K4me3 genes	# of genes containing H3K4me3 peaks within 5kb of the TSS	11299
H3K9ac genes	# of genes containing H3K9ac peaks within 5kb of the TSS	11961
transcribed genes	# of genes with RNA detected (>1 FPKM)	10418
H3K4me3 + H3K9ac	# of genes containing both H3K4me3 + H3K9ac modifications	10872
H3K4me3 + H3K9ac + RNA	# of transcribed genes containing both H3K4me3 + H3K9ac	9152
<p>Category of Annotated Genes: H3K4me3 + H3K9ac indicates the number of genes containing both H3K4me3 and H3K9ac enrichment within 5kb of gene body. H3K4me3 + H3K9ac + RNA indicates the number of genes expressed at the mRNA level plus both H3K4me3 + H3K9ac enrichment within 5kb of the gene body.</p>		

Consistent with carcinogen exposure, an average of 46 high multiplicity aberrant crypt foci (HM-ACF) were detected in the rats treated with AOM, while there were no detectable HM-ACF in rats injected with saline (**Figure 13.A**). Since HM-ACF formation precedes carcinoma and shares many of the histopathological characteristics of human CRC (201), it is a useful biomarker to determine the extent of colon cancer progression. A comparison of total H3K4me3 and H3K9ac nuclear levels in AOM and saline treated rat colonic epithelial cells by Western blotting show similar levels of histone tail modifications between the 2 treatments (**Figure 13.B**).

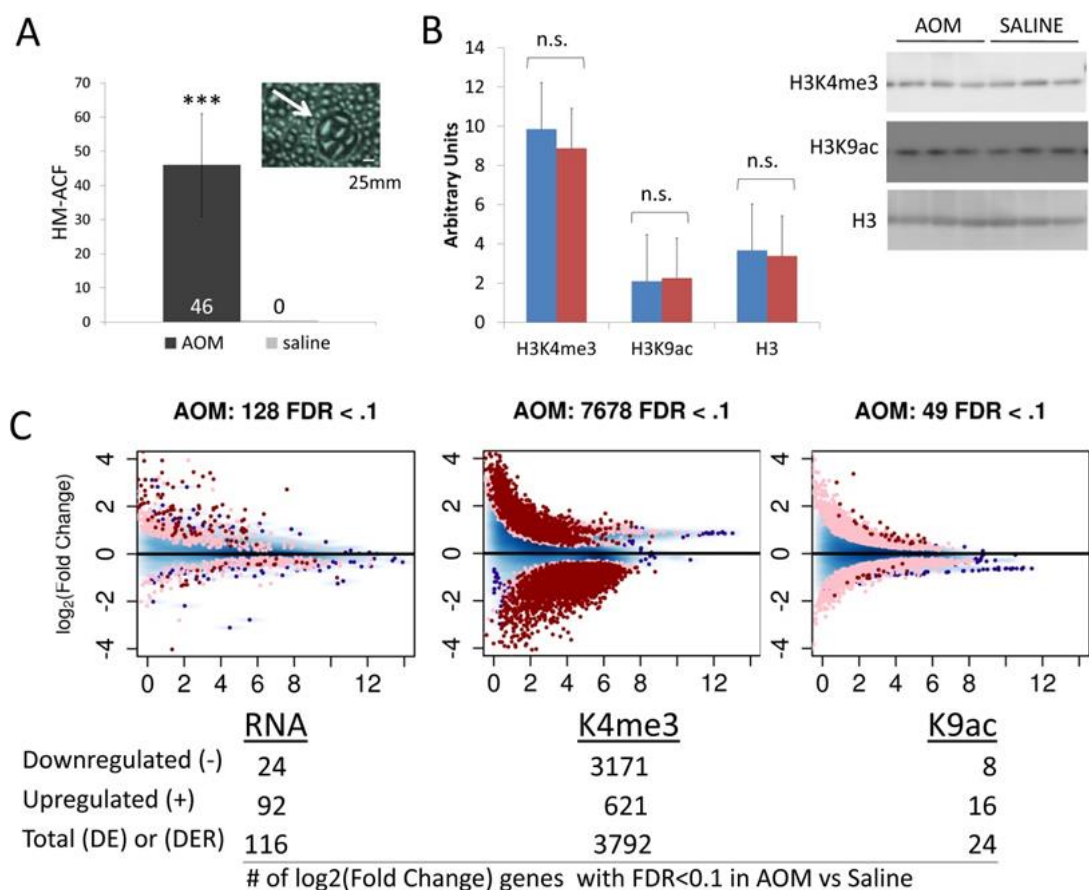


Figure 13. AOM induced pathophysiology and epigenetic changes during cancer progression. A, High multiplicity aberrant crypt foci (HM ACF) number, a precursor to tumorigenesis is shown. Data represent 43 rats at 10 weeks post AOM injection. No HM-ACF were observed in saline-injected animals. ***, p-val < 0.001 (one-way ANOVA). B, Levels of trimethylated H3K4 and acetylated H3K9 levels remain the same in AOM vs saline exposed colonic epithelial cells. The results of Western blot analysis of nuclear protein extracts from AOM and saline injected rat colon. Blue bars represent AOM treated and red bars saline injected colon. Quantification of band volumes using Quantity One software. Values are means \pm SE ($n = 7$). At least 2 independent assays were conducted. n.s., p-val > 0.05. C, H3K4me3 is highly sensitive to AOM exposure. MAplots indicate the differential expression of all transcribed genes or histone tail enriched regions (y-axis, log-ratio of difference in intensity of histone tail modifications enriched regions) vs their overall intensity of expression (x-axis, log-average of read counts) following AOM vs saline treatment. Pink represents differentially expressed (DE) transcripts and differentially enriched regions (DERs) with a p-value < 0.05, and genes FDR < 0.1 in red, all other detected genes are highlighted in blue

3.3.1 Global effects of AOM on transcription and histone tail modifications.

To identify DE transcripts and DERs with histone tail H3K9ac (K9ac) and H3K4me3 (K4me3) modifications, gene lists from treatment samples were filtered according to multiple criteria as described in the “Methods”. Briefly, DE genes and DER peaks with an FDR<0.1 and a minimal fold change $> |1.29|$ were compared across AOM and saline injected (control) rats. **Figure 13.C** shows the distribution of expression strength relative to the log-ratio of DE and DERs. Included above each MAplot are the total number of DE genes (including different isoforms) and the total number of DERs (annotated and un-annotated). **Ch.III Supplementary Table 2** contains a list of the transcripts and DERs with an FDR<0.1 in AOM vs saline treated animals. Enriched regions (peaks) further than 5kb from the transcription start site were considered “un-annotated” and classified with an ID representing the rn4 genomic location of that island.

The number of DER genes affected by AOM carcinogen treatment was significantly greater in K4me3 (3792), with the majority of DERs found at the gene transcription start site (TSS), 41 of which were located 1-1k bp and 71 were located 1k-5k bp from the gene TSS. Additionally, more genes with K4me3 DERs were downregulated (3171) vs upregulated (621) in AOM vs saline. This contrasted with the total number of differentially expressed genes, RNA (116) and K9ac DERs (24) which were less than 5% as abundant as K4me3 DERs (**Figure 13.C**). These data indicate that H3K4me3 exhibited greater sensitivity to AOM exposure as compared to K9ac DERs and RNA DEs. This pattern of enhanced sensitivity was also observed with respect to fold changes, e.g., K4me3 $\log_2(\text{fold change})$ DERs ranged from -7.02 to 5.61 as compared to K9ac DERs (-2.96 to 4.22) and RNA DEs (-4.03 to 5.20).

3.3.2 Poor correlation between AOM affected genes at the transcription and histone tail modification levels.

Based on previous work indicating that histone tail modifications regulate gene expression (178), we expected K4me3 and K9ac DER genes would correlate with RNA DE genes. Generation of a global plot of all the K4me3 and K9ac DER fold changes against RNA DE genes revealed poor correlation between these histone marks and RNA, regardless of p-values (the axes include number of genes with an FDR<0.1 (**Figure 14.A**)). Similarly, a poor correlation was observed between annotated K9ac and K4me3 DER.

Genes that were altered by AOM at multiple epigenetic stages and/or transcription are listed in **Table 12**. Only 3 genes exhibited a change in transcription and K4me3 and K9ac enrichment following AOM exposure (**Figure 14.B, Table 12**). This included the upregulated novel oncogene OASL2, an interferon-induced antiviral enzyme which plays a critical role in cellular antiviral response by degrading dsDNA in the cytosol (202). RTP4 was also upregulated by AOM exposure. This is noteworthy, because RTP4 is a receptor transporter protein chaperone that escorts GPCRs to the plasma membrane, and has been identified as a breast cancer biomarker correlated with poor patient survival (203). In addition, TPM2, a cytoskeleton-regulating protein necessary for cancer cell survival, was upregulated transcriptionally, but downregulated across K4me3 and K9ac DERs.

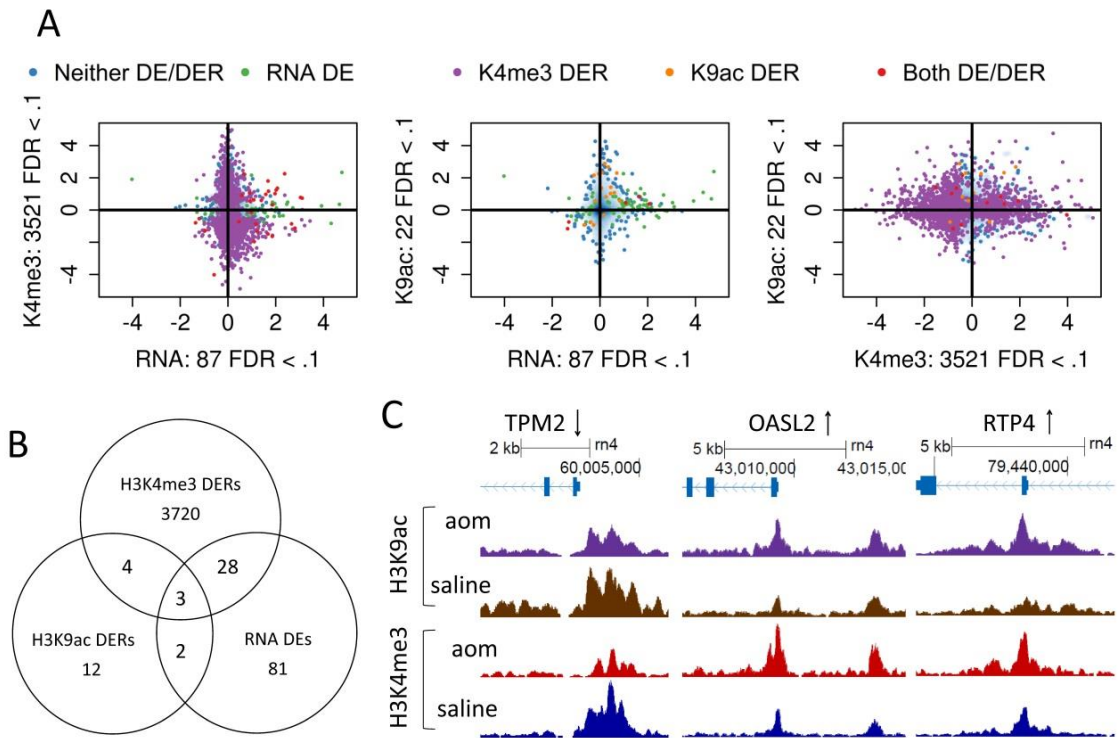


Figure 14. Correlation between transcriptional DE and histone tail DERs. A, Poor correlation is observed between DE transcripts and histone tail modifications with DERs. Scatterplots reveal low correlation between DE transcripts and histone tail modifications with DERs by comparing the $\log_2(\text{fold changes})$ from AOM vs saline treatments. Specific contrasts include RNA vs. K4me3, RNA vs. K9ac, and K4me3 vs. K9ac. B, Venn diagram illustrating the number of genes modulated by AOM in common between the various epigenetic stages (H3K4me3, H3K9ac and RNA). OASL2, RTP4, and TPM2 were affected across all measured epigenetic states. C, Chromatin signature of H3K9ac and H3K4me3 occupancy following AOM exposure is associated with upregulation of OASL2 and RTP4 and downregulation of TPM2. Representative UCSC genome browser images of all DNA reads mapped and grouped by genomic location in highly differentially enriched H3K4me3-regulated genes. H3K4me3 data from AOM treated animals are shown as red peaks, and saline control as blue peaks. H3K9ac data from AOM treated animals are shown as purple peaks, and saline control as brown peaks.

Table 12. List of AOM induced differentially expressed (DEs) and differentially enriched regions (DERs) detected across multiple epigenetic states.

Symbol	Entrez Gene Name	RNA	RNA	K4me3	K4me3	K9ac	K9ac	Location	Type(s)
		FDR	log2(FC)	FDR	log2(FC)	FDR	log2(FC)		
Anp32a	acidic (leucine-rich) nuclear phosphoprotein 32 family, member A	*	*	2.09E-02	-0.95	5.39E-02	0.50	Nucleus	other
APOB	apolipoprotein B	6.11E-03	1.40	5.91E-02	0.60	*	*	Extracellular Space	transporter
B4GALT1	UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase, polypeptide 1	1.01E-02	1.86	2.08E-04	-1.63	*	*	Cytoplasm	enzyme
Casp12	caspase 12	4.78E-02	2.07	1.51E-03	2.94	*	*	Cytoplasm	peptidase
CDH11	cadherin 11, type 2, OB-cadherin (osteoblast)	7.18E-08	3.04	1.16E-03	0.78	*	*	Plasma Membrane	other
Cxcl9	chemokine (C-X-C motif) ligand 9	1.05E-02	1.70	4.09E-02	1.17	*	*	Extracellular Space	cytokine
FIGN	fidgetin	9.74E-03	-1.31	*	*	3.38E-02	-0.74	Nucleus	other
Ifi27	interferon, alpha-inducible protein 27	9.15E-02	1.68	2.62E-05	1.80	*	*	Cytoplasm	other
Ifi47	interferon gamma inducible protein 47	8.06E-02	0.66	3.84E-05	1.14	*	*	Cytoplasm	other
IFIT2	interferon-induced protein with tetratricopeptide repeats 2	2.84E-03	1.58	4.10E-03	0.98	*	*	Cytoplasm	other
IFIT3	interferon-induced protein with tetratricopeptide repeats 3	8.85E-05	3.11	7.20E-02	0.74	*	*	Cytoplasm	other
IRF9	interferon regulatory factor 9	1.26E-05	0.91	6.79E-04	1.06	*	*	Nucleus	transcription reg
ISG15	ISG15 ubiquitin-like modifier	1.17E-03	1.87	9.02E-04	-1.17	*	*	Extracellular Space	other
NEGR1	neuronal growth regulator 1	3.81E-02	-1.36	*	*	3.85E-02	-0.68	Plasma Membrane	other
NFIA	nuclear factor I/A	*	*	4.57E-02	0.58	2.54E-03	0.47	Nucleus	transcription reg
OAS1	2'-5'-oligoadenylate synthetase 1, 40/46kDa	4.29E-04	1.97	3.97E-02	1.02	*	*	Cytoplasm	enzyme
Oasl2	2'-5' oligoadenylate synthetase-like 2	1.27E-07	2.07	2.39E-02	0.95	4.96E-04	0.78	Other	enzyme
PHLDA3	pleckstrin homology-like domain, family A, member 3	1.27E-03	2.66	9.09E-02	-0.86	*	*	Plasma Membrane	other
PRDM1	PR domain containing 1, with ZNF domain	8.86E-02	0.92	2.01E-02	0.55	*	*	Nucleus	transcription reg
RET	ret proto-oncogene	*	*	1.58E-02	1.16	7.05E-02	-0.93	Plasma Membrane	kinase
RRAGD	Ras-related GTP binding D	7.72E-03	-0.58	5.03E-04	-3.01	*	*	Cytoplasm	enzyme
RTP4	receptor (chemosensory) transporter protein 4	8.33E-03	1.08	4.08E-02	0.63	2.54E-02	0.45	Plasma Membrane	other
SRD5A2	steroid-5-alpha-reductase, alpha polypeptide 2 (3-oxo-5 alpha-steroid delta 4-dehydrogenase alpha 2)	6.58E-02	-1.20	4.84E-02	-0.54	*	*	Cytoplasm	enzyme
ST3GAL5	ST3 beta-galactoside alpha-2,3-sialyltransferase 5	5.44E-02	1.17	4.59E-02	-0.87	*	*	Cytoplasm	enzyme
TMEM50B	transmembrane protein 50B	2.80E-02	0.46	1.62E-02	-0.73	*	*	Plasma Membrane	other
Tpm2	tropomyosin 2, beta	4.68E-02	1.65	1.00E-01	-0.41	3.38E-02	-0.81	Cytoplasm	other
TTC7A	tetratricopeptide repeat domain 7A	*	*	8.56E-03	-0.85	3.65E-02	0.51	Plasma Membrane	other
VAMP7	vesicle-associated membrane protein 7	2.14E-02	-0.42	9.20E-02	1.04	*	*	Cytoplasm	transporter
ZBP1	Z-DNA binding protein 1	1.59E-03	2.02	4.32E-04	1.24	*	*	Cytoplasm	other

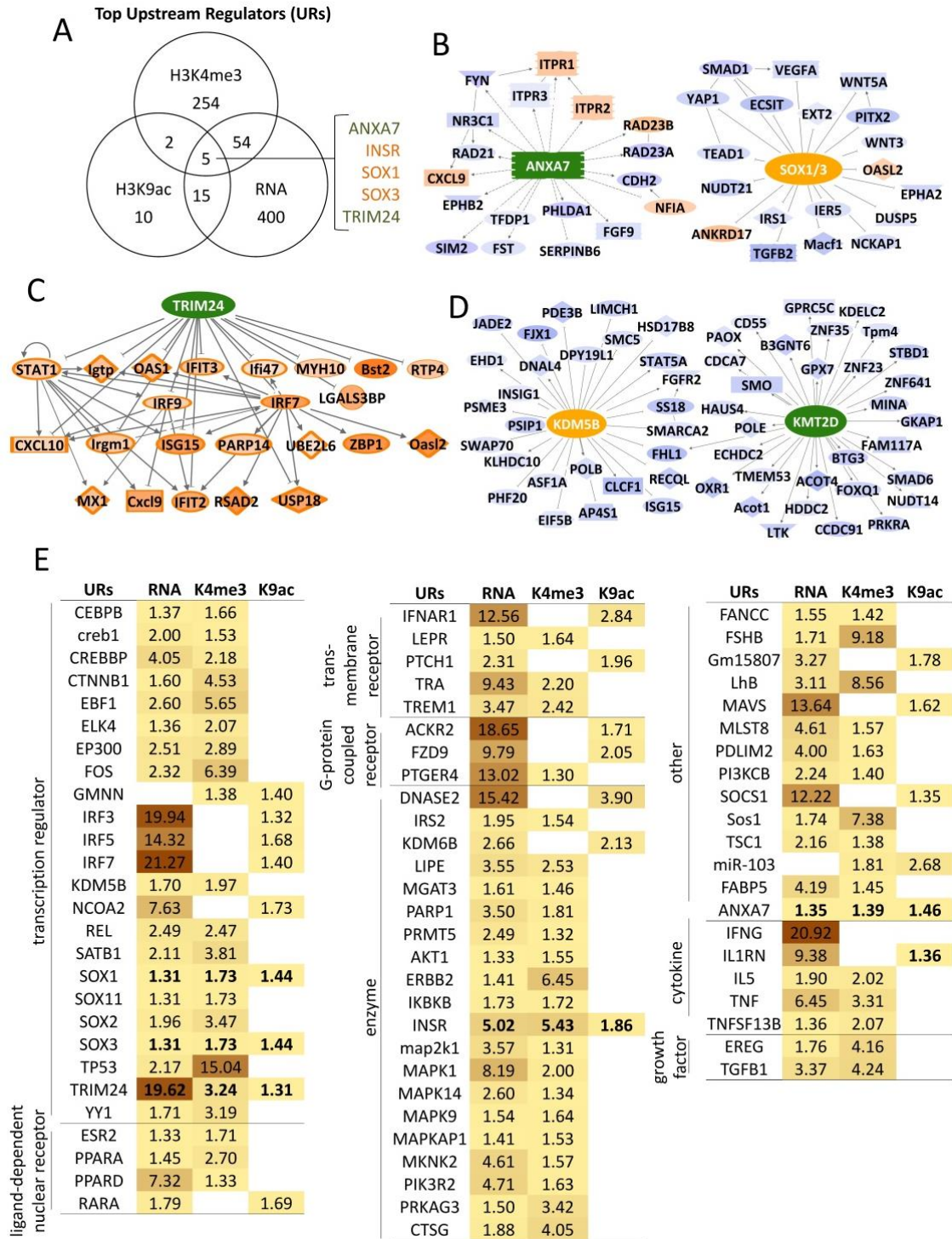
* FDR>0.1

3.3.3 Identification of upstream regulators perturbed by AOM.

To identify the cascade of transcriptional regulators linked to (can affect the expression of) AOM induced DE and DER genes, we utilized IPA Upstream Regulator (URs) analysis. This in silico analysis is based on prior knowledge of expected effects between transcriptional regulators and their target genes stored in the Ingenuity Knowledge Base. Initially, we quantified known targets of transcriptional regulators present in our dataset and compared their direction of change (over- or under-expression) to predict likely relevant regulators including transcription factors, nuclear receptors and enzymes. **Ch.III Supplementary Table 3** provides a summary Ingenuity Pathways Analysis of annotated RNA DEs and K4me3 and K9ac DER genes.

The number of AOM modulated URs in common between K4me3, RNA, and K9ac (**Figure 15.A**) was greater than the number of DER and DE genes in common between those same epigenetic levels (**Figure 14.B**). The top ranked AOM modulated URs affected at all levels (K4me3, RNA and K9ac) included the ion channel-ANXA7, insulin receptor-INSR, along with transcriptional regulators-SOX1/3 and TRIM24 (**Figures 15.A-D**). In addition to p-values, activation z-scores were used to infer likely activation states of URs based on comparison with a model that assigned random regulation directions (**Ch.III Supplementary Table 3 and Figure 15.A-D**). Using this approach, the inhibition of ANXA7 activity was linked to 17 K4me3 DERs, 2 RNA DE genes (upregulated CXCL9 and PHLDA1) and K9ac upregulated NFIA (**Figure 15.B and Ch.III Supplementary Table 3**). The activation of INSR was associated with 90 K4me3 DERs, 10 RNA DEs and 2 K9ac DEs (**Ch.III Supplementary Table 3**). Activated SOX1/3 was linked to 19 K4me3 DERs and 3 RNA DEs including OASL2, a gene affected at all measured epigenetic stages (RNA, K4me3 and K9ac).

Figure 15. Top Upstream Regulators (URs) affected by AOM across multiple epigenetic states. A, Venn diagram illustrating the total number of URs in common across epigenetic states (H3K4me3, H3K9ac and RNA). The 5 URs predicted to be affected by AOM are listed. P-values and activity (z-scores) of URs were determined by Ingenuity Pathway Analysis of genes with an FDR<0.1. B-D, Representative networks of genes regulated by top ranked upstream regulators (URs). Yellow fill represents the projected increase in UR activity, green fill represents projected decrease in UR activity. Blue fill indicates decreased gene activity (DE and/or DERs) and orange fill indicates increased gene activity (DE and/or DERs), deeper color hue indicates genes with greater $|\log_2(\text{fold change})|$. Solid lines represent direct and dashed lines represent indirect gene interaction. B, K4me3 DERs associated with activation of SOX1/3 development associated transcriptional regulators and inhibition of ANXA7, an ion channel with tumor suppressor function. C, Concurrent increase in mRNA DEs and K4me3 DERs predicted inhibition status of TRIM24 (Tif1 α), a tumor suppressor transcriptional inhibitor of proto-oncogenes that activate transcription, immune response and cell proliferation. D, Decrease in K4me3 DERs of key genes (blue) indicating AOM induced activation of KDM5B (histone K4me3 demethylase) associated with inhibition of KMT2D (histone K4 methylase). E, Heatmap of top URs in common between multiple epigenetic stages (H3K4me3, H3K9ac and RNA). Data represent $-\log(\text{Benjamini-Hochberg-adjusted p-value})$ with a score >1.29 indicating a B-H p-value<0.05. URs are grouped according to families, color scale ranges from yellow to brown, with decreasing p-value corresponding to deeper shades of brown.



Most notably, at the transcriptome level, activity inhibition of the master regulator TRIM24 was linked to 23 mRNA DEs (8 of which were also categorized as URs, including OASL2), all of which were detected within the 29 upregulated K4me3 DERs and the 1 K9ac DER (RTP4) (**Figure 15.C** and **Ch.III Supplementary Table 3**). At the H3K4 trimethylome level, only 2 of the many established K4me3 methylases and demethylases (URs) were detected as having predicted activity changes in response to AOM induced deregulation of K4me3 DERs. These histone tail modifiers were activated H3K4 demethylase KDM5B (z-score 2.2) and inhibited H3K4 methylase KMT2D (z-score -5.6) (**Figure 15.D**). Their predicted activity and inhibition, respectively, were consistent with the increased number of downregulated versus upregulated K4me3 DERs.

Among the URs detected at multiple epigenetic stages, 34 have been previously associated with adenocarcinoma (**Figure 15.E** and **Ch.III Supplementary Table 3**), the majority of which are transcriptional regulators and enzymes, especially kinases. We also cross-correlated top IPA predicted URs against all RNA DE genes to determine whether URs affected by AOM were being modulated at the early transcriptional stage (versus the translational or protein activity stages). Of the top URs predicted to affect the AOM induced transcription of DE genes, 13 were transcriptionally upregulated (APOB, CXCL10, CYBB, IRF7, IRF9, Irgm1, ISG15, SLC29A1, SPRY4, STAT1, TRAF3, USP18, ZBP1), and one was downregulated (EIF4B) (**Ch.III Supplementary Table 3**). Interestingly, these genes seem to play a role in both ‘cause’ and ‘effect’ of the transcriptional deregulation. Among K4me3 and K9ac associated URs, only FOXC1 (a K4me3 UR) was differentially expressed (**Ch.III Supplementary Table 3**).

3.3.4 Identification of pathways and networks perturbed by AOM.

To understand the biological relevance of the AOM-induced dysregulated genes, a functional analysis using the IPA algorithm was performed to identify canonical

pathways (CP). The score computed by IPA for each canonical pathway is derived from a p-value and indicates the likelihood of the genes being found together in that network by random chance (204). Output from this analysis revealed a minimal commonality between the canonical pathways affected at the K4me3 level (262 CPs) compared to RNA (28 CPs) and K9ac (2 CPs) (**Figure 16.A**). As expected, based on the number of AOM affected genes at each epigenetic stage, we observed fewer RNA DE and K9ac DER genes associated with each pathway category compared to K4me3. In the case of K9ac DERs, very few candidate genes/pathways were detected, making the analysis less reliable (**Ch.III Supplementary Table 3**).

At the transcriptome level (RNA-Seq), *Activation of IRF (interferon regulatory factor) by Cytosolic Pattern Recognition Receptors* and *Interferon Signaling* were identified as the most significantly enriched pathways (**Figures 17.A**). Furthermore, 38 of the 64 molecules that make up the *IRF Cytosolic Pattern Recognition Receptors* pathway were included in the URs associated with RNA-Seq differentially expressed genes (DEs) (**Figure 16.B**). We subsequently compared the top 2 enriched pathways with disease networks relevant to these data and noted that altered molecules in both pathways, including IRF7, ISG15, OAS1, STAT1 and TRAF3, were also primarily correlated with the top DE enriched disease network, *Infectious Disease, Antimicrobial Response, Inflammatory Response* (**Figures 17.B and Ch.III Supplementary Table 3**).

For K4me3, the top most activated and inhibited canonical pathways are shown in **Figures 17.A** ($p < 0.05$, $z\text{-score} > |1.4|$) and grouped by Signaling Pathway Categories to provide a broader overview of the effect AOM during the pre-adenomatous polyp stage of CRC. The top canonical pathway (*Molecular Mechanisms of Cancer*) mapped the greatest number (144) of genes with K4me3 DERs. Interestingly, none of these genes were differentially transcribed (DE) (**Ch.III Supplementary Table 3**). As expected, based on previous findings, *RAR Activation* and *Protein Kinase A Signaling* were also enriched (205,206).

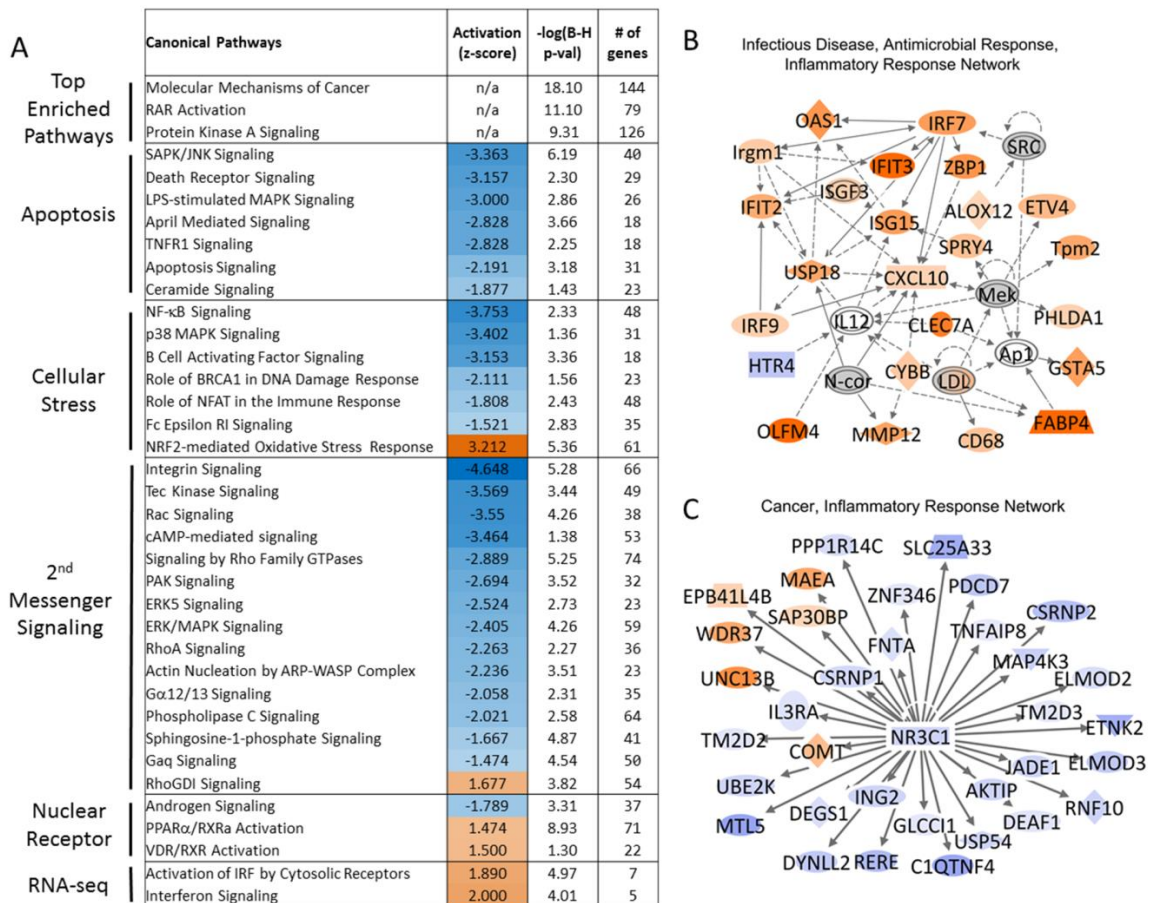


Figure 17. Top canonical pathways and networks affected by AOM. A, List of the top 3 enriched pathways, followed by top activated and inhibited pathways grouped by category. The RNA-Seq category includes the only 2 pathways that were activated/inactivated based on mRNA DE data analysis. All other canonical pathways listed were extrapolated from K4me3 DERs. Activation z-scores indicate whether an upstream transcription factor is significantly more “activated” or “inhibited” based on mRNA DE or K4me3 DER data. P-values are $-\log(\text{Benjamini-Hochberg-adjusted p-value})$ with any score >1.29 indicating a B-H p-value <0.05 . B-C, Blue fill indicates decreased gene activity (DE and/or DERs) and orange fill indicates increased gene activity (DE and/or DERs), deeper color hue indicates genes with greater fold change. Dashed lines indicate indirect interactions, solid lines indicate direct interactions. The arrow style indicates specific molecular relationships and the directionality of the interaction. B, Top network annotation of transcriptionally DE genes. C, Top network annotation of K4me3 DER genes.

The top 3 pathways in **Figures 17.A** lack an activation score because they include findings without associated directional attributes necessary for activity prediction by IPA. The disease network most associated with K4me3 DERs was *Cancer, Inflammatory Response*, centering around the glucocorticoid receptor (NR3C1) and included genes that are also part of the top affected canonical pathway (**Figures 17.C and Ch.III Supplementary Table 3**).

Examination of K4me3 DERs revealed that AOM downregulated multiple apoptotic pathways (**Figures 17.A**), however, a similar pattern was not observed at the transcriptome level. Interestingly, AOM also inhibited *Cellular Stress* associated pathways including the p38 MAPK signaling genes, MAPK12-14 and MAP2K4 (207) (**Figures 17.A and Ch.III Supplementary Table 3**).

Many second messenger signaling pathways were also inhibited by AOM induced cancer progression, including *RhoA Signaling* (**Figures 17.A**). Activated *RhoGDI Signaling* (Rho GDP-dissociation inhibitors) was the exception to this second messenger inhibitory pattern. Since RhoGDI negatively regulates Rho-family GTPases, the activation of this pathway corroborates the observed inhibition of *RhoA signaling* (**Figures 17.A**). Additionally, among the top activated diseases we found 1210 genes specifically associated with increased colorectal carcinoma, z-score=2.43 and p-value<0.0001 (**Ch.III Supplementary Table 3**). The majority of these genes were enzymes (427 genes), including upregulated G3BP1, and downregulated B3GNT6, MSH2, SOD1, MMP9, FGFR2, TGFB2, and FZR1; transcription regulators (199 genes) including downregulated H2AFX, NOTCH2, and SMAD3; and transporters (88 genes) including downregulated ABCB4, ATP2A2, and BAX (**Ch.III Supplementary Table 3**).

3.3.5 Cataloging annotated middle and long ncRNAs and un-annotated H3K4me3 and H3K9ac enriched regions.

For the purpose of assessing middle and long ncRNAs, we isolated annotated and un-annotated ncRNA reads. Based on technical restraints associated with standard RNA-Seq, in which transcripts <200 bp are eliminated prior to sequencing, the ncRNA genes reported in this section are confined to middle and long ncRNAs (**Ch.III Supplementary Table 2**). A total of 20 annotated middle and long ncRNAs were transcribed in the colonic epithelium, none of which were found to be differentially expressed. Out of the 46 annotated ncRNAs with K4me3 enriched regions, 19 contained DERs, all of which were lncRNAs except for Terc (**Table 13**). Only 1 of the 50 K9ac annotated ncRNAs were differentially enriched (**Table 13**). It is noteworthy that many of the ncRNA DERs detected have been previously reported to play a role in carcinogenesis (208). Additionally, there was no overlap between K4me3 and K9ac DERs of annotated ncRNAs induced by AOM, suggesting the same poor correlation between K9ac and K4me3 DERs and non-coding RNA DEs as observed with protein coding genes.

In an effort to extract further knowledge from the genome-wide data we also examined un-annotated genomic regions for AOM induced epigenetic changes and patterns. An assessment of K4me3 and K9ac enriched regions >5000 bp away from an annotated gene (un-annotated enriched region) exposed 2,514 DERs out of 18,477 K4me3enriched un-annotated regions and 20 K9ac DERs out of 25,426 un-annotated enriched regions (**Ch.III Supplementary Table 2**). Interestingly, there were similar numbers of upregulated (1,392) and downregulated (1,127) un-annotated K4me3 DERs, 6 of these DERs co-localized between K9ac and K4me3, 5 of which were upregulated by AOM. Thus, the originally described pattern of AOM induced K4me3 downregulation (**Figure 13**) was only observed in DERs corresponding to annotated genes. We also searched for putative middle and long noncoding RNA, de-novo generated transcripts

>150bp long that did not align to any annotated rat gene. These middle and long ncRNAs were compared for homology against annotated mouse and human genes and identified as “noncoding” by CPC to eliminate any potential un-annotated protein coding genes. A total of 242 genomic regions were identified as putative middle and long ncRNA transcripts (**Ch.III Supplementary Table 4**), however, none were differentially expressed by AOM treatment, and many of these transcripts were accompanied by K4me3 and K9ac occupancy (**Figure 18**). The majority of middle and long ncRNAs identified were categorized as small nucleolar RNA (snoRNA) by the RFAM covariance model (**Ch.III Supplementary Table 4 and Figure 18**), implying that the regulation of rRNA splicing and translation is linked to un-annotated functional RNAs in the colon.

3.4 Discussion

To our knowledge, this is the first in vivo study to globally assess the chromatin state and transcriptome profile of colonic crypt epithelial cells at a critical initial stage of cancer progression.

Our analyses were directed towards addressing 3 major biologically relevant questions: (1) are carcinogen-induced transcriptional profiles (DE) similar to DERs with respect to histone tail modifications associated with gene activity; (2), which molecules play major epigenetic regulatory roles in DE and DERs; and (3), what are the cellular functions of the DE genes and DERs?

Table 13. Description of lncRNA genes modulated by AOM at the H3K4me3, H3K9ac or transcription level.

