The Pennsylvania State University
The Graduate School

INVESTIGATING THE WNT/MYC AXIS IN THE PATHOGENESIS OF CROHN'S DISEASE

A Dissertation in

Biomedical Sciences

by

Stephen M. Matthews

© 2021 Stephen M. Matthews

Submitted in Partial Fulfilment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2021

The dissertation of Stephen M. Matthews was reviewed and approved by the following:

Gregory Yochum
Associate Professor of Biochemistry and Molecular Biology, & Surgery
Dissertation Advisor
Chair of Committee

David DeGraff
Assistant Professor, Department of Pathology and Laboratory Medicine
Division of Experimental Pathology
Assistant Professor, Department of Surgery

Walter Koltun
Professor of Surgery
Peter and Marshia Carlino Chair in Inflammatory Bowel Disease
Chief, Division of Colon and Rectal Surgery.

Vladimir Spiegelman
Professor Department of Pediatrics
IFC/Panhellenic Dance Marathon Chair for Pediatric Cancer Research

Ralph Keil
Associate Professor of Biochemistry and Molecular Biology
Chair of Biomedical Sciences Graduate Program

Abstract

A single layer of intestinal epithelial cells (IECs) protects underlying tissues from contents of the intestinal lumen and is composed of differentiated cells at the luminal surface, and non-differentiated, crypt structures embedded in the submucosa. Due to the harsh environment of the intestinal lumen, differentiated IECs are routinely lost through damage or controlled cell death. To maintain barrier integrity, IECs must be rapidly replaced by proliferating cells within the crypts. The primary driver for this proliferation is the Wnt/β-catenin signaling pathway. β-catenin is a transcriptional co-activator that activates gene expression through interaction with the TCF/LEF family of sequence specific transcription factors. Crohn's disease (CD) is one subtype of inflammatory bowel disease (IBD), a recrudescent and chronic inflammatory disease of the gastrointestinal (GI) tract. Through genome-wide association studies (GWAS), over 200 single nucleotide polymorphisms (SNPs) have been linked with predisposition for developing IBD. As the vast majority of these SNPs are intergenic, how these variants are associated with disease predisposition, and the relevant target genes impacted, represent significant gaps in knowledge in the IBD research community. We investigated one such CD-associated SNP, rs6651252, which maps to a gene desert on chromosome 8. Work presented within this dissertation demonstrate that rs6651252 resides within a novel Wnt responsive DNA enhancer element (WRE), and that the disease associated allele enhances binding of the TCF7L2 transcription factor to this region of DNA. Using CRISPR-mediated genomic editing and epigenetic modulation, we found that the

rs6651252-enhancer regulates expression of the *c-MYC* proto-oncogene (*MYC*). Furthermore, in patient-derived intestinal tissue, we saw a positive correlation between *MYC* transcript levels and the presence of rs6651252 disease-associated allele. These results suggest one mechanism by which Wnt/MYC signaling contributes to CD pathogenesis. Furthermore, our findings suggest patients with the disease-associated allele may benefit from therapies that target MYC, or its downstream target genes.

Table of contents

List of Figures viii

List of Tables x

List of Abbreviations xi

Acknowledgements xv

Chapter 1 Literature review 1

- Gastrointestinal tract 2
  - Organ function and organization 2
  - Organization and cellularity of the intestines 4
- Wnt signaling 8
  - Canonical Wnt signaling pathway 8
  - TCF family 12
  - Wnt in the IEC barrier 14
  - c-MYC as a target of Wnt signaling 15
- Inflammatory bowel disease 20
  - Definition and types of IBD 20
  - IBD treatments 22
  - Immune system in IBD 23
  - The intestinal epithelium in IBD 26
- Genetics of IBD 27
  - Genetic basis of IBD 27
  - DNA regulatory elements in IBD 32
- Dissertation overview 34

Chapter 2 Materials and methods 38

  - Cell lines 39

Plasmids ... 39
Chromatin immunoprecipitation (ChIP) ... 41
Luciferase reporter assay ... 41
CRISPR/Cas9-mediated deletion of the rs6651252 WRE in HCT116 cells . 42
DNA pull-down assay ... 42
Epigenetic silencing rs6651252-WRE through CRISPR/dCas9 ... 43
Quantitative reverse transcription PCR ... 44
Western blot ... 44
PAX gene® analysis ... 45
Patient samples ... 45
Statistics ... 46
Chapter 3 The Crohn's disease-associated SNP, rs6651252, impacts *MYC* gene expression in human colonic epithelial cells ... 49
Abstract ... 50
Introduction ... 50
Results ... 53
β-Catenin and TCF7L2 bind the rs6651252 locus in colonic epithelial cells 53
TCF7L2 binds DNA harboring disease-associated rs6651252 variant with stronger affinity ... 55
rs6651252 demarcates a Wnt responsive DNA enhancer element ... 57
The rs6651252 WRE regulates *MYC* gene expression ... 59
The rs6651252 C variant correlates with increased *MYC* expression in patient colonic tissues ... 62
Discussion ... 63
Supplemental ... 68
Chapter 4 Interplay of Wnt and MYC in Crohn's disease ... 70

Overview 71
Additional SNPs and targets of interest 72
Ch8q24 is a gene desert with many disease-associated SNPs 72
Other IBD-associated SNPs and β-cat ChIP-seq peaks 73
rs6651252 is within a putative $CD4^+$ T-cell DRE 76
Contribution of MYC to disease pathogenesis 77
MYC in restitution of the intestinal epithelium 77
MYC as a rheostat 80
Targeting MYC in disease 82
References 88

List of Figures

Figure 1-1. Diagram of the human gastrointestinal tract. 3

Figure 1-2. Organization of the intestinal epithelium. 5

Figure 1-3. Illustration of the canonical Wnt signaling pathway. 11

Figure 1-4. Domains in mammalian TCF family members. 13

Figure 1-5. Regulation of *c-MYC* by Wnt/β-cat signaling. 19

Figure 1-6. Photograph of total abdominal colectomy from patient with CD. 21

Figure 1-7. Conceptual representation of GWAS. 29

Figure 3-1. β-catenin and TCF7L2 bind the rs6651252 locus in colonic epithelial cells. 54

Figure 3-2. TCF7L2 binds DNA harboring disease-associated rs6651252 variant with stronger affinity. 56

Figure 3-3. rs6651252 demarcates a Wnt-responsive DNA enhancer element. 58

Figure 3-4. The rs6651252-WRE regulates *MYC* gene expression. 60

Figure 3-5. Epigenetic repression of rs6651252 reduces *MYC* gene expression. .. 61

Figure 3-6. The rs6651252 C variant correlates with elevated *MYC* expression in patient colonic tissues. 64

Supplemental; Figure 3-7. The rs6651252 element is repressive in HEK293T cells. 68

Supplemental; Figure 3-8. Depletion of rs6651252-WRE does not affect POU5F1B gene expression. 69

Figure 4-1. Overlap of IBD-associated SNPs with β-cat ChIP-Seq peaks. 74

Figure 4-2. The rs6651252 disease-associated allele does not affect *MYC* expression in patient blood lymphocytes. 78

Figure 4-3. Model of structural changes in intestinal mucosa structure with exposure to acute and chronic MYC overexpression. 84

List of Tables

Table 2-1 Primer sequences 47

List of Abbreviations

AKT – protein kinase B

AOM – azoxymethane

APC – adenomatous polyposis coli

ATG16L1 – autophagy-related 16 like 1

AXIN – axis inhibition protein

β-cat – β-catenin

BET – bromodomain and extra-terminal

bHLH-lz – basic-helix-loop-helix leucine zipper

Bmi1– B lymphoma Mo-MLV insertion region 1 homolog

βTrCP – beta-transducin-repeat-containing protein

CAC – colitis-associated cancer

CARD9 – caspase recruitment domain-containing protein 9

Cas9 - CRISPR-associated protein 9

CBC – column base columnar cells

CBP – CREB binding protein

CD – Crohn's disease

ChIP – chromatin immunoprecipitation

CK1 – casein kinase 1

CRC – colorectal cancer

CRISPR – clustered regularly interspaced palindromic repeats

dCas9 – catalytically dead Cas9

DRE – DNA regulatory element

DSS – dextran sodium sulfate

Dvl – dishevelled

E-box – enhancer-box

EMSA – electrophoretic mobility shift assay

ENCODE – encyclopedia of DNA elements

eQTL – expression quantitative trait loci

FAP – familial adenomatous polyposis

FMT – fecal microbial transplant

FZD – frizzled

GI – gastrointestinal

GCN5 – general control non-derepressible 5

gRNA – guide RNA

GSK3b – glycogen synthase kinase 3 beta

GWAS – genome-wide association studies

H3K27Ac– histone H3 lysine 27 acetylation

H3K4me1 – histone H3 lysine 4 mono-methylation

HAT – histone acetyltransferase

HDAC – histone deacetylase

HMT – histone methyl transferase

IBD – inflammatory bowel disease

IEC – intestinal epithelial cell

LDL – low-density lipoprotein

IL-10 – interleukin 10

Int1 – integration site 1

IRGM – immunity related GTPase M

KRAB – Krüppel associated box

LD – linkage disequilibrium

LEF – lymphoid enhancer factor

LGR5 – leucine-rich repeat-containing G-protein-coupled receptor 5

lncRNA – long non-coding RNA

LRP5/6 – low-density lipoprotein receptor-related protein

MAF – minor allele frequency

MAX – Myc-associated factor X

MIZ-1 – Myc-interacting zinc finger protein 1

MMTV – mouse mammary tumor virus

mRNA – messenger RNA

MYC – myelocytomatosis

NES – nuclear export signal

NF-κB – nuclear factor kappa-light-chain-enhancer of activated B cells

NLS – nuclear localization signal

NOD2 – nucleotide-binding oligomerization domain 2

NSAID – nonsteroidal anti-inflammatory drug

PBMC – peripheral blood mononuclear cell

PCR – polymerase chain reaction

PI3K – phosphoinositide 3-kinase

SC – stem cell

SMAD3 – mothers against decapentaplegic homolog 3

SNP – single nucleotide polymorphism

TA – transit-amplifying

TBE – TCF binding element

TCF – T-cell factor

TGFβ – transforming growth factor beta

Th1/2 – T-helper cell type 1/2

TIP60 – HIV-1 tat interactive protein, 60kDa

TJ – tight junction

TNF – tumor necrosis factor

TRRAP – transformation/transcription domain associated protein

TTP – tristetraprolin

UC – ulcerative colitis

VUS – variant of uncertain significance

WRE – Wnt-responsive DNA element

ZFP – zinc finger protein

4C-seq – circular chromosome conformation capture-sequencing

5-ASA – 5 aminosalicylic acid

Acknowledgements

I believe the old adage "it takes a village to raise a child" applies beyond rearing and well into adulthood. This is especially true for the trials and tribulations of graduate school. A great many people have supported and helped me get to where I am today. First and foremost, I need to thank my advisor and mentor, Dr. Gregory Yochum. Greg has invested substantial time and effort into my training, and helped me become a far better scientist, writer and presenter than I ever could have become on my own. I am most appreciative of his patience and determination in helping me during my time in working in his lab. I would also like to extend a thank you to my dissertation committee, Drs. Degraff, Koltun and Spiegelman. My project has twisted so many times over the years, and they were always there to offer advice and pertinent questions to help shape my work into the most interesting and impactful ways. Additionally, many thanks to my original mentors, Dr. Kalterider and Dr. Ihnat who have been steadfast supporters of mine and shaped me in ways prior to graduate school, which helped send me down the path that brought me to Penn State Hershey.

Our lab had a great many people who helped me throughout the years. When I first joined after my rotation, our lab moved down to the fourth floor where Leonard, Katie, Bryan and Sue were. I appreciate all the help, chatting and support you were able to provide me. The Yochum lab specifically gave me one of the best student mentors and mentees I could have asked for. Melanie was my older lab sister, always there to answer my million "quick questions," help me troubleshoot experiments, and also poke fun at me for anything and everything I

did that deserved it. When Carli joined, she fit into the lab easily, being a great banter buddy, idea bouncer and constant source of reminders on why "you should get a dog." I also wish to extend a thanks to my rotation students Erika and Olivia, who along with Carli, trusted me and gave me a chance to teach and pass on some of my skills.

I have never been great at being concise, but I could not make acknowledgements without giving credit to my family. My parents have mostly supported me and my scientific interests ever since I was a small child catching bugs. While I have not grown up to be an entomologist, I do hope I continue to make them proud. My sister Lauren and I grew much closer while I was in school here, and I always appreciated her support and care packages during stressful periods of time. I am also grateful to my extended family who have seen me largely fall off the face of the earth. I appreciate all the holiday cards, phone calls and texts over the last few years, and need to express that they all meant a great deal.

I would be remiss if I did not also offer thanks to my closest and most frequently utilized support network – my friends. Over the years I have met some truly amazing people who have helped me in more ways than I can count. No one person had a *small* impact on my time here not worth mentioning, though I do not have the space to acknowledge everyone. I specifically thank the people I have met in groups I have been in since 2015, The Triose, MedNotes, and DnD, as well as those I've spent less time in, such as GSA or LTS. Everything from mini-vacations, disc golf, hiking and game nights, to grabbing coffee or catching up

while passing in the hallway has helped keep my sanity in check, and I couldn't thank nor appreciate you all enough.

Lastly, I would also like to thank Peter and Marshia Carlino, for their efforts and assistance through the Peter and Marshia Carlino fund for Inflammatory Bowel Disease research, they made a large portion of my research possible.

Chapter 1 Literature review

Gastrointestinal tract

Organ function and organization

The gastrointestinal (GI) tract is responsible for digesting food, extracting nutrients, absorbing water and expelling solid waste matter (1). As the core of the digestive system, the GI tract forms a continuous tube that encompasses all connected organs between the mouth and the anus. The GI tract comprises the mouth, esophagus, stomach, small and large intestines, the rectum and anus (Fig. 1). The esophagus transports ingested food material into the stomach. The stomach contains a low pH environment due to the constant secretion of acid, and releases several enzymes to facilitate food breakdown (1). At the pylorus, or the base of the stomach, a pyloric sphincter controls flow of digested food and stomach juices into the small intestines. The small intestine is divided into the duodenum, which connects to the stomach, the jejunum and the ileum. Enzymes found in the pancreas are secreted into the small intestine to facilitate digestion (1, 2). In addition, bile secreted by the liver and stored in the gallbladder, further facilitates food breakdown in the small intestines. The terminal ileum of the small intestine connects to the large intestine, or colon, in the lower right quadrant of the abdomen. The right colon, or ascending colon, extends upwards towards the ribcage and across the abdomen in a section called the transverse colon. The descending colon, or left colon, projects downwards to the sigmoid colon. Finally, the sigmoid colon connects to the rectum and anus. Whereas the primary function of the small intestine is food digestion and nutrient absorption, the colon serves mainly to absorb water and to compact stool.

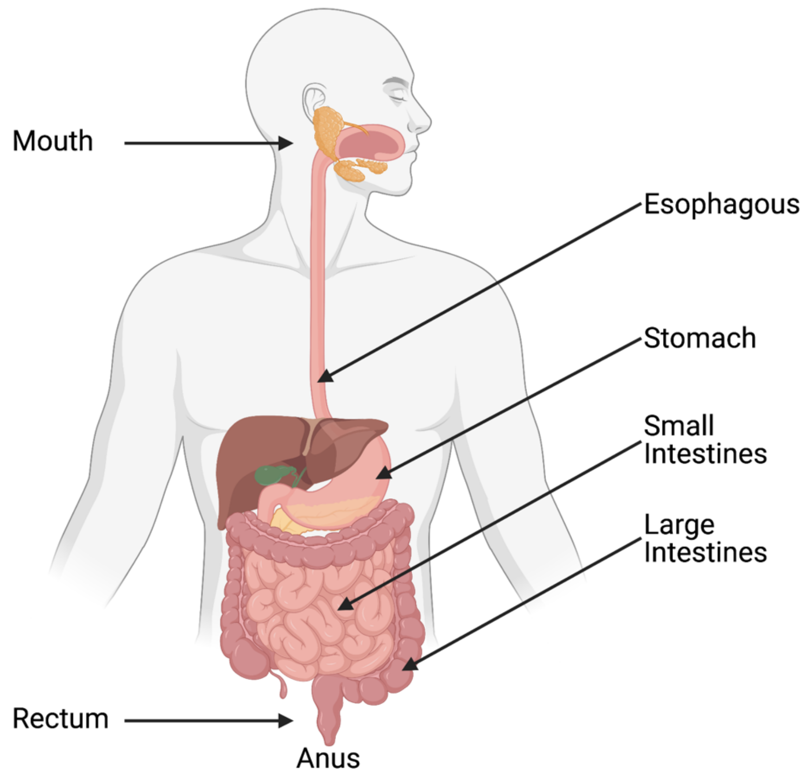


Figure 1-1. Diagram of the human gastrointestinal tract. Food enters the mouth and travels through the esophagus to the stomach. Following stomach emptying, it passes through the small and large intestines. Compacted stool is expelled through the anus. Also shown are organs of the digestive system including the liver (brown), pancreas (yellow) and gallbladder (green), which provide enzymes to the small intestines to facilitate food breakdown. Figure created with BioRender.com.

A single layer of epithelial cells line both the small intestine and colon (3). In the small intestine, this layer is arranged into ~1.0 mm long finger-like projections reaching into the lumen, known as villi. Millions of these villi increase the surface area of the small intestines approximately 6.5-fold to better facilitate nutrient absorption (4). In contrast, the colonic epithelium is smooth at the luminal surface to support its primary function to compact stool and reabsorb water. Additionally, in the small intestines and colon, the intestinal epithelium provides a physical barrier separating the underlying tissue from the microbiome and noxious contents of the lumen (2).

Organization and cellularity of the intestines

The intestinal epithelium is arranged into distinct zones of proliferative and differentiated intestinal epithelial cells (IECs)(3)(Fig. 1-2). In the small intestines, differentiated cells form the villi, while proliferative cells are found within invaginations embedded in the submucosa, known as crypts of Lieberkühn (3). In the colon, villi are absent and differentiated cells are instead located in the top half of the crypts, and across the top of the epithelium. The most abundant differentiated IECs are the enterocytes, (referred to as colonocytes in the colon). These cells are primarily responsible for absorbing water, metabolites and nutrients, but also play a role in sensing and responding to the microbiota (5, 6). To facilitate these roles, the apical side of each enterocyte is covered in smaller tendril-like projections called microvilli that further enhance the surface area of the intestines for absorption (4). In the large and small bowel, differentiated

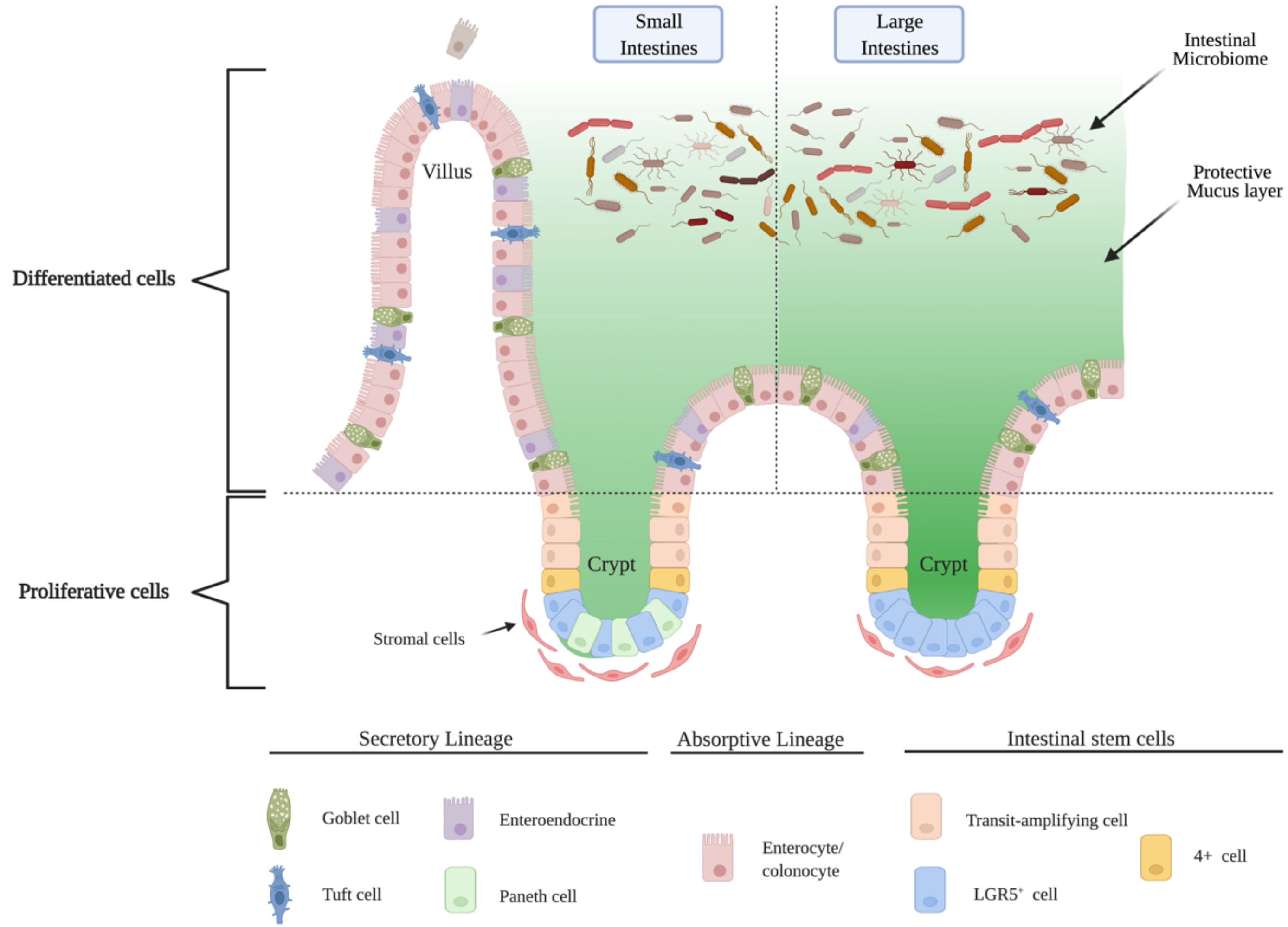

Small Intestines
Large Intestines
Intestinal Microbiome
Protective Mucus layer
Villus
Differentiated cells
Proliferative cells
Crypt
Crypt
Stromal cells
Secretory Lineage
Absorptive Lineage
Intestinal stem cells
Goblet cell
Enteroendocrine
Tuft cell
Paneth cell
Enterocyte/ colonocyte
Transit-amplifying cell
4+ cell
LGR5$^+$ cell

Figure 1-2. Organization of the intestinal epithelium. A single layer of IECs protects the submucosa from the contents of the intestinal lumen, and is composed of differentiated cells nearest the lumen, and non-differentiated, proliferative crypts. The IECs are separated from the microbiome through a thick mucus layer secreted by goblet cells. Goblet cells, other members of the secretory lineage, and the absorptive lineage are derived from LGR5+ stem cells at the base of the crypts. Figure created with BioRender.com.

enteroendocrine cells sense nutrients and secrete hormones to facilitate the digestive process. Paneth cells are found only in the small intestines and release anti-microbial peptides (7). Goblet cells are abundant in the colon and secrete a layer of mucus that protects the IECs and assists the passage of stool in the large intestines (8). Of all IECs, the two rarest and most poorly understood are the tuft and M cells. Derived from the secretory lineage, Tuft cells compose ~0.4% of IECs, but lack specific molecular markers making determining their function a challenge (9). However, it appears these cells play a role in immune signaling through secretion of IL-25 (10). Like the rest of the IECs, tuft cells are derived from the proliferative cells of the crypt, though express a more unique appearance, as cylindrical epithelial cells with apical microvilli (9, 10). A final absorptive cell type, M cells, fill an important role as an intermediary between the lumen and submucosal immune cells by delivering foreign luminal material to lymphoid tissues (11).

Due to the harsh environment of the GI lumen, differentiated IECs are routinely lost due to either damage or controlled cell death. To maintain the integrity of the IEC barrier, these cells must be rapidly replaced. In fact, the entire IEC layer is replaced every 5-6 days, making the intestines one of the most proliferative and regenerative organs in the body (3, 12). The replacement IECs are derived from intestinal stem cells (ISC) residing at the base of the intestinal crypts. Found in both the small and large intestines, the crypts provide protection from chemical and mechanical stress, and house a diverse and complex proliferative niche (13). Residing at the bottom of these crypts are crypt base

columnar cells (CBC) which serve as the main stem cell (SC) population (14, 15). Also known as LGR5$^+$ SCs, CBC cells express the leucine-rich repeat-containing G-protein-coupled receptor 5 (LGR5), an R-spondin receptor critical for their function as SCs (15). In response to extracellular ligands, CBCs rapidly proliferate and produce daughter cells. These cells renew the SC population, and push daughter cells up the crypt axis towards the lumen where they will selectively transition into transit-amplifying (TA) progenitor cells (3). As they migrate away from the crypts, cell signaling dictates terminal differentiation into IECs of the secretory and absorptive lineages (3). This process gives rise to another SC type in the crypt, the quiescent +4 SC, named after its location in the crypt as approximately four cells up from its base (16). Unlike other cells produced by the TA cells, the quiescent SC, along with the Paneth cells, migrate down to reside near the base of the crypt (3). At the base of the crypts, Paneth cells, CBCs and LGR5$^+$ cells play integral roles in maintaining homeostasis of the entire IEC barrier. The Wnt signaling pathway plays a critical role in driving proliferation of the SC niche.

Wnt signaling

Canonical Wnt signaling pathway

Wnt is a family of highly evolutionarily conserved ligands that function to stimulate cell growth and proliferation (17). Their discovery in mammals stemmed from work in the lab of Harold Varmus, who identified *Int1* (for, integration site 1) while searching for breast cancer oncogenes caused by modeling disease through integration of mouse mammary tumor virus (MMTV) (18, 19). Genetic

mapping studies eventually identified *Int1* as an ortholog of the *Drosophila* wingless (*wl*) gene discovered a decade earlier (20, 21). Through a mnemonic portmanteau, *Int1* was renamed *Wnt1,* and *wl*-related genes were renamed as other members of the Wnt family (22). Deletion experiments of mammalian *Wnt* genes in mice revealed developmental defects (e.g., Wnt1 (23), Wnt3a (24) Wnt4 (25)), much like the developmental phenotype of *wl* deletion in *Drosophila* (26). Together, the body of literature surrounding Wnts illustrates the importance of Wnt signaling in development across the animal kingdom.

Specifically, Wnt ligands are secreted glycoproteins that act in a short-range fashion, influencing cells neighboring those that are secreting the ligand (7). The Wnt family is composed of 19 *Wnt* genes in mammals, each of which produces a single Wnt protein (13, 27). The downstream outcome of each Wnt is determined by a complex signal transduction cascade that influences one of three pathways: the non-canonical Wnt/calcium signaling, cell-polarity, and the canonical Wnt/β-catenin (13, 28). All three pathways begin with the binding of a Wnt ligand to a heterodimeric receptor complex composed of a frizzled (FZD) receptor and low-density lipoprotein (LDL) receptor related 5 or 6 (LRP5/6). Downstream from FZD/LRP, the scaffold protein, Disheveled (Dvl), provides signal specificity to each of the Wnt pathways. In the context of binding the FZD receptor, there is limited functional selectivity between the various Wnt ligands and receptors, allowing for activation from different upstream pathways (29). However, the Wnt/β-catenin pathway differs from the noncanonical pathways

through its reliance on a primary regulator, β-catenin (β-cat), to transduce a nuclear gene expression response (13, 28).

β-cat is a transcriptional co-activator that drives gene expression through interaction with the T-cell factor/lymphoid enhancer factor (TCF/LEF; hereafter TCF) family of sequence specific transcription factors (30-32). In the absence of Wnt ligand, β-cat associates with a multi-protein "destruction complex" within the cytoplasm (Fig. 1-3). Here, adenomatous polyposis coli (APC) and axis inhibition protein 1 and 2 (AXIN1/2) serve as scaffolding proteins, and bind β-cat (13, 33). Also found associated within the destruction complex are the kinases, casein kinase 1 (CK1) and glycogen synthase kinase 3 beta (GSK3β). APC and AXIN1/2 position these kinases in close proximity to β-cat. CK1 phosphorylates β-cat at its Ser45 residue, followed by GSK3-mediated phosphorylation at Thr41, Ser37 and Ser33 residues (34). Phosphorylated β-cat is then bound by and subjected to β-transducin-repeat-containing protein (β-TrCP)-mediated ubiquitination, and subsequent proteasomal degradation (35, 36). This maintains a low level of cytoplasmic β-cat, and repression of downstream signaling. In the absence of Wnt, TCF binds Wnt-responsive DNA regulatory elements (termed WREs) and recruits Groucho/histone deacetylase complexes, to repress target gene expression (27, 37, 38).

Conversely, when Wnt is present, it binds the FZD receptor and its co-receptor, LRP5/6 (13, 27). The cytoplasmic tail of LRP is phosphorylated by several kinases, including GSK3, leading to the Dvl-mediated binding of AXIN1/2 (13). Recruitment of the destruction complex to the plasma membrane

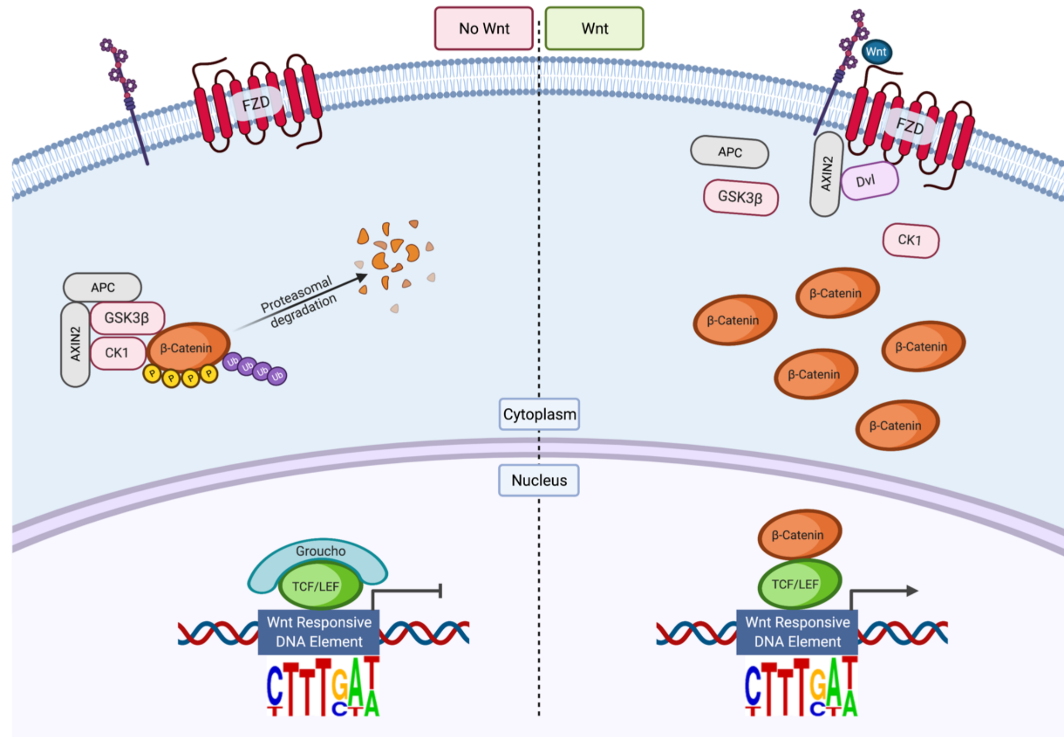


Figure 1-3. Illustration of the canonical Wnt signaling pathway. In the absence of Wnt signaling (left), β-catenin is targeted for degradation by the destruction complex (AXIN1/2, APC, GSK3β and CK1). Within the nucleus, Wnt responsive DNA elements are bound by the TCF/LEF family of transcription factors, which recruit corepressors such as Groucho to transcriptionally repress Wnt target genes. In the presence of Wnt (right), Dvl-mediated binding of AXIN1/2 to the cytoplasmic membrane stabilizes β-catenin. β-catenin translocates to the nucleus, binds TCFs and promotes expression of Wnt target genes. Figure created with BioRender.com.

inactivates it through mechanisms not fully understood. Nevertheless, β-cat is stabilized in the cytoplasm and subsequently translocates to the nucleus where it is recruited to DNA through interaction with TCFs (38). The TCF/β-cat complex binds CREB binding protein (CBP), and additional histone acetyltransferases (HATs) and histone methyltransferases (HMTs), to promote the recruitment of RNApol II and subsequent transcription of target genes (27).

TCF family

Whereas *CTNNB1* is the only gene coding for β-cat, there are four mammalian TCFs, which provide tissue specific and context-dependent response to Wnt signals (27). All TCF family members recognize a 5'-SCTTTGATS-3' motif with nanomolar affinity through a conserved high-mobility group (HMG) DNA binding domain (27, 31, 39). To fill their diverse roles and expression profiles, these family members, TCF7 (formerly TCF1), LEF1, TCF7L1 (formerly TCF3) and TCF7L2 (formerly TCF4) express several isoforms that lack domains, or domain fragments, broadening their specific roles in the nucleus (40-42) (Fig. 1-4). As such, simple expression of a family member is not a direct indication of their influence on transcription, nor can it be assumed they are functionally redundant.