DE middle and long ncRNAs annotated in H3K4me3 edgeR (total of 18 genes)						
assesion ID gene symbol	gene full name	log2FC (AOM/ Saline)	FDR	transcript length(bp)	cancer association	related annotated genes with DERs
NR_037147 FAM98C	family with sequence similarity 98, member C	-1.88	3.53E-05	1371	epithelial cancer	UBC
NR_027839 ATP2A2	ATPase, Ca++ transporting, cardiac muscle, slow twitch 2 (non-protein coding)	-1.5	2.43E-04	5711	colon cancer, epithelial cancer, etc.	KCNJ11, EGR1, HGS, TNF, MDM2, ATP2B1, etc.
NR_027983 LOC680254	hypothetical protein LOC680254	-1.09	3.61E-04	635	na	na
NR_038098 Ybx1-ps3	Y box protein 1 related, pseudogene 3	-2.89	3.25E-03	1541	na	STAT3, IL6ST, INSR, ERK1/2
NR_073057 NSMF	NMDA receptor synaptonuclear signaling and neuronal migration factor	-1.14	4.30E-03	2798	apoptosis	KIF5C, INA, RAN, SUPT5H, NFE2L2, MAPK3, etc.
NR_001567 Terc	telomerase RNA component	-0.81	6.05E-03	388	lymphoid cancer, hematologic cancer, etc.	na
NR_027235 RGD1559747	similar to Zinc finger and SCAN domain containing protein 2 (Zinc finger protein 29)	-1.77	4.76E-03	3298	na	na

Table 13 Continued						
assesion ID gene symbol	gene full name	log2FC (AOM/ Saline)	FDR	transcript length(bp)	cancer association	related annotated genes with DERs
NR_077057 PRRT2	proline-rich transmembrane protein 2 (non-protein coding)	-1.22	6.98E-03	1420	epithelial cancer, adenocarcinoma	na
NR_045202 Tmem80	transmembrane protein 80 (non-protein coding)	-0.61	1.27E-02	2288	na	na
NR_073042 HRH3	histamine receptor H3 (non-protein coding)	-2.34	1.49E-02	2788	breast cancer, melanoma cancer	CLIC4, GPCR, UCP1, CXCL3, MTOR, HRH, etc.
NR_073439 CREBZF	CREB/ATF bZIP transcription factor (non-protein coding)	-0.7	2.37E-02	5282	adenocarcinoma, epithelial cancer	ATF6B, MOV10, TP53, XBP1, ATF4, ELAVL1, etc.
NR_102354 PATZ1	POZ (BTB) and AT hook containing zinc finger 1	-0.8	3.32E-02	1088	lymphoid cancer, hematologic cancer	TP53, RNF4, BACH2, SOX2, AR, BCL6, etc.
NR_102345 SZRD1	SUZ RNA binding domain containing 1	1.09	3.47E-02	3205	na	ENO1, PRKCA, ETS1, UBC, ELAVL1,

Table 13 Continued						
assession ID gene symbol	gene full name	log2FC (AOM/ Saline)	FDR	transcript length(bp)	cancer association	related annotated genes with DERs
NR_051995 SMNDC1	survival motor neuron domain containing 1	-0.63	7.69E-02	2022	melanoma cancer	PRPF3,SNRPD 2, IFT99, DDX5, RPL3, PDCD7,
NR_073160 BOD1	biorientation of chromosome in cell div. 1	-0.71	8.58E-02	1474	tumorigenesis, neoplasia	MOV10, UBC, APP, NXF1
NR_027324 H19	imprinted maternally expressed transcript (non- protein coding)	0.55	8.68E-02	2369	tumorigenesis of intestinal polyp,	E2F6, PARP1, P53, ZFP57, DNMT1,.
NR_073134 AKNAD1	AKNA domain containing 1	1.21	8.84E-02	1952	na	APP
NR_038113 Rpl37a-ps1	ribosomal protein L37a, pseudogene 1	0.8	9.25E-02	359	na	na
DE lncRNA annotated in H3K9ac edgeR (total of 1)						
assession ID gene symbol	gene full name	log2FC (AOM/ Saline)	FDR	transcript length(bp)	cancer association	related annotated genes with DERs
NR_002597 LOC360231	MHC class I RT1.O type 149 processed pseudogene	1.13	1.45E-02	1621	na	na

A

Summary total of all detected un-annotated transcripts and H3K4me3 and H3K9ac enriched regions			Summary of annotated enriched regions
un-annotated peaks	total #	co-localized with putative lincRNA (208)	co-localized with putative incRNA (36)
K4me3 enriched regions	32,353	92	26
K4me3 DERs	2,514	29	4
K9ac enriched regions	27,031	94	2
K9ac DERs	20	1	0
K4me3 + K9ac enriched regions	10,760	79	26
K4me3 plus K9ac DERs	4	0	0

DERs= differentially enriched regions
lincRNA = long intergenic RNA
incRNA= intronic contained RNA

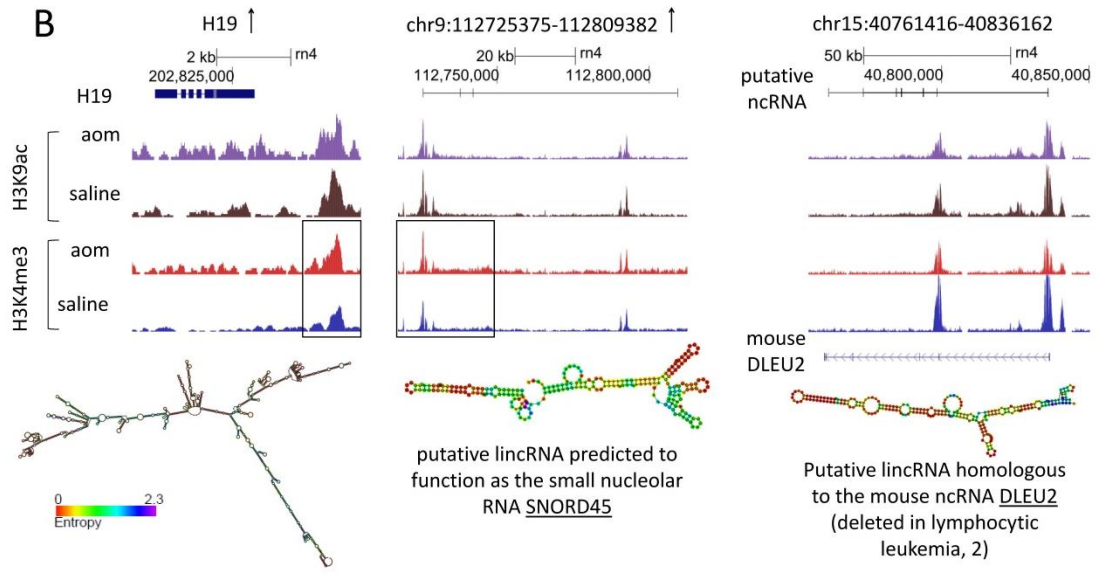


Figure 18. H3K4me3 and H3K9ac enriched regions co-localize with ncRNAs. A, Putative ncRNA loci with K4me3 and K9ac enriched regions. (B-C) UCSC genome browser snapshots of ncRNA, as well as their conserved, thermodynamically stable secondary structures predicted by RNAfold (below); putative ncRNAs also include predicted annotation. The entropy color scale represents the values at the weakest spots of the structure where 0 entropy means no deviations, and entropy >0 indicates some deviations; the higher the entropy, the more likely for the folding structure to deviate. B, Genomic occupancy and structure of annotated and putative ncRNAs with unregulated K4me3 DERs (black box) and putative ncRNA homologous to a mouse annotated lincRNA.

3.4.1 AOM selectively modulates H3K4me3 genome association.

Although carcinogen effects were identified at the RNA (116 DE genes), K9ac (49 DERs, including 24 genes) and K4me3 (7678 DERs, including 3792 genes) levels, the most notable overall changes were detected in the 3171 downregulated K4me3 genes. These novel findings indicate that AOM suppresses the enrichment of H3K4me3 in a very large number of genes at an early stage of colon cancer progression (**Figure 13**). We also noted that DE genes were poorly correlated with K4me3 (<28%) and K9ac (<24%) DERs (**Figure 14.B**).

In order to identify proteins that may affect the expression of genes detected in our experimental dataset, we employed IPA Upstream Regulator (UR) Analysis. We noted that carcinogen-induced changes in gene transcription and H3K4 trimethylation were more correlated at the UR pathway level (76 URs vs 37 DE/DERs in common) (**Figures 14.B & 15.A**). This is consistent with previous findings indicating that translational alterations are more extensive relative to transcriptional alterations in mediating AOM induced malignant transformation in the colon (115). Collectively, these data suggest that the lack of correlation between DEs and DERs may be attributed to distinct cofactors already positioned at corresponding transcription start sites.

We were also interested in identifying which of the many H3K4 methylases and demethylases (such as KDM2A, KDM7 and KMT3C-E) that regulate gene expression could be contributing to the bias towards downregulation of K4me3 DERs in our data. We noted that specific key downregulated DERs were associated with H3K4me3 demethylase KDM5B and methylase KM2TD (**Figure 15.D and Ch.III Supplementary Table 3**) both of which have been implicated in cancer development and proliferation (209,210) . These findings suggest that increased KDM5B and decreased KM2TD activity are responsible for the AOM induced overall downregulation of K4me3 DERs.

3.4.2 AOM induced transcriptional changes are strongly correlated with antimicrobial responses.

RNA-Seq results at the DE and pathway analysis levels, revealed a surprisingly strong link to interferon associated innate immune responses (**Figure 15.C and E and Figures 17.A-B and Ch.III Supplementary Table 3**). This included 58 interferon associated DE genes, the URs IFNG, IFNAR1, IRF-3,5,7, TRIM24, the top pathways *Interferon Signaling* and *Activation of IRF by Cytosolic Receptors*, and the *Infectious Disease, Antimicrobial Response, Inflammatory Response* network. Furthermore, recent evidence analyzing the effect of TRIM24 inhibition at various stages (week 5, week 14 and tumor) of hepatocarcinogenesis, indicated the presence of increased transcription of STAT1 along with many immune response associated DE genes (211). It is noteworthy, that many of these genes were also upregulated in our study (BST2, CXCL9, IFIT2, IFIT3, IFIT47, IGTP, IRF7, IRF9, IRGM, ISG15, MYH10, OASL2, RTP4 and USP18) and were uniquely expressed at the earlier stages of carcinogenesis (week 5) (211). Taken together, our results are consistent with recent findings indicating that cancer related changes in gut microbial composition may induce interferon associated responses that contribute to the early progression of colon cancer (202,212,213). Additionally, many key genes associated with microbial immune response were upregulated at the RNA plus K4me3 levels, hence, they may act as key modifiers of cancer progression (202,213,214).

One of the most prevalent URs detected from the analysis of both DE and DER genes was TRIM24. We propose a putative mechanism for future investigation in which decreased RXR/RAR activity at specific sites (in a context specific manner) during colon cancer (215) leads to a decrease in KDM5B binding (216). Since TRIM24 cannot bind to trimethylated H3K4 (217), while KDM5B can (218), an increase in K4me3 subsequently leads to a decrease in TRIM24 (211) binding, thus inducing colon cancer (214).

3.4.3 Carcinogen induced changes in K4me3 DERs may predict for future transcriptional events

Despite the fact that there were extensive links between K4me3 DERs and colon tumorigenesis (**Ch.III Supplementary Table 3**), these changes were not reflected at the RNA gene expression level during early cancer progression (10 weeks post AOM exposure). For example, among the K4me3 DERs, 1210 have been linked to colorectal carcinoma (**Ch.III Supplementary Table 3**). Biomarkers of colon tumors with K4me3 DERs, corroborated by previous findings, included: FOXO3, which regulates tumor suppressor genes associated with apoptosis (177), the G Protein GNAI2, NR3C1 (glucocorticoid receptor), PDCD4 (programmed cell death 4), VEGFA (an endothelial growth factor, and the DNA repair associated proteins H2AFX and MSH2 (180,219-221). Furthermore, we discovered a range of lncRNAs with K4me3 DERs (**Table 13**), many of which have been previously associated with cancer, including colon cancer associated ATP2A2, CREBZF, PRRT2, H19 (222). Based on these findings, we propose that K4me3 DERs are harbingers of future transcriptional inducers of carcinogenesis.

Following AOM exposure, the observed decrease in *Cellular Stress* associated pathways and genes, including *p38 MAPK Signaling* associated genes (**Figures 17.A**) is consistent with previous findings indicating that MAPK activity may be down-regulated in colorectal cancer (207). Further supporting evidence of this finding includes the association of K4me3 DERs with the downregulation of multiple apoptotic pathways, a common phenotype during tumorigenesis (176,223). Reports of decreased apoptotic pathway activity (and transcriptional decrease of apoptosis associated genes) in tumors but not in earlier stages of colon cancer (181,224), are consistent with our data where a decrease in apoptosis pathways was not detected transcriptionally at the HM-ACF stage, but was detected in the K4me3 DERs, a marker of potential future changes in gene activity.

In summary, we have documented for the first time the chromatin structure associated with gene expression profiles of an in vivo murine colonic tumorigenesis model. Our high-throughput sequencing approach revealed many expected changes at various regulatory stages of gene expression, plus unexpected insight into gene regulation during colon cancer progression. Specifically, we were able to show that AOM induced transcriptional deregulation was primarily associated with interferon-associated immune response genes, while K4me3 deregulation was linked to genes associated with colon tumorigenesis, perhaps acting as a harbinger of changes in gene activity. These findings emphasize the value of genome-wide analyses and may have important clinical relevance for future therapeutic targeting of histone demethylases (210) and microbial composition alteration (212).

CHAPTER IV
DIETARY FAT AND FIBER COMPOSITION ALTER GLOBAL HISTONE POST-
TRANSLATIONAL EPIGENETIC PROGRAMMING IN A RAT COLON CANCER
PROGRESSION MODEL

4.1 Introduction

”Nutri-epigenetics”- the influence of dietary components on mechanisms manipulating epigenetic programming, which is encoded by specific histone modifications (methylation and acetylation) has been identified as a promising new field for cancer prevention strategies (44,225). Recent evidence now indicates that epigenetic alterations contribute to cancer-related cellular defects. For example, epigenetic silencing of critical genes, e.g., detoxifying enzymes, tumor suppressor genes, cell cycle regulators, apoptosis-inducing and DNA repair genes, nuclear receptors such as PPARs, FXR, HNF4A, signal transducers and transcription factors, by promoter methylation and modification of histones and non-histone proteins including p53 and NF-κB, by acetylation or methylation, drives malignant transformation (62,225).

With respect to dietary chemoprevention, increasing numbers of published reports indicate a protective effect of fish oil and its bioactive components, n-3 polyunsaturated fatty acids (PUFA), against colon cancer (55,69,111,226,227). Similar reports exist focusing on the two main long n-3 PUFA found in fish oil, eicosapentaenoic acid (20:5Δ^{5,8,11,14,17}) and docosahexaenoic acid (DHA;22:6Δ^{4,7,10,13,16,19}) (79,82). In contrast, dietary lipids rich in n-6 PUFA [found in vegetable oils; e.g., linoleic acid (LA;18:2Δ^{9,12}) and arachidonic acid (20:4Δ^{5,8,11,14})] have been linked to an increase in colon tumor development (112,228,229).

DHA and EPA and their oxidative metabolites have been shown to interact with specific ligand dependent nuclear receptors including CAR, HNF4A, PPARG, PXR and RXRA (**Figure 1**) (46). Since the original description of dietary fat as a regulator of gene expression over a decade ago, many transcription factors have also been identified

as prospective indirect targets for n-3 PUFA regulation. For example, DHA can increase the activity of CREBBP, EP300, and MYC and decrease activity of NF- κ B (NFKB1), and STAT3 (46,225). In this fashion, n-3 PUFAs epigenetically regulate the function of nuclear receptors and transcriptional factors and thereby regulate transcriptional processes.

There are also many studies linking dietary fiber, gut microbiota and colon cancer prevention (230). The major metabolites produced by gut microbiota from readily fermentable fiber are short-chain fatty acids (SCFAs), such as butyrate, which has multiple beneficial effects at the intestinal and extra-intestinal level (231). As more dietary fiber is ingested, SCFAs production increases (232). Even though some controversies remain concerning the *in vivo* responses of colonocytes to butyrate exposure (233), butyrate has been shown to impact cell kinetics, lumen pH, and epigenetics which regulate development of colon carcinogenesis (103,109,234). There are at least 2 epigenetic mechanisms by which butyrate can increase histone acetylation in the. (235). (107) Bacteria in the gut generate SCFAs including butyrate, which can act as a histone deacetylase (HDAC) inhibitor or alternately, in colonocytes, butyrate can be metabolized into acetyl-CoA and used for energy, or transported to the nucleus and act as substrate by histone acetylases (HATs) (232).

Animals fed diets containing n-3 PUFA (fish oil) with pectin (which is fermented to butyrate) as a fiber source maximally promote apoptosis in the colon compared with diets high in other dietary lipids, e.g., n-6 (corn oil) (86,98,109,236). In a follow-up study, the administration of butyrate-containing pellets for targeted release in the colon was used to demonstrate that butyrate and fish oil work coordinately in the colon to trigger apoptosis (104,112,234). Subsequently, DHA and butyrate were shown to work synergistically to enhance apoptosis in colonocyte cultures compared with butyrate alone (100,101). A clinical example of the chemoprotective effects of fat x fiber interaction was recently reported by Orlich et al., which validated the pescovegetarian diet as the most protective regimen in terms of colorectal cancer prevention (118). Outside of human experimentation, the AOM chemical carcinogenesis model serves as one of the

most definitive means of assessing human colon cancer risk (128,237). We have previously demonstrated that at 10 weeks post AOM injection, the colonic mucosa is precancerous, e.g., high multiplicity aberrant crypt foci are apparent. Macroscopic tumors are not detectable until ~34 weeks post AOM injection (98). During carcinogenesis, major cellular functions and pathways, including drug metabolism, cell cycle regulation, DNA damage repair and targeted apoptosis, response to inflammatory stimuli, cell signaling, and cell growth control and differentiation become deregulated (3,209,238,239).

In this study, we employ novel technologies and bioinformatics algorithms, such as next-generation sequencing, in order to explore 'nutri-epigenomics' at a genome-wide level and determine the importance of epigenetic mechanisms related to chemoprevention. By determining 3 epigenetic levels of regulation (H3K4me3, H3K9ac and mRNA expression) in colonocytes from the same animals, we were able to gain a greater understanding of the chromatin state associated with the interaction of fish oil (rich in DHA/EPA n-3 PUFAs), pectin (a readily fermentable fiber) and AOM (colon carcinogen) treatments. Our data contribute to the understanding of the regulatory action of chemoprotective bioactive compounds found in fish oil and readily fermentable fiber (n-3 PUFAs and SCFAs) in terms of malignant transformation of the colonic crypt. Mechanistic insights provided by this study are needed to provide perspective with regard to recent human clinical and epidemiological findings (118,119).

4.2 Methods

See 3.2 Methods Section (Chapter 3) for details on methodology.

4.3 Results

Since we have previously demonstrated that dietary n-3 PUFA and pectin suppress colon tumorigenesis (66,101,112), experiments were designed to contrast the chemoprotective diet (fish oil and/or pectin) against the control diet containing non-chemoprotective oil and fiber (corn oil and cellulose) in the presence and absence of carcinogen. Three main biological comparisons were examined: 1) dietary fat effects (fish oil vs corn oil), 2) dietary fiber effects (pectin vs cellulose) and 3) dietary fat \times fiber interaction in the presence of carcinogen [fish oil + pectin + AOM (FPA) vs corn oil + cellulose + AOM (CCA)]. (APPENDIX A)

4.3.1 Enumeration of aberrant crypts

Whereas AOM-treated rats developed aberrant crypts, their saline-treated counterparts did not. Therefore, all aberrant crypt results represent AOM-injected groups only. Total HM-ACF were higher ($p < 0.0001$; data not shown) in the distal colon compared with the proximal colon, independent of the diet. The corn oil + cellulose treatment resulted in a greater number of total HM-ACF compared with other treatment groups. In addition, the fish oil + pectin-fed rats displayed the lowest incidence of HM-ACF compared with all other groups (**Figure 19**).

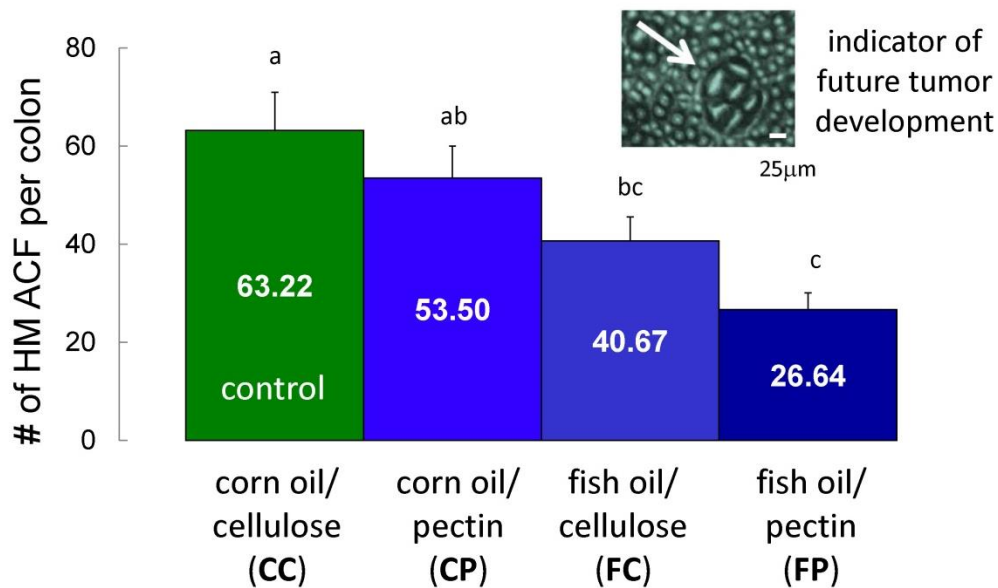


Figure 19. Fish Oil + Pectin synergistically suppress malignant transformation of the colon. High Multiplicity Aberrant Crypt Foci (HM ACF) incidence per rat colon is shown. Bars not sharing the same superscript letters are significantly different. $p < 0.05$.

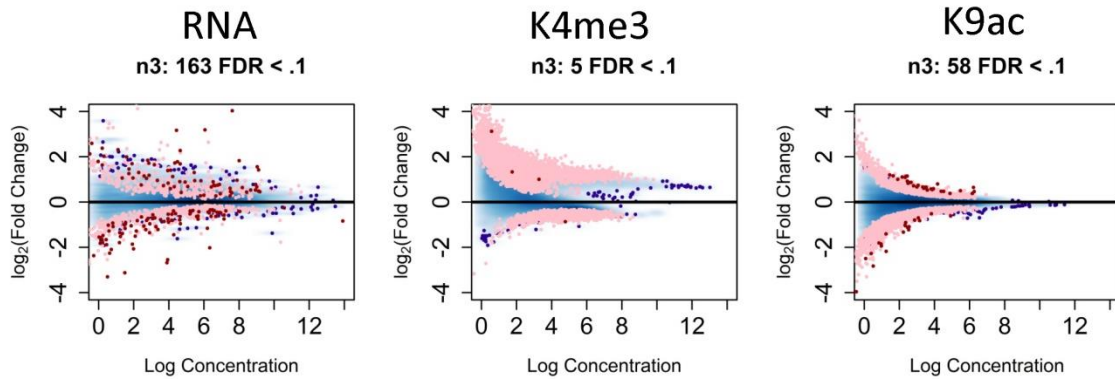
4.3.2 Overall effects of fish oil on transcription and histone tail modifications

In an effort to identify key, genome-wide, bioactive effects associated with fish oil and pectin, gene expression profiling by Next Generation Sequencing of RNA, H3K4me3 and H3K9ac ChIPed DNA was performed. Fish oil induced differentially expressed transcripts (DE) and differentially enriched chromatin regions (DERs) were determined by pooling sequence data from individual rats fed a fish oil diet across the various treatment groups in comparison to rats fed a corn oil diet (fcs+fps+fca+fpa versus ccs+cps+cca+cpa) (n=34 fish oil fed rats versus n=33 corn oil fed rats). Similarly pectin induced DEs and DERs were determined by pooling rats fed a pectin diet across the various treatment groups and comparing them against the rats fed a cellulose diet (cps+fps+cpa+fpa versus ccs+fcs+cca+fca) ((n=34 pectin fed rats versus n=33 cellulose fed rats).

Figure 20 shows the distribution of expression strength relative to the log-ratio of DE and DERs. Included above each MAplot are the total number of DE genes (including different isoforms) and the total number of DERs (annotated and un-annotated). **Table 14** provides a summary of genes with differentially expressed (DE) transcripts and differentially enriched peaks (DERs) with an FDR<0.1 and a p-value<0.01 in fish oil vs corn oil treated animals and readily fermentable (pectin) versus poorly fermentable (cellulose) fiber. **ChIV.Supplementary Table 1** details DE and DERs with FDR<0.1 and p<0.01. Enriched regions (peaks) further than 5kb from the transcription start site were considered “un-annotated” and classified with an ID representing the rat (rn4) genomic location of that island.

Fish oil feeding altered the transcription of 160 genes along with 32 K9ac and 3 K4me3 genes with DERs (FDR<0.1) affected by fish oil (**Figure 20 and Table 14**), with similar amounts of up- and downregulated genes detected at each level (**Ch.IV Supplementary Table 1**). Based on previous studies indicating that histone tail modifications regulate gene expression at the transcriptional level (178), we expected genes with K4me3 and K9ac DER to correlate with RNA differentially expressed (DE) genes. Generation of a global plot of all the K4me3 and K9ac DER fold changes against RNA DE genes revealed poor correlation between genes with these histone marks and RNA, regardless of p-values (the axes include number of genes with an FDR<0.1 (**Figure 21**)). Similarly, a poor correlation was observed between annotated K9ac and K4me3 DERs (**Figure 21**). There were no genes with an FDR<0.1 that were simultaneously modulated at the RNA, K9ac and K4me3 level by fish oil, and a more relaxed filtering parameter using genes with a Fisher’s Exact test p-value<0.01 revealed only 4 genes were simultaneously affected at all epigenetic levels (**Figure 22**). These genes were the upregulated tumor suppressors CDH11 (cadherin 11) and SCD2 (stearoyl-Coenzyme A desaturase 2) and downregulated oncogenes CERS4 (ceramide synthase 4) [Dlamini et al 2015] and PDE4B (cAMP-specific phosphodiesterase 4B) [Smith PG 2005].

Overall Fish Oil Effect



Overall Pectin (Readily Fermentable Fiber) Effect

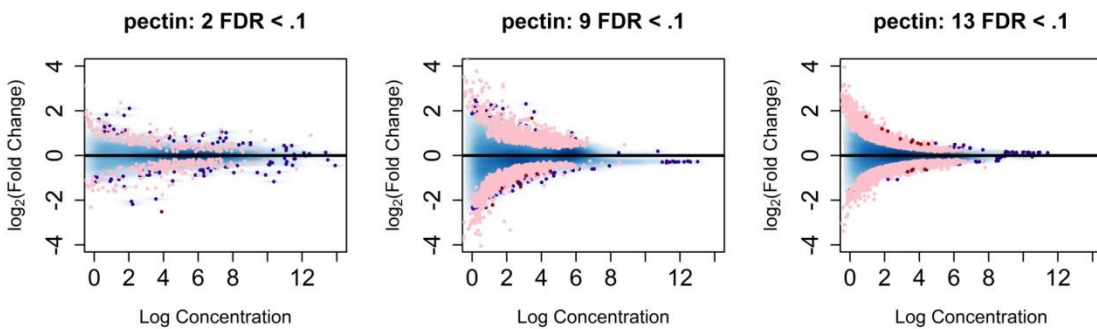
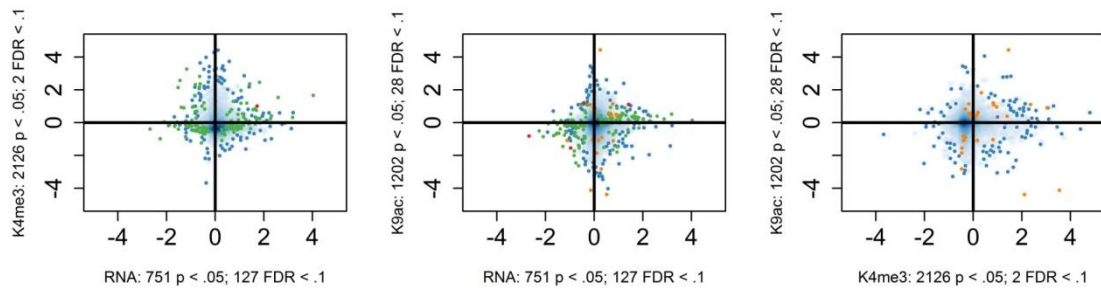


Figure 20. Differential expression (DE) of all transcribed genes or histone tail differentially enriched regions (DERs). MAplots indicate the differential expression of all transcribed genes or histone tail enriched regions (y-axis, log-ratio of difference in intensity of histone tail modifications enriched regions) vs their overall intensity of expression (x-axis, log-average of read counts) following AOM vs saline treatment. Pink represents differentially expressed (DE) transcripts and differentially enriched regions (DERs) with a p-value < 0.05. Genes with an FDR < 0.1 are highlighted in red, and all other detected genes are highlighted in blue.

Overall Fish Oil Effect



Overall Pectin (Readily Fermentable Fiber) Effect

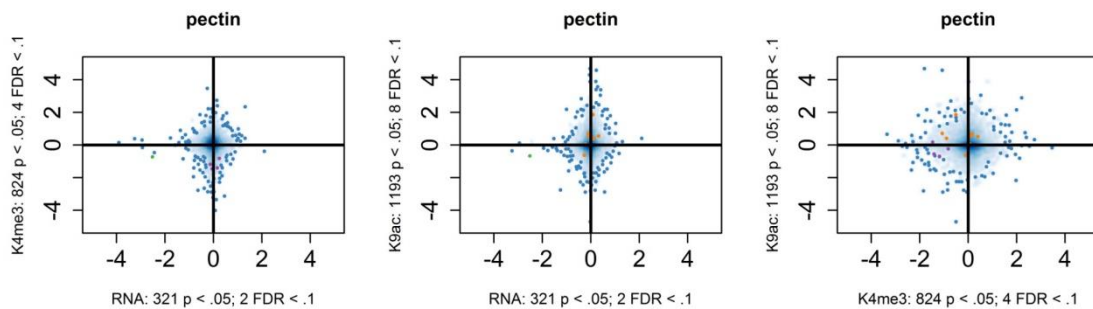


Figure 21. Poor correlation between DE transcripts and histone tail modifications with DERs. Scatterplots reveal low correlation between DE transcripts and histone tail modifications with DERs by comparing the $\log_2(\text{Fold Changes})$ from AOM vs saline treatments. Specific contrasts include RNA vs. K4me3, RNA vs. K9ac, and K4me3 vs. K9ac.

Among the fish oil (fcs+fps+fca+fpa versus ccs+cps+cca+cpa) DE genes, 8 have been previously associated with docosahexaenoic acid (DHA), a major bioactive component of fish oil, including the upregulated transporter FABP1, a fatty acid binding protein often downregulated in colon cancer and upregulated by DHA (240,241). Four mitochondrial enzymes previously shown to be modulated by DHA were also upregulated, acyl-CoA synthetase ACSBG1(242), carnitine palmitoyltransferases CPT1 and CPT2 (243), and the kinase MAP2K1 (66). Stearoyl-CoA desaturase SCD (244) was also upregulated and phospholipase A2 PLA2G1B was downregulated (245).

Overall Fish Oil Effect

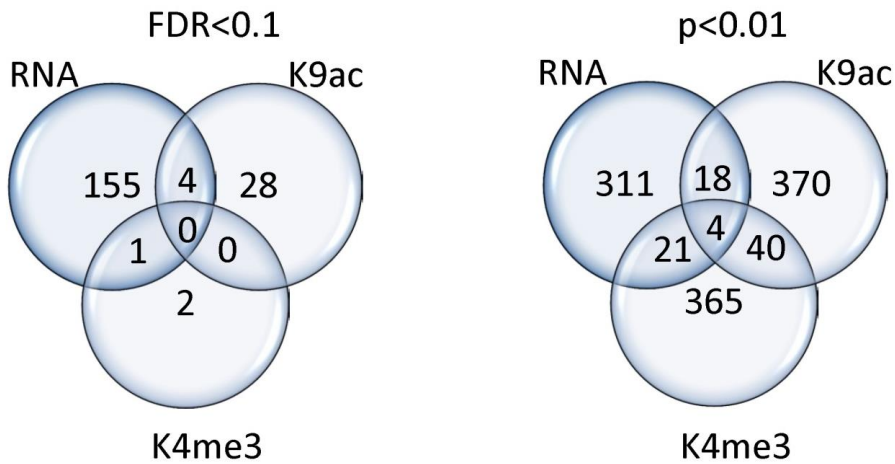


Figure 22. Poor correlation between DE transcripts and histone tail modifications with DERs modulated by fish oil. Overlap of differentially expressed (DE) or differentially enriched (DER) genes across multiple epigenetic levels (FDR<0.1 and p<0.01) is shown.

Table 14. Summary of diet effects on differentially expressed (DE) and differentially enriched (DERs) genes.

	K9ac	K9ac	K4me3	K4me3	RNA	RNA
	FDR<0.1	p<0.01	FDR<0.1	p<0.01	FDR<0.1	p<0.01
fish oil overall	32	217	3	430	160	367
pectin overall	9	432	6	222	2	102
fpa vs cca	0	400	21	228	83	278
fca vs cca	0	432	0	172	1	90
cpa vs cca	0	480	15	189	0	18
fps vs ccs	15	272	3	89	14	145
fcs vs ccs	4	303	0	133	63	302
cps vs ccs	24	703	0	38	0	55

Since genes typically function as part of intricate networks, prior biological knowledge greatly facilitates the meaningful interpretation of the gene-expression changes associated with large datasets. Therefore, we used pathway analysis to help interpret the data in the context of biological processes, pathways and networks. Functional analysis of DEs and DERs using the Ingenuity Pathway Analysis (IPA) software was performed in an effort to better understand the biological relevance of the fish oil dysregulated genes. Pathway analysis of the 32 genes with K9ac DERs (FDR<0.1) revealed 23 genes related to Metabolic Disease, Lipid Metabolism, and Cell Death and Survival networks (**ChIV.Supplementary Table 2**). Only 3 genes were differentially enriched at the K4me3 level (fcs+fps+fca+fpa versus ccs+cps+cca+cpa), e.g., downregulated metalloproteinase ADAMTS9A and upregulated peptidase PCSK9 and transcription regulator PBXIP1. Network analysis of our dataset using the less stringent filtering parameter ($p < 0.01$) revealed similar types of biological processes affected by fish oil at the RNA and K4me3 levels. Therefore, although the RNA and K4me3 networks were mostly comprised of different genes, the top affected metabolism pathways were correlated with Lipid Metabolism, Small Molecule Biochemistry, and Vitamin and Mineral Metabolism (**Figure 23 and ChIV.Supplementary Table 2**). Interestingly the top network affiliated with K9ac DER genes with a $p < 0.01$ were more closely related to Cell-to-Cell Signaling and Interaction (**Figure 23**).

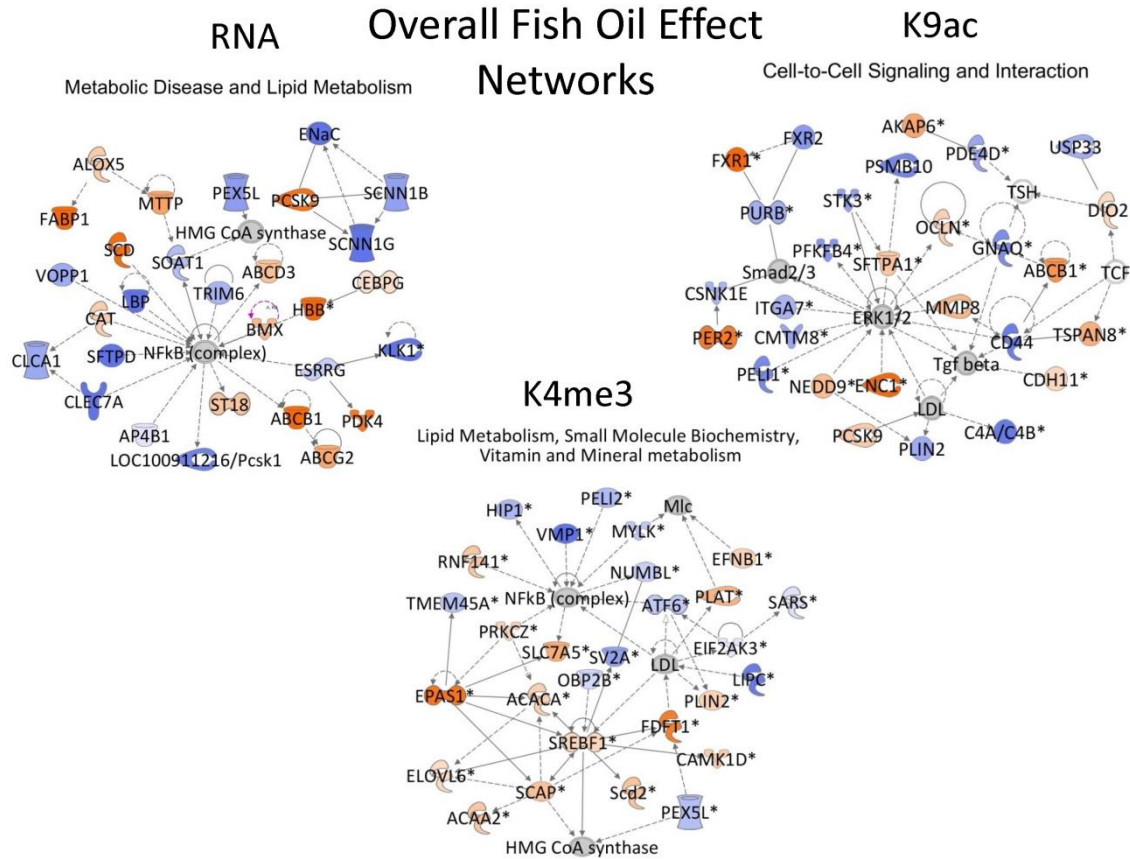


Figure 23. Top networks affected by fish oil. Network analysis of differentially expressed genes ($p < 0.01$) identified with Ingenuity Pathway Analysis (IPA) software in response to fish oil feeding (fcs+fps+fca+fpa versus ccs+cps+cca+cpa). Blue fill indicates decreased gene activity (DE and/or DERs) and orange fill indicates increased gene activity (DE and/or DERs). Deeper color hue indicates genes with greater fold change. Dashed lines indicate indirect interactions and solid lines indicate direct interactions. The arrow style indicates specific molecular relationships and the directionality of the interaction.

4.3.3 Overall effects of pectin on transcription and histone tail modifications

Very few genes with an $FDR < 0.1$ were modulated by pectin (cps+fps+cpa+fpa versus ccs+fcs+cca+fca) at any measured epigenetic levels (**Table 14**). When considering the genes using a $p < 0.01$ cutoff, the overall effects of the readily

fermentable fiber pectin were most noticeable at the K9ac level, with 432 K9ac gene DERs compared to 222 genes with K4me3 DERs and 102 RNA DEs ($p < 0.01$) (**Figure 24**). We detected a greater number of upregulated (250) as compared to downregulated (150) genes with K9ac DERs, while similar amounts of up- and downregulated genes were detected at the K4me3 and RNA level (**ChIV.Supplementary Table 1**). No genes with an $FDR < 0.1$ were simultaneously affected at the RNA, K9ac and K4me3 level by pectin, and only 2 genes with $p < 0.01$ were simultaneously affected between levels (**Figure 24**). Those genes that were upregulated include EGFLAM (EGF-like, fibronectin type III and laminin G domains) and downregulated oncogene ANXA3 (annexin 3) [Xie Oncol Lett. 2013].

Pathway analysis (IPA) of ($p < 0.01$) pectin modulated genes revealed the top network associated with differentially enriched transcripts (DEs) was related to Cell-to-Cell Signaling and Interaction, Cancer, Organismal Injury and Abnormalities. K4me3 DERs were linked to Lymphoid Tissue Structure and Development, Organ Morphology, Organismal Development, while K9ac DERs were related to Post-Translational Modification, Protein Folding, Cell Morphology (**Figure 25 and ChIV.Supplementary Table 2**).

4.3.4 Context specific epigenetic effects of fish oil and pectin

We compared the differentially transcribed genes (DEs) and K4me3 and K9ac differentially enriched (DERs) genes in common between saline and AOM injected rats fed the same diet (**Table 15**) to determine whether the bioactive compounds had the same effects on a healthy colon and a colon undergoing cancer progression (fps v ccs compared to fpa v cca, fcs v ccs compared to fca v cca, and cps v ccs compared to cpa v cca). Dietary effects differed in healthy (saline) vs carcinogenic (AOM) conditions at all epigenetic levels, e.g., few genes (~10%) were commonly dysregulated in rats fed the same diet but injected with saline instead of AOM. Only in fish oil + pectin fed rats, at the transcription level, did we observe a slightly higher number of genes in common,

with 42 genes differentially expressed in both saline and AOM treated animals (**Table 15**). Among these genes 17 were upregulated and 18 were downregulated in both datasets. In contrast, 7 additional genes were upregulated in AOM injected animals but downregulated in saline injected animals fed the same diet (**ChIV.Supplementary Table 1**).

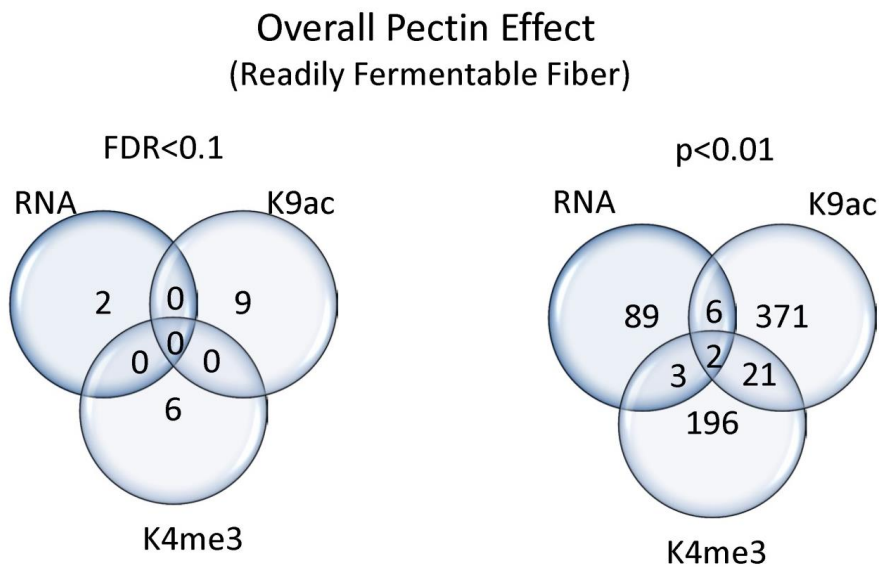


Figure 24. Poor correlation between DE transcripts and histone tail modifications with DERs dysregulated by pectin. Overlap of differentially expressed (DE) or differentially enriched (DER) genes across multiple epigenetic levels (FDR<0.1 and p<0.01).

Overall Pectin Effect

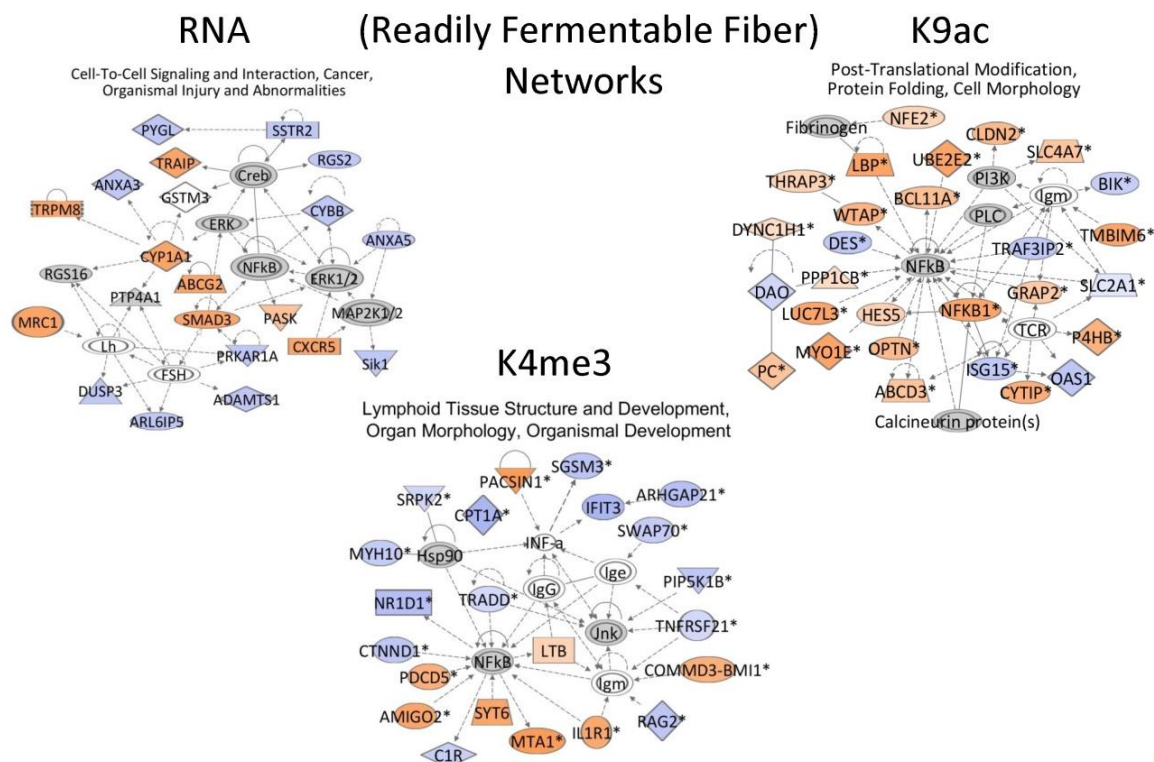


Figure 25. Top networks affected by pectin feeding. Network analysis of differentially expressed genes ($p < 0.01$) identified with Ingenuity Pathway Analysis (IPA) software in response to pectin. Blue fill indicates decreased gene activity (DE and/or DERs) and orange fill indicates increased gene activity (DE and/or DERs). Deeper color hue indicates genes with greater fold change. Dashed lines indicate indirect interactions, and solid lines indicate direct interactions. The arrow style indicates specific molecular relationships and the directionality of the interaction.

To assess the consequence of each bioactive compound in a context specific manner (fat x fiber interaction), we used the corn oil (rich in n-6 PUFAs) and cellulose (rich in poorly fermentable fiber) diets (111) under AOM and saline conditions as control (fpa v cca, fca v cca, cpa v cca, and fps v ccs, fcs v ccs, cps v ccs) (**Table 14**). For each of these treatment comparisons we observed poor correlation between the epigenetic levels (RNA, K4me3 and K9ac). Furthermore the epigenetic and functional response of the colonocytes to fish oil and pectin individually were often not reflected in

the combination fish oil+pectin diet (**Figures 25 and 26**). Among these 6 individual comparisons, the most noteworthy changes, with an FDR<0.1, were the transcriptional effects of fish oil + pectin under carcinogenic conditions (fpa vs cca) with 83 DEs (**Table 14**). Pathway analysis revealed 29 of those genes were associated with biological processes linked to lipid metabolism, especially increased beta-oxidation of fatty acids (**Figure 8**). Nineteen of these genes were linked to cellular functions associated with a decreased accumulation of lipids (such as cholesterol, acylglycerols and fatty acids) ($p<0.0001$, z-score -1.91) and 17 were linked to increased fatty acid metabolism ($p<0.001$, z-score 2.37). More specifically, these genes included ABCD3 and CPT2, two genes that are vital to the mitochondrial l-carnitine shuttling process, beta-oxidation acyl-CoA enzymes, ACSBG1 (which also plays a role in the activation of DHA) and ACADM. In addition, genes linked to the metabolism of acyl-CoA (DBI and ACOT1) and transporters that boost beta-oxidation of very long chain fatty acids (ABCD3 and FABP1) were detected (**Figure 26**). Examination of the 267 DE genes with a $p<0.01$ (fpa vs cca) revealed an increase in the number (68) of lipid metabolism affiliated genes (**Figure 26 and ChIV.Supplementary Table 4**).

Since the most chemoprotective diet was fish oil + pectin, we were also interested in which biological functions were most affected by this diet during carcinogenesis at the histone modification levels (K4me3 and K9ac). **Table 16** includes the category with the lowest multiple testing corrected p-value of predicted biological functions affected by fish oil + pectin + AOM (fpa) at every epigenetic level. Analysis was performed using DEs and DERs with $p<0.01$ (summarized in **Table 14**) from each pairwise comparison and highlighted cells representing the percentage of all DEs or DERs in each treatment are colored by epigenetic level, with deeper hues indicating a greater percentage of dysregulated genes were associated with specified biological functions.

Table 15. Summary of differentially expressed (DEs) and differentially enriched (DERs) genes in common between saline and AOM injected rats fed the same diet.

bioactive	unique saline	both & saline	aom	unique aom	epigenetic level	treatments compared
pectin	35	3		186	K4me3	cp vs cc
fish oil	127	6		166	K4me3	fc vs cc
fish oil + pectin	78	11		217	K4me3	fp vs cc
pectin	53	1		17	RNA	cp vs cc
fish oil	280	15		74	RNA	fc vs cc
fish oil + pectin	105	42		225	RNA	fp vs cc
pectin	641	62		418	K9ac	cp vs cc
fish oil	270	33		399	K9ac	fc vs cc
fish oil + pectin	244	28		372	K9ac	fp vs cc

4.3.5 Identification of upstream regulators perturbed by dietary fat and fiber interaction

Due to the enhanced chemoprotective activity of the bioactive fat x fiber diet during carcinogenesis, we utilized IPA Upstream Regulator (URs) analysis to identify the transcriptional regulators linked to fish oil + pectin (fpa vs cca) n=11 rats per treatment modulated DE and DER genes. Causal networks constructed from individual relationships curated from the literature were used to create mechanistic hypotheses that explained expression changes. We used a statistical approach to determine and score those Upstream Regulators whose connections to our dataset genes were unlikely to occur in a random model (246). Initially, we quantified known targets of transcriptional regulators present in our dataset and compared their direction of change (over- or under-expression) to predict likely relevant regulators including transcription factors, nuclear receptors and enzymes. This analysis was performed on the gene sets of all 6 individual comparisons mentioned in **Table 14** at each epigenetic level (total of 18 gene sets with $p < 0.01$) and also on the fpa vs cca DE genes with an $FDR < 0.1$.

Table 16. Top categories of biological functions affected by fish oil + pectin under carcinogenic conditions at each epigenetic level (parenthesis), for each treatment the listed category was predicted by pathway analysis with a certainty of $p < 0.05$.

Lipid Metabolism (RNA)				Gastrointestinal Disease (K4me3)				Cellular Movement (K9ac)			
treatment	epigenetic level	% of		treatment	epigenetic level	% of		treatment	epigenetic level	% of	
		DE/DERs	B-H adjp			DE/DERs	B-H adjp			DE/DER	B-H adjp
fpa	RNA	24%	7.73×10^{-9}	fpa	RNA	62%	8.13×10^{-4}	fpa	RNA	19%	2.1×10^{-2}
fca	RNA	22%	1.13×10^{-2}	fca	RNA	16%	4.27×10^{-2}	fca	RNA	11%	4.27×10^{-2}
cpa	RNA	11%	2.18×10^{-2}	cpa	RNA	28%	3.41×10^{-2}	cpa	RNA	6%	5.67×10^{-2}
fps	RNA	20%	3.11×10^{-2}	fps	RNA	12%	3.11×10^{-2}	fps	RNA	21%	3.11×10^{-2}
fcs	RNA	15%	4.51×10^{-4}	fcs	RNA	20%	2.25×10^{-4}	fcs	RNA	19%	2.99×10^{-4}
cps	RNA	25%	2.59×10^{-2}	cps	RNA	13%	2.35×10^{-2}	cps	RNA	15%	2.59×10^{-2}
fpa	K9ac	13%	7.41×10^{-3}	fpa	K9ac	60%	2.56×10^{-2}	fpa	K9ac	23%	2.27×10^{-4}
fca	K9ac	7%	1.46×10^{-2}	fca	K9ac	64%	4.31×10^{-4}	fca	K9ac	22%	3.93×10^{-5}
cpa	K9ac	9%	1.28×10^{-2}	cpa	K9ac	59%	9.24×10^{-4}	cpa	K9ac	18%	1.67×10^{-2}
fps	K9ac	18%	4.67×10^{-2}	fps	K9ac	63%	8.35×10^{-2}	fps	K9ac	21%	6.11×10^{-2}
fcs	K9ac	11%	4.08×10^{-2}	fcs	K9ac	58%	6.62×10^{-2}	fcs	K9ac	9%	1.23×10^{-1}
cps	K9ac	13%	3.58×10^{-4}	cps	K9ac	56%	3.97×10^{-3}	cps	K9ac	17%	3.09×10^{-5}
fpa	K4me3	4%	1.03×10^{-1}	fpa	K4me3	60%	6.91×10^{-3}	fpa	K4me3	15%	1.03×10^{-1}
fca	K4me3	6%	7.52×10^{-2}	fca	K4me3	63%	7.52×10^{-2}	fca	K4me3	17%	7.52×10^{-2}
cpa	K4me3	13%	4.52×10^{-2}	cpa	K4me3	52%	1.45×10^{-2}	cpa	K4me3	4%	7.22×10^{-2}
fps	K4me3	13%	1.65×10^{-2}	fps	K4me3	67%	9.93×10^{-3}	fps	K4me3	25%	2.86×10^{-2}
fcs	K4me3	15%	6.74×10^{-2}	fcs	K4me3	66%	6.74×10^{-2}	fcs	K4me3	20%	6.74×10^{-2}
cps	K4me3	3%	9.83×10^{-2}	cps	K4me3	18%	9.13×10^{-2}	cps	K4me3	5%	6.3×10^{-2}

B-H adjp = Benjamini-Hochberg Multiple Testing Correction p-value of biological functions predicted by IPA pathways analysis

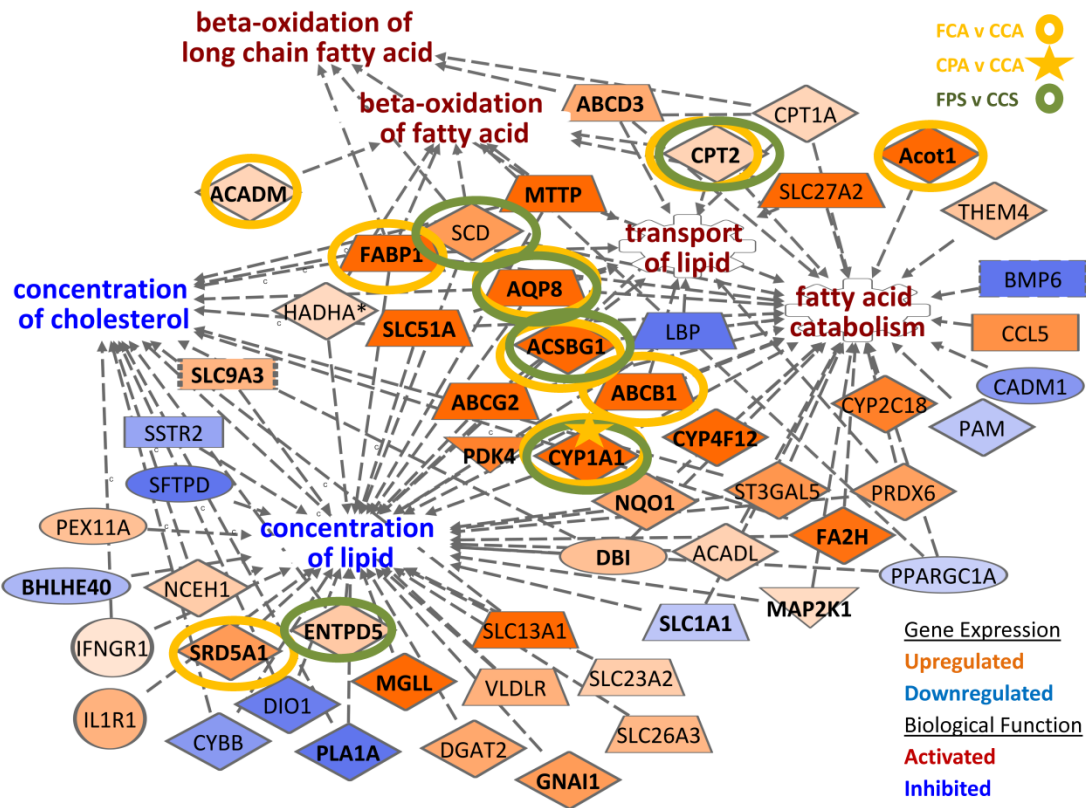


Figure 26. Fish oil + pectin transcriptionally enhance β -oxidation and decrease lipid concentration associated genes. Analysis of differentially expressed genes associated with lipid metabolism. Lines indicate the association between a biological process and each gene. Colors of biological functions indicate whether the differential expression of genes were associated with activation or inhibition of that biological process ($p < 0.001$). Differentially expressed genes with and $FDR < 0.1$ are bolded. Ovals and star indicate genes also dysregulated in diets/treatments different from fpa by any one variable including fat, fiber or carcinogen.

Ligand dependent nuclear receptors were prevalent among the top modulated URs connected to the greatest numbers of differentially transcribed (DE) genes in fpa vs cca (**Figure 27, 28** and **ChIV.Supplementary Table 3**). The nuclear receptors included activated PPARs alpha, delta, and gamma along with LXR (NR1H), FXR (NR1H4), PXR (NR1I2), GCR (NR3C1) and HNF4A. Some of these nuclear receptors were also

activated, to a lesser extent, in other fish oil containing diets under saline and carcinogenic conditions at the transcriptome and K9ac levels (**Figure 27 and 28**). Except for glucocorticoid receptor (NR3C1), HNF4A and FXR (NR1H4), the nuclear receptors were predicted to be activated only at the RNA and K9ac epigenetic levels. Furthermore, 39 of the 68 fpa vs cca DE genes associated with lipid metabolism were also part of the 58 genes directly connected to the nuclear receptors. The lipid metabolism related genes with K9ac differential enrichment (DERs) were different from lipid metabolism DE genes (**ChIV.Supplementary Table 2**).

4.4 Discussion

To our knowledge this is the first in vivo study to globally assess changes in histone post translational modifications and the transcriptome in colonocytes during colon cancer progression. There were two major novel findings from this study involving the chemoprotective effects of fish oil (rich in n-3 PUFAs) and readily fermentable fiber (a source of butyrate) during colon cancer progression: (i) a diet containing fish oil triggers the transcription of lipid metabolism-associated genes, including beta-oxidation, and is predicted to activate n-3 PUFA stimulated ligand dependent nuclear receptors; and (ii) administration of a fish oil + pectin diet during colon cancer progression synergistically modulates transcription more than diets containing fish oil or pectin separately.

Poor correlation was observed between differentially transcribed (DE) and enriched genes (DERs) at multiple epigenetic levels in fat x fiber dietary combinations and in the presence/absence of carcinogen (**Figures 21-27 and Tables 15-16**). We propose the chemoprotective influence of the bioactive compounds fish oil + pectin (rich in n-3 PUFAs and butyrate) in colon cancer progression are multifaceted and generate a specific epigenetic modification and transcriptional profile in a context specific manner.

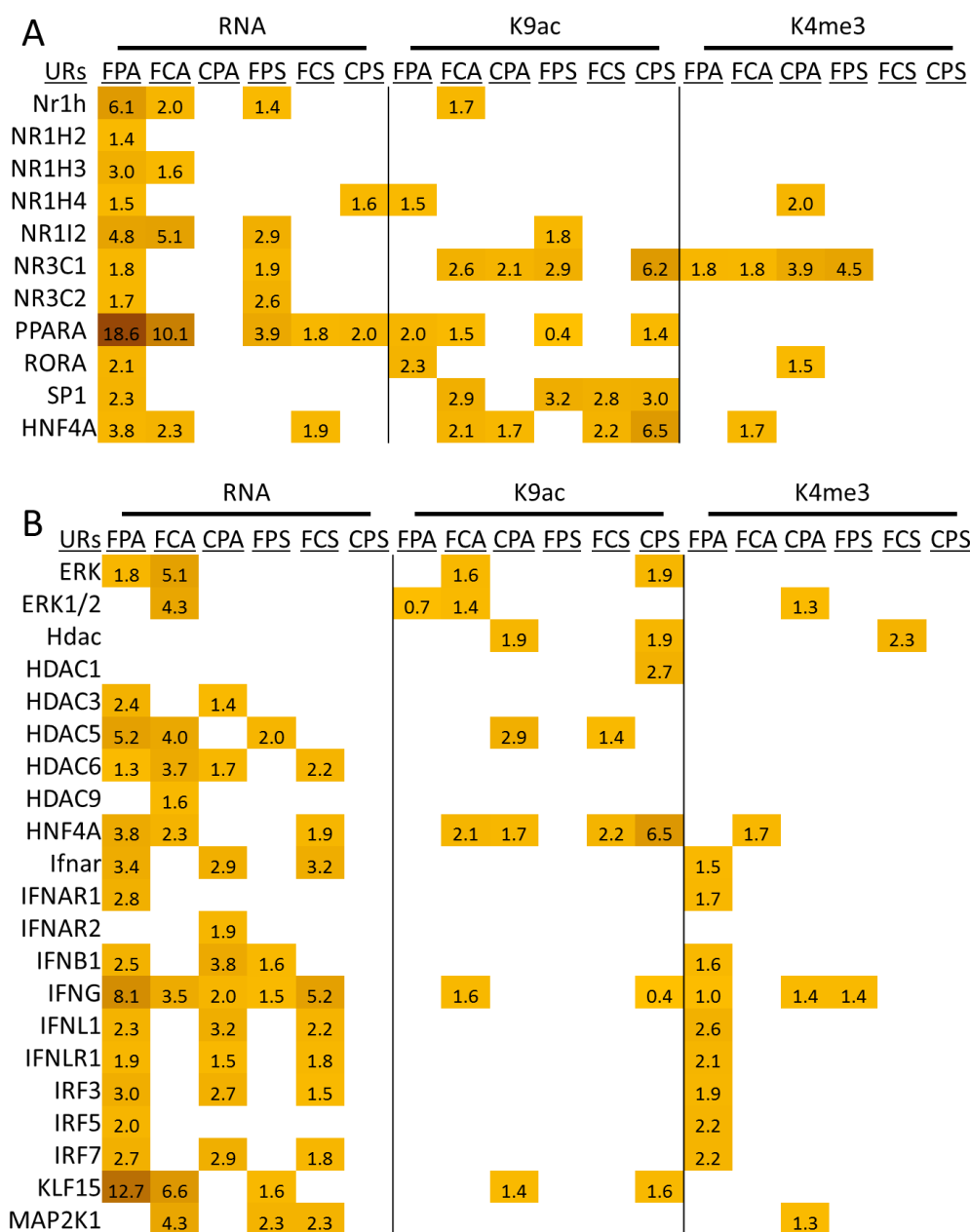


Figure 27. The chemoprotective effects of Pectin + Fish Oil during cancer progression included predicted activation of nuclear receptors. Heatmap of A. ligand-dependent nuclear receptors and B. chromatin modifiers (HDACs) and interferon associated URs, indicates the top Upstream Regulators in fish oil + pectin fed rats during cancer progression. Data represent $-\log(\text{Benjamini-Hochberg-adjusted } p\text{-value})$ with a score >1.29 indicating a B-H $p\text{-value} < 0.05$. URs are grouped according to families, color scale ranges from yellow to brown, with decreasing p-value corresponding to deeper shades of brown.

4.4.1 Pectin preferentially affects H3K9 acetylation of saline over AOM injected rats

Overall, when considering all pectin (readily fermentable fiber) containing diets (cps+fps+cpa+fpa) versus all cellulose (poorly fermentable fiber) containing diets (ccs+fcs+cca+fca) with a $p < 0.01$, we detected more upregulated (519) than downregulated (322) K9ac DERs. This result is generally consistent with previously published data (109,247).

Since poor correlation was observed between differentially transcribed (DE) and enriched genes (DERs) at multiple epigenetic levels in fat x fiber dietary combinations, we analyzed the effects of pectin on genes with K9ac differential enrichment in a context specific manner. A comparison of the number of upregulated versus downregulated K9ac DERs with $p < 0.01$ in each individual pectin containing diet versus the control corn oil + cellulose (fpa v cca, cpa v cca, fps v ccs, and cps v ccs) revealed 154 upregulated and 108 downregulated K9ac DERs in fps vs ccs, while the other pectin containing diets exhibited similar numbers of upregulated and downregulated K9ac DERs (**ChIV.Supplementary Table 1**). Previous studies have also described a global increase in histone acetylation in response to butyrate administration in saline but not AOM injected animals. Specifically, Crim et al. reported that mucosal levels of nuclear H4ac and fecal butyrate levels in fish oil + butyrate fed rats increased in comparison to corn oil + butyrate fed rats following saline but not AOM injection (112). These findings suggest that histone hyperacetylation, an epigenetic effect traditionally associated with increased levels of butyrate in the colon, is more commonly associated with 'healthy/normal' colonic epithelial cells compared to colonocytes undergoing cancer progression.

In general, our data are inconsistent with some previous studies in which histone hyperacetylation was enhanced following butyrate treatment in human adenocarcinoma cells exhibiting a Warburg-like metabolic profile (248,249). Our genome-wide analyses of histone 3 acetylated at lysine 9 (H3K9ac) enrichment did not detect upregulated K9ac

DERs at genes associated with the Warburg effect, including Fas, p21 and p27. There was also no general increase of H3K9 acetylation, in pectin fed rats, at the cancer progression stage, however it is still possible that a global assessment of histone 3 or histone 4 pan-acetylation (H3ac and H4ac) could produce results similar to Donohoe et al. (248). Furthermore, it is worth noting that histone hyperacetylation following butyrate treatment was assessed by Donohoe et al. at the tumorigenesis stage of colon cancer cells or was triggered under conditions of elevated glucose treatment, which was not the case in our investigation (248,250). This is noteworthy because our experimental design explored a different stage of colon cancer development and therefore provides additional mechanistic insight into a poorly studied time point (prior to polyp and tumor formation) in this multistep, multipath disease.

4.4.2 Fish oil triggers transcription of lipid metabolism associated genes

Mounting evidence indicates that dietary fish oil protects against colon cancer (111,251-253). Interestingly, we noted that fish oil feeding induced transcriptional changes in many genes associated with lipid metabolism (**Table 16**). Specifically, a large cluster of upregulated genes were associated with increased fatty acid catabolism and a decreased accumulation of lipid (**Figure 26**). The transcriptional changes we observed are supported by previous studies by Mori et al. who assessed the effect of fish oil on the small intestine of obese mice and identified many of the same lipid metabolism related genes modulated in our study, e.g., ACOT1, ACADM, CAT, CPT1A, CPT2, MGLL, PDK4, PEX11A (254,255) (**Figure 27**).

Upstream Regulators detected in fish oil containing diets under carcinogenic conditions were ligand dependent nuclear receptors (**Figure 27**). Previous reports supporting our predictions suggest that fish oil-derived n-3 PUFAs interact with the function of the nuclear receptors (NRs) presented in **Figure 28** and modulate colon cancer and colitis in a chemoprotective manner (62,225,265,266). Of the nuclear receptors known to be induced by n-3 PUFAs (PPARs, LXRs, FXRs, HNF4A and SREBPs), the majority were predicted to be induced by fish oil + pectin fed rats under carcinogenic conditions (**Figure 27 and 28**) (267).

4.4.3 Fish oil plus pectin synergistically modulate transcription

Even though the predicted activation of beta-oxidation has been previously associated with fish oil and butyrate individually (255,268-270), only ACADM, FABP1, and CPT2 were upregulated in *fca v cca*, and none of the beta-oxidation associated genes were upregulated by *cpa v cca*. The majority of beta-oxidation associated genes were only upregulated in the combinatorial diet (*fpa v cca*) (**Figure 26**). We also noted that, during carcinogenesis, the majority of the genes modulated by fish oil + pectin (*fpa v cca*) have been previously attributed to the effects of fish oil in other disease models (254,255), suggesting that pectin (in the form of short chain fatty acids, including butyrate) enhances the chemoprotective effects of fish oil during colon cancer progression (**Figure 26 and ChIV. Supplementary Table 1**) (64,270,271).

By considering the intersects from cross correlating the genes linked to the n-3 PUFA activated nuclear receptors described in **Figure 28** against the lipid metabolism associated genes in **Figure 26** we found all the beta-oxidation associated genes reported in **Figure 26** and ~50% of the genes associated with activated lipid catabolism (**ChIV. Supplementary Table 2**). These results imply the transcriptional activation of lipid metabolism genes is likely triggered by fish oil activated nuclear receptors.

The augmented beta-oxidation associated with fish oil + pectin (fpa) feeding may be explained by the ability of n-3 PUFA to activate ligand dependent nuclear receptors (**Figure 26 and 28**). In the fish oil only diet, we propose that n-3 PUFA activated NRs resulting in the transcription of beta-oxidation and lipid catabolism associated genes. This would result in the enhancement of beta-oxidation, promoting the catabolism of n-3 PUFAs in the mitochondria. In the combination fish oil + pectin diet, we observed an even greater increase in beta-oxidation and lipid catabolism associated genes. Under these conditions, we propose that enhanced butyrate entry into the mitochondria (from pectin), a preferential source of energy in colonocytes (250), provides additional substrate for beta-oxidation, allowing a greater percentage of the n-3 PUFAs (from fish oil) in the cell to act as nuclear receptor ligands, thereby inducing the transcription of even more beta-oxidation associated genes (**Figure 26 and 28**) (255,267,272,273). These results support our hypothesis that, in colonocytes, fish oil related effects are enhanced by the inclusion of pectin to the diet during the onset of carcinogenesis (268,274).

The majority of research examining cancer metabolism has been limited to a handful of metabolic pathways, mainly linked to glycolysis, glutaminolysis and fatty acid synthesis. Thus the relevance of beta-oxidation of fatty acids to cancer biology remains obscure (275). Current dogma suggests that fatty acid oxidation may promote carcinogenesis by providing extra substrate for the generation of ATP (275). However, we did not observe any evidence of increased transcription or activity of AMPK or ATP generating enzymes such as IDH1, ME1, G6PDH, 6PGLDH (**Figure 26 and ChIV. Supplementary Table 1**) nor did we observe a change in animal body weights of animals in any treatment group (APPENDIX B).

Studies addressing changes in transcriptional and histone tail modifications in colonic epithelial cells upon fish oil or fermentable fiber feeding typically detect a limited number of genes as assessed by qRT-PCR, in-vitro models, and/or the use of each bioactive fish oil or fiber individually (40,98,105,248,276). In contrast, in this study we applied a global approach that facilitated an examination of the entire genome

in an unbiased manner and concurrently compared changes in histone tail modifications to the transcriptome.

With respect to tissue collection, we chose to assess epigenetic changes among all the colonocytes in the distal colon for various reasons. First, in-vivo ChIP-seq of the colon requires a relatively high number of cells. Second, any dietary effects on the colon are likely to be present in all/any of the colonocytes and third, in order to detect fat and fiber “field effects” (a term which designates carcinogenic molecular abnormalities present in tissues that appear histologically normal) (277,278) during colon cancer progression.

In summary, we have documented for the first time the chromatin structure associated with a diet rich in fish oil (rich in n-3 PUFAs) and readily fermentable fiber (rich in short chain fatty acids) under normal (saline control) and carcinogenic conditions. We propose that the chemoprotective fish oil + pectin diet generates unique epigenetic modifications and transcriptional profiles in a context (AOM vs saline) specific manner. Genes associated with enhanced beta-oxidation were transcriptionally upregulated uniquely in fish oil + pectin (fpa v cca) fed rats. Only a small subset (3) of beta-oxidation genes previously linked to fish oil were upregulated in fish oil + cellulose fed rats (fca v cca). Upstream regulator analysis revealed the activation of ligand dependent nuclear receptors predominantly in the presence of pectin (fpa v cca). We therefore hypothesize that, in colonocytes, fish oil related epigenetic effects are enhanced by concurrent pectin feeding during the progression of carcinogenesis. A potential mechanism for this phenomenon may include the concurrent enhancement of beta-oxidation and the induction of futile mitochondrial respiration (proton leak) that leads to the uncoupling of ATP synthesis, resulting in nutrient wasting.

In conclusion, our data contribute to the understanding of the regulatory action of chemoprotective bioactive compounds found in fish oil and readily fermentable fiber (n-3 PUFAs and SCFAs) in colonic crypts and provide critical mechanistic insight into recently reported epidemiological findings (118).

CHAPTER V

SUMMARY AND FUTURE DIRECTIONS

Long n-3 PUFA from fish oil and fermentable fiber are ideal colon cancer chemotherapeutics because (1) they are toxicologically innocuous and free of safety problems intrinsic to drugs administered over long periods of time, (2) are relatively inexpensive, and (3) provide additional health benefits, such as reduction in mortality (Lien 2009, Cockbain, Toogood et al. 2012, Bell, Kantor et al. 2014). In addition, the simultaneous ingestion of fish oil and fermentable fiber can improve their efficacy in colon cancer prevention/therapy (Chang, Chapkin et al. 1998, Crim, Sanders et al. 2008, Davidson, Wang et al. 2009, Jia, Ivanov et al. 2011, Fenton and McCaskey 2013).

From an epigenetic perspective, there is still much to be discovered in terms of the effects of long n-3 PUFA and fermentable fiber in the colon at the transcriptional and chromatin state level. In this study, we employ novel technologies and bioinformatics algorithms, such as next-generation sequencing, in order to explore 'nutri-epigenomics' at a genome-wide level by using an in-vivo, pre-clinical model to better comprehend the importance of epigenetic mechanisms related to chemoprevention. Our ultimate objective is to contribute to the understanding of the regulatory action of chemoprotective bioactive compounds found in fish oil and readily fermentable fiber (n-3 PUFAs and SCFAs) in terms of malignant transformation of the colonic crypt. Mechanistic insights gained in this study are needed to interpret human clinical and epidemiological findings (O'Keefe, Li et al. 2015, Orlich, Singh et al. 2015).

In order to answer these questions we probed the epigenetics of the colon along its longitudinal axis (Aim 1), the effects of AOM, a colon carcinogen, (Aim 2), and the effects of fish oil (fat rich in n-3 PUFAs specifically DHA and EPA) and a readily fermentable fiber (pectin) on the distal colon (Aim 3).

Aim 1: Determine chromatin structure associated with differential expression and phenotypic responses in the distal versus proximal colon epithelial crypts by correlating ChIP-Seq to mRNA transcription data.

Aim 2: Monitor early pre-tumorigenic epigenetic molecular events driving colon cancer by generating high-resolution genome-wide transcriptional and “chromatin-state” maps for intestinal epithelial cell crypts in rats injected with a colon-specific carcinogen or saline.

Aim 3: Gain epigenetic insight into the mechanisms of nutritional chemoprevention by generating high-resolution transcriptional genome-wide “chromatin-state” maps for intestinal epithelial cell crypts in rats (i) injected with a colon-specific carcinogen or saline, and (ii) fed a diet containing fish oil (n-3 PUFA), (iii) pectin (butyrate) or (iv) a combination of both fish oil and pectin.

5.1 Aim 1 (Chapter II)

Many diseases of the colon exhibit an anatomical bias and proximal and distal colorectal cancers have been reported to exhibit distinct gene-specific methylation profiles, transcriptional profiles, and molecular and clinical characteristics (Deng, Kakar et al. 2008, Hiraoka, Kato et al. 2010, Mladenova, Daniel et al. 2011). Since there are no cell lines that uniquely represent the proximal or distal colon, the need to identify regional epigenetic differences at the chromatin level using an *in vivo* model remains unfulfilled because of the challenges associated with chromatin immunoprecipitation of organ tissues for high throughput sequencing.

Therefore, we determined the chromatin structure associated with gene expression profiles in the rat proximal and distal colon by globally correlating chromatin immunoprecipitation next generation sequencing analysis (ChIP-Seq) with mRNA transcription (RNA-Seq) data. This characterization provides the first available *in vivo* genome-wide map of histone PTMs in the rat colon with respect to the relationship

between chromatin alterations and gene transcription. For this purpose, we globally identified the transcriptome and sites of H3K4 trimethylation (H3K4me3) in colonic crypt epithelial cells. By assessing the correlation between histone PTMs and transcriptional data, we identified canonical pathways influenced by differential H3K4me3 occupancy and RNA expression. We were also able to catalogue the middle and long non-coding RNAs transcribed in the colon, including select long non-coding RNAs (lncRNAs) formerly only detected in the rat nervous system. In addition, in silico analysis of transcription factor (TF) activity revealed tumor suppressors and oncogenes unique to the distal and proximal colon. Knowledge of the proximal versus distal colon epigenetic landscape will assist in the improved detection, therapy and prognoses of colonic disease.

In summary, we documented for the first time the chromatin structure associated with gene expression profiles in the rat proximal and distal colon by correlating ChIP-Seq with RNA-Seq data. Globally, approximately 500 genes were differentially expressed between the proximal and distal colon. With regard to differentially expressed genes, a high correlation was observed between H3K4me3 occupancy and RNA-Seq data. Gene ontology analysis indicated that colonic crypt location significantly impacted both chromatin and transcriptional regulation of genes involved in cell transformation, lipid metabolism, lymphatic development and immune cell trafficking. We were also able to detect a range of lncRNAs that have not been previously reported in the colon. In addition, gene function analysis indicated that the PI3-Kinase signaling pathway was regulated in a site-specific manner. In conclusion, distinct combinatorial patterns of histone modifications exist in the proximal versus distal colon. These site-specific differences may explain the differential effects of bioactive chemoprotective agents on cell transformation in the ascending (proximal) and descending (distal) colon.

Our first aim also allowed us to refine and standardize our novel ChIP-seq and bioinformatics protocol and focus on the epigenetic differences of the colon with respect to anatomically different (distal vs proximal colon) regions. This was different from the

second and third aim which focused on how treating the cells with various compounds, carcinogenic and chemoprotective, would affect the distal colon.

5.2 Aim 2 (Chapter III)

During carcinogenesis major cellular functions and pathways, including drug metabolism, cell cycle regulation, potential to repair DNA damage or to induce apoptosis, response to inflammatory stimuli, cell signaling, and cell growth control and differentiation are dysregulated (18). Epigenetic alterations contribute to these cellular defects. (175-178).

Chromatin signatures are tightly linked to epigenetic regulation. For instance, transcriptionally active genes are characterized by active chromatin marks, such as trimethylated histone H3 lysine 4 (K4me3) and acetylated histone H3 lysine 9 (K9ac) (175,178). Alterations in these histone modifications can drive oncogenic processes, such as proliferation, invasion, angiogenesis, and dedifferentiation, by perturbing normal gene expression patterns (178). This is particularly relevant, because altered K4me3 levels are associated with the onset of colorectal cancer (179).

Few studies to date have integrated diverse epigenetic inputs in an in-vivo model of CRC. This type of analysis would improve the statistical and interpretative power of the changes in transcription and chromatin state during cancer progression. Since distinct genomic and epigenetic events drive the initiation, promotion and progression of colon cancer (18), in this study we integrated global chromatin immunoprecipitation sequencing (ChIP-Seq) and RNA-Seq data in order to explore the progression of colon cancer at a genome-wide epigenetic level.

In summary, we documented for the first time the chromatin structure associated with gene expression profiles of an in vivo murine colonic tumorigenesis model. Our high-throughput sequencing approach revealed many expected changes at various regulatory stages of gene expression, plus unexpected insight into gene regulation during

colon cancer progression (**Figure 29**). Specifically, we were able to show that AOM induced transcriptional deregulation was primarily concomitant with interferon-associated (IFN) immune response genes, while K4me3 deregulation was linked to genes associated with colon tumorigenesis, perhaps acting as a harbinger of changes in gene activity.

These findings emphasize the value of genome-wide analyses and may have important clinical relevance for future therapeutic targeting. For example early use of targeted histone demethylases (210) could inhibit colon cancer prior to the tumorigenesis stage by reversing K4me3 deregulation before the transcription of oncogenes. Additionally the therapeutic alteration of microbial composition (212) may counteract the interferon-associated ‘anti-microbial and anti-viral’ immune response genes that were upregulated by AOM. We suspect cancer progression alteration of the gut environment leads to a dysbiosis of ‘peace keeping’ bacteria and facilitates the colonization in the colon of pathogenic bacteria, therapeutic alteration of microbial composition would re-establish the ‘peace-keeping’ bacteria and help heal the colonic epithelial barrier (279,280).

5.3 Aim 3 (Chapter IV)

Animal fed diets containing n-3 PUFA (fish oil) with pectin (which is fermented to butyrate) as a fiber source maximally promote apoptosis in the colon compared with diets high in other dietary lipids, e.g., n-6 PUFA (corn oil) (86,98,109,236). In a follow-up study, the administration of butyrate-containing pellets for targeted release in the colon was used to demonstrate that butyrate and fish oil work coordinately in the colon to trigger apoptosis (104,112,234). Fish oil and butyrate have been shown to impact cell kinetics and epigenetics which regulate development of colon carcinogenesis (103,109,234). Butyrate can increase histone acetylation in the nucleus by at least two epigenetic mechanisms, by acting as a histone deacetylase inhibitor and as substrate for

histone acetylases (HATs) (232). Fish oil components DHA and EPA and their oxidative metabolites have been shown to interact with specific ligand dependent nuclear receptors including CAR, HNF4A, PPAR γ , PXR and RXRA (**Figure 1**) (46). In this fashion, n-3 PUFA regulate the function of nuclear receptors and their impact on transcriptional processes.

In this study, we employed novel technologies and bioinformatics algorithms, such as next-generation sequencing, in order to explore 'nutri-epigenomics' at a genome-wide level and determine the importance of epigenetic mechanisms related to chemoprevention. By determining 3 epigenetic levels of regulation (H3K4me3, H3K9ac and mRNA expression) in colonocytes from the same animals, we were able to gain a greater understanding of the chromatin state associated with the interaction of fish oil (rich in DHA/EPA n-3 PUFAs), pectin (a readily fermentable fiber) and AOM (colon carcinogen) treatments.

In summary, we documented for the first time the chromatin structure associated with a diet rich in fish oil (rich in n-3 PUFAs) and readily fermentable fiber (rich in short chain fatty acids) under normal (saline control) and carcinogenic conditions. We propose that the chemoprotective fish oil + pectin diet generates unique epigenetic modifications and transcriptional profiles in a context (AOM vs saline) specific manner. Genes associated with enhanced beta-oxidation were transcriptionally upregulated uniquely in fish oil + pectin (fpa v cca) fed rats. With only a small subset (3) of beta-oxidation genes previously linked to fish oil being upregulated in fish oil + cellulose fed rats (fca v cca). Upstream regulator analysis revealed the n-3 PUFA activation of ligand dependent nuclear receptors predominantly in the presence of pectin (fpa v cca). We therefore hypothesize that, in colonocytes, fish oil related epigenetic effects are enhanced by concurrent pectin feeding during the progression of carcinogenesis. A potential mechanism for this phenomenon may include the concurrent enhancement of beta-oxidation and the induction of futile mitochondrial respiration (proton leak) that leads to the uncoupling of ATP synthesis, resulting in nutrient wasting and apoptosis.