Despite the first discoveries of TCFs in lymphocytes, earning their namesakes, lymphoid enhancer factor (43), and T-cell factor (30), TCFs are expressed in the gut and play an important role in intestinal development and maintenance (44). TCF7L2 is the most well studied TCF family member in the intestines. Homozygous deletion of *Tcf7l2* in mice results in neonatal lethality and

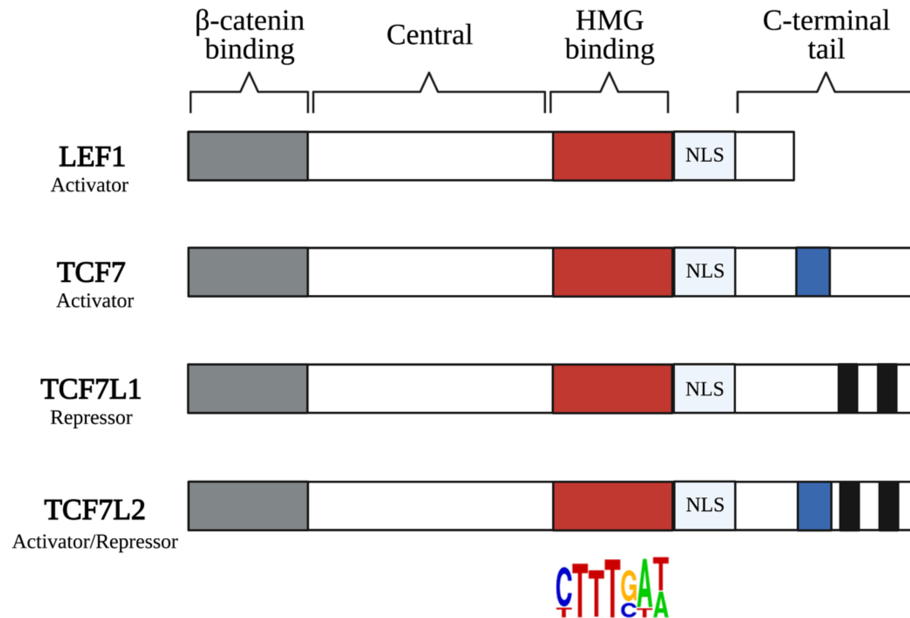


Figure 1-4. Domains in mammalian TCF family members. Structure of each TCF includes domains labeled across top. Alternative splicing to alter or remove conserved domains within the C-terminal tail, and β-cat binding domain account for some functional differences. TCF7 and TCF7L2 use C-clamp domains (blue) to bind an additional section of DNA during activation of target genes, while TCF7L1 and TCF7L2 use CtBP binding domains (black) in the absence of β-cat to recruit CtBP and repress target genes.

a severe intestinal defect (45). Moreover, *Tcf7l2* null mice develop an entire epithelial layer of differentiated cells with no proliferative crypts (45). As such, Tcf7l2 is not required for initial induction of the intestinal epithelium, yet maintenance was fully dependent upon its expression. Further work illustrated that the DNA binding domain of TCF7L2 is required for it to serve as an effector of the Wnt/β-cat pathway (46). Using an intestinal specific deletion of Tcf7l2's HMG domain, Van Es et al., observed proliferative defects so severe that intestinal crypts ceased to exist after 9 days (46). Additional data have shown maintenance of intestinal crypts is further dependent on signaling upstream of β-cat bound TCF7L2, as inhibition of Wnt signaling results in a loss of intestinal crypts (47, 48). Together these data demonstrate that proper crypt maintenance and homeostasis is reliant upon activation of Wnt/β-cat target genes driven by the binding of TCF7L2 to DNA, and recruitment of nuclear β-cat.

Wnt in the IEC barrier

As mentioned previously, Paneth cells of the small intestines are capable of producing a Wnt ligand that stimulates growth. Work by Hans Clevers' group illustrated *ex vivo* growth of CBCs can be maintained for an extended period of time as organoids (49), owing to the fact that these cells can produce the Wnt3a ligand, enabling self-renewal (50). Removal of Wnt3a production in these CBCs *in vivo* has no effect on growth, while organoids cultured *ex vivo* rely on Wnt3a for survival (51). It has since been confirmed that mesenchymal cells (or stromal SC) produce their own set of Wnt ligands, capable of compensating for the loss of CBC produced Wnt3a to maintain growth (51-53). Further redundancies exist

through the capabilities of enterocyte dedifferentiation, or B Lymphoma Mo-MLV Insertion Region 1 Homolog (Bmi1)-expressing cells, which can repopulate lost CBCs (54, 55).

Wnt ligands are only capable of moving cell-to-cell, and thus its impact is constrained to neighboring cells. Therefore, within the intestinal epithelium, a gradient of Wnt-driven proliferation is observed, where its expression is highest at the base of crypts (7). This gradient maintains the proliferative niche for SC renewal, while permitting TA cells to differentiate as they migrate further from the crypt. Indeed, due to the critical nature of Wnt signaling in development and maintenance of IECs, Wnt signaling is achieved and maintained by epithelial and non-epithelial cells in multiple redundant and well-tuned mechanisms.

c-MYC as a target of Wnt signaling

The downstream impacts of a pro-proliferative signaling pathway lend support to their involvement in a cell-type and developmental stage-dependent manner. As such, Wnt signaling plays a role in neural, immune, intestinal, skin, hair and early embryonic development (42, 44, 45, 56-59). The connection between Wnt signaling and disease was drawn from the identification of *APC* mutations as drivers in a hereditary form of colorectal cancer (CRC), familial adenomatous polyposis (FAP) (60, 61). Indeed, mutations in *APC* were identified in not only FAP, but also approximately 90% of sporadic CRC cases (62, 63). These findings strongly implicated deregulation of Wnt signaling as a critical underlying contributor to CRC, making it essential to identify gene targets of the pathway to understand disease pathogenesis. Using a differential gene expression

screen in CRC cell lines, an influential report by Bert Vogelstein's group identified the *c-MYC* proto-oncogene (hereafter referred to as "MYC") as a direct Wnt/β-cat target (64).

MYC is a member of the basic helix-loop-helix leucine zipper (bHLH-lz) family of transcription factors (65). It is a 439 amino acid nuclear protein with the bHLH DNA binding domain localized to the carboxy terminal third of the protein, and an acid rich transcriptional activator domain (TAD) localizing to the amino terminal third (66). Through MYC's bHLH-lz domain it dimerizes with MAX, which shares high homology with MYC in its bHLH-lz region, but lacks a transcriptional activator domain (67). MYC:MAX heterodimerizes to bind E-box motifs with the consensus of CAGGTG embedded within DNA enhancer elements (68). Through its TAD, MYC recruits histone modifying complexes such as transformation/transcription domain associated protein (TRAAP) / general control non-derepressible 5 (GNC5), HIV-1 tat interactive protein, 60kDa (TIP60) histone acetyltransferases, and chromatin remodeling complexes including SWI/SNF, to modify chromatin activity to facilitate RNA polymerase II (RNAP) binding and subsequent activation of gene expression (69-71).

The downstream effects of MYC activation are widespread and genome-wide expression profiles have shown approximately 15% of all known protein coding genes are regulated by MYC (72). MYC regulation of target genes is cell specific, and functions through direct binding to E-box domains, as well as indirect means. Through binding to the transcription factor MIZ-1, for example, MYC can be recruited to non-CACGTG sites where it represses MIZ-1 targets

(73, 74). As a transcription factor, MYC influences cell cycle regulation, apoptosis, protein synthesis, ribosome biogenesis, and cell adhesion (65, 72). These effects are widespread and can be challenging to clarify. Indeed, mouse fibroblasts lacking MYC have a near 3-fold reduction in protein synthesis and a lengthened G1 and G2 phase of the cell cycle (75). Identifying the difference between physiological targets and pathogenic targets remains a critical goal for MYC-related research (72).

Beyond the critical influence in CRC development mentioned above, elevated expression of MYC occurs in approximately 50% of all human cancers (76). This deregulation of MYC can be achieved through two mechanisms: modifications to *MYC's* locus through changes including gene amplification or translocations, or, the deregulation of upstream signaling pathways. The less common form of MYC deregulation occurs through translocations of the *MYC* locus, which are present in only a specific subset of cancers. The strongest example is Burkitt's lymphoma, where *MYC* translocations occur at a frequency of 100%, and is actually a means to diagnose the cancer (77-79). MYC deregulation through amplification of its locus is the most common form of sequence change to the *MYC* locus in solid tumors (80). However, *MYC* is often deregulated through manipulation of external signaling pathways that stabilize or overexpress MYC, while the protein coding sequence of *MYC* is not mutated. As alluded to earlier, mutations of APC, or other Wnt signaling pathway components, result in elevated expression of *MYC* and release its oncogenic potential in CRC (62, 63). Due to the frequency by which it is mutated, the accepted model of sporadic CRC

development posits a key set of mutations in the Wnt/β-cat signaling pathway drive disease pathogenesis (81). This model is supported by a comprehensive analysis of CRC mutations published by the cancer genome atlas (TCGA) outlining 80-90% of all sporadic cases contain mutations in the Wnt signaling pathway (82).

Several WREs have been identified that control *MYC* expression in CRC (Fig. 1-5). Initial discoveries by He et al. identified the 5'WRE located within the *MYC* proximal promoter (64). Using luciferase reporter assays and EMSA, three distinct TCF binding elements (TBEs) were identified as critical determinants of enhancer activity (64, 83). Through a β-cat ChIP-Seq screen in CRC cells many years later, a second WRE was identified (84, 85). Interestingly, this site overlapped a CRC-specific DNAse I hypersensitivity site identified in 1995 (86). Using chromatin confirmation capture (3C), Yochum et al. found that β-cat/TCF coordinated a 5'-3' chromatin loop that integrates WREs to control *MYC* expression (87). A third WRE, -335WRE, is located 335kb upstream of *MYC* and is also juxtaposed to the *MYC* promoter to influence its expression (88, 89). Importantly, an additional study showed that overexpression of β-cat and TCF7L2 increased the frequency of interaction, illustrating direct effects of upstream pathway overactivation (90). Lastly, research by Bottomly et al., not only confirmed binding of β-cat to all three of these WREs, but also identified a novel region, later identified as a cluster of WREs known as a "super enhancer," mapping approximately 500 kb upstream from the *MYC* transcription start site (91-93). This collection of WREs are found in regions of open chromatin, overlap

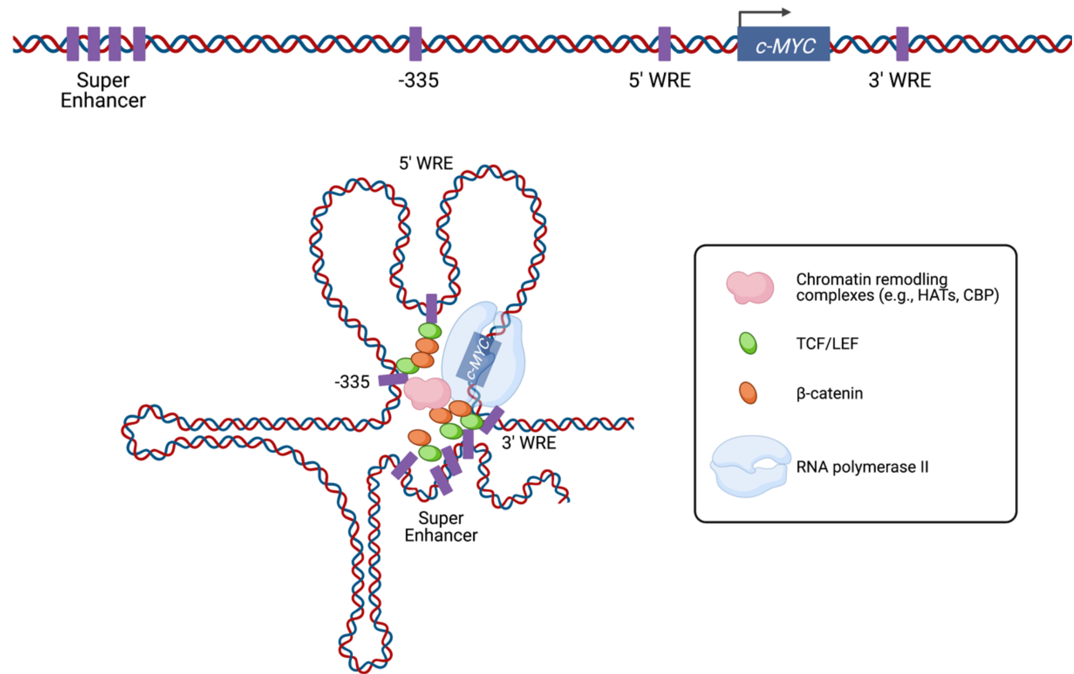


Figure 1-5. Regulation of *c-MYC* by Wnt/β-cat signaling.Top, diagram of WREs that control *MYC* expression. WREs are represented by purple boxes. Bottom, model of looping chromatin interactions that allow juxtaposition of distal WREs to *MYC*. Figure created with BioRender.com.

with TCF7L2 binding peaks, and function as enhancers (92). Critically, these data show how distal enhancers can affect *MYC* expression through binding of β-cat and TCFs, the effect of which can change depending on the amount of β-cat available. Thus, MYC expression is governed by multiple WREs to allow for "just right" levels of MYC expression to promote colorectal carcinogenesis (94). In addition, examination of patient samples, animal models and intestinal epithelial cell lines have shown the Wnt/Myc axis plays a role in inflammatory bowel disease (95-99).

Inflammatory bowel disease

<u>Definition and types of IBD</u>

Crohn's disease (CD) and ulcerative colitis (UC) are the two main classes of inflammatory bowel disease (IBD), and arise from chronic, recrudescent inflammation in the GI tract (100). CD and UC have a distinct yet overlapping presentation in patients, and can range in severity from virtually asymptomatic to intense pain and severe inflammation. UC presents as mucosal inflammation beginning at the rectum and proximally extends to include, at maximum, the entire large intestines (101). Inflammation can be mild, or present as characteristic mucosal ulcerations and may result in an inflamed lamina propria, but is typically restricted to the mucosa. In contrast to UC, which is restricted to the colon (102), CD presents as a non-contiguous inflammation and can occur anywhere along the GI tract (100, 103, 104). Illustrating the macroscopic changes undergone in a patient with CD, figure 1-6 shows a surgically resected colon from a CD patient,

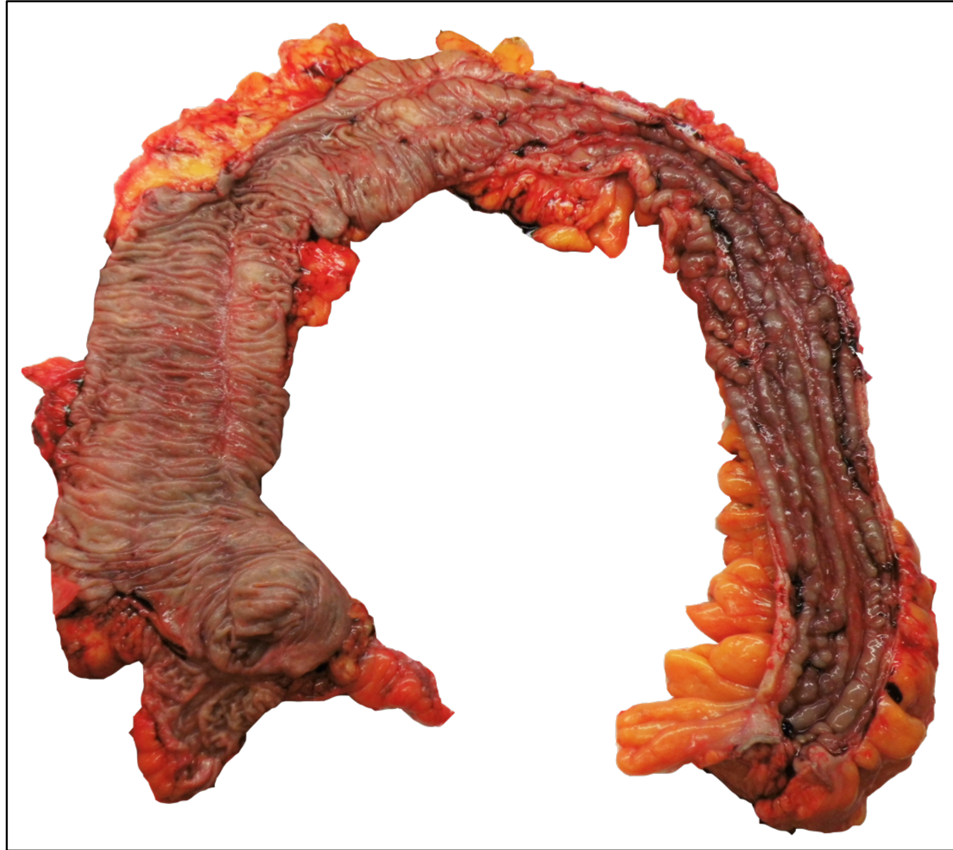


Figure 1-6. Photograph of total abdominal colectomy from patient with CD. Colon tissue is cut longitudinally to expose intestinal epithelium, and shown is the ascending, transverse and descending colon (clockwise starting on the left). Note the presence of normal colonic tissue on left which gradually transitions to severe inflammation marked by red tissue and coalesced ulcerations presenting in a "bear claw"-like appearance. Figure courtesy of the IBD and intestinal diseases biobank.

with clear distinction between the diseased segment on the right and the normal bowel on the left. Specifically, CD involves transmural inflammation, thickening of the bowel wall and presents as a "cobblestone" pattern of coalesced ulcerations (101). CD patients suffer from frequent macroscopic complications, that rarely present in UC, including strictures and fistulae which often require surgical intervention (101).

IBD treatments

There is no cure for IBD, however, several treatment regimens are available to help patients manage the symptoms of their disease. Unfortunately, IBD is a complicated disease to manage and patients often require several different mediations to induce and maintain remission over their lifetime. The method of treatment is determined through standardized protocols such as the Montreal classification of IBD, which is a scoring classification based on disease extent and severity (102, 104, 105). Medical management for UC and CD are similar, however due to differences in pathobiology, not all treatments are effective in both subsets of IBD (104). Mild disease can often be treated with aminosalicylates (5-ASA), which typically result in marked improvement in 60-80% of patients (106). However, 5-ASA therapies require supplemental steroids to control flare-ups, followed by progression to immunomodulatory agents in 5-ASA non-responders (102). Patients with moderate to severe disease frequently require hospitalization, but even in these cases, up to 40% of these patients will not respond to corticosteroids, and require more aggressive treatments (102). One of the most common immunomodulatory agents for severe disease is infliximab, an

anti-TNF (tumor necrosis factor) therapy with a 50-70% response rate in UC (107, 108), and ~86% response rate in CD (109).

Importantly, despite all the available treatments, up to two thirds of patients will not respond to medical management (102, 104). Moreover, disease remission is problematic in both CD and UC, which requires additional medications or interventions. If medications fail, patients may require prophylactic surgery to treat their disease. Indeed, within the first 10 years of diagnosis, the cumulative likelihood of undergoing surgery for UC is 15-25%, and ~47% for CD (110, 111). While surgery is curative for UC patients, and highly beneficial for CD patients, IBD patients receiving such treatments are prone to complications and suffer from lower quality of life. Therefore, a better understanding of disease pathogenesis, pathobiology and pharmacogenetics of UC and CD has been a long-term goal in the field of IBD research.

Immune system in IBD

IBD is a multifactorial disease and as a result, numerous external elements contribute to its pathogenesis. The phenotypic presentations of chronic colitis are underwritten by a complicated and unique combination of converging factors involving the innate and adaptive immune system, and their response to environmental factors, including the microbiome. The innate immune system includes macrophages and dendritic cells that sense luminal bacteria and antigens, and circulating neutrophils (100, 103). In general, the innate arm of the immune system senses pathogenic insults and signals to the adaptive arm consisting of T-cells and B cells (100, 103). A T-cell response is required to quell the initial innate

immune response and promote tissue repair. Unfortunately, in the chronic state of IBD, a deregulated T-cell response contributes to persistent epithelial barrier degradation in a feed-forward inflammatory response that never truly resolves. To aid in patient recovery, repression of this response may be necessary.

The highly effective response of many patients to immunomodulatory therapies like infliximab supports the critical influence of the immune system in IBD. As covered later in this chapter, several defects in pathogen recognition and autophagy pathways have causative and supportive roles in the development of IBD. Cytokines released by immune cells of the innate and adaptive variety play a critical role in inducing and maintaining an inflammatory response (103). Seminal work in the early 1990's illustrated that mutations in IL-10, IL-2 or T-cell receptors result in extensive intestinal inflammation reminiscent of chronic enterocolitis (112-114). Cytokine profiles largely overlap, but $CD4^+$ T cell activation and differentiation appears to present differently in UC vs CD. Upon presentation of antigens to naïve T-cells, the milieu of signaling molecules and antigens elicit a response within these cells initiating their differentiation into specific subtypes of $CD4^+$ T-cells (115). This stage of antigen sensing, and subsequent presentation of antigens, is a key interplay between the innate and adaptive immune response. While there is overlap between UC and CD in this regard, there is a bias of differentiated T cells, wherein T helper 1 cells (Th1) and Th17 are more common in CD, while a Th2 profile is more prevalent in UC (116). These lineages of $CD4^+$ cells respond to and elicit differing downstream inflammatory pathways and cytokine profiles, yet are critical for host immune

response to pathogens (115). Owing to the critical interplay between the immune response and the microbiome, intestinal colitis in immune-deficient murine models fails to develop in the absence of the resident intestinal microbiome (113, 117, 118). Additional supporting data indicates intestinal colitis involves overactivation of T-cell populations responding to native gut microflora (119-121). The "hygiene hypothesis" further supports the overactivation of the immune system. This theory posits that early exposure to pathogens is required to develop a healthy immune system (122), and failure to do so may result in excessive immune response when pathogens are introduced (123).

While an overactive immune response is critical in development in IBD, the microbiome itself can play a role through its dysregulation, or dysbiosis. Specifically, the gut houses tens of thousands of bacterial species, with the vast majority represented by the Bacteroides, Fermicutes, Proteobacteria, and Actinobacteria phyla (124). However, the relative ratio of these prevalent bacterial phyla starkly differs between patients with IBD to those without (125, 126). While determining "cause verses effect" is a challenge, the successes of fecal microbial transplantations (FMTs) offer support for modulating the microbiome to protect against IBD (127, 128). FMT has been used with success to manage *Clostridium difficile*-driven intestinal inflammation (129), which may provide some respite for other IBD patients as we better understand the role of microbial dysbiosis in disease pathogenesis.

The intestinal epithelium in IBD

It is clear that the IEC barrier is critical for separating luminal contents of the intestines and the submucosa, yet IECs require attachment to one another to maintain this barrier. IECs attach to one another through tight junction (TJ) proteins located on their lateral sides. TJ proteins are composed of occludens, claudins, junctional adhesion molecules, and zonula occludens which operate as a network that maintains firm cell-cell adhesion (130). Altered expression of IEC TJ proteins has also been reported in IBD, and are expected to lead to increased IEC permeability in these patients (131, 132). Indeed, murine models expressing deficiencies in TJ proteins result in elevated permeability of the IEC barrier, and a more severe colitis, but their defects alone do not induce spontaneous colitis (131, 133).

Another aspect of the innate immune system and IEC interplay is the protection offered by the goblet cells. Mucins secreted by goblet cells of the small and large intestines coat the IEC barrier in a thick layer that serves as another physical obstacle. Illustrating the importance of this protection, perturbations of mucus production and goblet cell prevalence has been reported in IBD (134). Further, knockout of the most prevalent intestinal mucus glycoprotein, *MUC2*, resulted in spontaneous colitis (135), and owing to the interplay in the IEC barrier, increased intestinal permeability (136). Additional work has also shown that mucus is not only protective against intestinal inflammation, but elevated levels of specific mucins may delay time to surgery in patients with UC (137).

The innate and adaptive immune systems are affected by, and respond to, more than just changes in the microbiome. Other environmental factors contributing to IBD pathogenesis include: geographic location, where physical location results in altered gut microbiomes, and incidence of IBD is higher in more developed countries (138-140); diet, which can affect microbiome composition and overall health (141, 142); and medications, where nonsteroidal anti-inflammatory drugs (NSAIDs) are associated with an increased risk of developing IBD (143).

Genetics of IBD

Genetic basis of IBD

While the precise influence and contribution of various environmental factors is debated, there is strong evidence for genetic predisposition for IBD. The generally accepted model for IBD pathogenesis stems from one or more environmental triggers converging in a genetically susceptible individual (103, 141, 144). Support for genetic predisposition stems from several observations including that 5-23% of IBD patients have a first degree relative with the disease (145), and incidence in developing IBD is even higher if both parents have IBD (146-148). Initial twin studies have also shown genetics plays a pivotal role, as phenotypic concordance in monozygotic twins is up to 50% in those with CD, but a much lower ~20% in UC (145). Family lineage tracing, twin and general heritability studies have significantly advanced our understanding of the genetic underpinnings of IBD, not only in relation to the importance of genetic inheritance, but also the close relationship between genetic and environmental

factors. However in IBD, which is a non-mendelian disease, twin and family studies are underpowered for identifying all polygenic sources of disease due to small sample size, low number of variants, and restrictions to one familial genetic background (149). Indeed, ethnicity and familial backgrounds contain their own unique risk factors. One well-established example is the higher propensity for Ashkenazi Jewish populations to develop IBD over non-Jewish Europeans regardless of geographic locale, despite holding similar genetic architecture (150). While still having their strengths, improving genetic linkage mapping through association studies could examine a larger population, and provide novel insights into the disease (149).

Recently, genome-wide association studies (GWAS) have enabled the genotype-phenotype comparison of various populations to better elucidate important genetic variations associated with disease predisposition. Conceptually, GWAS operate to identify genetic markers that are more prevalent in the genome of a population with a particular disease compared to those without (Fig. 1-7). Due to assembly of a reference genome and the annotation of common genetic variants by the Human Genome Project and HapMap Project, hundreds of thousands, or in some instances up to a few million, of these variants were mapped in thousands of individuals (151-154).

GWAS identify genetic variants most commonly through single nucleotide polymorphisms (SNPs) that occur at a known location within the reference panel. The prevalence of a SNP is defined through its minor allele frequency (MAF), which is the regularity by which the less common allele is present; wherein the

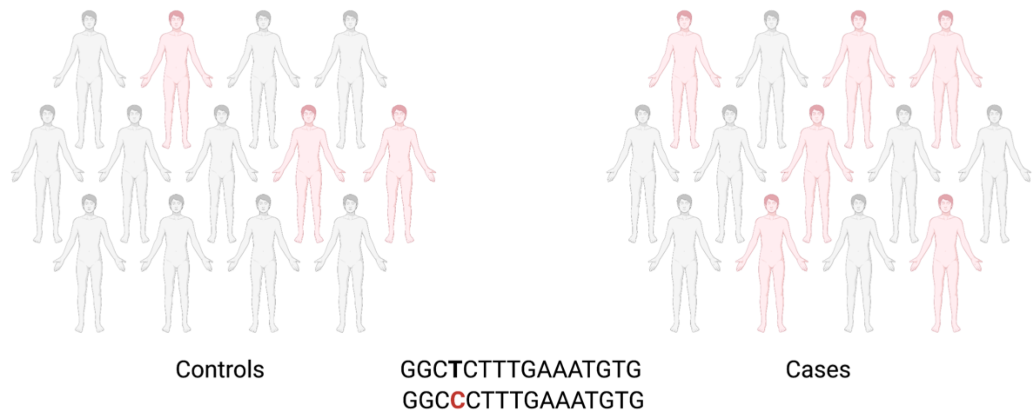


Figure 1-7. Conceptual representation of GWAS. GWAS aim to identify the prevalence of a particular SNP between two populations. Grey individuals represent those that possess black **T** allele, while red individuals possess the red **C** allele. By comparing controls to cases, there is an increased frequency by which the C allele is present, thus identifying C as a disease-associated allele correlating with the disease phenotype. Additional studies are required to determine if the SNP is causal (or merely in an LD block with the variant), and how the disease-associated variant influences disease. Figure created with BioRender.com.

lower the frequency, the greater the rarity of the SNP. A collection of SNPs are used to mark regions of chromosomes frequently inherited together in what are called haplotypes. Each SNP utilized in a GWAS is selected in regions that cluster together such that the genotype of one SNP results in an inference of the presence of those in the same block. These haplotypes can differ depending on the reference panel used, but marker SNPs are chosen to represent a block that historically has low frequency of recombination (149).

GWAS have identified SNPs associated with several common diseases including macular degeneration, diabetes, heart disease and IBD (144, 155, 156). Locations of these disease-associated SNPs are indicative of genomic regions that may contribute to disease. However, due to linkage disequilibrium (LD), the identified SNP is not necessarily the causal variant, rather merely indicates that locus is associated with the disease. As such, identifying the specific genetic mutation that associates with the disease may require additional fine mapping studies of patient haplotypes. Additionally, some SNPs returned from genetic testing are identified as variants of uncertain significance (VUS), which are genetic changes where the effect on health is unknown.

For polygenic diseases like IBD, identification of the most common disease-associated SNPs still provides only small increases of overall risk, with very little proportion of heritability and contribution to disease development. Indeed, in the many hundreds of GWAS conducted, the majority of SNPs carry risk ratios around 1.2 (157). Risk ratios are the measurement of association, wherein the greater the value is over 1.0 the more associated the variant is with

disease. By virtue of GWA study design, rare variants with larger penetrance, and common variants with low penetrance will be missed (157, 158). This is the foundation for the theory of "missing heritability," wherein complex traits and diseases have multiple factors that contribute to its heritability and cause, yet most remain undetected (159). Indeed, the large number of loci identified still accounts for less than 20% of the IBD disease phenotype (160, 161). Despite this limitation, GWAS have been very successful in identifying novel associations between traits and variants and novel mechanisms. Several examples are the discovery of CD-associated SNPs within the genes: nucleotide-binding oligomerization domain 2 *(NOD2),* autophagy-related 16 like 1 *(ATG16L1)* and immunity related GTPase M *(IRGM).* Several variants present in *NOD2* affects pathogen recognition, and contain risk ratios of up to 17.0 in Caucasian's homozygous for the mutation, labeling these variants as containing the strongest effect size of any currently known CD-associated SNP (161). Identification of *ATG16L1* and *IRGM* variants have lower risk ratios, yet provided novel insight into autophagy as a pathway involved in IBD.

Jostins et al., reported a meta-analysis of GWAS data from patients with sporadic and familial IBD, identifying 163 SNPs contributing to IBD susceptibility (162). Of these, 110 were common to both UC and CD, while 30 were specific for CD, and 23 for UC (162). Data generated in conjunction with the International IBD Genetics Consortium has now compiled a list of variants that exceeds 200 distinct loci (160-165). Unfortunately, risk loci identified by GWAS span a variable range, which can include between 0 and 50 genes, rendering the

exact effect of the variant unknown (166). A minority of these loci are demarcated by SNPs that reside within genes, resulting in a direct change in the sequence of a protein. Such examples include the intragenic variants in *NOD2*, *ATG16L1* and *IRGM* mentioned previously (161).

DNA regulatory elements in IBD

While these results clearly demonstrate the effectiveness of GWAS in identifying novel pathways and strong variants in polygenic diseases, approximately 80–90% of the identified alleles map to non-coding regions of the genome (162). Interpretating the impact of these intergenic SNPs is a significant challenge, particularly if the SNPs reside in gene-poor regions. In these instances, many SNPs have been mapped to the nearest gene, with assumptions made that those SNPs influence their expression due to their close proximity. However, this is not an accurate means of determining the target, and is biased by prior knowledge on current gene annotations, or the perceived relevance of its pathway on the disease in question. Methods such as expression quantitative trait loci (eQTL) infer a direct interaction between expression levels of a gene and a variant through correlation of gene expression with the presence of a variant (144, 166). This mapping technique allows for putative annotations of SNPs identified in GWAS. Presumably, non-coding SNPs impact target gene expression through the function of DNA regulatory element (DRE). However, the mere presence of a SNP in a non-coding region does not indicate the region is a DRE. First, DRE are cell-type specific and must be confirmed to function within cell types relevant to the disease. Additionally, interactions between a gene and a DRE can occur over

distances of up to 1Mb. Within this region numerous genes could be a target of the DRE, and lastly, any gene within that range could be a potential target.

Pivotal work from the encyclopedia of DNA elements (ENCODE) consortium and others have indicated the majority of common IBD-associated variants map to DREs within the genome, in regions known as enhancer elements (166-169). Epigenetic markers can be identified through chromatin immunoprecipitation sequencing (ChIP-Seq), and are enriched at critical regions of the genome including promoters (e.g., H3K4me3) and active DNA enhancers (e.g., H3K4me1 or H3K27Ac) (158, 170). In the context of IBD, Mokry et al. found that nearly 30% of IBD-associated SNPs directly overlapped with H3K27Ac marks (168). Further, H3K27Ac overlapped transcription factor binding sites, and when accounting for LD blocks, wherein the identified SNP may not be causative, the total number of SNPs associated with active enhancers increased to 56% of the 163 SNP identified at the time (168). Importantly, Mokry et al. and Huang et al. have shown that several IBD-associated SNPs directly alter transcription factor binding motifs in both intestinal and immune cell regulatory elements (167, 168).

Localization of a SNP to a region of open chromatin, and alteration of transcription factor motifs are critical first steps, but additional work is needed to confirm their impact on gene expression and disease. Regions identified in this fashion mark putative enhancers, which can act distally, upon multiple target genes and may even skip nearby targets to influence those up to 1 Mb away. One high-throughput method of identifying target genes of putative SNP-containing

enhancers is through circular chromosome conformation capture-sequencing (4C-seq). Using this technique, Meddens et al. has identified novel localization of disease-associated DRE to candidate genes in immune and IECs (171). While several IBD-associated SNPs were confirmed to demarcate enhancer elements (171), the precise impact of these variants on enhancer function, the upstream signaling pathways involved, and the relevant downstream transcription factors have not been adequately addressed. Furthermore, the target genes and influences of the majority of IBD-associated SNPs still remain undiscovered. Clarification of their targets is a critical and active area of research that holds potential to identify novel pathways, novel therapeutic targets, and to better clarify IBD pathogenesis.

Dissertation overview

The human gastrointestinal (GI) tract is an organ system responsible for digesting food, extracting nutrients and expelling solid waste matter. As the core of the digestive system, the GI tract forms a continuous tube that encompasses all connected organs between the mouth and the anus. The luminal contents of the intestines are separated from the underlying tissue through a protective single layer of intestinal epithelial cells (IECs), known as the intestinal epithelium (3). The IECs serve not only to protect the underlying tissue, but also act as the primary contact to the lumen through which nutrients, salts and water are sensed, secreted and absorbed. Due to the harsh environment of the GI lumen, IECs are frequently lost and must be rapidly replaced. Replacement cells are derived from intestinal stem cells (ISC) residing at the base of invaginations in the epithelial layer known as crypts of Lieberkühn (3). Driving the proliferation of the ISC is the

canonical Wnt signaling pathway. Wnt is an extracellular ligand that acts to stabilize a cytoplasmic protein, β-catenin. In the absence of Wnt, β-catenin is targeted for degradation, and under these conditions Wnt-responsive DNA elements (WRE) associate with factors to transcriptionally repress target genes. Conversely, in the presence of Wnt, stabilized β-catenin translocates to the nucleus, where it binds members of the TCF/LEF transcription factor family at WREs, and promote expression of Wnt target genes. One such target is the protooncogene *c-MYC*.