We were surprised to find no sign of upregulated H3K9ac enrichment in pectin fed, aom treated, rats (APPENDIX S). However, our findings are consistent with previous studies that have also described a global increase in histone acetylation in response to butyrate administration in saline but not AOM injected animals and therefore provides additional mechanistic insight into the chemotherapeutic effects of fish oil plus pectin during a poorly studied time point in this multistep, multipath disease.

In conclusion, our data contribute to the understanding of the regulatory action of chemoprotective bioactive compounds found in fish oil and readily fermentable fiber (n-3 PUFAs and SCFAs) in colonic crypts and provide critical mechanistic insight into recently reported epidemiological findings (**Figure 29**) (118).

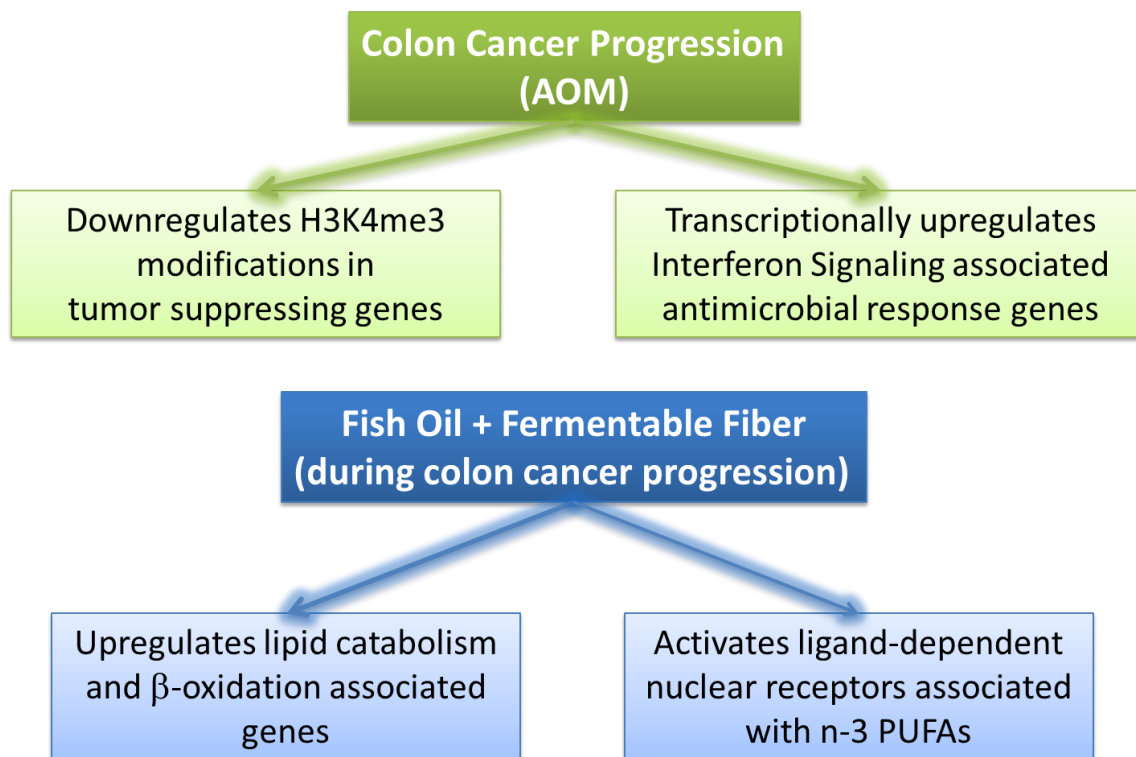


Figure 29. Summary of epigenetic findings from Aim 2 and 3.

5.4 Future studies

The research presented herein set the groundwork for a number of futures studies that could further delve into the effects of fish oil and pectin on nuclear receptor signaling and function, mitochondrial beta-oxidation and mitochondrial proton leak.

5.4.1 ChIP-seq and sub-cellular localization of upstream regulators

One question remaining is how fish oil is altering the localization of ligand-dependent nuclear receptors. The data presented in Aims 1-3, along with previous findings strongly support a correlation between nuclear receptors, colon cancer, fish oil and pectin. Figure 30 provides a list of candidate proteins and methodology for further analysis. A proposed mechanism for TRIM24 as a master regulator and tumor suppressor is depicted in Figure 31.

Future Studies Methodology:

- **Immunohistochemistry** (subcellular localization and protein levels)
- **ChIP-Seq of DNA binding proteins** (genome-wide DNA binding patterns)

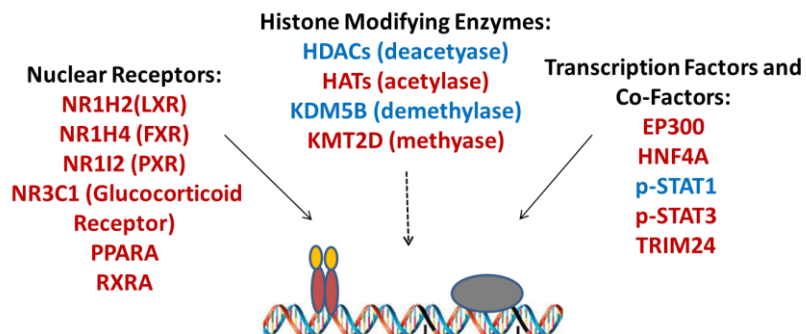


Figure 30. Future Studies Methodology and Protein of Interest. Red color indicates predicted activation and blue color predicted inhibition of transcription factors and co-factors, nuclear receptors, and histone modifying enzymes in response to chemoprotective treatment.

TRIM24 as a master transcriptional regulator, a putative mechanism

control (saline) [ccs] or chemoprotective fish oil diet (n-3 PUFA) [fca and fpa]
TRIM24 inhibits transcription

cancer progression (AOM) [cca]
TRIM24 cannot bind to regions with H3K4me3

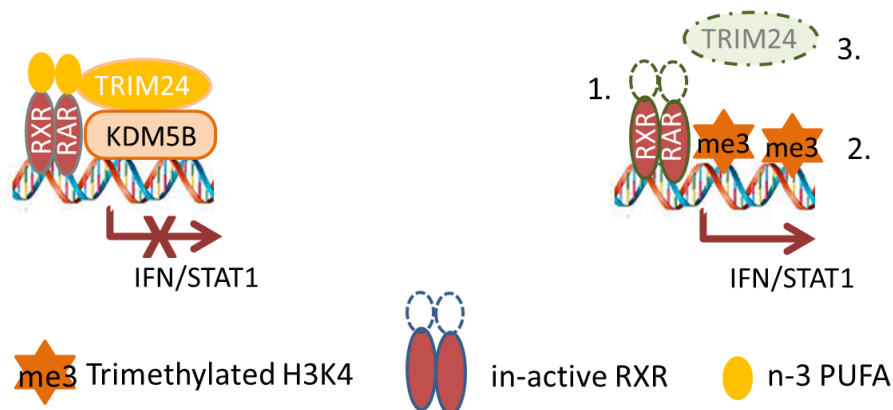


Figure 31. We propose a putative mechanism for future investigation in which **1.** decreased RXR/RAR activity at specific sites (in a context specific manner) during colon cancer (Tang and Gudas 2011) leads to a **2.** decrease in KDM5B binding (Zhang, Liang et al. 2014) therefore **3.** TRIM24 cannot bind to trimethylated H3K4 (Tsai, Wang et al. 2010). Thus an increase in K4me3 subsequently leads to a decrease in TRIM24 (Tisserand, Khetchoumian et al. 2011) binding, inducing IFN/STAT1 associated colon cancer genes (Kikuchi, Okumura et al. 2009).

5.4.2 Mitochondrial bioenergetic analysis

Our data suggests that fish oil enhances mitochondrial proton leak. Mitochondria couple respiration to ATP synthesis through an electrochemical proton gradient. Proton leak across the inner membrane allows adjustment of the coupling efficiency and may provide an attractive therapeutic target for diseases rooted in metabolic imbalance and oxidative stress, including colon cancer (Divakaruni and Brand 2011).

Previously reports in similar pre-clinical models have also found that fish oil derived n-3 PUFA can enhance mitochondrial proton leaks (Fan, Ran et al. 2011) We

moreover speculate, based on the results from Chapter IV, that the effects of fish oil are enhanced in the presence of pectin, most likely because it can be a source of butyrate in colonocytes and further promote an oxidation-reduction imbalance in the intestine, increasing proton leak across the mitochondrial inner membrane, contributing to a permissive environment for apoptosis. In order to validate this hypothesis mitochondrial bioenergetic profiles can be generated by using a Seahorse XF24 Extracellular Flux Analyzer (Seahorse Bioscience, Billerica, MA) as described by Fan et al. (Fan, Ran et al. 2011)

REFERENCES

1. Wright, N. A., and Alison, M. (1984) *The biology of epithelial cell populations*, Clarendon Press ; Oxford University Press, Oxford Oxfordshire, New York
2. Ricci-Vitiani, L., Fabrizio, E., Palio, E., and De Maria, R. (2009) Colon cancer stem cells. *Journal of molecular medicine* **87**, 1097-1104
3. Humphries, A., and Wright, N. A. (2008) Colonic crypt organization and tumorigenesis. *Nat Rev Cancer* **8**, 415-424
4. Palazzo, M., Balsari, A., Rossini, A., Selleri, S., Calcaterra, C., *et al.* (2007) Activation of enteroendocrine cells via TLRs induces hormone, chemokine, and defensin secretion. *J Immunol* **178**, 4296-4303
5. Specian, R. D., and Oliver, M. G. (1991) Functional biology of intestinal goblet cells. *The American journal of physiology* **260**, C183-193
6. Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E., *et al.* (2011) Global cancer statistics. *CA Cancer J Clin* **61**, 69-90
7. Siegel, R. L., Ward, E. M., and Jemal, A. (2012) Trends in colorectal cancer incidence rates in the United States by tumor location and stage, 1992-2008. *Cancer Epidemiol Biomarkers Prev* **21**, 411-416
8. Fernandez, E., Gallus, S., La Vecchia, C., Talamini, R., Negri, E., *et al.* (2004) Family history and environmental risk factors for colon cancer. *Cancer Epidemiol Biomarkers Prev* **13**, 658-661
9. Emmons, K. M., McBride, C. M., Puleo, E., Pollak, K. I., Marcus, B. H., *et al.* (2005) Prevalence and predictors of multiple behavioral risk factors for colon cancer. *Preventive medicine* **40**, 527-534
10. Lieberman, D. A., Prindiville, S., Weiss, D. G., Willett, W., and Group, V. A. C. S. (2003) Risk factors for advanced colonic neoplasia and hyperplastic polyps in asymptomatic individuals. *Jama* **290**, 2959-2967

11. Takayama, T., Katsuki, S., Takahashi, Y., Ohi, M., Nojiri, S., *et al.* (1998) Aberrant crypt foci of the colon as precursors of adenoma and cancer. *N Engl J Med* **339**, 1277-1284
12. Rudolph, R. E., Dominitz, J. A., Lampe, J. W., Levy, L., Qu, P., *et al.* (2005) Risk factors for colorectal cancer in relation to number and size of aberrant crypt foci in humans. *Cancer Epidemiol Biomarkers Prev* **14**, 605-608
13. Peters, U., Bien, S., and Zubair, N. (2015) Genetic architecture of colorectal cancer. *Gut* **64**, 1623-1636
14. Sugai, T., Habano, W., Jiao, Y. F., Tsukahara, M., Takeda, Y., *et al.* (2006) Analysis of molecular alterations in left- and right-sided colorectal carcinomas reveals distinct pathways of carcinogenesis: proposal for new molecular profile of colorectal carcinomas. *The Journal of molecular diagnostics : JMD* **8**, 193-201
15. Lengauer, C., Kinzler, K. W., and Vogelstein, B. (1997) Genetic instability in colorectal cancers. *Nature* **386**, 623-627
16. Georgiades, I. B., Curtis, L. J., Morris, R. M., Bird, C. C., and Wyllie, A. H. (1999) Heterogeneity studies identify a subset of sporadic colorectal cancers without evidence for chromosomal or microsatellite instability. *Oncogene* **18**, 7933-7940
17. Barber, T. D., McManus, K., Yuen, K. W., Reis, M., Parmigiani, G., *et al.* (2008) Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. *Proc Natl Acad Sci U S A* **105**, 3443-3448
18. Markowitz, S. D., and Bertagnolli, M. M. (2009) Molecular origins of cancer: Molecular basis of colorectal cancer. *N Engl J Med* **361**, 2449-2460
19. Kondo, Y., and Issa, J. P. (2004) Epigenetic changes in colorectal cancer. *Cancer metastasis reviews* **23**, 29-39
20. Young, J., Biden, K. G., Simms, L. A., Huggard, P., Karamatic, R., *et al.* (2001) HPP1: a transmembrane protein-encoding gene commonly methylated in colorectal polyps and cancers. *Proc Natl Acad Sci U S A* **98**, 265-270
21. Lao, V. V., and Grady, W. M. (2011) Epigenetics and colorectal cancer. *Nature reviews. Gastroenterology & hepatology* **8**, 686-700

22. Toyota, M., Ho, C., Ahuja, N., Jair, K. W., Li, Q., *et al.* (1999) Identification of differentially methylated sequences in colorectal cancer by methylated CpG island amplification. *Cancer research* **59**, 2307-2312
23. Hawkins, N., Norrie, M., Cheong, K., Mokany, E., Ku, S. L., *et al.* (2002) CpG island methylation in sporadic colorectal cancers and its relationship to microsatellite instability. *Gastroenterology* **122**, 1376-1387
24. Lesko, A. C., Goss, K. H., and Prosperi, J. R. (2014) Exploiting APC function as a novel cancer therapy. *Current drug targets* **15**, 90-102
25. Vazquez, M., Fabbrizzi, L., Taglietti, A., Pedrido, R. M., Gonzalez-Noya, A. M., *et al.* (2004) A colorimetric approach to anion sensing: a selective chemosensor of fluoride ions, in which color is generated by anion-enhanced pi delocalization. *Angewandte Chemie* **43**, 1962-1965
26. Iacopetta, B. (2002) Are there two sides to colorectal cancer? *Int J Cancer* **101**, 403-408
27. Meza, R., Jeon, J., Renehan, A. G., and Luebeck, E. G. (2010) Colorectal cancer incidence trends in the United States and United kingdom: evidence of right- to left-sided biological gradients with implications for screening. *Cancer research* **70**, 5419-5429
28. Lindbloom, E. J., and Chang, S. I. (2001) Clinical inquiries. What is the best test to diagnose urinary tract stones? *The Journal of family practice* **50**, 657-658
29. Bleeker, W. A., Hayes, V. M., Karrenbeld, A., Hofstra, R. M., Hermans, J., *et al.* (2000) Impact of KRAS and TP53 mutations on survival in patients with left- and right-sided Dukes' C colon cancer. *Am J Gastroenterol* **95**, 2953-2957
30. Chapkin, R. S., McMurray, D. N., Davidson, L. A., Patil, B. S., Fan, Y. Y., *et al.* (2008) Bioactive dietary long-chain fatty acids: emerging mechanisms of action. *The British journal of nutrition* **100**, 1152-1157
31. Chapkin, R. S., Seo, J., McMurray, D. N., and Lupton, J. R. (2008) Mechanisms by which docosahexaenoic acid and related fatty acids reduce colon cancer risk and

- inflammatory disorders of the intestine. *Chemistry and physics of lipids* **153**, 14-23
32. Reddy, B. S., Burill, C., and Rigotty, J. (1991) Effect of diets high in omega-3 and omega-6 fatty acids on initiation and postinitiation stages of colon carcinogenesis. *Cancer research* **51**, 487-491
33. Turk, H. F., and Chapkin, R. S. (2013) Membrane lipid raft organization is uniquely modified by n-3 polyunsaturated fatty acids. *Prostaglandins, leukotrienes, and essential fatty acids* **88**, 43-47
34. Pardini, R. S. (2006) Nutritional intervention with omega-3 fatty acids enhances tumor response to anti-neoplastic agents. *Chemico-biological interactions* **162**, 89-105
35. Kim, W., McMurray, D. N., and Chapkin, R. S. (2009) Chemotherapeutic Properties of n-3 Polyunsaturated Fatty Acids - Old Concepts and New Insights. *Immunology, endocrine & metabolic agents in medicinal chemistry* **9**, 38-44
36. Gomez Candela, C., Bermejo Lopez, L. M., and Loria Kohen, V. (2011) Importance of a balanced omega 6/omega 3 ratio for the maintenance of health: nutritional recommendations. *Nutricion hospitalaria* **26**, 323-329
37. West, N. J., Clark, S. K., Phillips, R. K. S., Hutchinson, J. M., Leicester, R. J., *et al.* (2010) Eicosapentaenoic acid reduces rectal polyp number and size in familial adenomatous polyposis. *Gut* **59**, 918-925
38. Moosheer, S. M., Waldschutz, W., Itariu, B. K., Brath, H., and Stulnig, T. M. (2014) A protein-enriched low glycemic index diet with omega-3 polyunsaturated fatty acid supplementation exerts beneficial effects on metabolic control in type 2 diabetes. *Primary care diabetes* **8**, 308-314
39. van Heusden, G. P., Bos, K., Raetz, C. R., and Wirtz, K. W. (1990) Chinese hamster ovary cells deficient in peroxisomes lack the nonspecific lipid transfer protein (sterol carrier protein 2). *The Journal of biological chemistry* **265**, 4105-4110
40. Shah, M. S., Schwartz, S. L., Zhao, C., Davidson, L. A., Zhou, B., *et al.* (2011) Integrated microRNA and mRNA expression profiling in a rat colon

- carcinogenesis model: effect of a chemo-protective diet. *Physiological genomics* **43**, 640-654
41. Hou, T. Y., Monk, J. M., Fan, Y. Y., Barhoumi, R., Chen, Y. Q., *et al.* (2012) n-3 polyunsaturated fatty acids suppress phosphatidylinositol 4,5-bisphosphate-dependent actin remodelling during CD4+ T-cell activation. *The Biochemical journal* **443**, 27-37
42. Goldberg, A. D., Allis, C. D., and Bernstein, E. (2007) Epigenetics: a landscape takes shape. *Cell* **128**, 635-638
43. Skinner, M. K. (2011) Environmental epigenomics and disease susceptibility. *EMBO reports* **12**, 620-622
44. Burdge, G. C., Hoile, S. P., and Lillycrop, K. A. (2012) Epigenetics: are there implications for personalised nutrition? *Current opinion in clinical nutrition and metabolic care* **15**, 442-447
45. Burdge, G. C., and Lillycrop, K. A. (2010) Nutrition, epigenetics, and developmental plasticity: implications for understanding human disease. *Annual review of nutrition* **30**, 315-339
46. Pegorier, J. P., Le May, C., and Girard, J. (2004) Control of gene expression by fatty acids. *The Journal of nutrition* **134**, 2444S-2449S
47. Ponferrada, A., Caso, J. R., Alou, L., Colon, A., Sevillano, D., *et al.* (2007) The role of PPARgamma on restoration of colonic homeostasis after experimental stress-induced inflammation and dysfunction. *Gastroenterology* **132**, 1791-1803
48. Zand, H., Rahimipour, A., Salimi, S., and Shafiee, S. M. (2008) Docosahexaenoic acid sensitizes Ramos cells to Gamma-irradiation-induced apoptosis through involvement of PPAR-gamma activation and NF-kappaB suppression. *Molecular and cellular biochemistry* **317**, 113-120
49. Edwards, I. J., and O'Flaherty, J. T. (2008) Omega-3 Fatty Acids and PPARgamma in Cancer. *Ppar Res* **2008**, 358052

50. Fan, Y. Y., Spencer, T. E., Wang, N., Moyer, M. P., and Chapkin, R. S. (2003) Chemopreventive n-3 fatty acids activate RXRalpha in colonocytes. *Carcinogenesis* **24**, 1541-1548
51. Sampath, H., and Ntambi, J. M. (2005) Polyunsaturated fatty acid regulation of genes of lipid metabolism. *Annual review of nutrition* **25**, 317-340
52. Possidonio, A. C., Miranda, M., Gregoracci, G. B., Thompson, F. L., Costa, M. L., *et al.* (2014) Cholesterol depletion induces transcriptional changes during skeletal muscle differentiation. *BMC genomics* **15**, 544
53. Robbins, D., and Chen, T. (2014) Tissue-specific regulation of pregnane X receptor in cancer development and therapy. *Cell & bioscience* **4**, 17
54. Zhang, X., Zhao, X. W., Liu, D. B., Han, C. Z., Du, L. L., *et al.* (2014) Lipid levels in serum and cancerous tissues of colorectal cancer patients. *World journal of gastroenterology : WJG* **20**, 8646-8652
55. Lo Sasso, G., Bovenga, F., Murzilli, S., Salvatore, L., Di Tullio, G., *et al.* (2013) Liver X receptors inhibit proliferation of human colorectal cancer cells and growth of intestinal tumors in mice. *Gastroenterology* **144**, 1497-1507, 1507 e1491-1413
56. Qiao, E. Q., Ji, M. H., Wu, J. Z., Ma, R., Zhang, X. H., *et al.* (2013) Expression of the PXR gene in various types of cancer and drug resistance (Review). *Oncol Lett* **5**, 1093-1100
57. Kuan, C. Y., Walker, T. H., Luo, P. G., and Chen, C. F. (2011) Long-chain polyunsaturated fatty acids promote paclitaxel cytotoxicity via inhibition of the MDR1 gene in the human colon cancer Caco-2 cell line. *Journal of the American College of Nutrition* **30**, 265-273
58. Ouyang, N., Ke, S., Eagleton, N., Xie, Y., Chen, G., *et al.* (2010) Pregnane X receptor suppresses proliferation and tumourigenicity of colon cancer cells. *British journal of cancer* **102**, 1753-1761

59. Ahn, S. H., Shah, Y. M., Inoue, J., Morimura, K., Kim, I., *et al.* (2008) Hepatocyte nuclear factor 4alpha in the intestinal epithelial cells protects against inflammatory bowel disease. *Inflammatory bowel diseases* **14**, 908-920
60. Cattin, A. L., Le Beyec, J., Barreau, F., Saint-Just, S., Houllier, A., *et al.* (2009) Hepatocyte nuclear factor 4alpha, a key factor for homeostasis, cell architecture, and barrier function of the adult intestinal epithelium. *Molecular and cellular biology* **29**, 6294-6308
61. Oshima, T., Kawasaki, T., Ohashi, R., Hasegawa, G., Jiang, S., *et al.* (2007) Downregulated P1 promoter-driven hepatocyte nuclear factor-4alpha expression in human colorectal carcinoma is a new prognostic factor against liver metastasis. *Pathology international* **57**, 82-90
62. Chellappa, K., Robertson, G. R., and Sladek, F. M. (2012) HNF4alpha: a new biomarker in colon cancer? *Biomarkers in medicine* **6**, 297-300
63. D'Archivio, M., Sczzocchio, B., Giammarioli, S., Fiani, M. L., Vari, R., *et al.* (2013) omega3-PUFAs exert anti-inflammatory activity in visceral adipocytes from colorectal cancer patients. *PloS one* **8**, e77432
64. Jakobsen, C. H., Storvold, G. L., Bremseth, H., Follestad, T., Sand, K., *et al.* (2008) DHA induces ER stress and growth arrest in human colon cancer cells: associations with cholesterol and calcium homeostasis. *Journal of lipid research* **49**, 2089-2100
65. Schonberg, S. A., Lundemo, A. G., Fladvad, T., Holmgren, K., Bremseth, H., *et al.* (2006) Closely related colon cancer cell lines display different sensitivity to polyunsaturated fatty acids, accumulate different lipid classes and downregulate sterol regulatory element-binding protein 1. *The FEBS journal* **273**, 2749-2765
66. Slagsvold, J. E., Pettersen, C. H., Storvold, G. L., Follestad, T., Krokan, H. E., *et al.* (2010) DHA alters expression of target proteins of cancer therapy in chemotherapy resistant SW620 colon cancer cells. *Nutrition and cancer* **62**, 611-621

67. Gerner, E. W., Ignatenko, N. A., Lance, P., and Hurley, L. H. (2005) A comprehensive strategy to combat colon cancer targeting the adenomatous polyposis coli tumor suppressor gene. *Annals of the New York Academy of Sciences* **1059**, 97-105
68. West, N. J., Clark, S. K., Phillips, R. K., Hutchinson, J. M., Leicester, R. J., *et al.* (2010) Eicosapentaenoic acid reduces rectal polyp number and size in familial adenomatous polyposis. *Gut* **59**, 918-925
69. Cockbain, A. J., Toogood, G. J., and Hull, M. A. (2012) Omega-3 polyunsaturated fatty acids for the treatment and prevention of colorectal cancer. *Gut* **61**, 135-149
70. Paulsen, J. E., Elvsaa, I. K., Steffensen, I. L., and Alexander, J. (1997) A fish oil derived concentrate enriched in eicosapentaenoic and docosahexaenoic acid as ethyl ester suppresses the formation and growth of intestinal polyps in the Min mouse. *Carcinogenesis* **18**, 1905-1910
71. Arango, D., Corner, G. A., Wadler, S., Catalano, P. J., and Augenlicht, L. H. (2001) c-myc/p53 interaction determines sensitivity of human colon carcinoma cells to 5-fluorouracil in vitro and in vivo. *Cancer research* **61**, 4910-4915
72. Calviello, G., Di Nicuolo, F., Serini, S., Piccioni, E., Boninsegna, A., *et al.* (2005) Docosahexaenoic acid enhances the susceptibility of human colorectal cancer cells to 5-fluorouracil. *Cancer chemotherapy and pharmacology* **55**, 12-20
73. Weber, C., Erl, W., Pietsch, A., Danesch, U., and Weber, P. C. (1995) Docosahexaenoic acid selectively attenuates induction of vascular cell adhesion molecule-1 and subsequent monocytic cell adhesion to human endothelial cells stimulated by tumor necrosis factor-alpha. *Arteriosclerosis, thrombosis, and vascular biology* **15**, 622-628
74. Mishra, A., Chaudhary, A., and Sethi, S. (2004) Oxidized omega-3 fatty acids inhibit NF-kappa B activation via a PPAR alpha-dependent pathway. *Arterioscl Throm Vas* **24**, 1621-1627
75. Martinez-Augustin, O., Lopez-Posadas, R., Gonzalez, R., Suarez, M. D., Zarzuelo, A., *et al.* (2009) Genomic analysis of sulfasalazine effect in experimental colitis

- is consistent primarily with the modulation of NF-kappaB but not PPAR-gamma signaling. *Pharmacogenetics and genomics* **19**, 363-372
76. Chapkin, R. S., Wang, N., Fan, Y. Y., Lupton, J. R., and Prior, I. A. (2008) Docosahexaenoic acid alters the size and distribution of cell surface microdomains. *Biochimica et biophysica acta* **1778**, 466-471
77. Habermann, N., Schon, A., Lund, E. K., and Gleib, M. (2010) Fish fatty acids alter markers of apoptosis in colorectal adenoma and adenocarcinoma cell lines but fish consumption has no impact on apoptosis-induction ex vivo. *Apoptosis : an international journal on programmed cell death* **15**, 621-630
78. Fasano, E., Serini, S., Piccioni, E., Toesca, A., Monego, G., *et al.* (2012) DHA induces apoptosis by altering the expression and cellular location of GRP78 in colon cancer cell lines. *Biochimica et biophysica acta* **1822**, 1762-1772
79. Giros, A., Grzybowski, M., Sohn, V. R., Pons, E., Fernandez-Morales, J., *et al.* (2009) Regulation of colorectal cancer cell apoptosis by the n-3 polyunsaturated fatty acids Docosahexaenoic and Eicosapentaenoic. *Cancer prevention research* **2**, 732-742
80. Kagan, V. E., Bayir, H. A., Belikova, N. A., Kapralov, O., Tyurina, Y. Y., *et al.* (2009) Cytochrome c/cardiolipin relations in mitochondria: a kiss of death. *Free radical biology & medicine* **46**, 1439-1453
81. Watkins, S. M., Carter, L. C., and German, J. B. (1998) Docosahexaenoic acid accumulates in cardiolipin and enhances HT-29 cell oxidant production. *Journal of lipid research* **39**, 1583-1588
82. Hong, M. Y., Chapkin, R. S., Barhoumi, R., Burghardt, R. C., Turner, N. D., *et al.* (2002) Fish oil increases mitochondrial phospholipid unsaturation, upregulating reactive oxygen species and apoptosis in rat colonocytes. *Carcinogenesis* **23**, 1919-1925
83. Siddiqui, R. A., Harvey, K., and Stillwell, W. (2008) Anticancer properties of oxidation products of docosahexaenoic acid. *Chemistry and physics of lipids* **153**, 47-56

84. Swamy, M. V., Cooma, I., Patlolla, J. M., Simi, B., Reddy, B. S., *et al.* (2004) Modulation of cyclooxygenase-2 activities by the combined action of celecoxib and decosahexaenoic acid: novel strategies for colon cancer prevention and treatment. *Molecular cancer therapeutics* **3**, 215-221
85. Yang, W. L., and Frucht, H. (2001) Activation of the PPAR pathway induces apoptosis and COX-2 inhibition in HT-29 human colon cancer cells. *Carcinogenesis* **22**, 1379-1383
86. Vanamala, J., Glagolenko, A., Yang, P., Carroll, R. J., Murphy, M. E., *et al.* (2008) Dietary fish oil and pectin enhance colonocyte apoptosis in part through suppression of PPARdelta/PGE2 and elevation of PGE3. *Carcinogenesis* **29**, 790-796
87. Hardman, W. E., Moyer, M. P., and Cameron, I. L. (2002) Consumption of an omega-3 fatty acids product, INCELL AAFA, reduced side-effects of CPT-11 (irinotecan) in mice. *British journal of cancer* **86**, 983-988
88. Agarwal, B., Swaroop, P., Protiva, P., Raj, S. V., Shirin, H., *et al.* (2003) Cox-2 is needed but not sufficient for apoptosis induced by Cox-2 selective inhibitors in colon cancer cells. *Apoptosis : an international journal on programmed cell death* **8**, 649-654
89. Boudreau, M. D., Sohn, K. H., Rhee, S. H., Lee, S. W., Hunt, J. D., *et al.* (2001) Suppression of tumor cell growth both in nude mice and in culture by n-3 polyunsaturated fatty acids: mediation through cyclooxygenase-independent pathways. *Cancer research* **61**, 1386-1391
90. Jho, D. H., Cole, S. M., Lee, E. M., and Espat, N. J. (2004) Role of omega-3 fatty acid supplementation in inflammation and malignancy. *Integrative cancer therapies* **3**, 98-111
91. Purasiri, P., Murray, A., Richardson, S., Heys, S. D., Horrobin, D., *et al.* (1994) Modulation of cytokine production in vivo by dietary essential fatty acids in patients with colorectal cancer. *Clinical science* **87**, 711-717

92. * Babeu, J. P., and Boudreau, F. (2014) Hepatocyte nuclear factor 4-alpha involvement in liver and intestinal inflammatory networks. *World journal of gastroenterology : WJG* **20**, 22-30
93. Cao, Y., Deng, C., Townsend, C. M., Jr., and Ko, T. C. (2006) TGF-beta inhibits Akt-induced transformation in intestinal epithelial cells. *Surgery* **140**, 322-329
94. Nowak, J., Weylandt, K. H., Habbel, P., Wang, J., Dignass, A., *et al.* (2007) Colitis-associated colon tumorigenesis is suppressed in transgenic mice rich in endogenous n-3 fatty acids. *Carcinogenesis* **28**, 1991-1995
95. Tang, B., Bottinger, E. P., Jakowlew, S. B., Bagnall, K. M., Mariano, J., *et al.* (1998) Transforming growth factor-beta1 is a new form of tumor suppressor with true haploid insufficiency. *Nature medicine* **4**, 802-807
96. Calviello, G., Resci, F., Serini, S., Piccioni, E., Toesca, A., *et al.* (2007) Docosahexaenoic acid induces proteasome-dependent degradation of beta-catenin, down-regulation of survivin and apoptosis in human colorectal cancer cells not expressing COX-2. *Carcinogenesis* **28**, 1202-1209
97. Toit-Kohn, J. L., Louw, L., and Engelbrecht, A. M. (2009) Docosahexaenoic acid induces apoptosis in colorectal carcinoma cells by modulating the PI3 kinase and p38 MAPK pathways. *The Journal of nutritional biochemistry* **20**, 106-114
98. Davidson, L. A., Nguyen, D. V., Hokanson, R. M., Callaway, E. S., Isett, R. B., *et al.* (2004) Chemopreventive n-3 polyunsaturated fatty acids reprogram genetic signatures during colon cancer initiation and progression in the rat. *Cancer research* **64**, 6797-6804
99. Chapkin, R. S., McMurray, D. N., and Lupton, J. R. (2007) Colon cancer, fatty acids and anti-inflammatory compounds. *Current opinion in gastroenterology* **23**, 48-54
100. Kolar, S. S., Barhoumi, R., Lupton, J. R., and Chapkin, R. S. (2007) Docosahexaenoic acid and butyrate synergistically induce colonocyte apoptosis by enhancing mitochondrial Ca²⁺ accumulation. *Cancer research* **67**, 5561-5568

101. Kolar, S. S., Barhoumi, R., Callaway, E. S., Fan, Y. Y., Wang, N., *et al.* (2007) Synergy between docosahexaenoic acid and butyrate elicits p53-independent apoptosis via mitochondrial Ca(2+) accumulation in colonocytes. *American journal of physiology. Gastrointestinal and liver physiology* **293**, G935-943
102. Chapkin, R. S., Clark, A. E., Davidson, L. A., Schroeder, F., Zoran, D. L., *et al.* (1998) Dietary fiber differentially alters cellular fatty acid-binding protein expression in exfoliated colonocytes during tumor development. *Nutrition and cancer* **32**, 107-112
103. Kolar, S., Barhoumi, R., Jones, C. K., Wesley, J., Lupton, J. R., *et al.* (2011) Interactive effects of fatty acid and butyrate-induced mitochondrial Ca(2)(+) loading and apoptosis in colonocytes. *Cancer* **117**, 5294-5303
104. Turk, H. F., Kolar, S. S., Fan, Y. Y., Cozby, C. A., Lupton, J. R., *et al.* (2011) Linoleic acid and butyrate synergize to increase Bcl-2 levels in colonocytes. *Int J Cancer* **128**, 63-71
105. Hofmanova, J., Hyrslova Vaculova, A., and Kozubik, A. (2013) Regulation of the metabolism of polyunsaturated Fatty acids and butyrate in colon cancer cells. *Current pharmaceutical biotechnology* **14**, 274-288
106. Donohoe, D. R., Garge, N., Zhang, X., Sun, W., O'Connell, T. M., *et al.* (2011) The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. *Cell metabolism* **13**, 517-526
107. Canani, R. B., Costanzo, M. D., Leone, L., Pedata, M., Meli, R., *et al.* (2011) Potential beneficial effects of butyrate in intestinal and extraintestinal diseases. *World journal of gastroenterology : WJG* **17**, 1519-1528
108. Scharlau, D., Borowicki, A., Habermann, N., Hofmann, T., Klenow, S., *et al.* (2009) Mechanisms of primary cancer prevention by butyrate and other products formed during gut flora-mediated fermentation of dietary fibre. *Mutation research* **682**, 39-53
109. Cho, Y., Kim, H., Turner, N. D., Mann, J. C., Wei, J., *et al.* (2011) A chemoprotective fish oil- and pectin-containing diet temporally alters gene

- expression profiles in exfoliated rat colonocytes throughout oncogenesis. *The Journal of nutrition* **141**, 1029-1035
110. Kachroo, P., Ivanov, I., Davidson, L. A., Chowdhary, B. P., Lupton, J. R., *et al.* (2011) Classification of diet-modulated gene signatures at the colon cancer initiation and progression stages. *Digestive diseases and sciences* **56**, 2595-2604
111. Chang, W. L., Chapkin, R. S., and Lupton, J. R. (1998) Fish oil blocks azoxymethane-induced rat colon tumorigenesis by increasing cell differentiation and apoptosis rather than decreasing cell proliferation. *The Journal of nutrition* **128**, 491-497
112. Crim, K. C., Sanders, L. M., Hong, M. Y., Taddeo, S. S., Turner, N. D., *et al.* (2008) Upregulation of p21Waf1/Cip1 expression in vivo by butyrate administration can be chemoprotective or chemopromotive depending on the lipid component of the diet. *Carcinogenesis* **29**, 1415-1420
113. Lien, E. L. (2009) Toxicology and safety of DHA. *Prostaglandins, leukotrienes, and essential fatty acids* **81**, 125-132
114. Bell, G. A., Kantor, E. D., Lampe, J. W., Kristal, A. R., Heckbert, S. R., *et al.* (2014) Intake of long-chain omega-3 fatty acids from diet and supplements in relation to mortality. *American journal of epidemiology* **179**, 710-720
115. Davidson, L. A., Wang, N., Ivanov, I., Goldsby, J., Lupton, J. R., *et al.* (2009) Identification of actively translated mRNA transcripts in a rat model of early-stage colon carcinogenesis. *Cancer prevention research* **2**, 984-994
116. Jia, Q., Ivanov, I., Zlatev, Z. Z., Alaniz, R. C., Weeks, B. R., *et al.* (2011) Dietary fish oil and curcumin combine to modulate colonic cytokinetics and gene expression in dextran sodium sulphate-treated mice. *The British journal of nutrition* **106**, 519-529
117. Fenton, J. I., and McCaskey, S. J. (2013) Curcumin and docosahexaenoic acid block insulin-induced colon carcinoma cell proliferation. *Prostaglandins, leukotrienes, and essential fatty acids* **88**, 219-226

118. Orlich, M. J., Singh, P. N., Sabate, J., Fan, J., Sveen, L., *et al.* (2015) Vegetarian dietary patterns and the risk of colorectal cancers. *JAMA internal medicine* **175**, 767-776
119. O'Keefe, S. J., Li, J. V., Lahti, L., Ou, J., Carbonero, F., *et al.* (2015) Fat, fibre and cancer risk in African Americans and rural Africans. *Nat Commun* **6**, 6342
120. Villeneuve, L. M., Reddy, M. A., Lanting, L. L., Wang, M., Meng, L., *et al.* (2008) Epigenetic histone H3 lysine 9 methylation in metabolic memory and inflammatory phenotype of vascular smooth muscle cells in diabetes. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 9047-9052
121. Brinkman, A. B., Roelofsen, T., Pennings, S. W., Martens, J. H., Jenuwein, T., *et al.* (2006) Histone modification patterns associated with the human X chromosome. *EMBO reports* **7**, 628-634
122. Fogue-Lafitte, M. E., Fabiani, B., Levy, P. P., Maurin, N., Flejou, J. F., *et al.* (2007) Abnormal expression of M1/MUC5AC mucin in distal colon of patients with diverticulitis, ulcerative colitis and cancer. *International journal of cancer. Journal international du cancer* **121**, 1543-1549
123. Ikeda, Y., Akagi, K., Kinoshita, J., Abe, T., Miyazaki, M., *et al.* (2002) Different distribution of Dukes' stage between proximal and distal colorectal cancer. *Hepato-gastroenterology* **49**, 1535-1537
124. Li, F. Y., and Lai, M. D. (2009) Colorectal cancer, one entity or three. *Journal of Zhejiang University. Science. B* **10**, 219-229
125. Deng, G., Kakar, S., Tanaka, H., Matsuzaki, K., Miura, S., *et al.* (2008) Proximal and distal colorectal cancers show distinct gene-specific methylation profiles and clinical and molecular characteristics. *Eur J Cancer* **44**, 1290-1301
126. Hiraoka, S., Kato, J., Horii, J., Saito, S., Harada, K., *et al.* (2010) Methylation status of normal background mucosa is correlated with occurrence and development of neoplasia in the distal colon. *Human pathology* **41**, 38-47

127. Mladenova, D., Daniel, J. J., Dahlstrom, J. E., Bean, E., Gupta, R., *et al.* (2011) The NSAID sulindac is chemopreventive in the mouse distal colon but carcinogenic in the proximal colon. *Gut* **60**, 350-360
128. Ahnen, D. J. (1985) Are animal models of colon cancer relevant to human disease. *Digestive diseases and sciences* **30**, 103S-106S
129. Hirose, Y., Yoshimi, N., Suzui, M., Kawabata, K., Tanaka, T., *et al.* (1997) Expression of bcl-2, bax, and bcl-XL proteins in azoxymethane-induced rat colonic adenocarcinomas. *Molecular carcinogenesis* **19**, 25-30
130. Takahashi, M., Fukuda, K., Sugimura, T., and Wakabayashi, K. (1998) Beta-catenin is frequently mutated and demonstrates altered cellular location in azoxymethane-induced rat colon tumors. *Cancer research* **58**, 42-46
131. Zhang, J., Wu, G., Chapkin, R. S., and Lupton, J. R. (1998) Energy metabolism of rat colonocytes changes during the tumorigenic process and is dependent on diet and carcinogen. *The Journal of nutrition* **128**, 1262-1269
132. Lee, T. I., Johnstone, S. E., and Young, R. A. (2006) Chromatin immunoprecipitation and microarray-based analysis of protein location. *Nature protocols* **1**, 729-748
133. Zang, C., Schones, D. E., Zeng, C., Cui, K., Zhao, K., *et al.* (2009) A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics* **25**, 1952-1958
134. Quinlan, A. R., and Hall, I. M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842
135. Gibbs, R. A., Weinstock, G. M., Metzker, M. L., Muzny, D. M., Sodergren, E. J., *et al.* (2004) Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature* **428**, 493-521
136. Bookout, A. L., and Mangelsdorf, D. J. (2003) Quantitative real-time PCR protocol for analysis of nuclear receptor signaling pathways. *Nuclear receptor signaling* **1**, e012

137. Pfaffl, M. W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic acids research* **29**, e45
138. Trapnell, C., Pachter, L., and Salzberg, S. L. (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105-1111
139. Roberts, A., Pimentel, H., Trapnell, C., and Pachter, L. (2011) Identification of novel transcripts in annotated genomes using RNA-Seq. *Bioinformatics* **27**, 2325-2329
140. Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., *et al.* (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols* **7**, 562-578
141. Mituyama, T., Yamada, K., Hattori, E., Okida, H., Ono, Y., *et al.* (2009) The Functional RNA Database 3.0: databases to support mining and annotation of functional RNAs. *Nucleic acids research* **37**, D89-92
142. Bu, D., Yu, K., Sun, S., Xie, C., Skogerbo, G., *et al.* (2012) NONCODE v3.0: integrative annotation of long noncoding RNAs. *Nucleic acids research* **40**, D210-215
143. Meyer, L. R., Zweig, A. S., Hinrichs, A. S., Karolchik, D., Kuhn, R. M., *et al.* (2013) The UCSC Genome Browser database: extensions and updates 2013. *Nucleic acids research* **41**, D64-69
144. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., *et al.* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research* **25**, 3389-3402
145. Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., *et al.* (2008) NCBI BLAST: a better web interface. *Nucleic acids research* **36**, W5-9
146. Wang, K. C., Yang, Y. W., Liu, B., Sanyal, A., Corces-Zimmerman, R., *et al.* (2011) A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **472**, 120-124

147. Zeitlinger, J., Stark, A., Kellis, M., Hong, J. W., Nechaev, S., *et al.* (2007) RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo. *Nature genetics* **39**, 1512-1516
148. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., and Wold, B. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods* **5**, 621-628
149. Minoo, P., Zlobec, I., Peterson, M., Terracciano, L., and Lugli, A. (2010) Characterization of rectal, proximal and distal colon cancers based on clinicopathological, molecular and protein profiles. *International journal of oncology* **37**, 707-718
150. Adams, H., Tzankov, A., Lugli, A., and Zlobec, I. (2009) New time-dependent approach to analyse the prognostic significance of immunohistochemical biomarkers in colon cancer and diffuse large B-cell lymphoma. *Journal of clinical pathology* **62**, 986-997
151. Su, W., Bush, C. R., Necela, B. M., Calcagno, S. R., Murray, N. R., *et al.* (2007) Differential expression, distribution, and function of PPAR-gamma in the proximal and distal colon. *Physiological genomics* **30**, 342-353
152. Deming, D. A., Leystra, A. A., Nettekoven, L., Sievers, C., Miller, D., *et al.* (2013) PIK3CA and APC mutations are synergistic in the development of intestinal cancers. *Oncogene*
153. Ogino, S., Lochhead, P., Giovannucci, E., Meyerhardt, J. A., Fuchs, C. S., *et al.* (2013) Discovery of colorectal cancer PIK3CA mutation as potential predictive biomarker: power and promise of molecular pathological epidemiology. *Oncogene*
154. Kwak, H., Fuda, N. J., Core, L. J., and Lis, J. T. (2013) Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science* **339**, 950-953
155. Yamauchi, M., Lochhead, P., Morikawa, T., Huttenhower, C., Chan, A. T., *et al.* (2012) Colorectal cancer: a tale of two sides or a continuum? *Gut* **61**, 794-797

156. Ohta, M., Sugimoto, T., Seto, M., Mohri, D., Asaoka, Y., *et al.* (2008) Genetic alterations in colorectal cancers with demethylation of insulin-like growth factor II. *Human pathology* **39**, 1301-1308
157. Pelsers, M. M., Namiot, Z., Kisielewski, W., Namiot, A., Januszkiewicz, M., *et al.* (2003) Intestinal-type and liver-type fatty acid-binding protein in the intestine. Tissue distribution and clinical utility. *Clinical biochemistry* **36**, 529-535
158. Ha, M., Hong, S., and Li, W. H. (2013) Predicting the probability of H3K4me3 occupation at a base pair from the genome sequence context. *Bioinformatics* **29**, 1199-1205
159. Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., *et al.* (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature biotechnology* **28**, 511-515
160. Kim, A., Kiefer, C. M., and Dean, A. (2007) Distinctive signatures of histone methylation in transcribed coding and noncoding human beta-globin sequences. *Molecular and cellular biology* **27**, 1271-1279
161. Galamb, O., Gyorffy, B., Sipos, F., Spisak, S., Nemeth, A. M., *et al.* (2008) Inflammation, adenoma and cancer: objective classification of colon biopsy specimens with gene expression signature. *Disease markers* **25**, 1-16
162. Garcia-Barcelo, M. M., Miao, X., Lui, V. C., So, M. T., Ngan, E. S., *et al.* (2007) Correlation between genetic variations in Hox clusters and Hirschsprung's disease. *Annals of human genetics* **71**, 526-536
163. Sodik, N. M., Chen, X., Park, R., Nickel, A. E., Conti, P. S., *et al.* (2006) Smad3 deficiency promotes tumorigenesis in the distal colon of ApcMin/+ mice. *Cancer research* **66**, 8430-8438
164. O'Hagan, H. M., Wang, W., Sen, S., Destefano Shields, C., Lee, S. S., *et al.* (2011) Oxidative damage targets complexes containing DNA methyltransferases, SIRT1, and polycomb members to promoter CpG Islands. *Cancer cell* **20**, 606-619

165. Diehl, D., Oesterle, D., Elmlinger, M. W., Hoeflich, A., Wolf, E., *et al.* (2006) IGF-II transgenic mice display increased aberrant colon crypt multiplicity and tumor volume after 1,2-dimethylhydrazine treatment. *Journal of carcinogenesis* **5**, 24
166. Renehan, A. G., Painter, J. E., O'Halloran, D., Atkin, W. S., Potten, C. S., *et al.* (2000) Circulating insulin-like growth factor II and colorectal adenomas. *The Journal of clinical endocrinology and metabolism* **85**, 3402-3408
167. Fussbroich, B., Wagener, N., Macher-Goeppinger, S., Benner, A., Falth, M., *et al.* (2011) EZH2 depletion blocks the proliferation of colon cancer cells. *PloS one* **6**, e21651
168. Tejpar, S., Bertagnolli, M., Bosman, F., Lenz, H. J., Garraway, L., *et al.* (2010) Prognostic and predictive biomarkers in resected colon cancer: current status and future perspectives for integrating genomics into biomarker discovery. *The oncologist* **15**, 390-404
169. Slattery, M. L., Lundgreen, A., Herrick, J. S., Caan, B. J., Potter, J. D., *et al.* (2011) Associations between genetic variation in RUNX1, RUNX2, RUNX3, MAPK1 and eIF4E and risk of colon and rectal cancer: additional support for a TGF-beta-signaling pathway. *Carcinogenesis* **32**, 318-326
170. Paziienza, V., Vinciguerra, M., and Mazzocchi, G. (2012) PPARs Signaling and Cancer in the Gastrointestinal System. *PPAR research* **2012**, 560846
171. Kelly, D. P. (2001) The pleiotropic nature of the vascular PPAR gene regulatory pathway. *Circulation research* **89**, 935-937
172. Cotsapas, C., Prokunina-Olsson, L., Welch, C., Saxena, R., Weaver, C., *et al.* (2010) Expression analysis of loci associated with type 2 diabetes in human tissues. *Diabetologia* **53**, 2334-2339
173. Schlezinger, J. J., Emberley, J. K., and Sherr, D. H. (2006) Activation of multiple mitogen-activated protein kinases in pro/pre-B cells by GW7845, a peroxisome proliferator-activated receptor gamma agonist, and their contribution to GW7845-induced apoptosis. *Toxicological sciences : an official journal of the Society of Toxicology* **92**, 433-444

174. Subbaramaiah, K., Lin, D. T., Hart, J. C., and Dannenberg, A. J. (2001) Peroxisome proliferator-activated receptor gamma ligands suppress the transcriptional activation of cyclooxygenase-2. Evidence for involvement of activator protein-1 and CREB-binding protein/p300. *The Journal of biological chemistry* **276**, 12440-12448
175. Suva, M. L., Riggi, N., and Bernstein, B. E. (2013) Epigenetic reprogramming in cancer. *Science* **339**, 1567-1570
176. Shoeb, M., Ramana, K. V., and Srivastava, S. K. (2013) Aldose reductase inhibition enhances TRAIL-induced human colon cancer cell apoptosis through AKT/FOXO3a-dependent upregulation of death receptors. *Free radical biology & medicine* **63**, 280-290
177. Bullock, M. D., Bruce, A., Sreekumar, R., Curtis, N., Cheung, T., *et al.* (2013) FOXO3 expression during colorectal cancer progression: biomarker potential reflects a tumour suppressor role. *British journal of cancer* **109**, 387-394
178. Fullgrabe, J., Kavanagh, E., and Joseph, B. (2011) Histone onco-modifications. *Oncogene* **30**, 3391-3403
179. Enroth, S., Rada-Iglesias, A., Andersson, R., Wallerman, O., Wanders, A., *et al.* (2011) Cancer associated epigenetic transitions identified by genome-wide histone methylation binding profiles in human colorectal cancer samples and paired normal mucosa. *BMC Cancer* **11**, 450
180. De Robertis, M., Massi, E., Poeta, M. L., Carotti, S., Morini, S., *et al.* (2011) The AOM/DSS murine model for the study of colon carcinogenesis: From pathways to diagnosis and therapy studies. *J Carcinog* **10**, 9
181. Vanamala, J., Leonardi, T., Patil, B. S., Taddeo, S. S., Murphy, M. E., *et al.* (2006) Suppression of colon carcinogenesis by bioactive compounds in grapefruit. *Carcinogenesis* **27**, 1257-1265
182. Arends, M. J. (2013) Pathways of colorectal carcinogenesis. *Appl Immunohistochem Mol Morphol* **21**, 97-102

183. Nakazawa, T., Kondo, T., Ma, D., Niu, D., Mochizuki, K., *et al.* (2012) Global histone modification of histone H3 in colorectal cancer and its precursor lesions. *Hum Pathol* **43**, 834-842
184. McLellan, E. A., and Bird, R. P. (1988) Aberrant crypts: potential preneoplastic lesions in the murine colon. *Cancer research* **48**, 6187-6192
185. Triff, K., Konganti, K., Gaddis, S., Zhou, B. Y., Ivanov, I., *et al.* (2013) Genome-wide analysis of the rat colon reveals proximal-distal differences in histone modifications and proto-oncogene expression. *Physiological genomics* **45**, 1229-1243
186. Zhang, Y., Liu, T., Meyer, C. A., Eeckhoutte, J., Johnson, D. S., *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137
187. Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140
188. Nikolayeva, O., and Robinson, M. D. (2014) edgeR for differential RNA-seq and ChIP-seq analysis: an application to stem cell biology. *Methods Mol Biol* **1150**, 45-79
189. Anders, S., and Huber, W. (2010) Differential expression analysis for sequence count data. *Genome Biol* **11**, R106
190. Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., *et al.* (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21
191. Bolger, A. M., Lohse, M., and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120
192. Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., *et al.* (2014) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks (vol 7, pg 562, 2012). *Nat Protoc* **9**, 2513-2513
193. Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J., and Prins, P. (2015) Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**, 2032-2034

194. Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., *et al.* (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* **7**, 562-578
195. Pauli, A., Valen, E., Lin, M. F., Garber, M., Vastenhouw, N. L., *et al.* (2012) Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis. *Genome Res* **22**, 577-591
196. Pruitt, K. D., Brown, G. R., Hiatt, S. M., Thibaud-Nissen, F., Astashyn, A., *et al.* (2014) RefSeq: an update on mammalian reference sequences. *Nucleic acids research* **42**, D756-763
197. Kong, L., Zhang, Y., Ye, Z. Q., Liu, X. Q., Zhao, S. Q., *et al.* (2007) CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic acids research* **35**, W345-349
198. Nawrocki, E. P., and Eddy, S. R. (2013) Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* **29**, 2933-2935
199. Lorenz, R., Bernhart, S. H., Honer Zu Siederdissen, C., Tafer, H., Flamm, C., *et al.* (2011) ViennaRNA Package 2.0. *Algorithms for molecular biology : AMB* **6**, 26
200. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., *et al.* (2009) BLAST+: architecture and applications. *Bmc Bioinformatics* **10**, 421
201. Kobaek-Larsen, M., Thorup, I., Diederichsen, A., Fenger, C., and Hoitinga, M. R. (2000) Review of colorectal cancer and its metastases in rodent models: comparative aspects with those in humans. *Comp Med* **50**, 16-26
202. Malilas, W., Koh, S. S., Srisuttee, R., Boonying, W., Cho, I. R., *et al.* (2013) Cancer upregulated gene 2, a novel oncogene, confers resistance to oncolytic vesicular stomatitis virus through STAT1-OASL2 signaling. *Cancer Gene Ther* **20**, 125-132
203. Saito, H., Kubota, M., Roberts, R. W., Chi, Q., and Matsunami, H. (2004) RTP family members induce functional expression of mammalian odorant receptors. *Cell* **119**, 679-691

204. Long, F., Liu, H., Hahn, C., Sumazin, P., Zhang, M. Q., *et al.* (2004) Genome-wide prediction and analysis of function-specific transcription factor binding sites. *In Silico Biol* **4**, 395-410
205. Bishop-Bailey, D., and Swales, K. E. (2008) The Role of PPARs in the Endothelium: Implications for Cancer Therapy. *Ppar Res* **2008**, 904251
206. Tachibana, K., Yamasaki, D., Ishimoto, K., and Doi, T. (2008) The Role of PPARs in Cancer. *Ppar Res* **2008**, 102737
207. Gulmann, C., Sheehan, K. M., Conroy, R. M., Wulfkuhle, J. D., Espina, V., *et al.* (2009) Quantitative cell signalling analysis reveals down-regulation of MAPK pathway activation in colorectal cancer. *J Pathol* **218**, 514-519
208. Prensner, J. R., and Chinnaiyan, A. M. (2011) The emergence of lncRNAs in cancer biology. *Cancer Discov* **1**, 391-407
209. Wilting, R. H., and Dannenberg, J. H. (2012) Epigenetic mechanisms in tumorigenesis, tumor cell heterogeneity and drug resistance. *Drug resistance updates : reviews and commentaries in antimicrobial and anticancer chemotherapy* **15**, 21-38
210. Rotili, D., and Mai, A. (2011) Targeting Histone Demethylases: A New Avenue for the Fight against Cancer. *Genes Cancer* **2**, 663-679
211. Tisserand, J., Khetchoumian, K., Thibault, C., Dembele, D., Chambon, P., *et al.* (2011) Tripartite motif 24 (Trim24/Tif1alpha) tumor suppressor protein is a novel negative regulator of interferon (IFN)/signal transducers and activators of transcription (STAT) signaling pathway acting through retinoic acid receptor alpha (Raralpha) inhibition. *The Journal of biological chemistry* **286**, 33369-33379
212. Yang, Y., and Jobin, C. (2014) Microbial imbalance and intestinal pathologies: connections and contributions. *Dis Model Mech* **7**, 1131-1142
213. Brodziak, F., Meharg, C., Blaut, M., and Loh, G. (2013) Differences in mucosal gene expression in the colon of two inbred mouse strains after colonization with commensal gut bacteria. *PloS one* **8**, e72317

214. Kikuchi, M., Okumura, F., Tsukiyama, T., Watanabe, M., Miyajima, N., *et al.* (2009) TRIM24 mediates ligand-dependent activation of androgen receptor and is repressed by a bromodomain-containing protein, BRD7, in prostate cancer cells. *Biochimica et biophysica acta* **1793**, 1828-1836
215. Tang, X. H., and Gudas, L. J. (2011) Retinoids, retinoic acid receptors, and cancer. *Annu Rev Pathol* **6**, 345-364
216. Zhang, Y., Liang, J., and Li, Q. (2014) Coordinated regulation of retinoic acid signaling pathway by KDM5B and polycomb repressive complex 2. *J Cell Biochem* **115**, 1528-1538
217. Tsai, W. W., Wang, Z., Yiu, T. T., Akdemir, K. C., Xia, W., *et al.* (2010) TRIM24 links a non-canonical histone signature to breast cancer. *Nature* **468**, 927-932
218. Klein, B. J., Piao, L., Xi, Y., Rincon-Arano, H., Rothbart, S. B., *et al.* (2014) The histone-H3K4-specific demethylase KDM5B binds to its substrate and product through distinct PHD fingers. *Cell Rep* **6**, 325-335
219. Hoang, B., Trinh, A., Birnbaumer, L., and Edwards, R. A. (2007) Decreased MAPK- and PGE2-dependent IL-11 production in Gialpha2^{-/-} colonic myofibroblasts. *American journal of physiology. Gastrointestinal and liver physiology* **292**, G1511-1519
220. Gruver-Yates, A. L., and Cidlowski, J. A. (2013) Tissue-specific actions of glucocorticoids on apoptosis: a double-edged sword. *Cells* **2**, 202-223
221. Allgayer, H. (2010) Pcd4, a colon cancer prognostic that is regulated by a microRNA. *Crit Rev Oncol Hematol* **73**, 185-191
222. Ye, L. C., Zhu, X., Qiu, J. J., Xu, J., and Wei, Y. (2015) Involvement of long non-coding RNA in colorectal cancer: From benchtop to bedside (Review). *Oncol Lett* **9**, 1039-1045
223. Iessi, E., Zischler, L., Etringer, A., Bergeret, M., Morle, A., *et al.* (2015) Death Receptor-Induced Apoptosis Signalling Regulation by Ezrin Is Cell Type Dependent and Occurs in a DISC-Independent Manner in Colon Cancer Cells. *PloS one* **10**

224. Wu, B., Iwakiri, R., Ootani, A., Tsunada, S., Fujise, T., *et al.* (2004) Dietary corn oil promotes colon cancer by inhibiting mitochondria-dependent apoptosis in azoxymethane-treated rats. *Experimental biology and medicine* **229**, 1017-1025
225. Triff, K., Kim, E., and Chapkin, R. S. (2015) Chemoprotective epigenetic mechanisms in a colorectal cancer model: Modulation by n-3 PUFA in combination with fermentable fiber. *Curr Pharmacol Rep* **1**, 11-20
226. Anti, M., Marra, G., Armelao, F., Bartoli, G. M., Ficarelli, R., *et al.* (1992) Effect of omega-3 fatty acids on rectal mucosal cell proliferation in subjects at risk for colon cancer. *Gastroenterology* **103**, 883-891
227. Cheng, J., Ogawa, K., Kuriki, K., Yokoyama, Y., Kamiya, T., *et al.* (2003) Increased intake of n-3 polyunsaturated fatty acids elevates the level of apoptosis in the normal sigmoid colon of patients polypectomized for adenomas/tumors. *Cancer Lett* **193**, 17-24
228. Chang, W. C., Chapkin, R. S., and Lupton, J. R. (1997) Predictive value of proliferation, differentiation and apoptosis as intermediate markers for colon tumorigenesis. *Carcinogenesis* **18**, 721-730
229. Whelan, J., and McEntee, M. F. (2004) Dietary (n-6) PUFA and intestinal tumorigenesis. *The Journal of nutrition* **134**, 3421S-3426S
230. Zeng, H., Lazarova, D. L., and Bordonaro, M. (2014) Mechanisms linking dietary fiber, gut microbiota and colon cancer prevention. *World journal of gastrointestinal oncology* **6**, 41-51
231. De Filippo, C., Cavalieri, D., Di Paola, M., Ramazzotti, M., Poullet, J. B., *et al.* (2010) Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc Natl Acad Sci U S A* **107**, 14691-14696
232. Shi, L., and Tu, B. P. (2015) Acetyl-CoA and the regulation of metabolism: mechanisms and consequences. *Current opinion in cell biology* **33**, 125-131
233. Lupton, J. R. (2004) Microbial degradation products influence colon cancer risk: the butyrate controversy. *The Journal of nutrition* **134**, 479-482

234. Honda, K., Ohba, Y., Yanai, H., Negishi, H., Mizutani, T., *et al.* (2005) Spatiotemporal regulation of MyD88-IRF-7 signalling for robust type-I interferon induction. *Nature* **434**, 1035-1040
235. Savitsky, D., Tamura, T., Yanai, H., and Taniguchi, T. (2010) Regulation of immunity and oncogenesis by the IRF transcription factor family. *Cancer immunology, immunotherapy : CII* **59**, 489-510
236. Sanders, L. M., Henderson, C. E., Hong, M. Y., Barhoumi, R., Burghardt, R. C., *et al.* (2004) An increase in reactive oxygen species by dietary fish oil coupled with the attenuation of antioxidant defenses by dietary pectin enhances rat colonocyte apoptosis. *The Journal of nutrition* **134**, 3233-3238
237. Reddy, B. S. (1994) Chemoprevention of colon cancer by dietary fatty acids. *Cancer metastasis reviews* **13**, 285-302
238. Young, J., Simms, L. A., Biden, K. G., Wynter, C., Whitehall, V., *et al.* (2001) Features of colorectal cancers with high-level microsatellite instability occurring in familial and sporadic settings: parallel pathways of tumorigenesis. *Am J Pathol* **159**, 2107-2116
239. Barone, M., Lofano, K., De Tullio, N., Licinio, R., Albano, F., *et al.* (2012) Dietary, endocrine, and metabolic factors in the development of colorectal cancer. *Journal of gastrointestinal cancer* **43**, 13-19
240. Lawrie, L. C., Dundas, S. R., Curran, S., and Murray, G. I. (2004) Liver fatty acid binding protein expression in colorectal neoplasia. *British journal of cancer* **90**, 1955-1960
241. Goichon, A., Chan, P., Lecleire, S., Coquard, A., Cailleux, A. F., *et al.* (2013) An enteral leucine supply modulates human duodenal mucosal proteome and decreases the expression of enzymes involved in fatty acid beta-oxidation. *Journal of proteomics* **78**, 535-544
242. Ohkuni, A., Ohno, Y., and Kihara, A. (2013) Identification of acyl-CoA synthetases involved in the mammalian sphingosine 1-phosphate metabolic pathway. *Biochem Biophys Res Commun* **442**, 195-201

243. Blouin, J. M., Bortoli, S., Nacfer, M., Collinet, M., Penot, G., *et al.* (2010) Down-regulation of the phosphoenolpyruvate carboxykinase gene in human colon tumors and induction by omega-3 fatty acids. *Biochimie* **92**, 1772-1777
244. Carmona-Antonanzas, G., Tocher, D. R., Martinez-Rubio, L., and Leaver, M. J. (2014) Conservation of lipid metabolic gene transcriptional regulatory networks in fish and mammals. *Gene* **534**, 1-9
245. Hanasaki, K., Yamada, K., Yamamoto, S., Ishimoto, Y., Saiga, A., *et al.* (2002) Potent modification of low density lipoprotein by group X secretory phospholipase A2 is linked to macrophage foam cell formation. *The Journal of biological chemistry* **277**, 29116-29124
246. Kramer, A., Green, J., Pollard, J., Jr., and Tugendreich, S. (2014) Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics* **30**, 523-530
247. Honda, K., Yanai, H., Mizutani, T., Negishi, H., Shimada, N., *et al.* (2004) Role of a transductional-transcriptional processor complex involving MyD88 and IRF-7 in Toll-like receptor signaling. *Proc Natl Acad Sci U S A* **101**, 15416-15421
248. Donohoe, D. R., Collins, L. B., Wali, A., Bigler, R., Sun, W., *et al.* (2012) The Warburg effect dictates the mechanism of butyrate-mediated histone acetylation and cell proliferation. *Molecular cell* **48**, 612-626
249. Donohoe, D. R., Curry, K. P., and Bultman, S. J. (2013) Microbial oncotarget: bacterial-produced butyrate, chemoprevention and Warburg effect. *Oncotarget* **4**, 182-183
250. Bultman, S. J., and Jobin, C. (2014) Microbial-derived butyrate: an oncometabolite or tumor-suppressive metabolite? *Cell host & microbe* **16**, 143-145
251. Hong, M. Y., Lupton, J. R., Morris, J. S., Wang, N., Carroll, R. J., *et al.* (2000) Dietary fish oil reduces O6-methylguanine DNA adduct levels in rat colon in part by increasing apoptosis during tumor initiation. *Cancer Epidemiol Biomarkers Prev* **9**, 819-826

252. Kim, S., Sandler, D. P., Galanko, J., Martin, C., and Sandler, R. S. (2010) Intake of polyunsaturated fatty acids and distal large bowel cancer risk in whites and African Americans. *American journal of epidemiology* **171**, 969-979
253. Granados, S., Quiles, J. L., Gil, A., and Ramirez-Tortosa, M. C. (2006) Dietary lipids and cancer. *Nutricion hospitalaria* **21 Suppl 2**, 42-52, 44-54
254. Mori, T., Kondo, H., Hase, T., Tokimitsu, I., and Murase, T. (2007) Dietary fish oil upregulates intestinal lipid metabolism and reduces body weight gain in C57BL/6J mice. *The Journal of nutrition* **137**, 2629-2634
255. van Schothorst, E. M., Flachs, P., Franssen-van Hal, N. L., Kuda, O., Bunschoten, A., *et al.* (2009) Induction of lipid oxidation by polyunsaturated fatty acids of marine origin in small intestine of mice fed a high-fat diet. *BMC genomics* **10**, 110
256. Ouellette, C., Rudkowska, I., Lemieux, S., Lamarche, B., Couture, P., *et al.* (2014) Gene-diet interactions with polymorphisms of the MGLL gene on plasma low-density lipoprotein cholesterol and size following an omega-3 polyunsaturated fatty acid supplementation: a clinical trial. *Lipids in health and disease* **13**, 86
257. Weng, H., Endo, K., Li, J., Kito, N., and Iwai, N. (2015) Induction of peroxisomes by butyrate-producing probiotics. *PloS one* **10**, e0117851
258. Mikula, M., Rubel, T., Karczmarski, J., Goryca, K., Dadlez, M., *et al.* (2010) Integrating proteomic and transcriptomic high-throughput surveys for search of new biomarkers of colon tumors. *Functional & integrative genomics*
259. Rudolph, U., Finegold, M. J., Rich, S. S., Harriman, G. R., Srinivasan, Y., *et al.* (1995) Ulcerative colitis and adenocarcinoma of the colon in G alpha i2-deficient mice. *Nat Genet* **10**, 143-150
260. Umoh, F. I., Kato, I., Ren, J., Wachowiak, P. L., Ruffin, M. T. t., *et al.* (2015) Markers of systemic exposures to products of intestinal bacteria in a dietary intervention study. *Eur J Nutr*

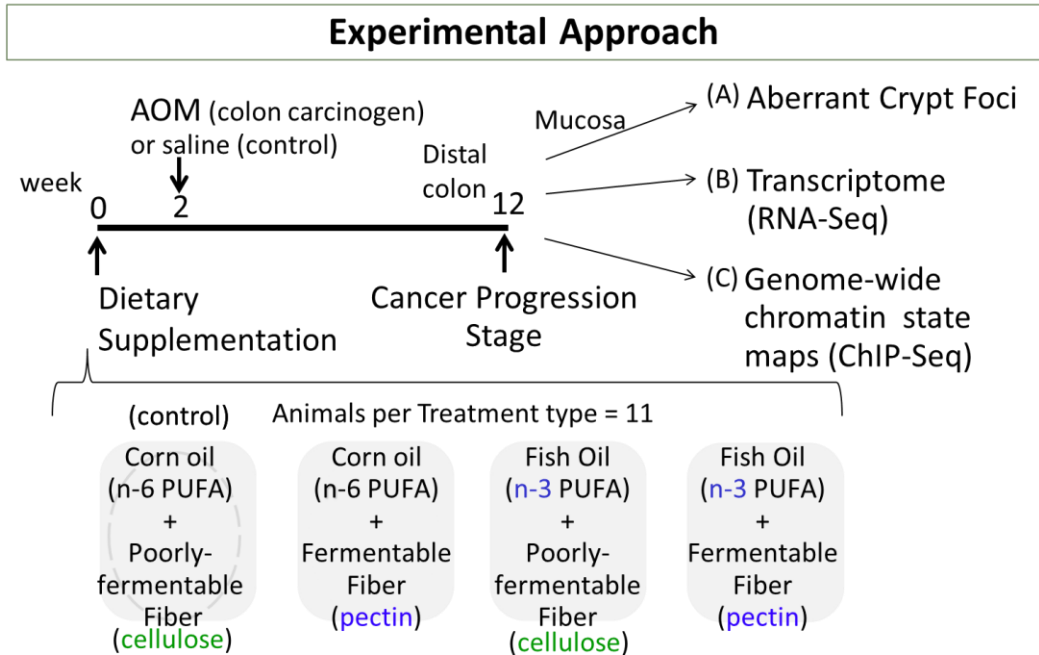
261. Nowinski, S. M., Solmonson, A., Rundhaug, J. E., Rho, O., Cho, J., *et al.* (2015)
Mitochondrial uncoupling links lipid catabolism to Akt inhibition and resistance
to tumorigenesis. *Nat Commun* **6**, 8137
262. Fan, Y. Y., Ran, Q., Toyokuni, S., Okazaki, Y., Callaway, E. S., *et al.* (2011)
Dietary fish oil promotes colonic apoptosis and mitochondrial proton leak in
oxidatively stressed mice. *Cancer prevention research* **4**, 1267-1274
263. Cardel, M., Lemas, D. J., Jackson, K. H., Friedman, J. E., and Fernandez, J. R.
(2015) Higher Intake of PUFAs Is Associated with Lower Total and Visceral
Adiposity and Higher Lean Mass in a Racially Diverse Sample of Children. *The
Journal of nutrition* **145**, 2146-2152
264. Correia, M., Casal, S., Vinagre, J., Seruca, R., Figueiredo, C., *et al.* (2014)
Helicobacter pylori's cholesterol uptake impacts resistance to docosahexaenoic
acid. *International journal of medical microbiology : IJMM* **304**, 314-320
265. Jacometo, C. B., Schmitt, E., Pfeifer, L. F., Schneider, A., Bado, F., *et al.* (2014)
Linoleic and alpha-linolenic fatty acid consumption over three generations exert
cumulative regulation of hepatic expression of genes related to lipid metabolism.
Genes & nutrition **9**, 405
266. Qi, Y., Jiang, C., Tanaka, N., Krausz, K. W., Brocker, C. N., *et al.* (2014)
PPARalpha-dependent exacerbation of experimental colitis by the hypolipidemic
drug fenofibrate. *American journal of physiology. Gastrointestinal and liver
physiology* **307**, G564-573
267. Petrescu, A. D., Huang, H., Martin, G. G., McIntosh, A. L., Storey, S. M., *et al.*
(2013) Impact of L-FABP and glucose on polyunsaturated fatty acid induction of
PPARalpha-regulated beta-oxidative enzymes. *American journal of physiology.
Gastrointestinal and liver physiology* **304**, G241-256
268. Takahashi, M., Tsuboyama-Kasaoka, N., Nakatani, T., Ishii, M., Tsutsumi, S., *et al.*
(2002) Fish oil feeding alters liver gene expressions to defend against
PPARalpha activation and ROS production. *American journal of physiology.
Gastrointestinal and liver physiology* **282**, G338-348

269. Bargut, T. C., Frantz, E. D., Mandarim-de-Lacerda, C. A., and Aguila, M. B. (2014) Effects of a diet rich in n-3 polyunsaturated fatty acids on hepatic lipogenesis and beta-oxidation in mice. *Lipids* **49**, 431-444
270. Rosignoli, P., Fabiani, R., De Bartolomeo, A., Fuccelli, R., Pelli, M. A., *et al.* (2008) Genotoxic effect of bile acids on human normal and tumour colon cells and protection by dietary antioxidants and butyrate. *Eur J Nutr* **47**, 301-309
271. Ajouz, H., Mukherji, D., and Shamseddine, A. (2014) Secondary bile acids: an underrecognized cause of colon cancer. *World journal of surgical oncology* **12**, 164
272. Mashek, D. G., Li, L. O., and Coleman, R. A. (2007) Long-chain acyl-CoA synthetases and fatty acid channeling. *Future lipidology* **2**, 465-476
273. de Vogel-van den Bosch, H. M., Bunger, M., de Groot, P. J., Bosch-Vermeulen, H., Hooiveld, G. J., *et al.* (2008) PPARalpha-mediated effects of dietary lipids on intestinal barrier gene expression. *BMC genomics* **9**, 231
274. Balaban, S., Lee, L. S., Schreuder, M., and Hoy, A. J. (2015) Obesity and cancer progression: is there a role of fatty acid metabolism? *BioMed research international* **2015**, 274585
275. Carracedo, A., Cantley, L. C., and Pandolfi, P. P. (2013) Cancer metabolism: fatty acid oxidation in the limelight. *Nat Rev Cancer* **13**, 227-232
276. Shah, M. S., Davidson, L. A., and Chapkin, R. S. (2012) Mechanistic insights into the role of microRNAs in cancer: influence of nutrient crosstalk. *Frontiers in genetics* **3**, 305
277. Chai, H., and Brown, R. E. (2009) Field effect in cancer-an update. *Annals of clinical and laboratory science* **39**, 331-337
278. Hawthorn, L., Lan, L., and Mojica, W. (2014) Evidence for field effect cancerization in colorectal cancer. *Genomics* **103**, 211-221
279. Peterson, L. W., and Artis, D. (2014) Intestinal epithelial cells: regulators of barrier function and immune homeostasis. *Nature reviews. Immunology* **14**, 141-153

280. Cerf-Bensussan, N., and Gaboriau-Routhiau, V. (2010) The immune system and the gut microbiota: friends or foes? *Nature reviews. Immunology* **10**, 735-744

APPENDIX A

EXPERIMENTAL APPROACH



These diets contained:

- Either 15% corn oil by weight or 3.5% corn oil + 11.5% fish oil (equal to 30% calories)
- Either 6% pectin or 6% cellulose (equal 30 g of fiber per day in humans)
- Traditional diets with similar fish oil nutritional levels: Inuits (Alaska) 9-34 g EPA/DHA, ; Greenland, Eskimos 7-10 g(long n-3 PUFAs), Japan 1.3 g
- Butyrate levels in the colon
~5mM in lumen
~5uM in blood
- n-3 PUFA levels in the colon
~10uM free fatty acid and
~ 100uM in the form of complex lipid

Composition of experimental diets

Ingredient	Corn oil diet	Fish oil diet
	<i>g/kg diet</i>	
Dextrose	510.6	510.6
Casein	223.5	223.5
DL-Methionine	3.4	3.4
Corn oil	150.0	35.0
Fish oil ¹	0.0	115.0
AIN-76 Mineral mix ¹⁻²	39.1	39.1
AIN-76 Vitamin mix ¹⁻²	11.2	11.2
Choline bitartrate	2.2	2.2
Cellulose ¹⁻³ or pectin ¹⁻⁴	60.0	60.0

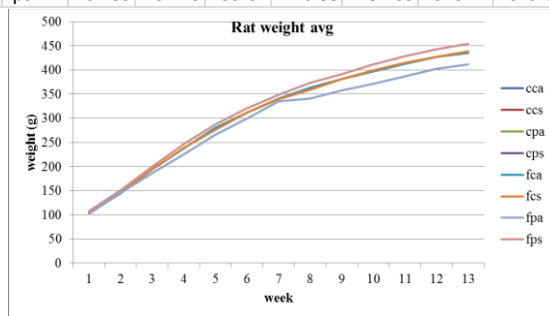
- F1-1 Vacuum-deodorized Menhaden fish oil (NIH Fish oil test material program, Southeast Fisheries Center, Charleston, SC). Both oils contained: α -tocopherol, 1.5 mg/g; γ -tocopherol, 1.0 mg/g; t butylhydroquinone, 0.025% as antioxidants.
- ¹F1-2 AIN 1977.
- ¹F1-3 Cellulose was microcrystalline.
- ¹F1-4 Citrus pectin

APPENDIX B

ANIMAL WEIGHT LOG

rat weight in grams													at kill
week	1	2	3	4	5	6	7	8	9	10	11	12	13
ID	4/19'	25-Apr	2-May	9-May	16-May	23-May	30-May	6/6/2011	6/13/2011	6/22/2011	6/27/2011	7/7/2011	7/14/2011
1101	113.1	161.5	203.3	249.4	288.2	319.7	336.4	353.1	377.2	394.3	405.7	417.1	428.1
1102	115.8	161.8	207.6	247.3	288.1	328	350.7	373.4	392.1	402	411.6	421.2	429.7
1103	108.4	148.4	190.1	235	271.8	315.9	338.5	361.1	380.1	400.1	410.5	420.9	430.7
1104	117.7	164.6	205.2	256.6	300	337.3	356.7	376.1	397.1	413.6	429.6	445.6	458.3
1105	106.9	154	195.9	240.6	272.8	305.7	323.4	341.1	356	361.5	368.85	376.2	385.2
2106	97.5	138.3	194.8	235.8	276.8	299.2	321.6	347.2	363.6	377.5	391.4	403.8	410.7
2107	95.5	130.1	179.1	219.85	260.6	282.25	303.9	319.2	331.9	340.25	348.6	364.2	370.7
2108	89.7	126.2	174.3	215.4	256.5	290	323.5	347.2	365.8	379.2	392.6	412.5	414.3
2109	89.3	126.2	173.3	213.85	254.4	275.25	296.1	314.2	325.8	337.95	350.1	369.4	371.5
2110	95.3	133.8	184.2	224.3	264.4	295.25	326.1	354.1	367.1	382.2	397.3	403.2	423.5
2111	101.4	141	194.8	230.8	266.8	292.8	318.8	333.8	345.7	358.45	371.2	386.1	393.3
cca	102.78	144.17	191.15	233.54	272.76	303.76	326.88	347.32	363.85	377.00	388.86	401.84	410.55
1501	105	153.5	195.2	243.5	280.15	316.8	334.7	352.6	373.5	389.4	400.25	411.1	420.3
1502	110.2	166	195.2	246.6	282.85	319.1	343.65	368.2	372.9	389.8	401.65	413.5	422.7
1503	113.1	154.6	198.3	249	290.5	332	347.4	362.8	389.2	414.3	428.45	442.6	449.3
2504	99	125.4	184.9	228.65	272.4	291.15	309.9	332.2	352.2	364.35	376.5	389.7	394.4
2505	97.1	135	187.8	235.5	283.2	305.65	328.1	350.9	361.5	376.55	391.6	395.5	400.1
2506	105.2	137.7	189.5	228.65	267.8	287.95	308.1	317.1	330.8	338.7	346.6	356.1	360.4
ccs	104.93	145.37	191.82	238.65	279.48	308.78	328.64	347.30	363.35	378.85	390.84	401.42	407.87
1201	110.1	157.6	200.7	245.8	288.8	318.7	336.55	354.4	370.2	383.8	397.2	410.6	417.8
1202	119	167.2	211	257.2	297.6	330	348.85	367.7	387.2	404.8	418.8	432.8	449.3
1203	104.5	153.1	191.4	237.7	274.5	306.8	328.65	350.5	370.7	375.5	389.5	403.5	413.4
1204	110.3	153.2	193.7	231.7	264.5	306.2	327.65	349.1	368.3	386.4	396	405.6	425.1
1205	109.2	160.8	203.9	254.3	298.4	331	354.55	378.1	393.4	415.5	428.35	441.2	453.7
2206	101.9	139.3	190.1	225.05	260	291.15	322.3	326.2	360.4	371.9	383.4	394.1	400.2
2207	99.7	142.1	190.7	230	269.3	291.5	313.7	332.2	344.1	355.1	366.1	378.1	383.7
2208	99.1	135.9	190	232.9	275.8	306.15	336.5	357.8	373.2	389.35	405.5	419.8	433.9
2209	99.6	139.3	182.1	227.9	273.7	302.55	331.4	353.3	364.8	386.45	408.1	422.6	426.5
2210	89.3	127.6	176.7	214.7	252.7	282.1	311.5	333	348	364.75	381.5	395.5	394.5
cpa	104.27	147.61	193.03	235.73	275.53	306.62	331.17	350.23	368.03	383.36	397.45	410.38	419.81
1601	108.6	152.3	188.1	231.7	270.75	309.8	324.75	339.7	358.3	379.9	391.25	402.6	414.1
1602	109.5	156.5	195.1	247.6	289.6	331.6	348.2	364.8	380	406.6	420.7	434.8	448.8
1603	105.2	150.2	186.4	237.1	273.35	309.6	331.45	353.3	372.6	390.4	406.55	422.7	431.1
2604	95.3	124.1	170.2	217	263.8	287.1	310.4	330.4	334.8	352.5	370.2	379.9	388.3
2605	94.4	129.4	161.6	205.8	250	276	302	318.6	331.7	344.4	357.1	378.1	375.8
2606	96.3	133	187.3	229.2	271.1	296.4	321.7	343.4	356.6	368.9	381.2	393.1	398.3
cps	101.55	140.92	181.45	228.07	269.77	301.75	323.08	341.70	355.67	373.78	387.83	401.87	409.40