Within the last decade, substantial progress has been made to better understand the genetic underpinnings of inflammatory bowel disease (IBD). Seminal work has identified 200 IBD-associated loci that contribute to disease development (161, 162). While several key genes have been identified as deregulated or modified resulting from direct structural changes, the majority of IBD associated single nucleotide polymorphisms (SNPs) map to non-coding regions leaving their function a mystery. Prior work by Mokry et al. overlapped several SNPs with markers of open chromatin structure in several cell types to determine whether such SNPs colocalize to DNA regulatory elements (168). While successfully identifying novel enhancers, the targets of many SNPs still remain unclear. Due to the importance of Wnt signaling in homeostasis and restitution of the IEC barrier, we hypothesized that **a subset of known IBD-associated SNPs demarcate Wnt-responsive elements that function in IECs.** To address this hypothesis, we developed and tested the following aims:

1. **Determine the role of a WRE demarcated by IBD-associated SNPs in IECs.** To identify these WREs, we co-localized β-cat ChIP-seq peaks with IBD-associated SNPs. Of several overlapping targets, rs6651252 mapped directly adjacent to a TCF/LEF binding motif. We conducted ChIP to analyze presence of histone marks associated with open chromatin, as well as binding of TCF7L2 to this region. To determine functionality of this region as a putative enhancer, we developed luciferase constructs containing a 550 bp fragment surrounding rs6651252. Due to the SNP's proximity to a TCF motif, we questioned if the SNP were causative and influenced binding of TCF7L2 to this region. A luciferase construct containing the disease-associated allele, and a biotinylated DNA pulldown was utilized to address preferential binding of TCF7L2 to the disease and ancestral alleles.

2. **Identify the targets of the rs6651252-WRE.** We utilized CRISPR-mediated genomic editing to knockout (KO) the enhancer in an intestinal epithelial cell lines, and confirmed our findings using a CRISPR/dCas9 approach to epigenetically silence the region surrounding rs6651252. Using RT-qPCR on mRNA isolated from KO, epigenetically silenced, and wt cells, we examined expression of several putative target genes.

3. **Examine the clinical relevance of rs6651252 within patient samples.** To determine whether rs665152 and its enhancer had an effect in patients, we examined the expression of putative target genes within genotyped patient tissues.

This work fills a critical gap in knowledge regarding the impact of intergenic disease-associated SNPs in the context of Crohn's disease (CD). By addressing these aims above, we determined that rs6651252 alters binding of TCF7L2 to a novel WRE that targets the *MYC* promoter, and thereby affirm that Wnt/MYC signaling contributes to CD pathogenesis. Additionally, this work suggests that patients harboring the disease-associated allele may benefit from therapies that target MYC or MYC-regulated genes

Chapter 2 Materials and methods

Cell lines

The HCT116 and DLD-1 cell lines were obtained from the American Type Culture Collection (cat. numbers CCL-221 and CCL-247) while HEK293T cells were obtained from Invitrogen. HCT116 and HEK293T cells were maintained in DMEM (Corning) supplemented with 10% FBS, 5 mM L-glutamine, and 1% penicillin/streptomycin. DLD-1 cells were maintained in RPMI-1640 supplemented with 10% FBS and 1% penicillin/streptomycin. All cells were cultured in an incubator at 37ºC in 5% CO_2.

Plasmids

To generate the rs6651252 WRE luciferase plasmid (rs6651252-luc), a 555 bp DNA segment containing rs6651252 was amplified by PCR from genomic DNA isolated from HCT116 cells using the DNAEasy kit (Qiagen) serving as the template. The PCR product was subcloned into the pGL3-promoter luciferase vector (Promega) as a KpnI-NheI fragment. Site-directed mutagenesis was conducted to convert the ancestral T variant to the disease-associated C variant using the QuickChange mutagenesis kit (Agilent) following the manufacturer's instructions. To generate rs6651252 WRE expanded luciferase construct [rs6651252 (exp)-luc], a 2.7 kb segment containing the rs6651252 WRE was amplified by PCR and the product was likewise subcloned as a KpnI-NheI fragment into the pGL3-promoter luciferase vector. Primer sequences used in the PCR reactions can be found in Table 2-1.

For the epigenetic silencing assay, the phU6-sgRNA plasmid and the pLV hU6-sgRNA hUbC-dCas9-KRAB-T2a-Puro plasmid encoding dCas9-KRAB fusion protein was obtained from Addgene (#53188 and #71236, respectively). Using the 2.7 kb segment of DNA, the CRISPR guide RNA design tool (crispr.mit.edu) was used to identify unique guide sequences. For each of eight guides, paired oligonucleotides were designed with BbsI overhanging restriction sites to facilitate subcloning into the phU6-gRNA plasmid vector. Complementary DNA oligonucleotides (10 μM) corresponding to each of the eight guide RNA sequences, were annealed in 1x T4 ligase buffer by heating to 95°C for 5 min and cooling 5°C/min to room temperature. To digest and anneal sgRNA, 25 ng of pU6-gRNA plasmid was mixed with 1 μl annealed guide pair, 1 μl of 10x T4 ligase buffer, 2.5U BbsI, 0.5 μl of T4 ligase and 6 μl H2O. The reactions were subjected to 25 cycles at 37°C for 5 min and 23°C for 5 min on a DNAengine thermocycler (BioRad). Primer sequences used to generate the guide RNAs can be found in Table 2-1.

To create the rs6651252 WRE deletion cell line, guide RNAs were designed to flank a 692 bp region surrounding rs6651252, and corresponding oligonucleotides were annealed as described above. The fragments were ligated into the pSpCas9(BB)-2A-GFP (PX458) CRISPR/Cas9 plasmid (Addgene, #48138), which was first digested with BbsI. Primer sequences used to generate guide RNAs and to assess the rs6651252 status are listed in Table 2-1. Sanger

sequencing was used to verify each plasmid insert and enhancer deletions in the knockout clones.

Chromatin immunoprecipitation (ChIP)

ChIP assays in this dissertation were performed as previously described (91), using approximately 6.0 x 10^6 cells per sample. Cells were crosslinked with exposure to 1.5% formaldehyde, and sonicated through use of a misonex probe sonicator. The cross-linked and sheared chromatin was precipitated with 3 μg of the following antibodies; TCF4/TCF7L2 (05-511 Millipore), β-catenin (610154, BD scientific), H3K27Ac (ab4729, Abcam) and FLAG (F1804-200UG, Sigma). Precipitated and purified DNA was analyzed through qPCR with primers listed in Table 2-1 at the end of this chapter. Reactions were run on a MyIQ real-time PCR machine (Biorad) using cycling parameters previously described (91).

Luciferase reporter assay

Luciferase assays were conducted as described previously (85). Briefly, approximately 2.5 x 10^4 cells were seeded per well in a 24-well plate. Transfections were conducted using Lipofectamine 2000 following manufacturer's guidelines. Each reaction contained 50 ng of the luciferase reporter plasmid, and 2 ng pLRL-SV40 Renilla, which served as a transfection control. Where indicated, 50 ng of pcDNA3.1-β-catenin S45F (172), 50 ng of pME18 Lef (172), and 50 ng of pLV hU6-sgRNA hUbC-dCas9-KRAB-T2a-Puro were added to the transfection. The phU6-sgRNA plasmids encoding the guide RNAs (25 ng each) were included as indicated. Total concentration of DNA was adjusted to 2 μg per

reaction using pBluescript (Stratagene). Transfection mixtures were incubated on cells for 6 hours, after which the media was replaced with normal growth media. Each reaction was conducted in quadruplicate. Twenty-four hours after transfection, cells were lysed in 200 μl passive lysis buffer and luciferase levels were measured using the dual luciferase assay kit (30005-2; Biotium) on a Glomax 20/20 single chamber luminometer (Promega).

CRISPR/Cas9-mediated deletion of the rs6651252 WRE in HCT116 cells

CRISPR/Cas9 modified clonal HCT116 cell lines were generated following the protocol outlined previously (173). Briefly, 500 ng of each CRISPR/Cas9 plasmid (PX458), encoding guide RNA sequences that flanked rs6651252, were transfected into HCT116 cells using lipofectamine 2000 (Invitrogen) for 6 h. After 24 h, the cells were harvested and a FACSDiva (Becton Dickinson) machine was used to seed 2 cells per each well of a 96-well plate. After expanding the clones, genomic DNA was isolated using the Lyse&Go kit (Pierce) and the rs6651252 region was amplified by PCR using the DreamTaq Green (Thermo) polymerase and primers listed in Table 2-1. The products were resolved on a 1% agarose gel, excised with a scalpel, and purified using the MinElute PCR Purification Kit (Qiagen). The products were sequenced to identify clones harboring rs6651252 WRE deletions.

DNA pull-down assay

The DNA pull-down assay was conducted as previously described with minor modifications (174). The annealed probes (15 μM) were incubated with 0.1 mg of

streptavidin coated magnetic beads (Promega Z5481) for 1.5 hours at room temperature. Prior to use, the beads were washed three times in 1X sodium citrate buffer (30 mM NaCl, 0.035 mM $C_6H_5O_7Na_3$) and then resuspended in 100 µl binding buffer (1M NaCl, 10 mM Tris, 1 mM EDTA, pH=8.0). To prepare the protein lysates, 2.5 x 10^7 HCT116 cells were harvested and lysed in RIPA buffer supplemented with freshly added protease inhibitors (1 mM PMSF, 10 µl/mL aprotinin, 10 µg/ml leupeptin). For each binding reaction, 200 µg of protein lysate was incubated at 4°C with 150 µg sonicated salmon sperm for 30 min to reduce non-specific interactions. This lysate was added to probes conjugated with the magnetic beads and the reactions were incubated for 2 hrs at room temperature on a rotating platform. The protein/DNA complexes were collected using a magnetic stand, washed three times in RIPA buffer, and eluted in 50 µl of 2x laemmli loading buffer. For competition assays, annealed and unlabeled probes were added concurrently in the binding reactions as concentrations of 250 ng, 500 ng, and 1500 ng.

Epigenetic silencing rs6651252-WRE through CRISPR/dCas9

In the epigenetic repression assays, 250 ng of the pLV hU6-sgRNA hUbC-dCas9-KRAB-T2a-Puro plasmid and 25 ng each of the phU6-gRNA plasmids harboring guide RNAs that tiled the 2.7 kb segment containing the rs6651252 WRE were electroporated into 1.0 x 10^7 HCT116 or HEK293T cells using the Amaxa Cell line Nucleofector Kit V (Lonza) and an Amaxa Nucleofector II electroporator following the manufacturers guidelines. *MYC* gene expression was assessed by RT-qPCR 72 h after electroporation.

Quantitative reverse transcription PCR

To assess *MYC* gene expression, RNAs were collected and cDNAs synthesized using protocols that were described previously (173). For experiments involving patient tissues, approximately 1 g of flash-frozen and full thickness colonic tissue was homogenized in 1 mL of TRIzol (Thermo Fisher) in an eppendorf tube using a disposable micropestle (175). Following a 5 min incubation at room temperature, 200 µL of chloroform was added and the mixture shaken vigorously before centrifugation at 12,000 x g for 15 min at 4°C. The upper layer was removed, mixed with an equal volume of 100% ethanol, and the RNAs were purified using an RNeasy Mini kit (Qiagen). For both cell line and patient sample experiments, cDNA was synthesized from 500 ng of RNA using the iScript cDNA synthesis kit (BioRad), following manufacturer protocol. Data are presented as relative levels using the $2^{\Delta CT}$ method with GAPDH and TUBB3 serving as the reference genes. Primer sequences used in the RT-qPCR experiments are listed in Table 2-1.

Western blot

The proteins eluted from the biotinylated primers above, were resolved on an 8% polyacrylamide gel and standard western blot analysis as previously described (93). The blots were incubated overnight with anti-TCF7L2 antibodies (05-511 Millipore, 1:750 dilution) followed by incubation with HRP conjugated anti-mouse secondary (Jackson Immunoresearch, 1:5000 dilution) for two hours prior to ECL treatment and exposure to film.

PAX gene® analysis

PreAnalytiX PAXgene® blood RNA kit (Cat. 762164) was used to isolate RNA from PAXgene® Blood RNA tubes as per manufacturer instruction. Briefly, tubes containing patient blood and stored at -20C were equilibrated to RT before lysing blood cells for 2 h. Lysate was centrifuged to remove cellular debris and the pellet was resuspended, and treated with Proteinase K, before RNA was isolated and purified through silica column purification and DNAse treatment. cDNA was synthesized from 500 ng of RNA using the iScript cDNA synthesis kit (BioRad), following manufacturer protocol. Data are presented as relative levels using the $2^{\Delta CT}$ method with GAPDH and TUBB1 serving as the reference genes.

Patient samples

The samples evaluated in this manuscript were collected as part of a retrospective study performed at the Pennsylvania State University College of Medicine with Institutional Review Board approval (HY98 057EP.A, W Koltun, PI). Patients are recruited into our Colorectal Disease Biobank under informed consent and allelic status of IBD associated SNPs is assessed using a custom oligonucleotide array(176). Patients for PAXgene RNA analysis were first genotyped using a custom array (Thermofisher Cat # 4351379), and also with a taqman probe with the sequence found in Table 2-1.

Statistics

Each experiment was performed at least three times. For ChIP-qPCR and RT-qPCR, each sample was amplified in quadruplicate reactions per experiment. A Student's T-test was used to calculate statistical significance.

Table 2-1 Primer sequences

Construction of rs6651252 luciferase constructs		
555 bp rs6651252 fragment	F	GGGGTACCCCAACAGTGATTCCTTTCTCAAGCG
	R	CTAGCTAGCTAGGCAATATACGGGAAACAGAAGTG
site-directed mutagenesis	F	CTGCTCACATTTCAAAGGGCCCACTTTTCCTCCTA
	R	TAGGAGGAAAAGTGGGCCCTTTGAAATGTGAGCAG
2.7-kb rs6651252 fragment	F	GGGGTACCCCTTGCACAGAGCAGAACATTTTAGCCAT
	R	CTAGCTAGCTAGGAAGCTTTGAGCTTGAAACCCTGA

CRISPR/Cas9 deletion of rs6651252 enhancer		
Upstream guide RNA	F	CACCGTAGGCTGAGTTACCTATGCT
	R	AAACAGCATAGGTAACTCAGCCTAC
Downstream guide RNA	F	CACCGAGCATTTTGAGCTTGCCGAC
	R	AAACGTCGGCAAGCTCAAAATGCTC
PCR analysis of clonal cell lines	F	CCCAATTCACTTTCATCCTTGATG
	R	GTGTCCCAGGTAGAGGGAATAG

Probes for DNA binding assay		
"T" containing probe	F	5’-(biotin)-AAAGTGGGCTCTTTGAAATGTGAGCAG
	R	5-'CTGCTCACATTTCAAAGAGCCCACTTT
"C" containing probe	F	5’-(biotin)-AAAGTGGGCCCTTTGAAATGTGAGCAG
	R	5'-CTGCTCACATTTCAAAGGGCCCACTTT
mutant TBE	F	5’-(biotin)-AAAGTGGGCTCGCTGATATGTGAGCAG
	R	5'-CTGCTCACATATCAGCGAGCCCACTTT

Chromatin Immunoprecipitation/qPCR		
rs6651252	F	TCTGTGCCCAATGTCACCTGGA
	R	GCTGTGCAACTATTAGCTGATCCCAC
TUBULIN control	F	AGAGAGGTGGTGACTCCGCTC
	R	ACAACGAGGCGCTCTACGACATCT

RT-qPCR gene expression analysis		
MYC	F	AACTTGACCCTCTTGGCAGCA
	R	GCAAACCTCCTCACAGCCCAC

TUBULIN B3	F	GTGTCTAAACCCCCGGAGCCATCT
	R	TGGGGAGGACGAGGCCATAAATAC
TUBULIN B1	F	CAGCACTCCAAAACCCAGTCTGC
	R	ACCTGTCTCCCTCAGTCCCTGTGT
GAPDH	F	CAAGGGGTCTACATGGCAACTGTG
	R	CCAGCAAGAGCACAAGAGGAAGAG

CRISPR/dCas9 tiling rs6651252 locus		
set 1	F	CACCGACCTAGCATCCTGTAAAAGC
	R	AAACGCTTTTACAGGATGCTAGGT
set 2	F	CACCGAGCTATAGCTACCTCTGTA
	R	AAACTACAGAGGTAGCTATAGCTC
set 3	F	CACCGAAAGTGAACAGGTCACGCGC
	R	AAACGCGCGTGACCTGTTCACTTTC
set 4	F	CACCGAGCATTTTGAGCTTGCCGAC
	R	AAACGTCGGCAAGCTCAAAATGCTC
set 5	F	CACCGAATAGCCTATGCAATATAC
	R	AAACGTATATTGCATAGGCTATTC
set 6	F	CACCGCCACTAAATGGTTGTAGAGC
	R	AAACGCTCTACAACCATTTAGTGGC
set 7	F	CACCGCAATCGTCTTTGCCCTTGAA
	R	AAACTTCAAGGGCAAAGACGATTGC
set 8	F	CACCGTAGAAAACTCGGCATACAGC
	R	AAACGCTGTATGCCGAGTTTTCTAC

Custom Taqman array for Genotyping

TTGACACATAGGAGGAAAAGTGGGC[C/T]CTTTGAAATGTGAGCAGAGCACCGA
where C/T are rs6651252, and marked Vic/FAM

Chapter 3 The Crohn's disease-associated SNP, rs6651252, impacts *MYC* gene expression in human colonic epithelial cells

These results have been published in *PLoS One* 2019. 14(2):e0212850.
https://doi.org/10.1371/journal.pone.0212850

Stephen M. Matthews, Melanie A. Eshelman, Arthur S. Berg, Walter A. Koltun, Gregory S. Yochum

Author Contributions: SMM, GSY: conceived and designed experiments; SMM, ASB and GSY curated the data; SMM, MAE, ASB and GSY analyzed the data; WAK acquired funding, SMM wrote the original draft; SMM, MAE, WAK and GSY reviewed and edited the manuscript.