1301	106.8	154.8	202.4	254.7	288.4	322.1	344.6	367.1	388.9	404.2	417.85	431.5	446.1
1302	105.7	151.7	194.5	241.5	291.1	318.6	342.55	366.5	380.3	401.9	416.75	431.6	445.2
1303	117	165.2	207.5	251.4	292.8	322	341.75	361.5	376.6	387.6	401.85	416.1	413.3
1304	103.2	148.4	190.6	233.1	276.6	308.5	336	363.5	387.3	397.9	412.85	427.8	439.7
1305	106.9	151.7	190.6	233.9	276.7	302.6	326.35	350.1	364.1	383.1	396.45	409.8	421
2306	100.2	140.2	192.2	242.65	293.1	327.25	361.4	385.7	413.1	423.85	434.6	456.1	463.4
2307	100.6	136.6	194	233.3	272.6	304.4	336.2	360.3	376.4	392.5	408.6	428.2	430.6
2308	105	145	200.9	247	293.1	328.55	364	388.9	404.7	427.55	450.4	461.9	466.5
2309	94.6	125.5	178.7	215.75	252.8	285.95	319.1	330.1	345.5	362.8	380.1	385.5	389.5
2310	103.9	137.3	193.1	236.35	279.6	316.95	354.3	374.9	394	412.95	431.9	444.4	445.4
2311	98.2	133.9	183.6	226.35	269.1	294.95	320.8	343.4	359	376.7	394.4	410.9	423.6
fca	103.83	144.57	193.46	237.82	280.54	311.99	340.64	362.91	380.90	397.37	413.25	427.62	434.94
1701	113.6	160.2	203.6	253.7	290.65	327.6	352	376.4	401.5	419.1	426.8	434.5	444.6
1702	112.8	163.3	206.7	257.8	296.9	336	358.1	380.2	401.4	423.2	437.5	451.8	463.9
1703	116.7	167.1	206.9	259.2	295.4	331.6	354.95	378.3	398.5	413.6	427.7	441.8	457
2704	92.5	128.4	182.7	212.5	242.3	294.3	346.3	364	386.2	404	421.8	433.2	449.6
2705	95.8	130.9	181.1	222.5	263.9	288.2	312.5	326.8	342.2	358.65	375.1	388.2	391.7
2706	95.6	130.3	183.9	224.4	264.9	291.05	317.2	325.1	357.3	377.65	398	413.8	426.7
fcs	104.50	146.70	194.15	238.35	275.68	311.46	340.18	358.47	381.18	399.37	414.48	427.22	438.92
1401	106.9	166.5	194.6	232.5	267.5	282.2	311.05	339.9	347.2	354.5	371.7	389.4	400.3
1402	109.1	166.7	200.3	243	283.8	315.2	345.9	376.6	398.3	417.5	435.6	453.7	463.7
1403	95.7	138.6	179.4	219.7	260.1	288.1	319.4	350.7	373.1	388.6	406.3	424	432
1404	105.8	154.4	192.3	237.3	286.7	311.4	332	352.6	370.2	388.2	400.3	412.4	428.2
1405	109.2	162	198.9	237.9	290.5	325.1	348.55	372	392.5	407.9	426.85	445.8	450
1406	116.7	159.2	206.1	247.1	286.2	293.3	332.7	372.1	390.1	405.7	421.2	436.7	445.1
2407	99.7	132.4	172.4	205.25	238.1	309.8	381.5	288.1	302.7	316.8	330.9	332.7	337.6
2408	98.2	124.5	165.3	198.25	231.2	302.5	373.8	287.4	299	314.25	329.5	339.7	346.9
2409	96.2	135.5	180.8	220.8	260.8	290.25	319.7	342.5	360.2	375.15	390.1	401.7	411.9
2410	90.2	125.8	170.8	213.45	256.1	282.8	309.5	324.1	335.3	352.55	369.8	384.8	390.5
2411	94.4	132.8	182.7	224.2	265.7	292.25	318.8	340.2	362.3	367	371.7	407.8	418.1
fpa	102.01	145.31	185.78	225.40	266.06	299.35	335.72	340.56	357.35	371.65	386.72	402.61	411.30
1801	111.8	160	202.2	251.8	292.35	332.9	354.65	376.4	397.3	414.1	429.25	444.4	452.5
1802	115	166.2	209	253.4	298.3	343.2	363.4	383.6	405.9	425.5	442.25	459	475
1803	117.2	166.8	211	262.8	284.1	305.4	344.95	384.5	399.2	420	430.75	441.5	456.2
2804	106.2	151	208.9	255.45	302	331.95	361.9	386.6	401.8	421.3	440.8	457.9	463.3
2805	93.4	128.1	174.7	219.1	263.5	292.5	321.5	340.2	357.5	379.7	401.9	410.9	420.5
2806	100.5	136.6	191.9	238.7	285.5	316.7	347.9	370.2	389.7	408.95	428.2	441.6	458.6
fps	107.35	151.45	199.62	246.88	287.63	320.44	349.05	373.58	391.90	411.59	428.86	442.55	454.35



APPENDIX C

AZOXYMETHANE (AOM) INJECTION OF RATS

AZOXYMETHANE (AOM) INJECTION OF RATS

Azoxymethane

Preparation for injection

Dose for Rat: 15 mg AOM/kg Body Wt.

AOM from Sigma (Cat # A2853)

Preparation of Stock Solution

100 mg AOM/bottle (in 60 ml solution of 15% EtOH/acetic acid).

Add 940 ml sterile saline (Sigma S-8776).

100 mg AOM / 1000ul = 100 mg/ml

Filter sterilize solution into sterile serum vial using 0.2 micron filter.

Aliquot to several sterile vials to avoid frequent freeze/thaws.

Store solution in -20°C freezer up to 1 year.

Preparation of Working Solution

Dilute AOM solution 1:20 in sterile saline to give final concentration of 5 mg/ml. Use sterilized volumetric flask.

Aliquot AOM into sterile serum vials, store at -20°C for up to 6 months.

Once thawed, do not re-freeze and re-use AOM.

AOM Injection of Rats

Injection conditions

Time: Perform injections between 1 and 3 pm

Light Cycle: Alternating 12 h light/dark.

Concentration: 15 mg/kg body weight.

Injection volume: 700 – 1000 ul depending on weight of rat.

Injection site: Subcutaneously on right flank.

Needle: 27 G 1/2" needle, luer-lock syringe

AOM: Concentration 5mg/ml, warmed to room temperature.

Calculation: (rat wt in g)(15 mg/kg)(ml/5 mg)= injection volume.

250 g rat = 0.75 ml injection volume

APPENDIX D

LARR ANIMAL HANDLING

LARR Animal Handling

- Bring with you: key card, room key, diet if small freezer needs restocking (transport diet in container with cold packs and lid) To keep in room: rolling cart, writing equipment and paper, weigh sheet, rat ID list etc, scale, weighing container, plastic cup for diets.
- Wear protective equip inside facility.
- Go get cages from the clean cages area. (Rm 334)
- Sign up at room door.
- Door should be kept closed if other people are opening cages in the room.
- Be quiet/ soothing tones.
- Remove diet from freezer (**only handle 1 diet at a time**)
- Weigh out food into clean containers (40 g)
- Remove old bowl/ add new bowl, check water bottle.
- Return cage to original position, check that lid is all the way in
- Check bottle isn't leaking
- Replace diet in freezer (make sure bag is airtight)
- Wipe old bowls to remove food, place in large containers with lids, (must have lids),
- Wipe all room surfaces used
- Recheck corresponding diet to treatment.
- Recheck cages have the right diet and lids are secured and water bottles not leaking
- Take dirty bowls (In covered containers) to cage washing area in other side of building (don't ever go to clean cages with dirty bowls)
- If you need to re enter cubicles with animals and open cages after going to dirty cage area change booties.
- On weekends bowls are not washed, therefore on Friday feeding must be done early enough to have bowls washed (enough bowls for 2 days feeding)
- Remember all dirty bowls must be turned to dirty cage area on Sunday so they can be washed by Monday feeding

Handling Tricks

- Note: one trick is to eat a mint before interacting with animals, it has a soothing effect on them and it adds another control to their interaction/exposure with handler.
- Be relaxed, use decisive movements
- Allot enough time for non rushed animal handling
- Petting makes rats less stressed (reduces variables in physical response to diet, and makes it easier to weigh them)

APPENDIX E

LIPID ANALYSIS OF DIETS USING GC

Lipid analysis of diets using GC

Purpose: To examine total lipid of diets to confirm the mixing process.

Reagents:

Dry Reagents	Wet Reagents
25 mL screw cap glass tube	Acetone
12 mL leak-proof screw cap glass tube	1X PBS
4 mL Teflon-coated screw cap glass vial	MeOH – gas chromatography grade
Glass transfer pipette	Folch reagent (CHCl ₃ :MeOH 2:1 v/v)
Rubber bulb (for pipette)	Cold 0.1 M KCl
Disposable Pasteur pipette (located under mail slots)	6% HCl-MeOH
	Hexane
	Dichloromethane (CH ₂ Cl ₂)

Procedures:

Preparation of Leak Proof Tubes

1. Put 1 mL of acetone in 12 mL glass tubes with black caps. Screw tightly. Make ~40 tubes and half may end up being airtight.
2. Draw a line at the meniscus of the acetone.
3. Evaporate the tubes for 1 h at 80°C in the oven.
4. Check to see if acetone has evaporated. Only keep the tubes that are leak-proof. Do not mix up the caps as this could affect whether the tubes are leak-proof or not.
5. Discard the acetone in the leak-proof tubes and wash the tubes with methanol by squirting in methanol using a squirt bottle and discarding the methanol. Leave the lids off in the fume hood to let the methanol/residual acetone evaporate.

Extraction of Total Lipids

1. Obtain 10 to 100 mg of diet per sample in a 12 mL glass tube. These glass tubes are not leak-proof.
2. Perform subsequent steps on ice.
3. Resuspend the diet in 1 mL of cold 0.1 M KCl and 5 mL of Folch reagent. Vortex for 1 min. (babysit vortexer)
4. Centrifuge at 3000 rpm for 5 min at 4°C. Make sure the glass tubes are inside rubber tube holder to prevent tube from breaking. During this time clean N-vac needles with MeOH

Transfer the lower phase to the leak-proof 12 mL glass vial. Apply positive pressure on the way down to avoid any upper phase going into tube and exclude 2 drops after take pipette out of both layers. Dry under N₂. (give ~10 psi for 30-60 min, check on them every 15 min)

5. Dissolve the dried lipid with 3 mL of 6% HCl-MeOH. Flush the sample with N₂ before closing the cap tightly. (~15 s just enough to enter liquid) Wipe needles with MeOH to clean when finished.

6. Vortex 1 min as before and incubate at 76°C in the oven overnight (approximately 15 hrs).

7. Prepare 4 mL Teflon-coated screw cap (green cap) glass vials.

8. Turn on the GC overnight to allow for stabilization.

Extraction of FAME

1. Extract FAME by adding 1 mL of 0.1 M KCl and 2 mL of hexane to the methylated sample. Vortex for 1 min.

2. Centrifuge at 3000 rpm for 5 min at 4 °C.

3. Transfer the upper layer to 4 mL Teflon-coated screw cap (green cap) glass vials.

4. Dry FAME under N₂. (Wipe needles with MeOH prior to drying)

5. Redissolve the sample in 100 µL of CH₂Cl₂.

6. Vortex the Teflon-coated screw cap glass vials to dissolve the FAME.

7. Transfer 25 µL of sample into GC vial. Flush the remaining sample in the 4 mL Teflon-coated screw cap glass vial under N₂ and store at -20°C for future analysis.

8. Set up the GC vials for GC analysis. 2 vials labeled A1 and WB in the machine need to be filled with CH₂Cl₂ prior to run.

Notes:

- Be careful of cross-contamination.
- Organic solvents must be handled in the fume hood.

References:

- Folch J. *et al.* (1957) A simple method for the isolation and purification of total lipids from animal tissues. *J. Biol. Chem.* **226**: 497-509.

APPENDIX F

ACF BY RAT ID

Rat ID
info

experimental diet start date: set 1 - 4/19/2011 set 2 -
4/26/2011

HM ACF (>2)

							1	2	3	4	5	6			
							ACF	ACF	ACF	ACF	ACF	ACF	total	avg	st dev
		ID	diet	injection	set #	ID									
CCA	KT	1101	corn oil + cellulose	AOM	1	1101	34	35	27	17	9	4	57	60.45	22.17
CCA	KT	1102	corn oil + cellulose	AOM	1	1102	15	25	14	13	6	6	39		
CCA	KT	1103	corn oil + cellulose	AOM	1	1103	15	17	20	11	8	0	39		
CCA	KT	1104	corn oil + cellulose	AOM	1	1104	23	30	29	21	9	8	67		
CCA	KT	1105	corn oil + cellulose	AOM	1	1105	38	54	36	29	13	3	81		
CCA	KT	1106	corn oil + cellulose	AOM	1	1106	14	25	27	17	12	6	62		
CCA	KT	2107	corn oil + cellulose	AOM	2	2107	16	15	14	11	7	1	33		
CCA	KT	2108	corn oil + cellulose	AOM	2	2108	14	24	22	13	4	9	48		
CCA	KT	2109	corn oil + cellulose	AOM	2	2109	42	38	30	8	8	4	50		
CCA	KT	2110	corn oil + cellulose	AOM	2	2110	30	37	35	37	18	14	104		
CCA	KT	2111	corn oil + cellulose	AOM	2	2111	58	40	34	27	11	13	85		
CCS	KT	1501	corn oil + cellulose	Saline	1	1501									
CCS	KT	1502	corn oil + cellulose	Saline	1	1502									
CCS	KT	1503	corn oil + cellulose	Saline	1	1503									
CCS	KT	2504	corn oil + cellulose	Saline	2	2504	2								
CCS	KT	2505	corn oil + cellulose	Saline	2	2505									
CCS	KT	2506	corn oil + cellulose	Saline	2	2506									

ACF by Rat ID Cont

experimental diet start date: set 1 - 4/19/2011 set 2 - 4/26/2011

							1	2	3	4	5	6			
							ACF	ACF	ACF	ACF	ACF	ACF	total	avg	st dev
CPA	KT	1201	corn oil + pectin	AOM	1	1201	34	39	37	15	7	3	62	53.5	21.49
CPA	KT	1202	corn oil + pectin	AOM	1	1202	23	9	23	6	4	3	36		
CPA	KT	1203	corn oil + pectin	AOM	1	1203	22	29	29	16	11	6	62		
CPA	KT	1204	corn oil + pectin	AOM	1	1204	19	21	13	12	2	0	27		
CPA	KT	1205	corn oil + pectin	AOM	1	1205	15	13	20	10	6	9	45		
CPA	KT	2206	corn oil + pectin	AOM	2	2206	24	25	17	18	4	1	40		
CPA	KT	2207	corn oil + pectin	AOM	2	2207	16	31	22	16	5	3	46		
CPA	KT	2208	corn oil + pectin	AOM	2	2208	48	58	42	20	22	10	94		
CPA	KT	2209	corn oil + pectin	AOM	2	2209	24	27	23	12	3	2	40		
CPA	KT	2210	corn oil + pectin	AOM	2	2210	14	31	28	18	23	14	83		
CPA	KT	2211	corn oil + pectin	AOM	2	2211									
CPS	KT	1601	corn oil + pectin	Saline	1	1601									
CPS	KT	1602	corn oil + pectin	Saline	1	1602									
CPS	KT	1603	corn oil + pectin	Saline	1	1603	2	1							
CPS	KT	2604	corn oil + pectin	Saline	2	2604									
CPS	KT	2605	corn oil + pectin	Saline	2	2605	1	1							
CPS	KT	2606	corn oil + pectin	Saline	2	2606									

ACF by Rat ID Cont

experimental diet start date: set 1 - 4/19/2011 set 2 - 4/26/2011

							1 ACF	2 ACF	3 ACF	4 ACF	5 ACF	6 ACF	total	avg	st dev
FCA	KT	1301	fish oil + cellulose	AOM	1	1301	5	11	8	1	4	3	16	51.3	36.71
FCA	KT	1302	fish oil + cellulose	AOM	1	1302	33	16	19	10	2	2	33	40.66	15.61
FCA	KT	1303	fish oil + cellulose	AOM	1	1303	40	36	25	13	12	5	55		
FCA	KT	1304	fish oil + cellulose	AOM	1	1304	31	33	9	6	4	2	21		
FCA	KT	1305	fish oil + cellulose	AOM	1	1305	29	29	19	17	11	14	61		
FCA	KT	2306	fish oil + cellulose	AOM	2	2306	23	24	25	17	11	2	55		
FCA	KT	2307	fish oil + cellulose	AOM	2	2307	32	26	17	19	8	2	46		
FCA	KT	2308	fish oil + cellulose	AOM	2	2308	41	47	39	31	32	45	147		
FCA	KT	2309	fish oil + cellulose	AOM	2	2309	bad sample								
FCA	KT	2310	fish oil + cellulose	AOM	2	2310	27	27	19	19	5	1	44		
FCA	KT	2311	fish oil + cellulose	AOM	2	2311	38	32	17	16	2	0	35		
FCS	KT	1701	fish oil + cellulose	Saline	1	1701									
FCS	KT	1702	fish oil + cellulose	Saline	1	1702									
FCS	KT	1703	fish oil + cellulose	Saline	1	1703									
FCS	KT	2704	fish oil + cellulose	Saline	2	2704									
FCS	KT	2705	fish oil + cellulose	Saline	2	2705									
FCS	KT	2706	fish oil + cellulose	Saline	2	2706	2								

ACF by Rat ID Cont

experimental diet start date: set 1 - 4/19/2011 set 2 - 4/26/2011

							1 ACF	2 ACF	3 ACF	4 ACF	5 ACF	6 ACF	total	avg	st dev
FPA	KT	1401	fish oil + pectin	AOM	1	1401	15	18	17	3	2	1	23	26.63	11.47
FPA	KT	1402	fish oil + pectin	AOM	1	1402	18	21	12	5	2	2	21		
FPA	KT	1403	fish oil + pectin	AOM	1	1403	13	15	9	5	2	0	16		
FPA	KT	1404	fish oil + pectin	AOM	1	1404	18	8	12	8	4	4	28		
FPA	KT	1405	fish oil + pectin	AOM	1	1405	22	16	9	8	3	1	21		
FPA	KT	1406	fish oil + pectin	AOM	1	1406	18	25	22	6	6	2	36		
FPA	KT	2407	fish oil + pectin	AOM	2	2407	15	10	5	4	4	0	13		
FPA	KT	2408	fish oil + pectin	AOM	2	2408	20	16	16	13	10	13	52		
FPA	KT	2409	fish oil + pectin	AOM	2	2409	26	20	18	5	4	2	29		
FPA	KT	2410	fish oil + pectin	AOM	2	2410	5	6	11	4	1	1	17		
FPA	KT	2411	fish oil + pectin	AOM	2	2411	24	21	18	7	6	6	37		
FPS	KT	1801	fish oil + pectin	Saline	1	1801									
FPS	KT	1802	fish oil + pectin	Saline	1	1802	1	2							
FPS	KT	1803	fish oil + pectin	Saline	1	1803									
FPS	KT	2804	fish oil + pectin	Saline	2	2804									
FPS	KT	2805	fish oil + pectin	Saline	2	2805									
FPS	KT	2806	fish oil + pectin	Saline	2	2806	0								

APPENDIX G

NUCLEI ISOLATION FROM COLON CRYPTS

Protocol: Nuclei Isolation

Purpose: Isolate nuclei to detect histone modifications and antibody specificity (Chromatin Immunoprecipitation (ChIP) Ab Validation with Western Blot)

Materials

Silicon scrapers (BD Falcon, 353089)
50 mL and 15 mL conical tubes, 1.5 mL epi-tubes
Ice filled bucket
Micropipettes

Solutions

Make stock first, just add PI freshly (details below)

TBS (Tris-Biffered Saline)

Lysis Buffer 1 Add just before use, filter (0.2um) and keep cold. Consists of 50 mM HEPES-KOH with pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40, 0.25% Triton X-100, 1× protease inhibitors

Lysis Buffer 2 Add protease inhibitors just before use, filter and keep cold. Consists of 10 mM Tris-HCl, pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 1× protease inhibitors

Lysis Buffer 3 Add protease inhibitors just before use, filter and keep cold. Consists of 10 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na-Deoxycholate, 0.5% N-lauroylsarcosine, 1× protease inhibitors

Proteinase inhibitors

- 100X Protease Inhibitor Cocktail (Sigma P8340)
 - Sodium-Butyrate (Na-But) must be made fresh , 22 mg/0.1ml gives 2 M solution in MilliQ water add 10ul/ml LB (final 20 mM)
 - AEBSF concentration of 12 mg/ml in MeOH add 3 ul per ml buffer
- Add Protease inhibitors (PI) to all Lysis buffers and last TBS wash solution

Methods

pre-cool centrifuge to 4°C.

- Wash cells (175 ml flask) with cold 1X TBS 25 ml twice
Scrape adherent cells from culture dish to 50 mL conical tube on ice
- Rinse the adherent cells culture dish with (20 ml) TBS, add the remaining cells to the 50 mL tube and mix (basically use TBS to pool all cells into tube.

- Centrifuge cells at 1500 x g for 5 min at 4°C, discard supernatant
- Resuspend pellet in 1 ml of TBS per 10⁷ cells (which is ~ 2X the number of cells (YAMC?) in a confluent 175 ml flask) with gentle inversion (this is an estimation, no need to count)
- Spin at 1500g for 5 min at 4°C and aspirate the supernatant
- Cells can be used immediately or snap frozen straight in the tube in liquid nitrogen and stored in -80°C freezer indefinitely
- Remove frozen cell pellets from -80 °C and resuspend each pellet of ~10⁸ cells in 10 ml of Lysis Buffer 1. [10⁷ cells in 1 ml] Rotate at 4°C in rotator for 10 min.
- Spin at 1,350g for 5 min at 4 °C in a table-top centrifuge Discard supernatant.
- Resuspend each pellet in 10 ml of Lysis Buffer 2 [10⁷ cells in 1 ml]. Rotate at room temp in rotator for 10 min
- Pellet nuclei in table-top centrifuge by spinning at 1,350g for 5 min at 4°C. Discard supernatant.
- Resuspend each pellet in each tube in 1ml Lysis Buffer 3 per 10⁷ cells. Yields ~1ng nuclei protein per ul.
- *Store at -20°C until ready to use*

References:

Abcam Histone Extraction Protocol www.abcam.com

Chromatin immunoprecipitation and microarray-based analysis of protein location.

Tong Ihn Lee, Richard Young Nature Protocols Vol 1 No 2. 2006, PMID: 17406303

Solutions

Lysis Buffer 1	stock	50 ml
50 mM HEPES-KOH pH 7.5	1M HEPES-KOH pH 7.5	2.5
140 mM NaCl	5M NaCl	1.4
1 mM EDTA	500 mM EDTA	0.1
10% glycerol	50% glycerol	5
0.5% NP-40	10% NP-40	0.25
0.25% Triton X-100	10% Triton X-100	0.125
Lysis Buffer 2	stock	50 ml
10 mM Tris-HCl, pH 8.0,	1M Tris-HCl, pH 8.0	0.5
200 mM NaCl	5M NaCl	2
1 mM EDTA	500 mM EDTA	0.1
0.5 mM EGTA	500 mM EGTA	0.05

Lysis Buffer 3	stock	50 ml
10 mM Tris-HCl, pH 8.0	1M Tris-HCl, pH 8.0	0.5
100 mM NaCl	5M NaCl	1
1 mM EDTA	500 mM EDTA	0.1
0.5 mM EGTA	500 mM EGTA	0.05
0.1% Na-Deoxycholate	5% Na-Deoxycholate	0.05
0.5% <i>N</i> -lauroylsarcosine	5% <i>N</i> -lauroylsarcosine	0.25
TBS	stock	50 ml
150 mM NaCl	5 M NaCl	1.5
10 mM Tris pH8.0	1 M Tris-HCl, pH 8.0	0.5

Invitrogen/Gibco; Cat. no.: 15630-080)	1 M HEPES-KOH, pKa 7.55
(Sigma S-5150)	5 M NaCl
(Invitrogen/Gibco; Cat. no: 15575-038)	500 mM EDTA
(Sigma, Cat no: E3889)	500 mM EGTA
Sigma) B5887	Sodium butyrate Na-But
(Sigma; Cat. no.: G5516	50% glycerol
(Sigma; Cat. no.: I8896)	10% Igepal CA-360
Sigma; Cat. no.: T8787)	10% Triton X-100
Invitrogen/Gibco; Cat. no.: 15568-025	1 M Tris-HCl, pH 8.0
(Sigma; Cat. no.: D6750)	5% Na-Deoxycholate
Sigma; Cat. no.: L7026)	5 M LiCl
Invitrogen/Gibco; Cat. no.: L4390	10% SDS
(Sigma) P8340	100X Protease Inhibitor Cocktail
Sigma A8456	ABSF
Sigma 61743	N-Lauroylsarcosine

If you need to make your own then Aprotinin inhibits serine proteases; Leupeptin inhibits serine and thiol proteases; Pepstatin inhibits aspartic proteases. Make **1 mg/ml** each in MilliQ water, aliquot and store at -20°C. It is also possible to use commercially

*Protease inhibitors (PI) to all Lysis buffers and last TBS wash solution

-Aprotinin, Leupeptin and Pepstatin concentration of **1 µg/ml** add 40 ul per ml buffer ,

- AEBSF concentration of 12 mg/ml in MeOH add 3 ul per ml buffer

APPENDIX H

REAL TIME QPCR PROTOCOL FOR CHIP-QPCR

Before getting started:

Check availability of primers and primer concentrations (prepare primer mix if necessary), availability of samples and all reagents. Label tubes and draw a qPCR plate loading plan.

Preparation:

- Prepare enough mix for 2 reactions of each DNA sample, a standard curve (at least 3 dilutions), plus 2 reactions for no template control (NTC).
- For a whole 96-well optical plate, prepare reaction mix for 100 reactions in total (allow 4 extra reactions for pipetting errors).
- The reaction mix contains one set of primers (either Enriched or End); each set of primers contains both forward and reverse primers.
- Design plate layouts.
- Calculate the amount of reaction mix needed (See table below for calculations):

	Enriched primers Volume (μ L) per well	End primers Volume (μ L) per well	48 for wells each set, μ L (example)
2X SYBR master mix (ABI 4309155)	7.5	7.5	375
4 μ M (10X) primer mix	1.5	1.5	75
ddH ₂ O	3	3	150
Total reaction mix	12	12	600
DNA template (range .2-100ng)	3	3	Add 3 μ L directly to each well
Total well volume	15	15	

Procedures:

1. Load plate with reaction mix (adding DNA last): 12 μ L reaction mix per well + 3 μ L template. (Watch out for bubbles and check the volume of each well.)
2. Place the optical adhesive cover over the plate with caution (Do not to fingerprint the cover!)
3. Smooth with seal applicator; make sure all wells are securely sealed to avoid loss of sample during PCR.
4. Centrifuge briefly (1000k, 20sec, room temp. in centrifuge Jouan CR 412). Make sure there are no bubbles in any wells (will distort readings).

Note: Leave the centrifuged plate in a dark place before setting up Bio-Rad machine. Be careful not to disturb the content when taking out the plate from centrifuge.

Bio-Rad Machine:


1. Assign in for Bio-Rad usage on calendar and turn on 7900HT Fast Real Time PCR Machine.
2. Open SDS 2.2.4 software on Desktop. (Login: Username=administrator; no Password needed; click OK)
3. Create New Document under File tab. Set Assay: Standard Curve (AQ)/ Container: 96 well clear plate/ Template: Blank Template. Click OK
4. On the right side of the screen: In the Instrument tab, under *Thermal Cycler Protocol* choose Standard Mode and adjust Sample Volume to 15ul, set the *Thermal Profile* as described in the chart below.

<i>Parameters</i>	
Stage 1	1 cycle
95°C	10min
Stage 2	35 cycles (set <u>repeats</u> to 35)
95°C	15sec
53°C (optimal annealing Temp is primer specific)	30sec
Stage 3 (dissociation curve)	1 cycle
95°C	15sec
53°C (optimal Temp)	15sec
95°C	15sec

5. In Setup tab select Add Detector to specify primers and assign them to the corresponding wells in the plate.
6. Select the wells on the left side of the screen and add sample names accordingly.
7. Save the *.sds. file with your primer abbreviations, your initial and date in an appropriate folder.
8. Click Instrument - Connect and Open the machine's plate tray. Insert the plate into machine, with A1 well positioned at the upper left corner, then Close.
9. Go through all the setups before hitting the start bottom. Select Start and check back in 5-10min to make sure the Run has started properly.

Analysis for ChIP-qPCR:

- **In your saved sds. file :**

1. Press the Analyze bottom, and  then go to Analysis – Analysis Settings. Under the *Detector tap*, choose Manual Ct, click OK.

Under **Results** tap, select the standard wells of each primer and compare the amplification plots to determine if the lines are separated according to dilution increments. (i.e. a 10 fold dilution is about $2^{3.3}$ fold difference.)

2. *To adjust **Standard Ct in the same primer sets***: first pick the right detector; check if the green line is positioned at the steepest slope of the plot lines. (If not, move the green line towards the steepest slope, around the 24-34 Ct range limit and where the lines of duplicated samples are closest to one another.)
3. *To equalize **Standard Ct values for the sample gene in different primer sets***: first pick only one detector (not all detectors) to show the green line and adjust its Ct towards the other detector's Ct (usually, pick the less organized one to match the more organized one; remember to watch out for Ct range limit.)
4. Try to match the Ct values of standard with the most similar concentration to experimental samples.
5. **Save as *.txt** and **Export** to your USB. Re-open as Excel and work on calculation. Save both the original data and final data.

• **On Excel Spreadsheet:**

1. Re-organize columns and delete the less important columns. (See example below)
2. Calculate the values of *Ct primerA1-Ct primerA2* and *Ct mean primerA1-Ct mean primerA2*
3. Check and make sure the formulas and values are right after calculation.

	A	B	C	D	E	F	G	H	I
1	Filename	141022_ChIP_16ac_EP_e_C_M							
2									
3			CSNK2B 4k	CSNK2B 4k	CSNK2B 178	CSNK2B 178	Ct(CSNK2B 4k)-Ct(CSNK2B 178)		
4		Sample Name	Ct	Ct Mean	Ct	Ct Mean	Ct	Ct Mean	
5		16ac Libr(1/10)	25.78	25.76	25.04	24.90	0.74	0.86	
6		16ac Libr(1/10)	25.74	25.76	24.77	24.90	0.97	0.86	
7		16ac Libr(1/100)	28.39	28.53	28.25	28.29	0.14	0.23	
8		16ac Libr(1/100)	28.67	28.53	28.34	28.29	0.33	0.23	
9		16ac_cca_2106	30.52	30.64	30.11	29.81	0.41	0.83	
10		16ac_cca_2106	30.76	30.64	29.50	29.81	1.26	0.83	
11		16ac_fpa_1402	25.49	25.39	26.65	26.67	-1.16	-1.29	
12		16ac_fpa_1402	25.29	25.39	26.70	26.67	-1.41	-1.29	
13		EB	34.40	33.98	33.50	35.52	0.90	-1.54	
14		EB	33.56	33.98	37.55	35.52	-3.98	-1.54	
15		i(1/10)	27.16	27.49	25.79	25.74	1.37	1.75	
16		i(1/10)	27.82	27.49	25.68	25.74	2.13	1.75	
17		i(1/100)	28.22	28.17	28.22	28.19	0.00	-0.02	
18		i(1/100)	28.11	28.17	28.15	28.19	-0.04	-0.02	
19		i(1/1000)	32.18	31.52	31.48	31.76	0.70	-0.24	
20		i(1/1000)	30.85	31.52	32.03	31.76	-1.18	-0.24	

Analysis Summary:

$$dCt (\text{sample}) = Ct [\text{IP sample-control primer}] - Ct [\text{IP sample}]$$

fold ChIP enrichment = 2^{dCt} = level of DNA protein bound to gene site in ChIP sample compared to control sample

$$\text{ddCt} = \text{dCt (sample 1)} - \text{dCt (sample 2)}$$

$$2^{\text{ddCt}} = \text{fold induction of gene in sample 2 compared to sample 1.}$$

Notes for Reagents:

- Master Mix: SYBR green I qPCR mixture from ABI 4309155 (this mixture includes dNTP, Taq DNA polymerase, reaction buffer and the fluorescent dyes)
- Primers: custom-made gene specific primers from Integrated DNA Technology dissolved in ddH2O
- Plastic: white 96-well qPCR plates (MicroAmp® Optical 96-well Reaction Plate, Life Technologies, PN N801-0560) with optical clear seal sheets (MicroAmp® Optical Adhesive Film Kit, Life Technologies, PN 4313663) and press applicator
- DNA template: range .2-100ng as measured by picoGREEN quantification, if measured by A260 the concentration given is often higher

References:

Applied Biosystems Protocol for SYBR® Green PCR Master Mix and SYBR® Green RT-PCR Reagents Kit

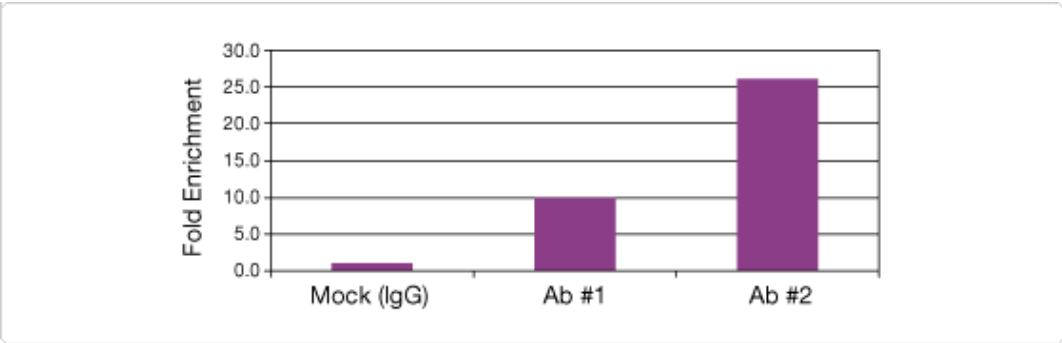
http://www3.appliedbiosystems.com/cms/groups/mcb_support/documents/generaldocuments/cms_041053.pdf

ChIP-qPCR Fold Enrichment Method

This normalization method is also called 'signal over background' or 'relative to the no-antibody control'. With this method, the ChIP signals are divided by the no-antibody signals, representing the ChIP signal as the fold increase in signal relative to the background signal. The assumption of this method is that the level of background signal is reproducible between different primer sets, samples, and replicate experiments. Background signal levels do vary between primer sets, samples, and experiments. An example is shown below.

To calculate fold enrichment:

Example	Raw Ct	Step 1	Step 2
		Nonspecific adjustment (Ct IP) - (Ct mock)	Fold enrichment (2-DDCt)
Mock (IgG)	34.6	0	1.0
Antibody #1	31.3	-3.3	9.8
Antibody #2	29.9	-4.7	26.0



APPENDIX I

BCA PROTEIN QUANTIFICATION ASSAY

BCA Protein Quantification Assay Pierce (23225)

A. Preparation of Diluted Albumin (BSA) Standards

std	Volume of Diluent (μL)	Volume BSA (μL)	Source of BSA	Final BSA Concentration (μg/mL)
A	0		Stock	2000
B	40	120	Stock	1500
C	100	100	Stock	1000
D	100	100	of vial B dilution	750
E	100	100	of vial C dilution	500
F	100	100	of vial E dilution	250
G	100	100	of vial F dilution	125
H	400	100	of vial G dilution	25
X	400		I 400 0 0 = Blank	

B. Preparation of the BCA Working Reagent (WR)

For Working Reagent use a 50:1 ratio of Reagent A: Reagent B. Use the following formula to determine the total volume of WR required:

$(\# \text{ standards} + \# \text{ unknowns}) \times (\# \text{ replicates}) \times (\text{volume of WR per sample}) + 2 \text{ extra} = \text{total volume WR required}$

Example: for the standard test-tube procedure with 3 unknowns and 2 replicates of each sample:

$(9 \text{ standards} + 3 \text{ unknowns}) \times (2 \text{ replicates}) \times (0.2 \text{ mL}) + 2 = 5 \text{ mL WR required (that's 50 ml Reagent A + 1 ml Reagent B)}$

C. Microplate Procedure (Sample to WR ratio = 1:8)

1. Pipette 25 μL of each standard or unknown sample replicate into a microplate well (working range = 20-2000 $\mu\text{g/mL}$)

(e.g., Costar 96-Well Plates, Product No. 9017).

Note: If sample size is limited, 5 μL of each unknown sample and standard can be used (sample to WR ratio = 1:40).

However, the working range of the assay in this case will be limited to 250-3000 $\mu\text{g/mL}$.

2. Add 200 μL of the WR to each well and mix plate thoroughly on a plate shaker for 30 seconds.

3. Cover plate and incubate at 37°C for 10-30 minutes.

4. Cool plate to RT. *Measure the absorbance at or near 562nm on a plate reader. (See protein measurement instructions in Section D)*

5. Subtract the average 562 nm absorbance measurement of the Blank standard replicates from the 562 nm measurements of all other individual standard and unknown sample replicates.

6. Prepare a standard curve by plotting the average Blank-corrected 562 nm measurement for each BSA standard vs. its concentration in $\mu\text{g/mL}$. Use the standard curve to determine the protein concentration of each unknown sample.

D. Protein Measurement using SpetraMax

1. Software set ups:

- Go to *SoftMax Pro5.2* on computer desk top.
- Click on “?” icon, select reader-*SepctraMax 190* and connection-*COM4*, click *OK*.
- Click on $\leftarrow \rightarrow$ to Open/Close the loading deck to check if *COM4* is connected.
- In the *Protocols* tap, select Protein Quant-->*BCA*-->*Template*; enter sample identities to wells. Click *OK*.
- Save your set ups in an appropriate place in *People's folder* on desktop.

2. Measurement:

- Click on $\leftarrow \rightarrow$ to Open the loading deck and load your plate.
- Hold the “*Shake*” icon for 10 seconds.
- Click *Read* to start sample reading.
- Export data after the reading is done.

Notes:

For more detailed information please consult product manual Pierce™ BCA Protein Assay Kit.