Abstract

Crohn's disease (CD) is a debilitating inflammatory bowel disease (IBD) that arises from chronic inflammation in the gastrointestinal tract. Genome-wide association studies (GWAS) have identified over 200 single nucleotide polymorphisms (SNPs) that are associated with a predisposition for developing IBD. For the majority, the causal variant and target genes affected are unknown. Here, we investigated the CD-associated SNP rs6651252 that maps to a gene desert region on chromosome 8. We demonstrate that rs6651252 resides within a Wnt responsive DNA enhancer element (WRE) and that the disease associated allele augments binding of the TCF7L2 transcription factor to this region. Using CRISPR/Cas9 directed gene editing and epigenetic modulation, we find that the rs6651252 enhancer regulates expression of the *c-MYC* proto-oncogene (*MYC*). Furthermore, we found *MYC* transcript levels are elevated in patient-derived colonic segments harboring the disease-associated allele in comparison to those containing the ancestral allele. These results suggest that Wnt/MYC signaling contributes to CD pathogenesis and that patients harboring the disease-associated allele may benefit from therapies that target MYC or MYC-regulated genes.

Introduction

Crohn's disease (CD) and ulcerative colitis (UC) are the two main classes of inflammatory bowel disease (IBD), and arise from chronic inflammation in the gastrointestinal (GI) tract (100). CD can present anywhere along the GI tract whereas UC is confined primarily to the colon (100, 177). In a generally accepted view, IBD results from one or more environmental triggers in a genetically

susceptible individual (141, 144). While the precise environmental exposure is debatable, that fact that 5–23% of IBD patients have a first-degree relative that is also afflicted with disease is supportive of a genetic inheritance (145).

Numerous genome-wide association studies (GWAS) have been conducted to identify genetic variants associated with IBD predisposition (144). In a landmark study, Jostins et al. reported results from a meta-analysis of GWAS data generated from sporadic and familial IBD patients (162). Altogether, 163 single nucleotide polymorphisms (SNPs) that conferred IBD susceptibility were identified, of which 110 were associated with both CD and UC (162). Of the remaining 53, 30 were specific for CD and 23 were specific for UC (162). Several disease variants produced missense mutations within the nucleotide oligomerization domain two (NOD2) protein that senses bacterial peptidoglycans, and the autophagy protein ATG16L1. Other variants were also identified in the IL23 receptor and the HLA locus, and together, these findings substantiate the notion that IBD manifests, in part, through a deregulated immune response to commensal or pathogenic bacteria in the gut lumen (144, 162). Since that time, the list of variants associated with IBD has grown to include over 200 distinct loci (161, 162, 164, 177). While these results clearly demonstrate the power of GWAS to inform on disease pathogenesis, approximately 80–90% of the identified alleles map to non-coding regions of the genome, many of which are in gene-poor regions (162). How these non-coding variants are associated with IBD is largely unknown, and this remains a significant obstacle in our understanding of the genetic basis for disease pathogenesis.

Studies from the encyclopedia of DNA elements (ENCODE) consortium (169, 178), and from others (166, 179), indicate the majority of common disease-associated variants map to gene regulatory regions of the genome, also known as enhancer elements. Indeed, Mokry et al. found that 27% of the 163 IBD associated SNPs directly overlapped a putative enhancer element, as defined by regions containing elevated levels of histone H3 that is acetylated on lysine 27 (H3K27Ac) (168). Consideration of regions in linkage disequilibrium increased the total number of SNPs associated with active enhancer elements to 56% of the 163 SNPs (168). While several IBD-associated SNPs were confirmed to demarcate enhancer elements, the precise impact of these variants on enhancer function, the upstream signaling pathways involved and the relevant downstream transcription factors have not been adequately addressed.

The intestines are lined by a single layer of epithelial cells that protect the underlying mucosa and sub-mucosa from toxic contents of the gut lumen and the microbiota (3). Because epithelial cells are subjected to damage, the entire epithelial layer is replaced every five to six days making the intestines one of the most highly regenerative organs in the body(3, 12). The Wnt/β-catenin signaling pathway contributes to this regenerative process by driving cellular proliferation (13). In the presence of Wnt, the β-catenin transcriptional co-activator translocates to the nucleus and associates with members of the T-cell factor/ lymphoid enhancer factor (TCF/LEF; hereafter, TCF) family of transcription factors (37, 38). TCF7L2 is the predominant TCF family member expressed in intestinal epithelial cell lines (45, 180). β-Catenin/TCF complexes bind to Wnt-responsive

DNA regulatory elements (WREs) and primarily increase expression of target genes (38). One critical target in the intestines is the *c-MYC* (*MYC*) proto-oncogene(64, 84, 85). MYC is a transcription factor that promotes cellular growth and proliferation through increasing expression of genes involved in the cell cycle, ribosome biogenesis and metabolism (65, 72). While several WREs localize to the proximal promoter of the *MYC* gene, many also map tens to hundreds of kilobases away and are juxtaposed to *MYC* through long-range chromatin loops (181).

In this study, we investigate the CD-associated SNP, rs6651252, which maps to a gene-poor region on chromosome 8 (162). We demonstrate that this SNP impacts a novel WRE that controls *MYC* expression in intestinal epithelial cells, and that patient intestinal tissues harboring the disease-associated allele display increased levels of *MYC* transcripts. These findings offer one explanation for how Wnt/MYC signaling may contribute to CD pathogenesis and raise the possibility that patients harboring this allele may benefit from MYC targeted therapies.

Results

β-Catenin and TCF7L2 bind the rs6651252 locus in colonic epithelial cells

By overlapping β-catenin ChIP-Seq peak regions with positions of IBD SNPs, we identified the CD-associated SNP, rs6651252, as a leading candidate (Fig. 3-1A). Analysis of ENCODE data deposited on the genome browser (http://genome.ucsc.edu/) found that elevated levels of H3K27 acetylated histones, which are a marker of active enhancers, colocalized to this region (182, 183). To confirm these findings, we conducted ChIP-qPCR assays in the HCT116 and

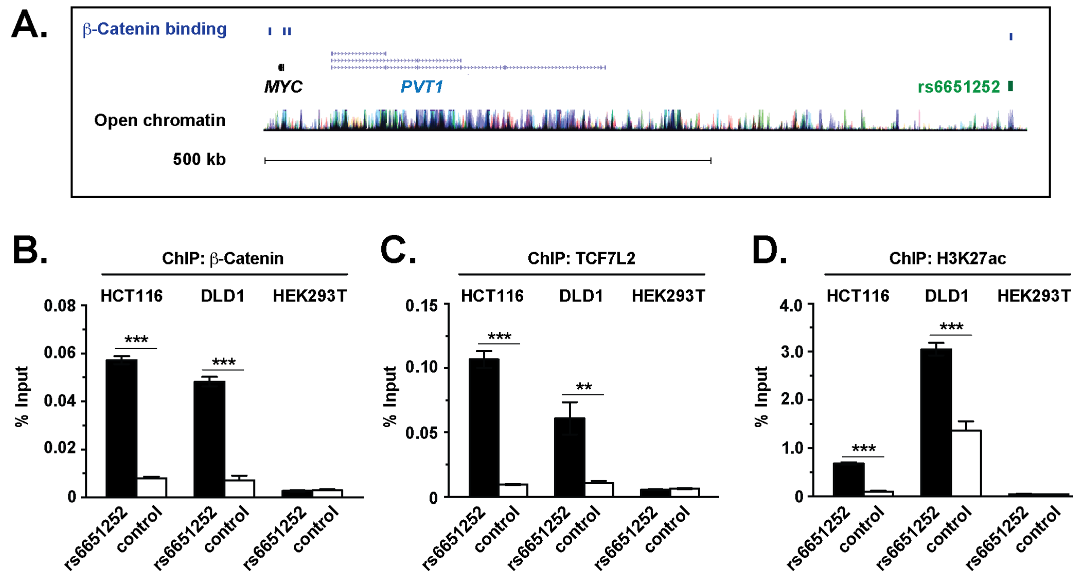


Figure 3-1. β-catenin and TCF7L2 bind the rs6651252 locus in colonic epithelial cells. (A). Diagram of the chromosomal *MYC* locus. Shown is the position of rs6651252 relative to *MYC* and *PVT1* genes on chromosome 8. Blue rectangles above depict positions of ChIP-Seq peak regions for β-catenin in HCT116 cells. Clustered vertical lines at bottom indicate H3K27 acetylated regions of chromatin that were downloaded from the UCSC Genome Browser, build hg18 (http://genome.ucsc.edu/). (B). qPCR analysis of DNA fragments precipitated with α-β-catenin antibodies in ChIP assays conducted in HCT116 and DLD-1 colonic epithelial cell lines, and in HEK293T cells. Oligonucleotides used in the PCR reactions flanked rs6651252 or a region in *TUBULIN* as a negative control. (C) and (D)., as in (B) except α-TCF7L2 or α-H3K27Ac antibodies were used in the ChIP assays, respectively. The data are presented as percent of input with error bars representing SEM (**P <0.01; ***P <0.001).

DLD-1 colonic epithelial cell lines and used the non-intestinal cell line, HEK293T, as a control. Using primers that flanked rs6651252, we detected robust β-catenin and TCF7L2 binding to this region in the colonic epithelial cell lines and not HEK293T (Fig. 3-1B and C). Very little signal was detected using primers that annealed to a region in the *TUBULIN* gene, which attests to the specificity of our ChIP assays. Furthermore, we detected elevated levels of H3K27 acetylated histones at rs6651252 in the intestinal cell lines and not HEK293T, suggesting that this CD-associated SNP may localize to an active and cell-type specific regulatory DNA enhancer element (Fig. 3-1D)(182, 183).

<u>TCF7L2 binds DNA harboring disease-associated rs6651252 variant with stronger affinity</u>

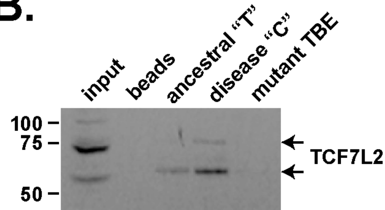
The rs6651252 SNP maps immediately adjacent to a consensus TCF binding motif (Fig. 3-2A) (91). The ancestral allele at this position is a T, while the disease-associated allele is a C with MAF of 0.14 (164). We conducted DNA pull-down assays to determine whether TCF7L2 bound to this fragment of DNA and to test whether the rs6651252 allelic variants impacted its association. Nuclear protein lysates from HCT116 cells were incubated with biotinylated DNA probes, the complexes were precipitated using streptavidin-conjugated magnetic beads and eluted proteins were subjected to western blot analysis. We found that TCF7L2 bound to this element and that mutating the TCF consensus motif blocked its association (Fig 3-2B). Moreover, the probe containing the C variant precipitated more TCF7L2 compared to the probe containing the T variant (Fig. 3-2B). In

A.

AAGTGGGCTCTTTGAAATGTGA
TTCACCCGAGAAACTTTACACT

Underline: TCF motif
Star: rs6651252

B.



C.

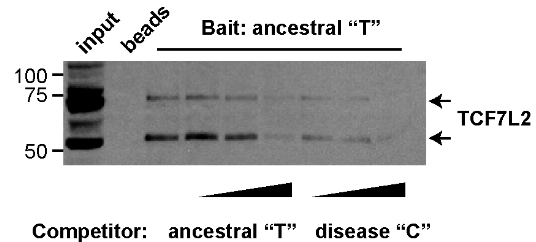


Figure 3-2. TCF7L2 binds DNA harboring disease-associated rs6651252 variant with stronger affinity. (A). DNA sequence flanking rs6651252 (asterisk) and the adjacent TCF consensus motif (underlined). (B). Western blot analysis of TCF7L2 from HCT116 protein lysates that were incubated with the biotinylated DNA probes indicated and subsequently precipitated with streptavidin conjugated magnetic beads. A probe containing mutations in the consensus TCF binding element (mutant TBE) was used as a negative control. (C). as in (B) except the biotinylated probe harboring the ancestral T variant was used in all reactions. As indicated below, increasing concentrations of unlabeled probes containing T or C were included in the binding reactions prior to precipitation.

addition, we found that adding increasing amounts of unlabeled C probe more effectively competed with TCF7L2 binding to the biotinylated T probe in comparison to reactions containing equivalent amounts of unlabeled T probes (Fig. 3-2C). Therefore, the C variant of rs6651252 potentiates binding of TCF7L2 to this DNA element.

rs6651252 demarcates a Wnt responsive DNA enhancer element

We next used heterologous luciferase reporter assays to determine whether rs6651252 was embedded within a DNA enhancer element. Using genomic DNA isolated from a human colonic epithelial cell line, we PCR amplified a 555 bp segment containing rs6651252 and inserted it upstream of the minimal SV40 promoter in the pGL3-luciferase vector (Fig. 3-3A). We refer to this plasmid as rs6651252-luc and upon sequencing the insert, it contained the T ancestral allele. In comparison to HCT116 and DLD-1 cells transfected with the vector backbone alone, rs6651252-luc drove higher levels of luciferase (Fig. 3-3B and C). In HEK293T cells, rs6651252-luc produced lower levels of luciferase relative to the control vector (Supplemental Fig. 3-1). Next, we used site-directed mutagenesis to substitute the rs6651252 T variant with the disease-associated C variant (Fig. 3-3D). In both HCT116 and DLD-1 cells, this plasmid drove higher levels of luciferase relative to rs6651252 containing the T variant (Fig. 3-3E and F). These results demonstrate that a DNA segment harboring rs6651252 is a cell-type specific enhancer element and that the C variant augments its activity.

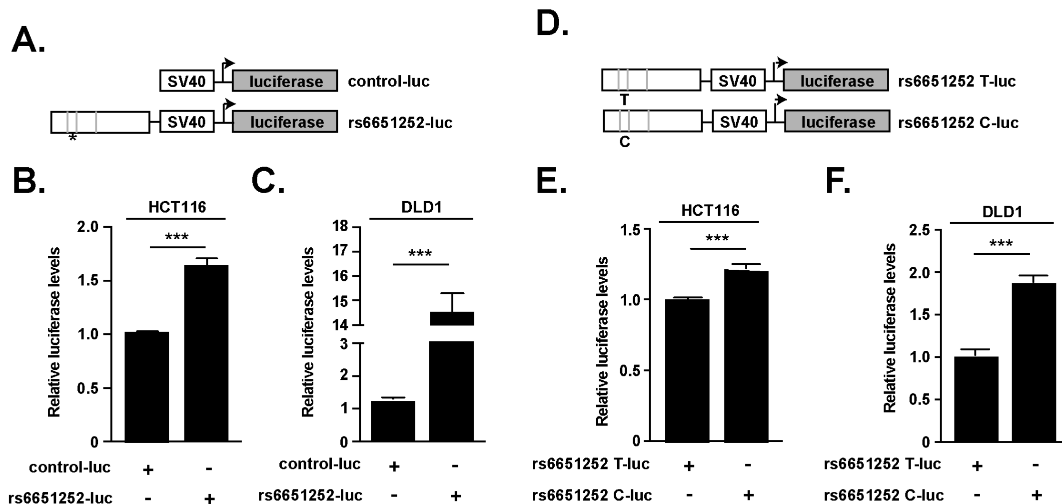


Figure 3-3. rs6651252 demarcates a Wnt-responsive DNA enhancer element. (A). Diagram of luciferase reporter vectors used in (B) and (C). The pGL3-promoter vector backbone served as a control (control-luc). A 555 bp DNA fragment encompassing rs6651252 was inserted upstream of the SV40 promoter in the pGL3-promoter plasmid vector. Grey vertical lines depict consensus TCF motifs with an asterisk depicting rs6651252. (B) and (C). Luciferase assays conducted in the indicated cell lines with values obtained from rs6651252-luc transfected cells normalized to values obtained from cells transfected with control-luc (D). Diagram of rs6651252 luciferase vectors used in (E) and (F). The T variant of rs6651252 was changed to the C variant using site-directed mutagenesis. (E) and (F). As in (B) and (C), except the indicated cells were transfected with rs6651252 harboring either ancestral T or disease-associated C variant. Error bars are SEM (***P <0.001).

The rs6651252 WRE regulates *MYC* gene expression

In a recent study, Meddens et al. used 4C-seq to identify the gene targets of enhancer elements containing embedded IBD-associated SNPs (171). In that report, rs6651252 was found juxtaposed to the *MYC* and *POU5F1B* gene loci on chromosome 8 through long-range chromatin loops (171). Whether the rs6651252 enhancer regulated *MYC* or *POU5F1B* expression was not demonstrated. To test if these genes were regulated targets, we used CRISPR/Cas9 to delete the rs6651252 enhancer in the HCT116 cell line (Fig. 3-4A)(173). After propagating independent clonal lines, we used PCR to assess rs6651252-enhancer status. We successfully identified multiple lines with either heterozygous or homozygous deletions (Fig. 3-4B). The PCR-based genotyping was confirmed using Sanger sequencing. In comparison to a control clone that lacked deletions, two independent knockout clones displayed reduced levels of *MYC* expression as assessed by RT-qPCR (Fig. 3-4C). Deletion of the rs6651252 WRE did not significantly impact *POU5F1B* gene expression levels in HCT116 cells (Supplemental Fig. 3-2)

To confirm that *MYC* gene expression is regulated by the distal rs6651252 WRE, we used a CRISPR/Cas9-based approach to epigenetically repress the function of this DNA regulatory element (184). In this assay, guide RNAs are used to recruit a mutant Cas9 (dCas9, lacking endonuclease activity) fused to the KRAB transcriptional repressor. We generated eight guide RNAs that tiled a 2.7 kb segment to target dCas9-KRAB to the rs6651252 locus (Fig. 3-5A). Prior to using this system, we performed various control experiments to validate the approach. First, we conducted luciferase assays using the same vector backbone

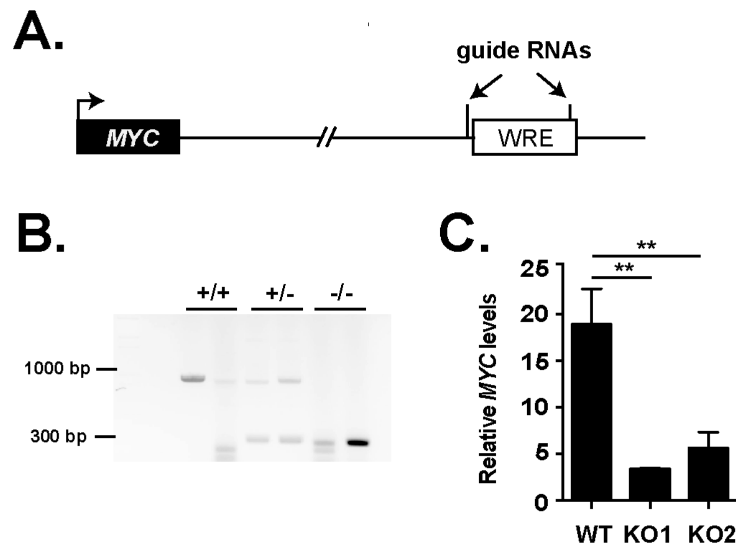


Figure 3-4. The rs6651252-WRE regulates *MYC* gene expression. (A). Diagram of the MYC chromosomal locus with the position of the rs6651252 WRE located approximately 750 kb downstream. The positions of the guide RNAs used for CRISPR/Cas9-mediated deletion are indicated. (B) Agarose gel of PCR products generated from genomic DNAs isolated from clones of HCT116 cells engineered to delete the rs6651252 WRE using CRISPR/Cas9-mediated gene editing. Primers that flanked the rs6651252 WRE were used in PCR reactions. The +/+ indicates clones with non-edited rs6651252 WREs, whereas +/- and -/- indicate heterozygous and homozygous deleted clones, respectively. (C). RT-qPCR analysis of MYC mRNA levels in rs6651252 WRE wildtype (WT) and knockout (KO1 and KO2) clones. Values are normalized to average *GAPDH* and *TUBULIN* levels. Error bars denote SEM (**P <0.01).

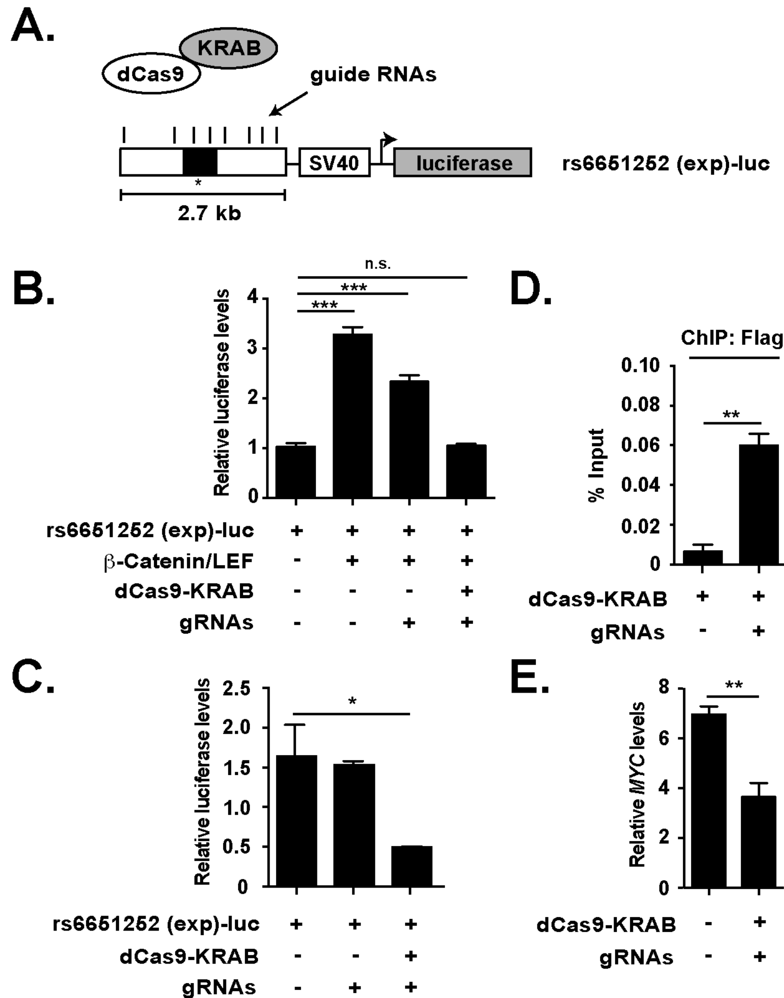


Figure 3-5. Epigenetic repression of rs6651252 reduces *MYC* gene expression. (A). Diagram of a luciferase reporter vector containing a 2.7-kb DNA fragment encompassing rs6651252. The vertical lines above depict the positions of multiple guide RNAs designed to target the element. These guide RNAs recruit the dCas9-KRAB fusion protein to rs6651252 in transfected cells. (B). Luciferase reporter assays conducted in HEK293T cells. Where indicated, plasmids encoding β-catenin S45F, LEF, dCas9-KRAB and plasmids expressing the guide RNAs (gRNAs) were co-transfected. Data are normalized to luciferase levels in cells receiving the reporter vector alone. (C). Relative luciferase reporter activity in HCT116 cells transfected with the plasmids indicated. (D). qPCR analysis of DNA fragments precipitated with a α-FLAG antibody in ChIP assays conducted in HEK293 cells expressing gRNAs alone or gRNAs and FLAG-dCas9-KRAB. The oligonucleotides used for detection amplified a region encompassing rs6651252. (E). RT-qPCR analysis of MYC transcript levels in HCT116 cells expressing gRNAs and the dCas9-KRAB fusion. Non-transfected cells served as a control. Values are normalized to *TUBULIN*. In (B-E), error bars are SEM (*P <0.05, **P <0.01, ***P <0.001).

used previously in Fig. 3-3, except that the 2.7 kb segment containing rs6651252 was inserted upstream of the SV40 promoter. In transfected HEK293T cells, plasmids containing β-catenin and LEF (a TCF family member) cDNAs activated expression of luciferase (Fig. 3-5B). Whereas inclusion of plasmids encoding the guide RNAs alone resulted in a slight decrease, co-transfecting guides along with dCas9-KRAB blocked β-catenin/LEF-driven luciferase activity (Fig. 3-5B). Similarly, gRNA/dCas9-KRAB complexes effectively reduced rs6651252 (exp)-driven luciferase levels in HCT116 cells (Fig. 3-5C). Using ChIP-qPCR assays in transfected HEK293T cells, we found that inclusion of the guide RNAs increased levels of dCas9-KRAB recruited to the rs6651252 locus (Fig. 3-5D). Having validated the system, we introduced the guide RNAs and dCas9-KRAB into HCT116 cells and this significantly reduced *MYC* expression in these cells (Fig. 3-5E). Together, these experiments demonstrate that the rs6651252 enhancer regulates *MYC* expression in colonic epithelial cells.

The rs6651252 C variant correlates with increased *MYC* expression in patient colonic tissues

To determine whether rs6651252 impacts *MYC* gene expression *in vivo*, we obtained colonic segments that were surgically resected from CD patients. As controls, we obtained intestinal tissues from patients that underwent resection for non-IBD related issues such as slow transit or volvulus. We genotyped rs6651252 in genomic DNAs isolated from collected blood and found that while all control patients were homozygous for the ancestral T variant, half of the CD patients were TT and the other half were heterozygous TC. We then isolated RNA from flash-

frozen tissues and assessed *MYC* gene expression using RT-qPCR. This analysis found that tissues from CD patients that were heterozygous for the risk variant (TC) contained higher levels of *MYC* transcripts compared to tissues from CD patients harboring homozygous alleles (TT) or controls (Fig. 3-6).

Discussion

GWAS have identified over 200 SNPs that are associated with a predisposition for developing IBD (144, 161, 162, 164). While some of these are found within protein-coding regions of the genome, most map to intergenic and gene-poor loci (144). Recent work has shown that many of these non-coding SNPs are found within regions of accessible chromatin, demarcated by elevated levels of H3K27ac (168). Most often, these SNPs are assumed to impact the nearest gene promoter (144, 177). However, it is known that enhancers are capable of impacting more than one gene and can bypass the nearest gene to influence expression of a neighboring gene (185). Due to the inherent difficulty in studying non-coding regions of DNA, particularly those that may function in a cell-type specific manner, the causative impact of these non-coding SNPs on gene function largely remains a mystery. This current study focused on the CD-associated SNP, rs6651252 that maps to the 8q24 locus. This locus is a large non-coding region of the genome that contains numerous SNPs that have been shown to impact ovarian, prostate and colorectal cancers (186).

We demonstrate that rs6651252 demarcates a WRE and that the disease-associated allele potentiates enhancer activity through higher affinity binding of

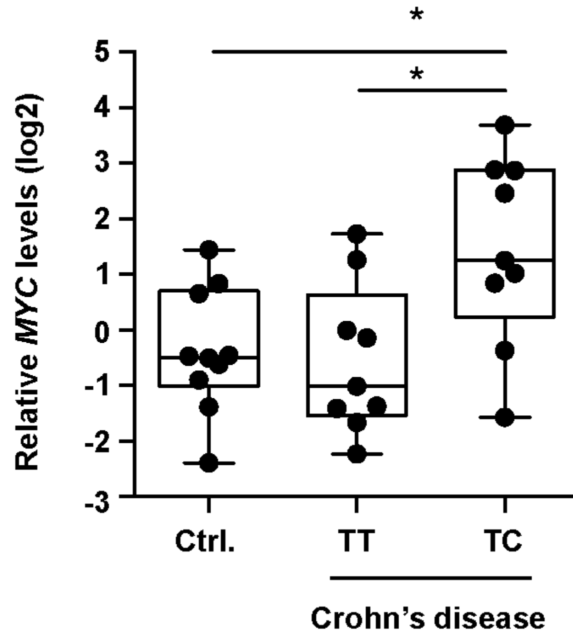


Figure 3-6. The rs6651252 C variant correlates with elevated *MYC* expression in patient colonic tissues. RT-qPCR analysis of *MYC* transcript levels in resected full-thickness colonic tissues from control and CD patients. For CD patients, the genotype of rs6651252 is indicated. Values are normalized to *GAPDH*. Error is SEM (*P <0.05).

TCF7L2. In an earlier study, Meddens et al. reported evidence that this rs6651252 enhancer is juxtaposed to the *MYC* promoter, implicating *MYC* as a direct target of this enhancer (171). Our work confirms and extends this finding as we found that either deletion or epigenetic silencing of this element reduces *MYC* expression in human colonic epithelial cells. Moreover, in our survey of human colonic tissues resected from IBD patients, we find that the disease-associated rs6651252 increases *MYC* expression. Together, these findings are similar to those reported for the cancer-associated SNP rs6983267, which also resides within this gene desert region on chromosome 8 (88, 187, 188). While the rs6983267 disease-associated allele augments TCF binding and enhancer activity, whether it differentially impacts *MYC* gene expression in normal tissues or cancers is still a matter of debate (189).

While much attention has been given to understand the role of MYC in colorectal carcinogenesis, much less is known about its role in IBD (190). Using the dextran sodium sulfate (DSS) model of acute colitis in mice, we reported that slight elevation of MYC (~2.5 fold) promotes restitution of the colonic epithelium (191, 192). Furthermore, we found that lithium treatment, which inhibits glycogen synthase kinase and stabilizes MYC, confers a favorable response to colonic regeneration after acute DSS-induced damage in mice (95). These findings indicate that short-term MYC stabilization may provide favorable outcomes in IBD by promoting restitution of the epithelial monolayer (95, 191-193). However, an earlier study reported that levels of *MYC* transcripts are elevated in intestinal tissues isolated from IBD patients in comparison to controls (194). Furthermore,

higher levels of MYC protein was found in inflamed IBD intestinal tissue in comparison to control, non-IBD colonic segments (194). In addition, MYC expression is elevated, and the *MYC* chromosomal locus is frequently amplified, in colitis-associated cancer (CAC) (195-197). Therefore, while transient stabilization of MYC is beneficial, long-term and sustained MYC expression is likely detrimental in IBD and CAC.

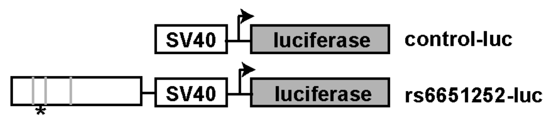
Our analysis of surgically resected colonic tissues indicates that CD patients harboring the rs6651252 risk allele display elevated levels of *MYC* transcripts relative to levels found in tissues from CD patients that are homozygous for the ancestral allele or controls (Fig. 3-6). Because we have a limited number of control tissues in our Colorectal Disease Biobank, we were unable to identify control colonic tissues that were heterozygous for the risk allele. This precludes interpretations that elevated *MYC* may contribute to disease onset. However, our analysis of genotyped diseased tissues suggests that elevated *MYC* may contribute to disease pathogenesis in CD patients harboring the risk allele. It follows that patients who are homozygous for the disease variant may have a more pronounced phenotype, and as we continue to recruit patients into our biorepository, we hope to identify such patients and will explore this possibility in a future study.

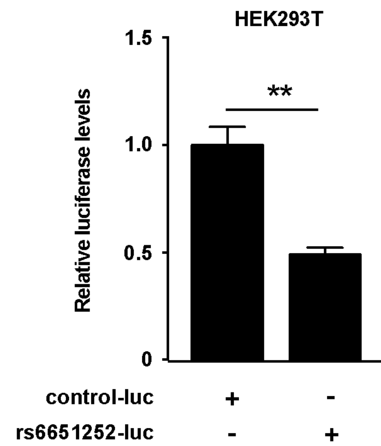
One limitation of our study of human colonic tissues is that full-thickness specimens were analyzed. While we favor a model whereby the rs6651252 disease associated allele elevates *MYC* expression in the colonic epithelium, we cannot dismiss the possibility that other cell types, such as resident lymphocytes, could be

contributing to the overall differences we detect. In support of this possibility, Mokry et al. demonstrated that rs6651252 resides in a region of accessible chromatin in specific $CD4^+$ sub-cellular populations(168). Due to the importance of T-cells in IBD pathologies (198), and the role of Wnt signaling in T-cell biology (199), it is possible that the rs6651252 WRE influences *MYC* expression in cells that function in the adaptive cellular immune response. Purification of specific population of cells from genotyped samples and assessing *MYC* gene expression in these cells is required to fully understand how rs6651252 is impacting CD pathogenesis.

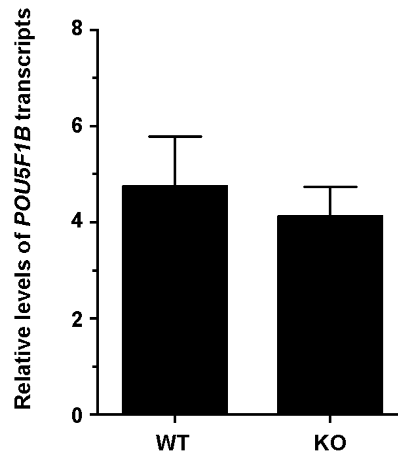
In summary, our work shows that rs6651252 demarcates a WRE within the 8q24 locus. The CD-associated allele facilitates stronger TCF7L2 binding to the WRE, potentiates enhancer activity, and increases *MYC* gene expression. While additional work is needed to further define the cell types in which this rs6651252 WRE operates, and the constellation of genes whose expression it impacts, these findings suggest that CD patients harboring this allele may benefit from therapies that target MYC or MYC-regulated genes. Along these lines, *MYC* gene expression is sensitive to inhibitors that target bromodomain and extra-terminal family of proteins (iBETs), such as JQ1 (200, 201). Our findings presented here support the idea that JQ1 should be further evaluated in pre-clinical mouse models of IBD with the hope that someday it could be used to augment current treatment strategies for IBD patients.

Supplemental





Supplemental; Figure 3-7. The rs6651252 element is repressive in HEK293T cells. Luciferase assays conducted in HEK293T cells transfected with the indicated constructs above. Error bars depict SEM. **P < 0.01



Supplemental; Figure 3-8. Depletion of rs6651252-WRE does not affect *POU5F1B* gene expression. RT-qPCR analysis of *POU5F1B* expression in parental (wild-type, wt) and rs6651252 WRE knockout (KO) HCT116 cell lines. *POU5F1B* levels are normalized to *GAPDH*. Error bars depict SEM.

Chapter 4 Interplay of Wnt and MYC in Crohn's disease

Overview

GWAS have identified over 200 SNPs associated with a predisposition for IBD (144, 161, 162, 164). Several key SNPs localize to protein coding genes, but the majority map to intergenic loci (144). The function of most of these intergenic SNPs remain unknown, but they localize frequently to regions of accessible chromatin suggesting that they may impact the function of DREs (168). Despite the fact that DNA enhancer elements can bypass the nearest gene to influence expression of a neighboring gene, or multiple genes (185), the non-coding SNPs identified to date are often assumed to impact the nearest gene promoter (144, 177). Work in this dissertation aimed to elucidate the function of one intergenic SNP, and contributes to our understanding of how non-coding SNPs impacts IBD pathogenesis. Work presented in chapter 3 demonstrated that a CD-associated SNP, rs6651252, demarcates a WRE whose disease-associated allele potentiates stronger binding of TCF7L2 to the enhancer. We confirm an earlier report implicating *MYC* as a target of this particular DRE (171), and build on this finding by identifying the region as a WRE and that either deletion or epigenetic repression of the WRE reduces expression of *MYC* in human colonic epithelial cells (202). We also observe a positive correlation between expression of the disease-associated allele and *MYC* expression in resected intestinal tissue from CD patients. Here, I will discuss implications of this research and provide several future avenues of study that can further our understanding of IBD pathogenesis.