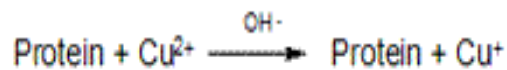
- Because plate readers use a shorter light path length than cuvette spectrophotometers, the Microplate Procedure requires a greater sample to WR ratio to obtain the same sensitivity as the standard Test Tube Procedure. If higher 562nm measurements are desired, increase the incubation time to 2 hours.
- Increasing the incubation time or ratio of sample volume to WR increases the net

562nm measurement for each well and lowers both the minimum detection level of the reagent and the working range of the assay. As long as all standards and unknowns are treated identically, such modifications may be useful.

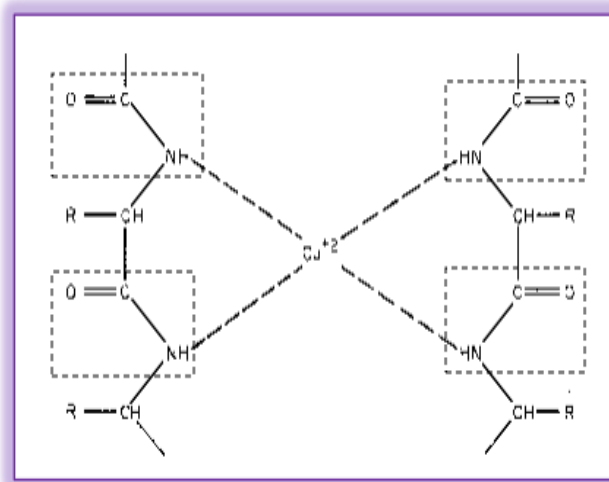
Note: If using curve-fitting algorithms associated with a microplate reader, a four-parameter (quadratic) or best-fit curve will provide more accurate results than a purely linear fit. If plotting results by hand, a point-to-point curve is preferable to a linear fit to the standard points.

- Wavelengths from 540-590nm have been used successfully with this method.
- Color change in sample wells (green to purple) results from reduction of Cu ion from 2+ to 1+. (reaction shown below)

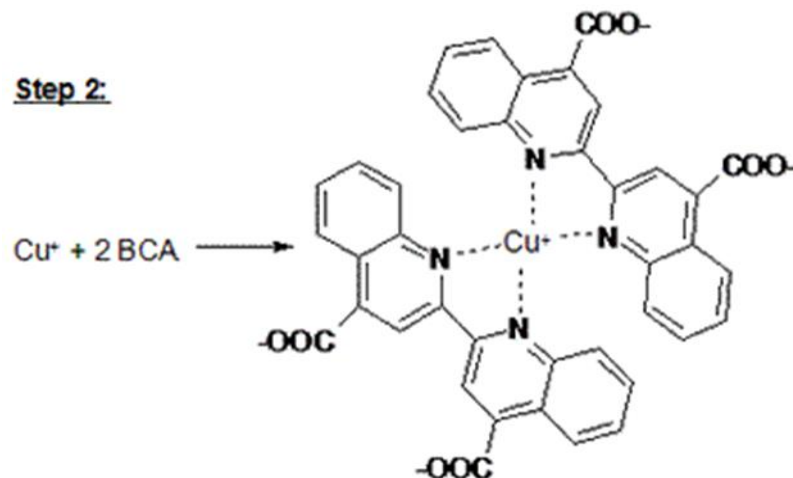
Step 1:



Cu^{2+} - peptide complex



Step 2:



APPENDIX J

QUBIT PROTOCOL FOR MEASURING DNA CONCENTRATION

Qubit® dsDNA HS (High Sensitivity) Assay Kit Protocol

Description

The Qubit® (previously known as Quant-iT™) dsDNA HS (High Sensitivity) Assay Kit is designed specifically for use with the Qubit® 2.0 Fluorometer (Q32866), but may be used with any fluorometer or fluorescence plate reader. The assay is highly selective for double-stranded DNA (dsDNA) over RNA and is designed to be accurate for initial sample concentrations from 10 pg/μl to 100 ng/μl. The kit provides concentrated assay reagent, dilution buffer, and pre-diluted DNA standards. Simply dilute the reagent using the buffer provided, add your sample (any volume between 1 μl and 20 μl is acceptable), and read the concentration using the Qubit® 2.0 Fluorometer. Common contaminants, such as salts, free nucleotides, solvents, detergents, or protein are well tolerated in the assay.

Notes:

1. All Qubit® assay kits can also be used with the Qubit® 1.0 Fluorometer.
2. 500 μL thin-walled PCR tubes (Q32856) are required but not included. (Qubit® assay tubes are 500 μl thin-walled polypropylene tubes for use with the Qubit® 2.0 Fluorometer (Q32866). 500 tubes per package.
3. Best way to fully understand the protocol is to read the official Qubit Manual <http://www.lifetechnologies.com/order/catalog/product/Q32854>
5. Accuracy is best when concentration of sample is between 1-10 ng/ul. This means if the concentration of the initial sample is higher, it should be diluted to fall under these parameters. (For example if Qubit reading yields 12 ng/ul in sample, a dilution of that sample, recommend a 1(sample):5(EB) dilution, should be performed, and concentration reassessed.

Tube Loading Guide

For any 'input DNA' samples

<u>sample</u>	Sample Volume (ul)	Working solution (ul)
Stock standard (10 ng/ul)	1	99
Input DNA A	1	99
Input DNA B	1	99

For ChIPed samples

sample	Sample Volume (ul)	Working solution (ul)
Stock standard (10 ng/ul)	5	95
ChIPed DNA A	5	95
ChIPed DNA B	5	95

If dilutions are necessary

sample	Sample Volume (ul)	Dilution Solution=EB (ul)
Input DNA A	2	8
Input DNA B	2	8

Full protocol

The protocol below assumes you will be preparing standards for calibrating the Qubit® 2.0

Fluorometer. If you plan to use the last calibration performed on the instrument, you will need fewer tubes (step 1.1) and less working solution (step 1.3). More detailed instructions on the use of the Qubit® 2.0 Fluorometer (corresponding to steps 1.9–1.15 and 2.1–2.6) can be found in the user manual accompanying the instrument.

1.1 Set up the number of 0.5 mL tubes you will need for standards and samples. The Qubit™ dsDNA HS assay requires 2 standards.

Note: Use only thin-wall, clear 0.5 mL PCR tubes. Acceptable tubes include Qubit™ assay tubes (500 tubes, Cat. no. Q32856) or Axygen PCR-05-C tubes (VWR, part no. 10011-830).

1.2 Label the tube lids.

1.3 Make the Qubit™ working solution by diluting the Qubit™ dsDNA HS reagent 1:100 in Qubit™ dsDNA HS buffer. Use a clean plastic tube each time you make Qubit™ working solution. Do not mix the working solution in a glass container. Note: The final volume in each tube must be 100 µL. Each standard tube will require 95 µL of Qubit™ working solution, and each sample tube will require anywhere from 90 µL to 99 µL. Prepare sufficient Qubit™ working solution to accommodate all standards and samples. For example, for 8 samples, prepare enough working solution for the samples and 2 standards: ~100 µL per tube in 10 tubes yields 1 mL of working solution (5 µL of Qubit™ reagent plus 995 µL of Qubit™ buffer). 1.4 Load 95 µL of Qubit™ working solution into each of the tubes used for standards.

1.5 Add 5 µL of each Qubit™ standard to the appropriate tube and mix by vortexing 2–3 seconds, being careful not to create bubbles. Note: Careful pipetting is critical to ensure that exactly 5 µL of each Qubit™ dsDNA HS standard is added to 95 µL of

Qubit™ working solution. It is also important to label the lid of each standard tube correctly as calibration of the Qubit® 2.0 Fluorometer requires that the standards be introduced to the instrument in the right order. 1.6 Load Qubit™ working solution into individual assay tubes so that the final volume in each tube after adding sample is 100 µL.

Note: See 'Tube Loading Guide' for ChIP specific sample concentration assessment. A more general rule of the protocol allows for sample volume used to be anywhere between 1 µL and 10 µL, therefore, load each assay tube with a volume of Qubit™ working solution anywhere between 90 µL and 99 µL.

1.7 Add each of your samples to assay tubes containing the correct volume of Qubit™ working solution (prepared in step 1.6) and mix by vortexing 2–3 seconds. The final volume in each tube should be 100 µL

1.8 Allow all tubes to incubate at room temperature for 2 minutes.

1.9 On the Home Screen of the Qubit® 2.0 Fluorometer, press DNA, and then select dsDNA High Sensitivity as the assay type. The Standards Screen is automatically displayed.

Note: If you have already performed a calibration for the selected assay, Qubit® 2.0 Fluorometer will prompt you to choose between reading new standards and using the previous calibration. See Calibrating the Qubit® 2.0 Fluorometer above for calibration guidelines.

1.10 On the Standards Screen, press Yes to run a new calibration or press No to use the last calibration.

1.11 If you pressed No on the Standards Screen, proceed to step 1.12. If you selected Yes to a run new calibration, follow instructions below.

Running a New Calibration

Insert the tube containing Standard #1 in the Qubit® 2.0 Fluorometer, close the lid, and press Read. The reading will take approximately 3 seconds.

Remove Standard #1.

Insert the tube containing Standard #2 in the Qubit® 2.0 Fluorometer, close the lid, and press

Read. Remove Standard #2.

1.12 If you pressed No on the Standards Screen, the Sample Screen will be automatically displayed. Insert a sample tube into the Qubit® 2.0 Fluorometer, close the lid, and press Read

1.13 Upon the completion of the measurement, the result will be displayed on the screen. Note: The value given by the Qubit® 2.0 Fluorometer at this stage corresponds to the concentration after your sample was diluted into the assay tube. You can record this value and perform the calculation yourself to find out the concentration of your original sample (see Calculating the Concentration of Your Sample, below) or the Qubit® 2.0 Fluorometer performs this calculation for you (see Dilution Calculator, next page).

1.14 To read the next sample, remove the sample from the Qubit® 2.0 Fluorometer, insert the next sample, and press Read Next Sample. 1.15 Repeat sample readings until all samples have been read

Calculating the Concentration of Your Sample

The Qubit® 2.0 Fluorometer gives values for the Qubit™ dsDNA HS assay in ng/mL. This value corresponds to the concentration after your sample was diluted into the assay tube. To calculate the concentration of your sample, use the following equation:

First determine dilution factor

Dilution Factor = QF value/10

(Rationale: Concentration of ChIP standard (10ng/ul) = QF value * dilution factor)

Second determine concentration of your sample

Sample Concentration= QF value / Dilution Factor

where:

QF value = the value given by the Qubit® 2.0 Fluorometer

Other calculation notes

By using the same number of ul for both the tubes for the ChIP standard and the tubes for the samples.

x = the number of microliters of sample you added to the assay tube

APPENDIX K

WESTERN BLOT PROTOCOL

Protocol: Western Blot

Purpose: Detect histone modification and antibody specificity (Chromatin Immunoprecipitation Ab Validation)

Protein volume and Antibodies for CHIP Western blot

- 1.20 ul of the nuclei sample in CHIP lysis Buffer 3 from each of the treatments are run on a 4-20% polyacrilamide gel (Invitrogen, EC6028BOX) and blotted onto a PVDF membrane (Immobilon, IPVH00010)
- 2.All Abs were polyclonal Rabbit and were used at 1:20000 dilution
- 1.List 1° Ab and dilution (H3ac Upstate 06-599, H3 Abcam ab1791, H3K4me3 Active Motif 39159)
- 2.List secondary antibody and dilution (2ry anti-rabbit Ab used at 1:15,000 dilution)KPL 047-1509

Preparation

- Thaw samples on ice
- Label 0.6 ml epi tubes for each sample (2 per sample to run in duplicate)
- Prepare the western template sheet
- Cut Immobilon membrane (**Immobilon, IPVH00010**) and filter papers (**Midsci, 6MW-2020**) ready

Materials:

-Micropipettes (200ul, 10ul) and tips, ice bucket, .6ml tubes, thermocycler or other 98°C heat source, pre-made gel (4-20%) 4°C, molecular weight marker (ladders) (**MagicMark,XP LC5603**) -20°C, 5X Pyronine room temp, gel running unit, gel transfer unit, electrophoresis power supply, stirrer, shaker, PVDF membrane, forceps, roller, filter paper Ab incubation containers, sponge/filterpaper/membrane soaking container, acetate sheet (**C-line Products, No.0010**), chemiluminescent super signal kit (**Pierce, 34095**) 4°C,

Sample Preparation

- 1.Thaw the samples on ice
- 2.Dye used for the sample dilution is 5X Pyronine. Use 1X of the dye based on the total volume of sample required (usually 25 ml total for 10 well).
- 3.Set the PCR machine (*heating source*) temperature to 98°C (takes about 1 min).
- 4.With the aid of western template sheet make the necessary dilution and add the calculated amount of dye and water to the samples and standard.
- 5.Give the samples a quick spin on the table top centrifuge.

6. Boil the samples for 5-10 min depending on the volume of the samples (25 ml volume boil for 10 min). *Do not boil the marker.*
7. Repeat the quick spin of the samples on the tabletop.

Gel unit set up

8. Take the pre-made gel (4-20%) and carefully rip off and discard the white tape and the comb. Mark the lanes on the plate
9. Attach the gel to the gel rack with the shorter side facing in, align 3rd with the lower gasket and clamp the unit. (Note that the red clip should have the broad end facing you, broad ends face outside on all 4 clips).
10. Pour running buffer into middle of chamber to fill the stand and the trough up to the top.
11. Use 10 ml tips and load the complete sample volume.
12. Close the unit with the lid and check the leads and make sure black-to-black and red-to-red.
13. In the cold room run the gel at 200V for 30 min. This is protein size dependent, bigger protein use 200V, smaller proteins use 150V (Big proteins >100 kDa)
14. Check after approximately 5 min to ensure the sample is moving.
15. After about 20 min check every 5 min to ensure the sample does not run off the gel.
16. Stop the gel when it has run 4/5ths of the way down the gel.

Gel transfer

17. Take the gel transfer unit and a staining trough and pour transfer buffer into both the tray and the unit.
18. Take a cassette and lay it open with black side down.
19. Put the thicker sponge on the black side of the cassette and place a filter paper on top of it.
20. Pour transfer buffer into trough above the level of the cassette and filter paper to keep it wet.
21. Take the membrane out and let it soak in methanol for 15 seconds. Soak in a tray with transfer buffer.
22. Take the gel plate out of the running unit and transfer the running buffer into the bottle for reuse.
23. Crack open the plate and keep the front of the gel down.
24. Cut the gel just above the bottom to remove the foot of the gel. Also cut off the wells at the top of the gel
25. Carefully separate the gel from the plate and cut the gel at lane one to identify the side (right end).
26. Place the gel on the filter paper with the front of the gel facing down.
27. Cut the right hand top corner of the membrane to identify the side.
28. Place the membrane on the gel and roll with the roller from bottom to top gently

to eliminate any air bubbles.

29. Wet another piece of filter paper and place on the membrane. Roll again with the roller.
30. Place the thinner sponge on top and close the white side of the cassette and clip it.
31. Place the cassette in the transfer unit with the hinges facing the top and black side facing back.
32. Put a stir bar into the transfer unit.



33. Make sure transfer buffer in unit covers the hinges of the cassette. Check the terminals black to black correspond.
34. Place it on the cold room stir plate.
35. Connect black-to-black and red-to-red and set current to 400 mAmps and let it transfer for 100 min. (this wasn't long enough for 3 gels to transfer proteins over ~100 kDa, for 3-4 gels run 120-150 min)

Blocking

36. At the end of transfer- make 25 ml TBS-T + 1g BSA (Roche, 03116956001) for BSA/TBS/Tween Mix gently by inverting.
37. Pour the milk into a dish and keep ready to transfer the membrane into it.
38. After the transfer is complete- open the gel unit and transfer the transfer buffer into the bottle.
39. Use a pair of forceps to take the membrane and place the membrane into the milk dish (with the side facing gel-protein side now facing up)
40. Place it on the shaker for 1-3 hr at room temperature.

Primary antibody

41. Discard blocking solution.
42. Now, add the appropriate volume of the primary antibody in new blocking solution. (based on the dilution and add it into the dish). H3, H3ac, IgG (abcam) H3K4me3 (active motif) all 1:20,000 in 4% BSA

43. Close the lid of the dish and shake it gently on the cold room shaker overnight.

Washing

44. The next day take the membrane and give a quick wash with TBS –Tween.
45. Transfer membrane to new dish with fresh TBS- Tween and keep on the shaker at room temp for 5 min. Let it shake vigorously.
46. Repeat the wash 2 more times for **5 min, then 10 min.**

Secondary antibody

47. Make 30 ml milk or BSA/TBS/Tween and pour into the dish after the second wash.
48. Add the required volume of secondary antibody based on the dilution. (If dilution is 1:10,000, so use 3 μ l secondary antibody in 30 ml of milk or BSA)
49. Set on shaker for 1 h at room temperature.
50. Repeat washing with TBS –Tween 3 times as described above
51. While the 1 wash of 2^o antibody is going on turn on the imager and set focus.

Developing

52. After the 3rd wash is complete
53. Cut an acetate sheet into 2 halves and remove the black sheet
54. Mix 0.3 ml of chemiluminescent super signal reagent A with 0.3 ml of reagent B in an eppendorf tube.
55. Mix gently by inversion.
56. Now transfer the membrane between the layers of the acetate sheet and squirt the developing solution on the top.
57. Slowly close the top layer so that the solution gets evenly distributed on the membrane.
58. Expose for 5 min and then transfer the membrane on to the clean acetate sheet
59. Transfer it into the Bio-Rad imager for imaging immediately.

Imaging

60. Turn on switch and make sure the lever on the hood is at chemiluminescence.
61. Select Quantity One program on desktop.
62. Select scanner – click on chemidoc.xrs.
63. Step 1- option is chemiluminescences.
64. Step 2 – live focus. Focus with a printed sheet and set the iris as you need for brightness. Zoom and focus, as you need for clarification.
65. Freeze. Put the gel in the imager and zoom and freeze again. Close the door.
66. Click on live acquire.
67. Starting exposure time – for Bcl2 imaging - use 30 sec.
Total exposure time – 180 sec

Number of exposures – 6

68. Save the images and analyze using the Quantity One software.

Recipes

Tris-buffered saline - Tween (TBS-T, 1X) store room temp

Dissolve the following in 800 ml of distilled H₂O

- 8.8 g of NaCl
 - 0.2 g of KCl
 - 3 g of Tris base
2. Add 500 ul of Tween-20
 3. Adjust the pH to 7.4
 4. Add distilled H₂O to 1L

For all other western blotting solutions use “Western Blotting solutions Yang-Yi Fan 11-21-97”

Lysis Buffer 3 (keep cold). Consists of 10 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na-Deoxycholate, 0.5% *N*-lauroylsarcosine

Notes:

Recommended controls to include

- Positive control lysate: protein lysate known to contain protein of interest
- Negative control : “blank” sample containing lysate without protein
- Loading control: If comparing different sources of protein, use a **“house-keeping protein” such as B-actin, or proteins that are expressed at equivalent levels in almost all tissues and cells.**
- Colorimetric and HRP ladder: the colorimetric ladder will allow visualization on how far multiple protein sizes have traveled on a running gel and whether the whole protein was transferred from gel to membrane. The HRP ladder will be visible during membrane developing and confirm the protein band size.
- If using an antibody for the first time alongside an antibody already sanctioned as high quality in the lab. This can be a loading control Ab.

References: Abcam Histone Western blotting www.abcam.com protocol section, and HT Western Blot protocol

Expected Results:

Use controls to make sure: positive control contains the appropriate size band (it may contain extra bands if it's a whole cell lysate), negative control contains no band, isolated nuclei sample should show only nuclear proteins and clear bands at expected size (for histones ~17 kDa) It is possible the sample will contain protein dimmers and multiple bands will be noticed (example ~17 kDa plus ~35 kDa)

If there are many bands present in a lane with no prevalent band of the expected size there might have been a problem with procedure or reagents. (Example: degraded antibody, incorrect washing, solutions not made properly)

APPENDIX L
CHIP PROTOCOL

ChIP Method

Day 0: Steps A (approximately 1–4h, depending on numbers of cells or tissues being collected)

Day 1: Steps B-C (approximately 8 h to multiple days,)

Day 1,2: Steps D (approximately 2 h and overnight incubation then 1 hr)

Day 2: Steps E-F (approximately 2 h and overnight incubation)

Day 3: Steps G-H (approximately 4 h and overnight incubation)

Day 4: Steps H-I (approximately 6 h)

Day 5: Steps J-K (approximately 1 h and 2 h)

See Crypt isolation protocol for extracting crypts

A. Crosslink crypts Timing: Day 0, 2hr

Pre-cool centrifuge to 4°C

A.1. Add 1/10 volume of fresh 11% Formaldehyde Solution to tube. Swirl briefly. (To 20 ml cell suspension add 2 ml)

A.2. Incubate cells with Formaldehyde Solution in rotation for 15 min at room temperature.

A.3. Add 1/10 volume of 3.5 M glycine to quench formaldehyde. (To 20 ml cell suspension add 2.2 ml)

A.4. Pool cells in required number of 50 ml conical tubes and spin at 1500g (300 rpm, setting 3 of Juoan centrifuge) for 5 min at 4 °C. Discard supernatant.

A.5. Resuspend cells in 30 ml of 1× TBS w PIs (see reagents for components) with gentle inversion or swirling (cells may stick to pipettes at this stage so don't use pipettes). Spin at 1500g for 5 min at 4 °C to pellet cells. Discard supernatant. Repeat once more. Discard supernatant.

A.6. Flash-freeze cells in liquid nitrogen and store pellets at –80°C.

Pause Point Once cells are crosslinked, they may be stored frozen at –80°C indefinitely.

B. Nuclei isolation Timing: Day 1, 2-3hr

Add PIs to Lysis Buffers right before use (PIs: use 10 ul/ml of 100X Sigma PI, 10 ul/ml of 2 M NaBut, 2 ul/ml 500 mM AEBSF)

B.1 Remove frozen cell pellets from -80°C and resuspend each pellet of distal or proximal crypts (~2x10⁷ cells in 5 ml of Lysis Buffer 1, [~10⁷ cells in 1 ml]) Rotate at 4 °C in rotator for 10 min.

B.2 Spin at 1,350g for 5 min at 4 °C in a table-top centrifuge. Discard supernatant.

Resuspend each pellet in 5 ml of Lysis Buffer 2 [10⁷ cells in 1 ml]. Rotate at room

temp in rotator for 10 min.

B.3 Pellet nuclei in table-top centrifuge by spinning at 1,350g for 5 min at 4°C. Discard supernatant.

Resuspend each pellet in each tube in 3 ml Lysis Buffer 3 (or 1 ml per 10⁷ cells).

Yields ~1 ng nuclei protein per ul or 30 ng DNA per ul.

Store at -80°C until ready to use (sonication)

C. Sonication Covaris (40 min per sample) **Timing: Day 2, 6 h per day**

Always keep samples on ice when not sonicating

Instrument:

-Covaris S2

- 3 ml Covaris TC16 tubes (16x100 mm) Part #520011

Settings:

-Duty Cycle 20% -Intensity 8 -Cycles per burst 200

-Processing time 40 min

C.1. Thaw samples on ice and transfer to labeled Covaris tubes (3 ml sonication volume) Screw on Covaris plastic cap.

C.2. Place in holder, and sonicate with settings indicated above.

C.3. Remove tube and return to ice.

Pause Point: Lysates can be frozen at -80°C and used at a later date.

C.4. Add 1/10th volume of 10% Triton (300 ul for 3 ml)

C.5. Split sample into 2 x 1.5 ml microcentrifuge labeled tubes

C.6. Spin down 4°C 1min at 14000g.

C.7. Remove supernatant and place in new tubes

C.8. Save 50 µl of cell lysate supernatant from each sample as reference DNA. Store at -20°C. At least one DNA aliquot should be kept per batch of sonicated lysate. Note that the DNA concentration, the effects of the sonication and the resulting distribution of fragment sizes can only be checked after crosslink reversal and purification of DNA.

Store rest of sample at -80°C until ready to use

C.9. Extract reference DNA from 50 ul aliquot by following Steps H to I (reverse crosslink and purify DNA) in order to determine how much lysate equals 10 ug chromatin.

Pause Point Lysates can be frozen at -80°C and used at a later date.

D. Preparing magnetic beads **Timing: Day 1 and 2, 2 h o/n (ChIP A)**

Critical step Steps keep all reagents, rack, and tubes on ice whenever possible

Day 1 Conjugate beads to antibody (Ab details below)

D.1. Aliquot 100 µl of Dynal beads to a 1.5 ml microcentrifuge tube. Set up 1 tube per 50 ul or 100 ul beads. 50 ul beads per 1 immunoprecipitation so 100 ul of beads is good for 2 immunoprecipitations with the same Ab. Add 1 ml ice cold Block Solution.

D.2. Collect beads using Dynal MPC. Place tubes in rack. Allow beads to collect on side of tube. This should take approximately 25 s. Invert rack once or twice to help collect all beads. Remove supernatant with aspirator or pipettor.

D.3. Add 1.5 ml Block Solution (see reagents) and gently resuspend beads. This can be done by removing the magnetic strip from the rack and inverting the rack, with the tubes still in place — either 10–20 times or until the beads are evenly resuspended. Collect beads as above (Step D.2.). Remove supernatant with aspirator or pipettor.

D.4. Wash beads in 1.5 ml Block Solution, as in Step D.3., one more time.

D.5. Resuspend beads in 250 μ l Block Solution and add 10 μ l or 10 μ g of antibody for each Ab (note: volume and concentration may vary per antibody). Incubate at 4°C overnight (for a minimum of 6 h) on a rotator.

Day 2 Wash beads just prior to starting ChIP

D.6. Wash beads three times in 1 ml Block Solution, as described in Step 3.

D.7. Resuspend each aliquot of beads in original volume (50 μ l or 100 μ l) Block Solution.

E. Chromatin Immunoprecipitation from Lysate **Timing: Day 2, 2 h and overnight incubation (ChIP B)**

E.1. Save 50 μ l of cell lysate from each sample as input DNA. (this is done just prior to immunoprecipitation, as a confirmatory measure of the DNA concentration previously obtained in C.9.)

E.2. Set up 1.5 ml microcentrifuge tubes on ice, one tube for each immunoprecipitation

E.3. Add 50 μ l antibody/magnetic bead mix from Step D.7 to 2.5 μ g DNA worth of cell lysates (supernatant) from Step C.8. Incubate overnight on rotator at 4°C.

F. Wash Beads **Timing: Day 3, 2hr (ChIP C)**

Critical step: All wash steps are performed at 4°C. Buffers should be kept cold.

F.1. Collect beads using Dynal MPC. Place tubes in rack. Allow beads to collect on side of tube. This should take approximately 25 s. Invert rack once or twice to help collect all beads. Remove supernatant with aspirator or pipettor, changing tips between samples.

F.2. Add 1 ml Wash Buffer (RIPA) to each tube and gently resuspend beads. This can be done by removing the tubes from the rack and inverting the tubes 10–20 times or by removing the magnetic strip from the rack and inverting the rack, with tubes still in place — 10–20 times or until the beads are evenly resuspended. Collect beads. Remove supernatant by aspirator or pipettor. Repeat this wash four more times, changing tips between washes.

F.3. Wash once with 1 ml TE + 50mM NaCl.

F.4. Spin at 960g for 3 min at 4 °C and remove any residual TE buffer.

Proceed directly to Elution

G. Elution **Timing: Day 3, 1.5 h (ChIP D)**

G.1. Add 110 µl of Elution Buffer and elute material from beads by incubating tubes in a 65°C heat block for 15 min with constant shaking (1400 rpm). Vortex briefly 30sec.. (Careful to keep tubes secure by holding the cap while vortexing, since the high temp mixed with vortexing tends to pop the cap open). Spin and place tubes in r.t. magnet. Transfer supernatant to new tube.

G.2. Repeat G.1

G.3. Spin down beads at 16,000g for 1 min at room temperature. Place tubes in r.t. magnet rack.

G.4. Remove 200µl of supernatant and transfer to new tube.

SOME ISSUES WITH K4 and ac P3

Pause Point Material can be frozen at -20°C and stored overnight.

H. Crosslink Reversal **Timing: Day 3 to 4, 1 h then overnight (ChIP D)**

H.1. Reverse crosslink the immunoprecipitated DNA from Step G.3 by incubating at 65°C for a minimum of 6hr, can be done in waterbath or heat block. Vortex and spin down at least once during incubation to avoid improper crosslinking due to condensation

H.2. Thaw 50 µl of input DNA reserved after Covaris sonication (Step E.1), add 150 µl (3 volumes) of elution buffer and mix. Reverse crosslink this input DNA by incubating at 65°C as in Step H.1. From this point, every tube of immunoprecipitation or input DNA is considered to be a separate tube or sample for later processing steps.

Pause Point. If the wash was performed early in the day, the crosslink reversal can be stopped after a minimum of 6 h and the material can be frozen at -80°C and stored overnight. If the wash was performed later in the day, the crosslink reversal can be extended up to 15 h and performed overnight with little effect on the reaction.

Critical note: Longer times of crosslink reversal (18 h or more) usually result in increased noise in DNA analysis

I. Purification of DNA **Timing: Day 4, 6 h (ChIP E)**

I.1. Add 200 µl of TE to each tube to dilute SDS in Elution Buffer. We have found that high levels of SDS can inhibit RNase activity.

I.2. Add 2 µl of 10 mg/ml RNaseA (0.05 mg/ml final concentration), mix by inverting the tube several times and incubate at 55°C for 1 h. *Pause Point Material can be frozen at -20°C and stored overnight.*

I.3. Add 4 µl of 20 mg/ml Proteinase K (0.2 µg/ml final concentration) and mix by inverting the tube several times and incubate at 55°C for 2 h.

I.4. Add 400 µl phenol:chloroform:isoamyl alcohol (P:C:IA), vortex 30 sec on high speed (careful to hold cap to avoid tube opening and separate phases by centrifuging in refrigerated microcentrifuge at 20000g for 15min 4°C.

Clean up DNA using the DNA PCR purification columns (Follow Qiagen PCR

purification kit instructions for washing and eluting)

Centrifuge Spins are all at room temp. 17900g 1 min in microcentrifuge.

I.5. Add 3x (1200 ul) Qiagen Buffer PB (PCR cleanup kit 28104) to each extracted supernatant and vortex

I.6. Add ~800 ul of the solution to a Qiagen PCR Cleanup column, centrifuge, and then rerun the flowthrough through the same column, then discard flow through.

I.7. Repeat with the rest of same sample to bind the total DNA from sample to a Qiagen column. (Run 800 ul, then the same 800 ul, then discard the flowthrough. Repeat step with the other 800 ul)

I.8. Wash the column with 750 ul Qiagen Buffer PE, centrifuge, empty, and centrifuge the column again, which contains the DNA, to dry.

I.9. Elute the DNA from the column with 50 ul aliquots of warmed (~55°C) Qiagen Buffer EB.

J. Measure DNA concentration with picogreen

K. Measure DNA size with Bioanalyzer DNA 7500.

References:

Abcam Histone Extraction Protocol www.abcam.com

Chromatin immunoprecipitation and microarray-based analysis of protein location.

Tong Ihn Lee, Richard Young Nature Protocols 729-748 Vol 1 No 2. 2006, PMID: 17406303

Follow with DNA size selection by gel extraction and NuGen library building for DNA sequencing

Ab details

Ab	Company	ul per 100 ul of beads	Amount chromatin/IP
H3	Abcam	10	1 ug
H3K9me3	Abcam ab8898	10	10 ug Prox, 5 ug Dist
H3K4me3	ActiveMotif 39159	10	2.5 ug
RNA Pol II	Ab5408	10	2.5 ug
IgG	Abcam ab46540	10	1 ug

APPENDIX M

CHIP REAGENTS AND EQUIPMENT

Note: All solutions are dissolved in ddH₂O and filtered with 0.2 µm syringe filter (Like Whatman) unless otherwise noted, and stored at room temp unless otherwise stated.

Reagents

- 1 M HEPES-KOH, pKa 7.55 (Invitrogen/Gibco; Cat. no.: 15630-080) 4°C
- 5 M NaCl (Sigma S-5150)
- 500 mM EDTA (Invitrogen/Gibco; Cat. no: 15575-038)
- 500 mM EGTA (Sigma, Cat no: E3889) 4°C
- 37% formaldehyde (J.T. Baker; Cat. no: 2106-01)
Caution: Formaldehyde is flammable; highly toxic by inhalation, contact with skin or if swallowed; causes burns; and is potentially carcinogenic. Formaldehyde should be used with appropriate safety measures, such as protective gloves, glasses and clothing, and adequate ventilation. Formaldehyde waste should be disposed of according to regulations for hazardous waste.
- 3.5 M glycine (Sigma; Cat. no: G8790) (must pH to 8 in order for glycine to dissolve)
- 10× Dulbecco's phosphate buffered saline (PBS) (Invitrogen/Gibco; Cat. no.: 14200-075)
- Dynabeads G Protein coupled (Dyna) (Invitrogen 10004D) 4°C
- Block Solution: 1× PBS, 0.5% bovine serum albumin (BSA) (Sigma; Cat. no: A7906)
Critical: Should be made fresh and kept cold.
- Primary antibody of choice
- 100× solution Complete Protease Inhibitor Cocktail: (can store aliquots at -20 °C) (Sigma)
- Sodium butyrate Na-But (Sigma) B5887
- 50% glycerol (Sigma; Cat. no.: G5516)
- 10% Igepal CA-360 (also known as NP-40) (Sigma; Cat. no.: I8896)
- 10% Triton X-100 (Sigma; Cat. no.: T8787)
- 1 M Tris-HCl, pH 8.0 (Invitrogen/Gibco; Cat. no.: 15568-025)
- 5% Na-Deoxycholate (Sigma; Cat. no.: D5760)
- 5% N-lauroylsarcosine (Fluka; Cat. no.: 61743)
- 5 M LiCl (Sigma; Cat. no.: L7026)
- 10% SDS (Invitrogen/Gibco; Cat. no.: 15553-035)
- 10 mg/ml RNaseA (Sigma; Cat. no.: R6513) -20°C
- Proteinase K solution (Invitrogen; Cat. no.: 25530-049) -20°C
- Phenol:chloroform:isoamyl alcohol (Fluka; Cat. no.: 77617) **CRITICAL:** If this solution is old or is at low pH, there will be degradation of DNA. 4°C
Caution: Phenol is toxic when in contact with skin or if swallowed; causes burns; and is irritating to eyes, the respiratory system and skin. Chloroform is harmful by inhalation or if swallowed; is irritating to skin; and is potentially carcinogenic. Isoamyl alcohol is

flammable; and is irritating to eyes, the respiratory system and skin.

The phenol:chloroform:isoamyl alcohol should be used with appropriate safety measures, such as protective gloves, glasses and clothing, and adequate ventilation. The phenol:chloroform:isoamyl alcohol waste should be disposed of according to regulations for hazardous waste.

- Double-distilled water
- Formaldehyde Solution (see REAGENT SETUP)
- Lysis Buffer 1 (see REAGENT SETUP) 4°C
- Lysis Buffer 2 (see REAGENT SETUP) 4°C
- Lysis Buffer 3 (see REAGENT SETUP) 4°C
- Wash Buffer (RIPA) (see REAGENT SETUP)
- Elution Buffer: 50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 1.0% SDS
- TE: 10 mM Tris-HCl, pH 8.0, 1 mM EDTA
- TE + NaCl: 10 mM Tris-HCl, pH 8.0, 1 mM EDTA, 50 mM NaCl (keep cold)
- PCR Purification Kit Qiagen 28004

Equipment

- Rotator (e.g., Fisher Hematology/Chemistry Mixer)
- Magnetic particle collector (MPC; Dynal)
- Sonicator (Covaris S2)
- NanoDrop ND-1000 spectrophotometer (This is highly useful as it allows you to assay DNA concentrations and dye incorporations by using low volumes (1.5 µl) of sample)
- Heat Block (with 1.5 ml tube insert)
- Swinging bucket centrifuge, variable temperature
- Liquid nitrogen and appropriate container
- 50 mL and 15 mL conical tubes, 1.5 mL epi-tubes
- Ice filled bucket
- Micropipettes

REAGENT SETUP

- **Formaldehyde Solution** Formaldehyde should be added right before use. Consists of 50 mM HEPES-KOH, pH 7.5, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA and 11% formaldehyde.
- **Lysis Buffer 1** Add protease inhibitors just before use, and keep cold. Consists of 50 mM HEPES-KOH, pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40, 0.25% Triton X-100, 1× protease inhibitors
- **Lysis Buffer 2** Add protease inhibitors just before use, keep cold. Consists of 10 mM Tris-HCl, pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 1× protease inhibitors
- **Lysis Buffer 3** Add protease inhibitors just before use, keep cold. Consists of 10 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na-Deoxycholate,

0.5% *N*-lauroylsarcosine, 1× protease inhibitors

- **Wash Buffer (RIPA)** Keep cold. Consists of 50 mM HEPES-KOH, pKa 7.55, 500 mM LiCl, 1 mM EDTA, 1.0% NP-40 (Igepal), 0.7% Na-Deoxycholate
- **Protease inhibitors (PIs)** Keep cold. Add Protease inhibitors (PI) to all Lysis buffers and last TBS wash solution
- 100X Protease Inhibitor Cocktail (Sigma P8340)
- Sodium-Butyrate (Na-But) must be made fresh , 22 mg/0.1ml gives 2 M solution in MilliQ water add 10 ul/ml LB (final 20 mM)
- AEBSF concentration of 12 mg/ml in MeOH add 2 ul per ml buffer

Solutions

Formaldehyde Solution		50 ml
50 mM HEPES-KOH pH 7.5	1M HEPES-KOH pH 7.5	2.5
Or 50 mM HEPES native	Or 1M HEPES	
100 mM NaCl	5M NaCl	1
1 mM EDTA	500 mM EDTA	0.1
0.5 mM EGTA	500 mM EGTA	0.05
11% formaldehyde	1.1g formaldehyde/10ml	5.5
ddH ₂ O		40.85

Lysis Buffer 1		50 ml
50 mM HEPES-KOH pH 7.5	1M HEPES-KOH pH 7.5	2.5
140 mM NaCl	5M NaCl	1.4
1 mM EDTA	500 mM EDTA	0.1
10% glycerol	50% glycerol	5
0.5% NP-40	10% NP-40	2.5
.25% Triton X-100	10% Triton X-100	1.25
1× protease inhibitors	100X PI	0.5
0.02 M Na-But	2M Na-But	0.5
ddH ₂ O		37.25

Lysis Buffer 2		50 ml
10 mM Tris-HCl, pH 8.0,	1M Tris-HCl, pH 8.0	0.5
200 mM NaCl	5M NaCl	2
1 mM EDTA	500 mM EDTA	0.1
0.5 mM EGTA	500 mM EGTA	0.05
1× protease inhibitors	100X PI	0.5
0.02M Na-But	2M Na-But	0.5
ddH ₂ O		47.35

Lysis Buffer 3		50 ml
10 mM Tris-HCl, pH 8.0	1M Tris-HCl, pH 8.0	0.5
100 mM NaCl	5M NaCl	1
1 mM EDTA	500 mM EDTA	0.1
0.5 mM EGTA	500 mM EGTA	0.05
0.1% Na-Deoxycholate	5% Na-Deoxycholate	1
0.5% N-lauroylsarcosine	5% N-lauroylsarcosine	5
1× protease inhibitors	100X PI	0.5
0.02M Na-But	2 M Na-But	0.5
ddH ₂ O		42.35

Wash Buffer (RIPA)		50 ml
50 mM HEPES-KOH, pH 7.5	1M HEPES-KOH pH 7.5	2.5
500 mM LiCl	5M LiCl	5
1 mM EDTA	500 mM EDTA	0.1
1.0% NP-40	10% NP-40	5
0.7% Na-Deoxycholate	5% Na-Deoxycholate	7
ddH ₂ O		30.4

Elution Buffer		50 ml
50 mM Tris-HCl, pH 8.0	1M Tris-HCl, pH 8.0	2.5
10 mM EDTA	500 mM EDTA	1
1.0% SDS	10% SDS	5

TE + NaCl:		50 ml
10 mM Tris-HCl pH 8.0	1M Tris-HCl, pH 8.0	0.5
1 mM EDTA	500 mM EDTA	0.1
50 mM NaCl	5M NaCl	0.5
ddH ₂ O		48.9
TE		50 ml
10 mM Tris-HCl, pH 8.0	1M Tris-HCl, pH 8.0	0.5
1 mM EDTA	500 mM EDTA	0.1
ddH ₂ O		49.4

Proteinase K	Solution w/o proK	5 ml
---------------------	-------------------	------

1 mM CaCl ₂	0.5M CaCl ₂	0.002
50 mM Tris-HCl, pH 8.0	1M Tris-HCl, pH 8.0	0.25
25% glycerol	50% glycerol	1.25
ddH ₂ O		3.498
20 mg/ml proteinase K	prot K powder	Add 5 ml to 100 mg bottle

TBS 10X		50 ml
1.5 M NaCl	5 M NaCl	15
100 mM Tris pH8.0	1 M Tris-HCl, pH 8.0	5
ddH ₂ O		30

TBS		50 ml
150 mM NaCl	5M NaCl	1.5
10 mM Tris pH8.0	1M Tris-HCl, pH 8.0	0.5
ddH ₂ O		48
1× protease inhibitors		0.5
0.02M Na-But	2M Na-But	0.5

Elution Buffer ChIP		50 ml
50 mM Tris-HCl pH8.0	1M Tris-HCl, pH 8.0	2.5
10 mM EDTA	500 mM EDTA	1
1.0% SDS	10% SDS	5
ddH ₂ O		41.5
Dilution Buffer (reChIP)		50 ml
15.5 mM Tris-HCl, pH 8.0,	1M Tris-HCl, pH 8.0	0.775
174 mM NaCl	5M NaCl	1.74
0.58 mM EGTA	500 mM EGTA	0.058
1.16% Triton X-100	10% Triton X-100	5.8
ddH ₂ O		41.63

Low-Salt wash buffer		50 ml
20 mM Tris-HCl, pH 8.0,	1M Tris-HCl, pH 8.0	1
150 mM NaCl	5M NaCl	1.5
1 mM EDTA	500 mM EDTA	0.1
0.5 mM EGTA	500 mM EGTA	0.05
0.1% SDS	10% SDS	0.5

1% Triton X-100	10% Triton X-100	5
-----------------	------------------	---

High-Salt wash buffer		50 ml
20 mM Tris-HCl, pH 8.0,	1M Tris-HCl, pH 8.0	1
500 mM NaCl	5 M NaCl	5
1 mM EDTA	500 mM EDTA	0.1
0.5 mM EGTA	500 mM EGTA	0.05
0.1% SDS	10% SDS	0.5
1% Triton X-100	10% Triton X-100	5

LiCl wash buffer		50 ml
10 mM Tris-HCl, pH 8.0,	1M Tris-HCl, pH 8.0	0.5
0.25M LiCl	5M LiCl	2.5
1 mM EDTA	500 mM EDTA	0.1
0.5 mM EGTA	500 mM EGTA	0.05
1% Na-Deoxycholate	5% Na-Deoxycholate	10
1.0% NP-40	10% NP-40	5
ddH ₂ O		31.85

5 M LiCl		
42.39 g/mol	21.19 g LiCl	100 ml ddH ₂ O

2M NaBut		
110 g/mol	2.2 g NaBut	10 ml ddH ₂ O
aliquot and freeze right away		

2.5M Glycine		
75.07 g/mol	18.76 g Glycine	100 ml ddH ₂ O
adjust pH to 8.0 in order to dissolve glycine		

**Reasons for Dilution Buffer
[composition]**

Dilution Buffer	Elution Buffer	Final reChIP buffer
15.5 mM Tris-HCl, pH 8.0,	50 mM Tris-HCl, pH 8.0,	20 mM Tris-HCl, pH 8.0,
174 mM NaCl		150 mM NaCl
	10 mM EDTA	1.43 mM EDTA

0.58 mM EGTA		0.5 mM EGTA
1.16% Triton X-100		1% Triton X-100
	1% SDS	0.13% SDS

Citations

Lee and Young

Nat Protoc. 2006; 1(2): 729–748.

Truax and Greer

If you need to make your own 25X PI cocktail then use: Aprotinin inhibits serine proteases; Leupeptin inhibits serine and thiol proteases; Pepstatin inhibits aspartic proteases. Make **1 mg/ml** each in MilliQ water, aliquot and store at -20°C. It is also possible to use commercially available proteinase inhibitor tablets

*Protease inhibitors (PI) to all Lysis buffers and last TBS wash solution

-Aprotinin, Leupeptin and Pepstatin concentration of **1 µg/ml** add 40 ul per ml buffer ,

- AEBSF concentration of 12 mg/ml in MeOH add 2 ul per ml buffer

APPENDIX N

ISOLATION OF COLONIC CRYPTS

Isolation of Colonic Crypts

Note: Temp during isolation steps depends on what the crypts will be used for. To maintain active metabolizing crypts, may want to keep all steps warm. For termination of crypts for enzyme assays, chromatin immunoprecipitation, etc. keep solutions cold after crypt removal from mucosa.

- ___ Turn on shaking water bath (37°C)
 - ___ Thaw Gln, EDTA and BSA aliquots

 - ___ Make buffer 1, and 2
 - ___ Warm buffer 1 to 37°C (Keep buffer 2 on ice)
 - ___ Euthanize rat and remove colon (clean fatty residues and other tissues off)
 - ___ Open colon and swish in cup of PBS. Repeat with fresh PBS.
 - ___ Place tissue into 50 ml flask containing 50 ml buffer 1
 - ___ Incubate for 15 min with shaking at 100 rpm at 37°C
 - ___ Place colon in sterile petri dish on ice with 30ml new buffer 1. Remove any large chunks of floating fat.
 - ___ Gently scrape mucosal side with sterile rubber policeman to remove crypts.
 - ___ Transfer contents into 50 ml conical tube (keep on ice). Use 20 ml buffer 1 to rinse out petri dish and retrieve rest of all crypts (From this point on all steps should be performed on ice or 4°C centrifuge)
 - ___ Centrifuge at 3000 x g, 10 min at 4°C.
 - ___ Discard sup and resuspend pellet in 35 ml cold buffer 2 to wash cells and remove DTT & EDTA
 - ___ Centrifuge at 3000 x g, 10 min at 4°C.
 - ___ Discard supernatant.
 - ___ Resuspend pellet in 24 ml (or desired volume) HBSS or buffer of choice
- Samples are now ready for aliquoting portions for RNA and protein extraction and crosslinking.

Buffer 1:

HBSS –CaMg +30 mM EDTA +5 mM DTT + 1 mM Glutamine + 0.1% BSA (final concs)

- Need 100 ml per sample (2 per colon, distal - proximal)
- 1,134 ml - 94.5 ml HBSS – CaMg (Cellgro 21-021-CV)
- 48 ml - 4 ml 750 mM EDTA (made in HBSS)
- 6ml - 0.5 ml 200 mM Glutamine
- 12 ml - 1 ml 10% BSA (made in HBSS)
- 924mg - 77 mg DTT

Adjust to pH 7.4 at 37°C

Buffer 2: HBSS + 1 mM Glutamine + 0.1% BSA

12 x Need 40 ml per colon

567ml - 49.25 ml HBSS + CaMg – phenol red – bicarb (Gibco 14025-092)

2.88ml - 0.25 ml 200 mM Glutamine

5.76ml - .5 ml 10% BSA (made in HBSS)

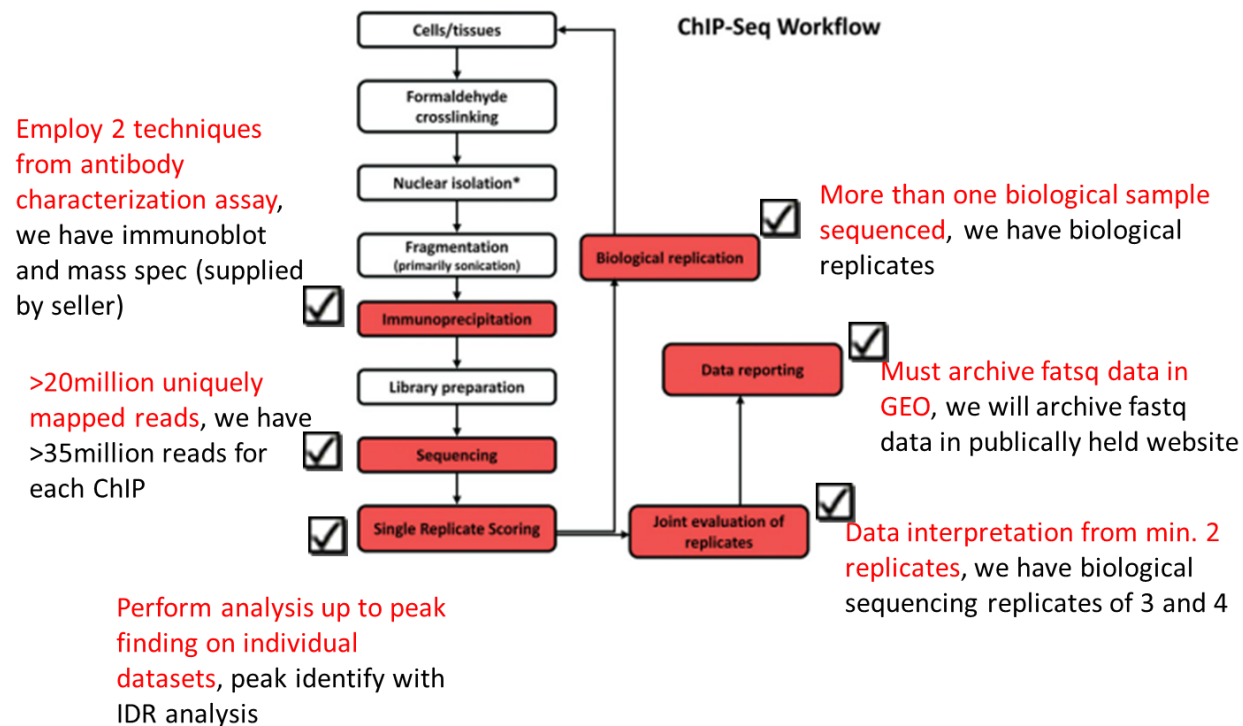
APPENDIX O

QUALITY CONTROL REQUIREMENTS FOR CHIP-SEQ COMPARED TO ENCODE GUIDELINES

All the steps outline in this workflow have been troubleshooted for quality control purposes prior to ChIP-Seq in Aims 1-3.

Quality Control Required by ENCODE compared to Chapkin Lab

Specific ENCODE Guidelines in red

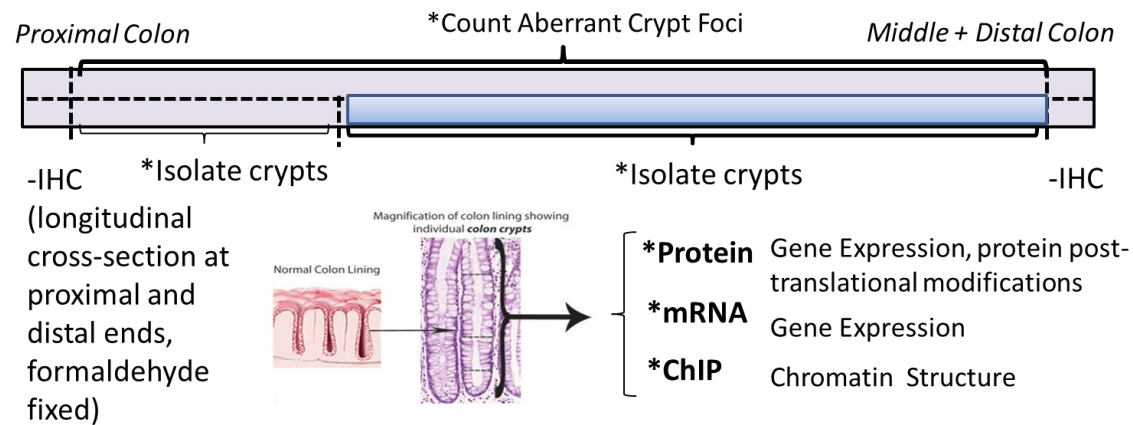


APPENDIX P

SAMPLE PROCESSING; TYPES OF TISSUES COLLECTED

Methods –Sample Processing, all tissue types with * were used in Aim 1-3, other samples are stored for potential future studies

-Fecal collection at week 0, 5 and 10 (prior) to termination 10 weeks post AOM or saline injection



Crosslink the crypts for ChIP analysis (100ug of DNA)



ChIP with each antibody and the amount of sample needed



ChIP-qPCR for QC and RNA-seq to determine changes in chromatin structure in each experimental group



ChIP-Seq to determine genome-wide changes in chromatin structure in each experimental group

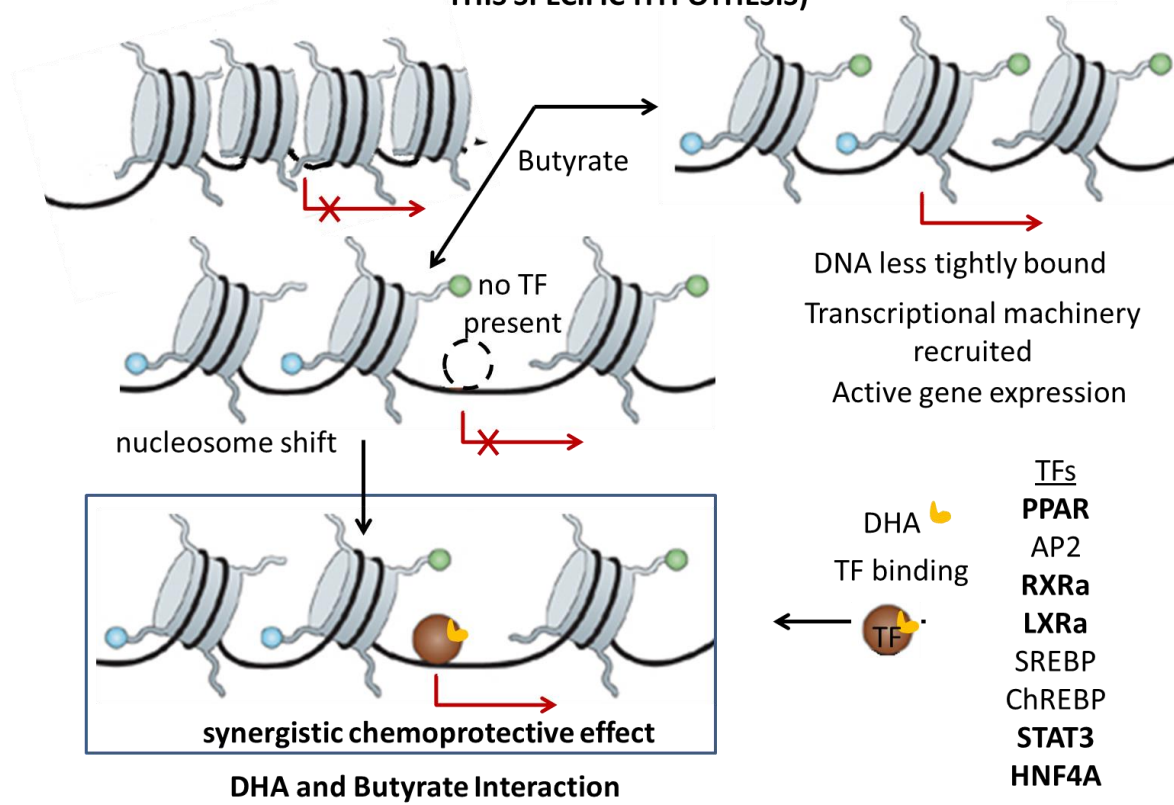
Extract 1ug RNA for transcriptional analysis



RT-qPCR for QC and RNA-seq to determine changes in transcription of selected genes in distal compared to proximal colon

APPENDIX Q

Proposed Epigenetic Mechanism (Aim 3) of DHA and Butyrate Interaction (WE FOUND NO EVIDENCE OF H3K9ac DIFFERENTIAL ENRICHMENT TO CORROBOORATE THIS SPECIFIC HYPOTHESIS)



APPENDIX R

SAMPLE POOLING OF ANIMALS FOR SEQUENCING

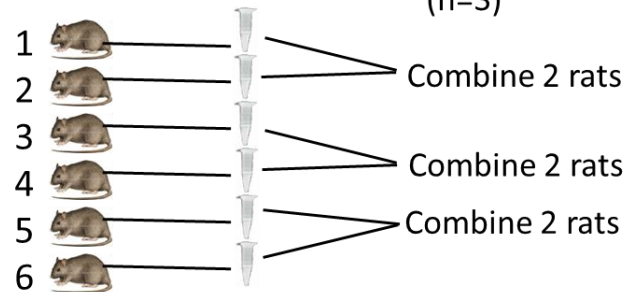
Animal study: Sample processing prior to sequencing of saline injected animals
n=6 per diet treatment

ChIP of individual rat colons, *then* pool all sample from that specific treatment and perform sequencing

ChIP for
-H3K4me3
-H3K9ac

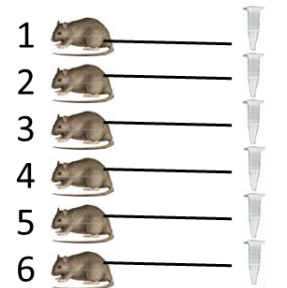
AOM injected rats. Separated the samples into 3 groups all of which represent that treatment's H3K4me3

(n=3)



Extract RNA from individual samples and sequence individually

RNA-seq on 6 of the rats individually (n=6)



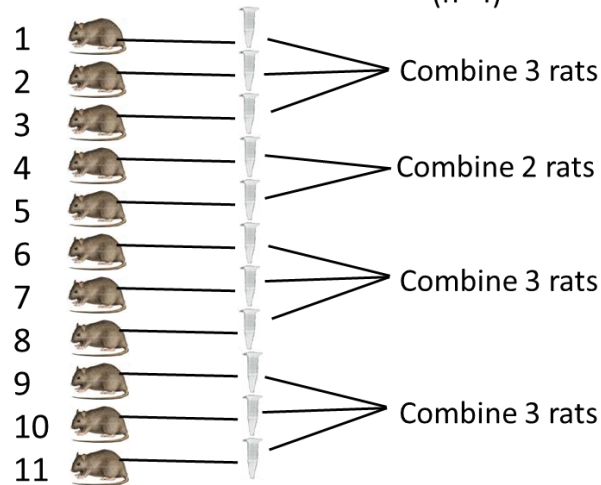
Animal study: Sample processing prior to sequencing of AOM injected animals

n=11 per diet treatment

ChIP of individual rat colons, *then* pool all sample from that specific treatment and perform sequencing

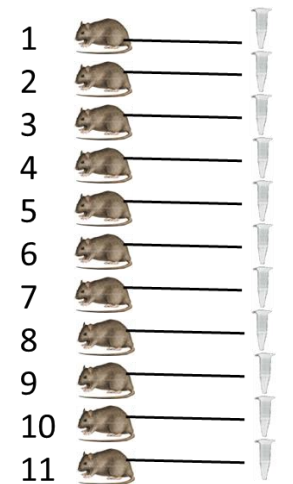
ChIP for
-H3K4me3
-H3K9ac

AOM injected rats.
Separate the samples into 4 groups all of which represent that treatment's H3K4me3 (n=4)



Extract RNA from individual samples and sequence individually

RNA-seq on 6 of the rats individually (n=6)

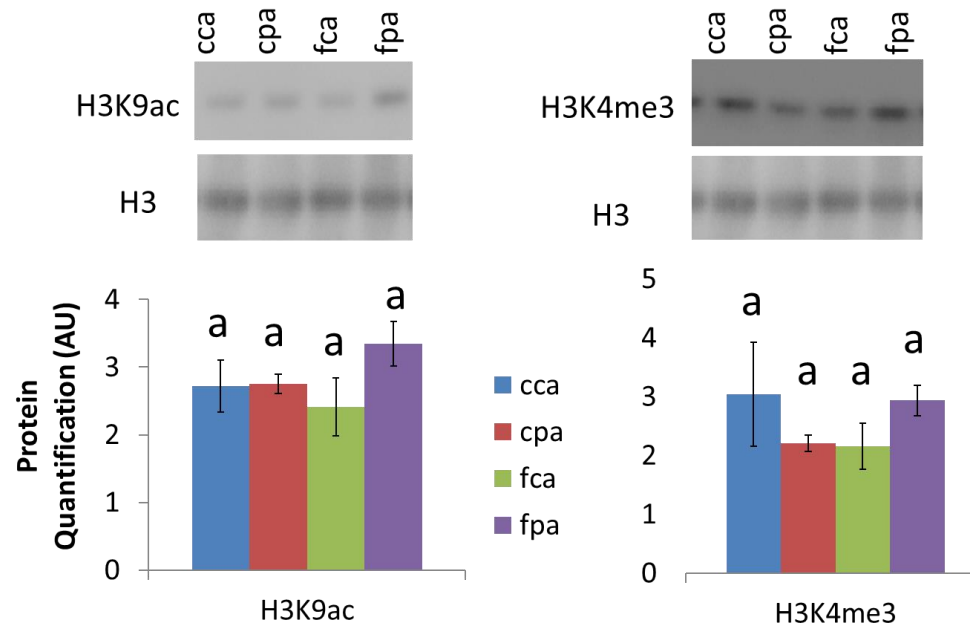


APPENDIX S

NUCLEAR PROTEIN LEVELS OF H3K4ME3 AND H3K9AC ARE NOT ALTERED BY DIET IN AOM INJECTED

ANIMALS

Cellular protein levels of modified histone tails are not altered by diet in AOM injected animals

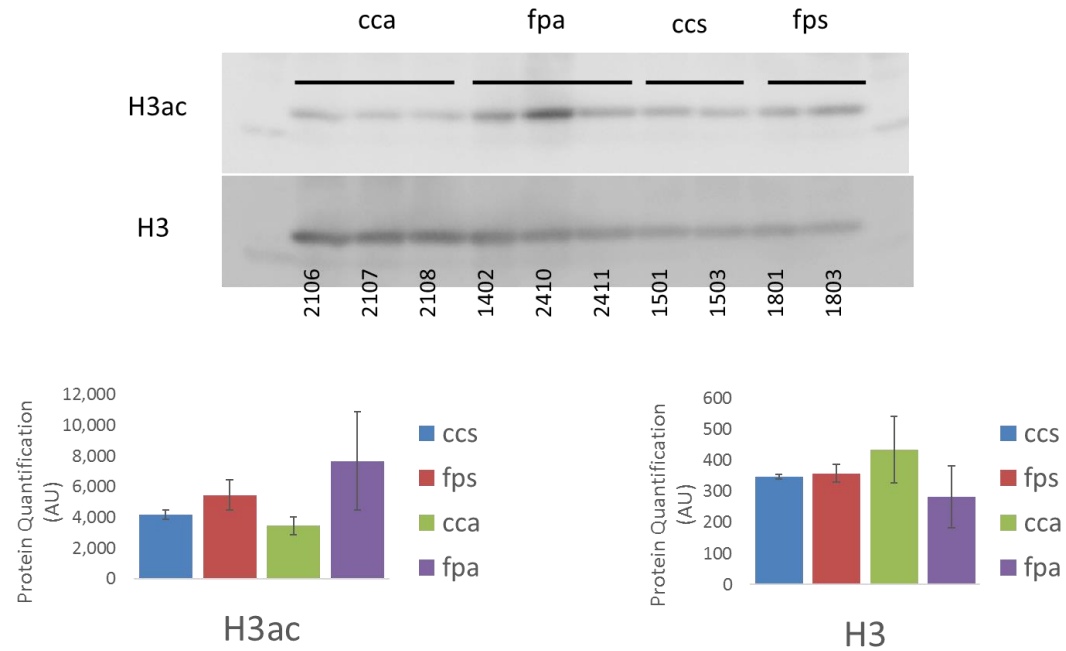


Western blot analysis of H3K4me3 and H3K9ac shows no variation between treatments (Bars not having the same superscript letters are significantly different, n=3 per treatment)

APPENDIX T

NUCLEAR LEVELS OF PAN-ACETYLATED H3 ARE NOT ALTERED BY A FISH OIL + PECTIN DIET

Cellular protein levels of modified histone tails are not altered by fish oil + pectin diet



Western blot analysis of H3ac shows no variation between treatments (Bars not having the same superscript letters are significantly different, n=3 per treatment in at least 1 WB)

APPENDIX U

SUMMARY OF DIFFERENTIAL EXPRESSION AND DIFFERENTIAL ENRICHMENT OUTPUT

	9ac	9ac	4me3	4me3	RNA	RNA
	DER(#genes)		DER(#genes)		(#genes)	
	padj<0.1	P<0.01	padj<0.1	p<0.01	padj<0.1	p<0.01
n3:pectin:aom	8	862	1	190	0	19
n3:pectin	8	782	0	141	0	94
n3:aom	1	733	0	268	0	76
pectin:aom	12	1421	1	184	0	14
cca_ccs	17 (11)	1205	868 (530)	2455	31(22)	134
cpa-cps	25(11)	1339	14	1135	15	107
fca-fcs	9	800	1	588	1	92
fpa-fps	0	510	7	378	2	80
fpa_cca	0	866	21	488	86	278
cpa-cca	0	1024	15	451	0	18
fca-cca	1	861	0	488	1	90
fish oil overall	58	868	5	1156	163	367
pectin overall	13	849	9	432	2	103
aom overall	49(27)	1135	7678 (3576)	6192 (4150)	128	288
fps-ccs	6	579	3	201	14	145
fcs vs ccs	11	642	0	348	64	302
cps vs ccs	49	1193	0	80	0	55

DER: Differentially Enriched Regions (ChIP peaks with differential binding between treatments) -(#genes) details number of genes, removing duplicates if multiple splice variants were detected in RNA or if multiple ChIP DERs were present in a gene.

APPENDIX V

ANOVA TEST FOR RAT HM-ACF

Anova test for rat ACF (used in Aim 3 HM-ACF analysis)
compare diets with outlier removed (174)

The ANOVA Procedure

Class Level Information		
Class	Levels	Values
diet	4	CCA CPA FCA FPA

Number of Observations Read	39
Number of Observations Used	39

Dependent Variable: count

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	7587.75796	2529.25265	7.50	0.0005
Error	35	11796.60101	337.04574		
Corrected Total	38	19384.35897			

R-Square	Coeff Var	Root MSE	count Mean
0.391437	40.61222	18.35881	45.20513

Source	DF	Anova SS	Mean Square	F Value	Pr > F
diet	3	7587.757964	2529.252655	7.50	0.0005

Tukey's Studentized Range (HSD) Test for count

Alpha	0.05
Error Degrees of Freedom	35
Error Mean Square	337.0457
Critical Value of Studentized Range	3.81400
Minimum Significant Difference	22.503
Harmonic Mean of Cell Sizes	9.682152

Note: Cell sizes are not equal.

Means with the same letter are not significantly different.				
Tukey Groupin		Mean	N	diet
g				
	A	63.222	9	CCA
	A			
B	A	53.500	10	CPA
B				
B	C	40.667	9	FCA
	C			
	C	26.636	11	FPA

APPENDIX W

INFORMATION ON THE DATA IN EACH OF THE SUPPLEMENTARY FILES

(IDENTIFIED BY BULLET AND BOLDED TEXT) IS PROVIDED BELOW

- **ChII.Distal_v_Proximal_Supplemental-Data-1 -DEs DERs**

As detailed in section 2.2.5

SICER program was employed to select regions that exhibited significant differences (DERs) in the tag counts between the two colonic regions (P-value < 0.00001; FC > 2). The nearest gene to each island, i.e., within 2 kb of the island, was identified using closestBed from the BEDTools software suite (Quinlan and Hall 2010) and the refGene table downloaded from the UCSC Genome Browser for the Baylor 3.4/rn4 assembly (Gibbs, Weinstock et al. 2004).

Sheets B_H3K4me3-seq and C_H3K9me3-seq

island signal directionality: gives information indicating whether calculated fold change was decreased or increased in proximal versus distal colon.

gene symbol: gene annotated to that genomic region (island)

fold change: fold change increase or decrease of island intensity in proximal compared to distal colon

island start: genomic start position of island

island end: genomic end position of island

island chr: chromosome where island is located

accession: accession number for annotated gene

distance to TSS as % of gene length: island distance (in bp) to TSS as % of total gene length (in bp)

annotation source: Database source of annotated genes

gene description: definition of annotated gene

Sheets A_RNA-seq

As detailed in section 2.2.8 Differential expression (DE) analyses were performed using cuffdiff2 and cummeRbund (Trapnell, Roberts et al. 2012).

gene: gene symbol annotated to that transcript

locus: genomic location of transcript

status: Cuffdiff parameters indicating whether FPKM<0 in either of the 2 samples

Proximal_FPKM: intensity of transcript from proximal colon

Distal_FPKM: intensity of transcript from distal colon

log2(fold_change): Differential expression (DE) as log2(distal/proximal)

p_value: extremal probability for calculated log2(fold_change) under the null

hypothesis, not accounting for multiple comparisons
q_value: extremal probability for calculated $\log_2(\text{fold_change})$ under the null hypothesis, accounting for multiple comparisons
significant: 'yes' if $q\text{-value} < 0.05$

- **ChII.Distal_v_Proximal_Supplemental-Data-2 – ncRNA**

As detailed in section 2.2.8

The cuffcompare program of cufflinks2 pipeline (reference annotation based method) was used to classify the assembled transfrags into different classes. The transfrags annotated with class codes u (unknown intergenic transcript) and i (a transfrag falling entirely within a reference intron) were selected to annotate against known lncRNA's. A custom perl script was written to fetch sequences for the coordinate information produced by cuffcompare. NONCODE v3, UCSC, and fRNAdb v3.0 databases were downloaded and only non-coding RNA (ncRNA) sequences belonging to the rat genome whose length was greater than 100 bp were used to create a unified database (Mituyama, Yamada et al. 2009, Bu, Yu et al. 2012, Meyer, Zweig et al. 2013). The fetched sequences from UCSC were then queried for ncRNA sequence similarity with blastn by thresholding at $1e-10$ (Altschul, Madden et al. 1997, Johnson, Zaretskaya et al. 2008). High confidence hits, % with high-scoring segment pairs (HSPs) length greater than 100 and whose identity was greater than 95% were retained. Whenever possible, BLAST results were further filtered to remove hits that did not correspond to known ncRNA coordinate information. Due to the consistent lack of coordinate information for the known ncRNA sequences, the results include matches to multiple genomic locations.

- **ChII.Distal_v_Proximal_Supplemental-Data-3 - IPA pathway analysis**
- **ChIII.Supplementary Table 3 - IPA Analysis Full Summary**
- **ChIV.Supplemental Table 2 (networks)**
- **ChIV.Supplementary Table 3 (URs)**
- **ChIV.Supplemental Table 4 (BF)**

As detailed in section 2.2.9 and 3.2.12.

To perform IPA analysis, all differentially expressed genes (adjusted $P < 0.005$) in the distal or proximal colon were uploaded into three columns for the purpose of generating Illumina probe ID, t-value (fold change) and adjusted P-value (FDR) data. By convention, genes that were up-regulated in the distal colon are shown in red and genes that were down-regulated are shown in green. By default, during IPA analysis, only

molecules from the data set associated with the Ingenuity Knowledge Base repository (Ingenuity Systems Inc.) were considered. Functional Analysis identified the biological functions (BFs) and/or diseases that were most significant to the data set. The significance of the association between the data set and the specific pathways of interest was determined in three ways: (1) ratio of the number of molecules from the data set that mapped to the pathway divided by the total number of molecules that mapped to the Ingenuity Knowledge Base pathway, (2) Fisher's exact test was used to calculate a P value determining the probability that the association between the genes in the data set and the pathway of interest could be explained by chance alone, and (3) activation state ("Increased" or "Decreased") was inferred by the activation z-score. The derivations of the z-scores are based on relationships in the molecular network that represent experimentally observed causal associations between genes and those functions. "Canonical pathway" analysis was used to identify networks from the IPA library that were most significantly modulated across anatomical sites. Significance of the association between each data set and the canonical pathway was measured in 2 ways: (1) A ratio of the number of molecules from the data set that mapped to the pathway divided by the total number of molecules that mapped to the canonical pathway, and (2) Fisher's exact test was used to calculate p-values determining the probability that the association between genes in the dataset and each canonical pathway was explained by chance alone. "Upstream regulator" (UR) analysis was based on prior knowledge of expected associations between transcriptional regulators and their target genes stored in the Ingenuity® Knowledge database, and significance for each TF was measured in 2 ways: (1) Fisher's Exact Test (p-value) was used to identify differentially expressed genes from the RNA-Seq data set which overlapped with genes known to be regulated by a TF. Since the regulation direction ("activating" or "inhibiting") of an edge is not taken into account for the computation of overlap p-values, the underlying network also included findings without associated directional attribute, such as protein-DNA (promoter) binding. In addition, the activation score (z-score) was used to infer the status of predicted transcriptional regulators by comparing the observed differential regulation of genes ("up" or "down") in the dataset relative to the literature-derived regulation direction which can be either "activating" or "inhibiting".

- **ChIII.Supplementary Table 1 - AOM DEs**
- **ChIV.Supplemental Table 1 - DE DER (DEs)**

As detailed in section 3.2.10

Regions showing differences in histone modification were identified using the edgeR package (Robinson, McCarthy et al. 2010, Nikolayeva and Robinson 2014) for the R software environment (Zhang, Liu et al. 2008, Robinson, McCarthy et al. 2010). Read counts per gene were normalized using the scaling factor method of Anders and Huber (Anders and Huber 2010). Differential expression testing of genes was performed using likelihood ratio tests on the negative binomial GLMs estimated by edgeR (Robinson, McCarthy et al. 2010, Nikolayeva and Robinson 2014). Regions with FDR < 0.1 and p-val < 0.01 and minimal threshold of one count per million mapped reads in at least four samples were selected as differentially enriched regions (DERs).

Accession: gene accession id annotated to that transcript

Pval_aom: extremal probability for calculated $\log_2(\text{fold_change})$ under the null hypothesis, not accounting for multiple comparisons

AdjPVal (FDR) aom: extremal probability for calculated $\log_2(\text{fold_change})$ under the null

hypothesis, accounting for multiple comparisons

$\log_2(\text{FoldChange})$: Differential expression (DE) as \log_2

FCsign: positive or negative one is assigned based on whether calculated fold change was positive or negative.

locus: genomic location of transcript

- **ChIII.Supplementary Table 2 - AOM DERs**
- **ChIV.Supplemental Table 1 - DE DER (DERs)**

As detailed in section 3.2.8

To determine differential enrichment (DERs) The peak caller programme MACS (version 1.4.1) (Zhang, Liu et al. 2008) was used to identify peaks. Islands were defined (using merge function of BEDTools (Quinlan and Hall 2010)) as the genomic areas enriched with the ChIPed protein (peaks) in at least one sequenced sample, and reads were quantified using coverageBed function of BEDTools (Quinlan and Hall 2010). The UCSC Genome Browser was used to visualize bigwig data tracks. The nearest gene to each island, i.e., within 5 kb of the island was identified using closestBed from the BEDTools software suite (Quinlan and Hall 2010) and the refGene table downloaded from the UCSC Genome Browser for the Baylor 3.4/rn4 assembly files (Gibbs, Weinstock et al. 2004).

Regions showing differences in histone modification were identified using the edgeR package (Robinson, McCarthy et al. 2010, Nikolayeva and Robinson 2014) for the R software environment (Zhang, Liu et al. 2008, Robinson, McCarthy et al. 2010). In order to increase the statistical power of our analysis (higher number of samples per

treatment) and detect key AOM effects associated with cancer progression, rats were pooled across the various diet groups described under the methods section “Chromatin Immunoprecipitation”. Read counts per gene were normalized using the scaling factor method of Anders and Huber (Anders and Huber 2010). Differential expression testing of genes was performed using likelihood ratio tests on the negative binomial GLMs estimated by edgeR (Robinson, McCarthy et al. 2010, Nikolayeva and Robinson 2014). Regions with FDR < 0.1 and minimal threshold of one count per million mapped reads in at least four samples were selected as differentially enriched regions (DERs).

accession: gene accession number or genomic locus id of island/peak

Pval_aom: extremal probability for calculated $\log_2(\text{fold_change})$ under the null hypothesis, not accounting for multiple comparisons

AdjPVal (FDR) aom: extremal probability for calculated $\log_2(\text{fold_change})$ under the null hypothesis, accounting for multiple comparisons

$\log_2(\text{FoldChange})$: Differential enrichment (DER) as \log_2

FCsign: positive or negative one is assigned based on whether calculated fold change was positive or negative.

- **ChIII.Supplementary Table 4 – ncRNA**

As detailed in section 3.2.10

Non-coding RNAs were identified using the lncRNAPipe (<https://github.com/biocoder/Perl-for-Bioinformatics/releases>) software.

Pipeline steps with further technical details:

A total of approximately 1.25 billion (including all samples and replicates) 50bp single-end Illumina reads were obtained from a multiplexing run.

- a. Reads were trimmed for low quality bases (phred score <Q20) and any adapter sequences using Trimmomatic (Bolger, Lohse et al. 2014). Any reads less than 25bp in length were discarded after trimming.
- b. A total of approximately 1.2 billion reads survived trimming which were then aligned against rn4 rat assembly (from UCSC) using tophat (Trapnell, Roberts et al. 2014) with multi-hit filtering for each replicate of a treatment.
- c. A total of approximately 590 million reads uniquely mapped to the transcriptome and resulting alignments for each treatment were merged using Sambamba (Tarasov, Vilella et al. 2015).

- d. Merged read alignments for each treatment were processed with cufflinks (Trapnell, Roberts et al. 2014) to perform reference annotation based transcript assembly with bias and multi-read correction.
- e. Cuffcompare (Trapnell, Roberts et al. 2012) was run to compare assembled transcripts with reference annotation (rn4) and assign various class codes for transcript assembly of each treatment.
- f. Any transcripts that belonged to class codes “j”, “i”, “o”, “u” and “x” were considered and categorized into 5 categories: long intergenic lncRNAs (LincRNAs), intronic contained lncRNAs (Incs), partially overlapping lncRNAs (Poncs), completely overlapping lncRNAs (Concs) and exonic overlaps (LncRNAs with sense or antisense overlap with reference exon) (Pauli, Valen et al. 2012) by comparison to known RefSeq (Pruitt, Brown et al. 2014) gene annotation (rn4). When calculating exonic overlaps with reference exon boundaries, only cufflinks assembled transcripts whose exon (at least one) overlaps with reference exon by at least 80% of the exon length was considered. Single exon transcripts were also retained.
- g. Since RefSeq gene annotation consists both protein-coding and non protein-coding genes, the resulting catalog of categorized ncRNAs were compared to identify novel ncRNAs in rn4 i.e., any assembled transcripts from cufflinks which were not overlapping with RefSeq genes were marked as potential novel ncRNAs.
- h. FASTA sequences were created for the potential novel ncRNAs to run CPC (Coding Potential Calculator) (Kong, Zhang et al. 2007) with UniRef90 protein database, and any transcripts which were flagged as “coding” by CPC were discarded.
- i. Cmscan from INFERNAL1.1 (Nawrocki and Eddy 2013) was run with default E-value cutoff of 10 on the resulting transcripts using Rfam covariance models (CMs). Annotation from cmscan result was applied to putative non-coding RNA if at least 10% of query sequence was matched with the resulting best hit.
- j. Finally, RNAfold (Lorenz, Bernhart et al. 2011) was run to predict minimum free energy structure of putative novel ncRNAs.

After the pipeline was run on each treatment, the individual putative novel lncRNA assemblies for each treatment were merged with cuffmerge (Trapnell, Roberts et al. 2014) to create unified ncRNA transcript catalog for this experiment. BLAST (Camacho, Coulouris et al. 2009) homology search was performed against hg38, mm10 and rn6 RefSeq genes for the merged transcripts with e-value cutoff of 1e-5, 95% sequence

identity and 90% query sequence coverage thresholds to eliminate any similar or newly identified ncRNAs in closely related species to rat. A final list of potential novel long non-coding RNAs (transcript length ≥ 150) was created. Cuffdiff (Trapnell, Roberts et al. 2014) was run using the final list of putative novel lncRNAs as reference between AOM and saline (control) conditions to identify differentially expressed putative novel lncRNAs. A total of 324 transcripts were identified as putative novel lncRNAs of which, 266 have been assigned RNA families by INFERNAL with at least 10% query sequence coverage and E-value cut off of 10. None of the putative novel lncRNAs were significantly differentially expressed.