Additional SNPs and targets of interest

Ch8q24 is a gene desert with many disease-associated SNPs

My work focused on rs6651252, which maps to the gene poor region of Ch8q24 (162). We were able to confirm earlier work identifying MYC as the target of rs6651252 in IECs. However, this prior work by Meddens et al. used three cell types involved in IBD, a colorectal cancer cell line and PBMC-derived lymphocytes and monocytes, to identify the targets of rs6651252 (171). While this provided insight into novel DREs and gene targets, inherited or environmentally-driven epigenetic changes can result in differential gene expression within patients, that may be different than the expression profiles of healthy patients used in the study (203). While only *POU5F1B* and *MYC* have been identified as direct, juxtaposed targets of rs6651252 (171), ch8q24 ranges approximately 4.1 Mb and encompasses several other genes including *PVT1, FAM84, GSDMC, ASAP1* and *PRNCR1* (186, 204) that may be additional context-dependent targets.

The genomic region 8q24 itself is of great interest in the context of non-coding functionality, because it contains several SNPs linked with ovarian, prostate and colorectal cancers (186). Our findings that rs6651252 impacts TCF binding is similar to other work on the CRC-associated SNP rs6983267, which also resides within ch8q24 (88, 187, 188). As is the case for rs6651252, co-localization of the epigenetic marker H3K27Ac with rs6983267 coincided with enhanced *MYC* expression (88, 188). Moreover, the disease-associated allele for rs6983267 potentiated stronger binding of TCF7L2, augmented its function as an enhancer (187, 188), and increased the frequency of this region to loop to the *MYC*

promoter (88, 89, 188). Intriguingly, beyond a direct juxtaposition of rs6983267 and *MYC*, Ling et al. identified a distinct mechanism of modulation, through which rs6983267 represses *MYC* (205). rs6983267 resides within the lncRNA, colon-cancer associated transcript two (*CCAT2*), a Wnt-target gene whose expression is elevated in colonic tumors and promotes tumorigenesis when overexpressed (205). Moreover, *CCAT2* stabilizes β-cat/TCF transcription complexes. However, whether the disease-associated variant of rs6983267 impacts these phenotypes is unknown. Critically, a poorly understood lncRNA, *linc00824* is expressed from a region overlapping rs6651252. LncRNAs can modulate gene expression through diverse mechanisms, including quenching miRNA or proteins, and impacting transcription, translation or splicing of target gene transcripts (204). Indeed, in the case of *CCAT2,* its knockdown has previously been shown to reduce the invasiveness of a CRC cell line (205). Therefore, whether *linc00824* expression is directly regulated by the rs6651252-WRE is an important area of future study that may further our understanding of how this SNP impacts IBD pathogenesis.

<u>Other IBD-associated SNPs and β-cat ChIP-seq peaks</u>

Beyond rs6651252, colocalization of IBD-associated SNPs and β-cat ChIP-seq peaks identified several other potential targets. In fact, nearly 40% of all examined IBD-associated SNPs were within 250 kb of a β-cat peak (Fig. 4-1), and further work is necessary to determine whether these SNPs demarcate novel WREs relevant to IBD. Two particular SNPs identified in close proximity of β-cat peaks are rs194749 and rs17293632. rs194749 is an IBD-associated SNP found

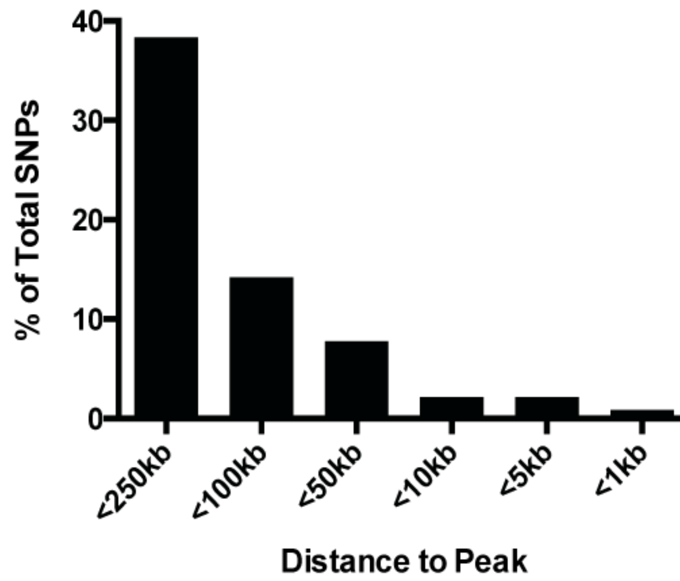


Figure 4-1. Overlap of IBD-associated SNPs with β-cat ChIP-Seq peaks. Proximity of IBD associated SNPs to β-cat binding regions in HCT116 cells. Percentage of total IBD-associated SNPs (n=231).

near zinc finger protein 36 ring finger like protein 1 (*ZFP36L1*), one of three ZFP36 family members of RNA binding proteins (162). These regulatory proteins bind the 3'UTR of select mRNAs and promote their degradation (206). Despite little being known about *ZFP36L1,* another family member known as Tristetraprolin (TTP, encoded for by *ZFP36)* is a potent repressor of pro-inflammatory cytokines and plays a role in IBD. Evidence suggests TTP expression may be indirectly repressed by Wnt signaling (207, 208), but thus far, no evidence has suggested *ZFP36,* or other family members, are direct Wnt-target genes. However, TTP and its family members remain important targets of interest due to their critical role in inflammation and intestinal homeostasis. To this end, an IEC-specific KO of *Zfp36* within intestinal crypts expanded the zone of proliferation, and increased the number of goblet cells (209). Moreover, mice lacking TTP expression in IECs are protected from dextran sodium sulfate (DSS)-induced colitis (209), while full body KO results in a severe inflammatory syndrome resulting from over-expression of TNFα (210). Identification of a novel WRE that regulates *ZFP36L1* expression could provide insight into another inflammatory-driven mechanism for contribution of an IBD-associated SNP to IBD pathogenesis.

The second SNP of interest, rs17293632, resides within an intron encoding the tumor suppressor signaling protein, SMAD family member 3 (*SMAD3*). SMAD3 functions within the transforming growth factor beta (TGFβ) pathway to promote cellular proliferation (211, 212). Even if this SNP is not within a WRE, it could alter splicing of SMAD3, and is therefore a reasonable and intriguing target

to pursue as well. In fact, a few IBD-associated SNPs have been previously identified to alter transcript splicing, including the protective variant within caspase recruitment domain-containing protein 9 (*CARD9)* (213).

rs6651252 is within a putative $CD4^+$ T-cell DRE

While the results presented in this dissertation promotes a model wherein the disease associated allele for rs6651252 elevates *MYC* expression in the colonic epithelium, it is possible that this SNP could function within other cell types that contribute to IBD disease pathogenesis. In support of this possibility, Mokry et al. demonstrated that rs6651252 is within an open region of chromatin in specific subsets of $CD4^+$ T-cells (168). As T-cells are essential for IBD pathogenesis (198), and Wnt signaling is heavily involved in T-cell biology (199), it is possible that the rs6651252-WRE controls *MYC* expression in $CD4^+$ T-cells to regulate the adaptive immune response.

To test this hypothesis, we isolated circulating lymphocytes from CD patients. Using frozen blood RNA samples from patients previously genotyped for rs6651252 using a customized Illumina genotyping platform, we identified CD patients that were heterozygous for the disease (TC) allele, or homozygous for the ancestral (TT) allele. Because patients homozygous for the disease-associated allele (CC) are rare, we are unable to include these patients in our study. We isolated patient RNA from PAXgene® blood RNA tubes, which stabilize RNA from all peripheral blood mononuclear cells (PBMCs) in the patient's sample. We assessed *MYC* expression using RT-qPCR (primers found in table 2-1) from these samples. Through this analysis, expression of *MYC* does not significantly correlate

with allelic status, nor does it differ from the control patients (Fig. 4-2). While we did not observe a significant change in *MYC* expression, this may be a result of examining all PBMCs, not just CD4+ T-cells where we anticipate rs6651252 functions. Assessing MYC from purified T-cells of blood or intestinal tissue can provide important insight into how rs6651252 may impact CD pathogenesis, and is an open area for future study.

Contribution of MYC to disease pathogenesis

<u>MYC in restitution of the intestinal epithelium</u>

The work presented here used surgically resected intestinal tissue to correlate expression of *MYC* with CD patients harboring the disease-associated allele for rs6651252. Limited sample size and a lack of patents homozygous for the disease-associated allele limits interpretations of how strongly elevated *MYC* contributes to disease onset. However, the role of MYC itself in IBD is an important avenue of research. Prolonged inflammation in mice results in recruitment of mesenchymal cells (214), which have previously been shown to supplement Wnt signaling (51-53), and promotes crypt regeneration (214). Indeed, Wnt3a promotes expression of *MYC* which stimulates repair of the intestinal epithelium (215) and, a ~2.5-fold increase in MYC promotes repair of the colonic epithelium (191, 192). Importantly, a previous report using the murine DSS model of acute colitis indicates that epithelial regeneration following inflammatory damage can be improved by inhibition of GSK3 and stabilization of MYC through pharmacological treatment with lithium (95). Together, these reports argue that

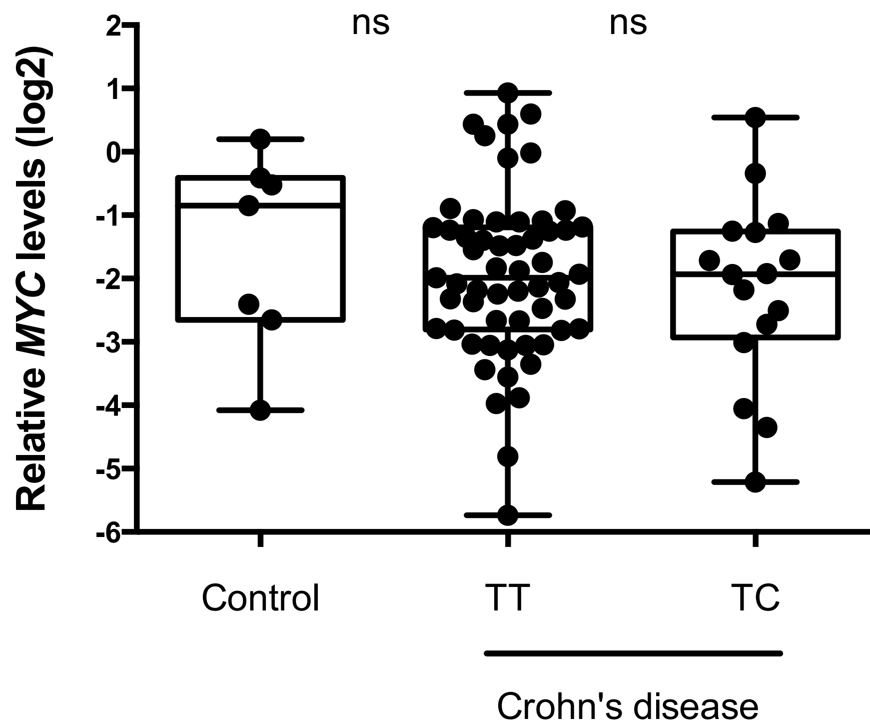


Figure 4-2. The rs6651252 disease-associated allele does not affect *MYC* expression in patient blood lymphocytes. RT-qPCR analysis of *MYC* transcript levels derived from patient blood samples (n=94). For CD patients, the genotype is indicated. Values are normalized to average of *GAPDH* and *TUBULIN B1*. Error is SEM.

acute elevation of MYC expedites restitution, and may promote a better response to damage. However, sustained and elevated MYC expression for 9 months in mice leads to hyperplasia (191). As I will address in greater detail below, these initial findings suggest that where stabilizing MYC expression transiently can promote epithelial restitution, long-term strategies to manage chronic cases of IBD through these mechanisms is not recommended.

Patients with IBD are at an elevated risk to develop a specific form of CRC known as colitis-associated cancer (CAC). In fact, patients with IBD are at a 5-8-fold increased risk for developing CAC (216). Due to the importance of Wnt and MYC in development of sporadic CRC, the role of Wnt and MYC in the development of CAC was tested as well. Several reports have illustrated that within human tissue samples, inflamed intestinal tissue has higher levels of MYC protein and mRNA, over that of control tissue (194). In CAC, MYC expression is not only elevated, but its genomic locus is frequently amplified (195-197). Intriguingly, CAC and CRC have distinct mechanisms of development due to differences in initiation, wherein CAC stems from chronic inflammation. Indeed, mice with colitis developed colonic tumors faster and with fewer doses of a carcinogen than mice without colitis (217, 218). Additional models of CAC has shown that with elevated MYC, treatment with azoxymethane (AOM) prior to DSS-induced colitis resulted in an increased tumor burden, ranging up to 20-fold more depending on the region in the intestines (219). These results, and others, have shown that the overall influence of IBD on CAC depends on disease severity,

age of diagnosis, therapies and management strategies, and appears to rely on many aspects of inflammation (220).

Importantly, the genetic mutations observed in sporadic CRC are largely still present in CAC, but occur at different rates and stages of cancer progression (103, 195, 221). Notably, the most commonly mutated component of the Wnt signaling pathway in sporadic CRC, APC, is found mutated at substantially lower rates in CAC (195, 221). Despite this, the nuclear levels of β-cat are elevated in over 40% of human CAC samples irrespective of APC mutations (195), as well as within animal models of IBD (222), and IBD patient tumors (223, 224). Additional work has suggested that Wnt/β-cat networks are epigenetically modified in IBD, and modified to a greater extent as the intestinal tissue progresses to CAC (225). Indeed, activation of nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB), protein kinase B (AKT) (226) and phosphoinositide 3-kinase (PI-3K) (227) have been shown to stabilize β-cat and play a role in CAC progression, which does not require mutation of Wnt/β-cat signaling pathway components. Therefore, the initial driver mutations of APC in sporadic CRC may be annulled in CAC through stabilization of β-cat via non-traditional mechanisms, such as mutations in the Wnt signaling pathway outside of APC, or the inflammatory mediators prevalent in IBD.

MYC as a rheostat

Nearly 30 years ago, a cellular safeguard against *MYC* over-expression was identified. A seminal discovery by Askew et al. identified that forced expression

of MYC in arrested myeloid cells resulted in robust apoptosis (228). In these myeloid cells, as IL-3 is removed they repressed *MYC,* underwent cell cycle arrest, and ultimately apoptosed. However, ectopic re-expression of MYC did not resume cellular proliferation, but rather hastened the rate of apoptosis in these cells (228). Furthermore, forced expression of MYC resulted in apoptosis in several other instances where cell growth is "inadvisable" for the cell, including cells grown at a high density (229), or while undergoing serum deprivation (230). Importantly, while normal untransformed cells can sense and respond to an elevation in MYC by undergoing apoptosis, it appeared that this response could be turned off within transformed cells (228, 231).

Several years later, an intriguing observation was made within FAP adenocarcinoma samples supporting precise modulation of MYC expression in cancer (232). By sequencing patient samples, Albuquerque et al. demonstrated that if the germline mutation in *APC* resulted in retention of one or two β-cat binding motifs in the expressed protein, the somatic mutation resulted in a truncation removing all β-cat binding from the second allele (232). Conversely, if the germline mutation resulted in no β-cat binding motifs, the somatic mutation resulted in conservation of one or two β-cat binding motifs (232). This argued that the location of the germline mutation within *APC* inherited by patients with FAP influenced the somatic "second-hit" mutation. Atypical for most two-hit models of carcinogenesis, this observation argued some control over β-cat expression by APC was advantageous to tumorigenesis (233). A convergence of these discoveries is illustrated in the "just right" model of tumorigenesis, wherein levels

of MYC must be maintained at a specific threshold to maintain tumorigenesis, but not so high as to promote apoptosis (94). The implications of this model are broad, and suggest Wnt signaling as a whole functions as a rheostat to allow proper expression of its target genes to promote proliferation.

By applying these observations to IBD, a similar model can be inferred, where short-term stabilization of MYC is initially favorable, promoting restitution of the epithelium (95, 191-193), but transitions to a detrimental role in long term disease by promoting apoptosis (Fig. 4-3). The mechanism of rs6651252-WRE disease-associated allele, which elevates MYC, can theoretically induce a similar result. That is, it may assist restitution in acute disease, but later promotes apoptosis and destruction of the IEC barrier, prolonging inflammation and disease. Further work will need to be completed to examine the accuracy of this model. However, targeting and repressing MYC to acceptable levels within these patients may prevent the activation of pro-apoptotic target genes.

Targeting MYC in disease

Due to the critical role of MYC in sporadic CRC and most other cancers, the design of therapies targeting MYC has been a goal in disease treatment for decades (66). Several proof-of-concept compounds have been identified that target and inactivate various aspects of MYC function, such as inhibition of MYC:MAX binding to DNA, or interfering with MYC:MAX dimerization (66). Recent studies have focused on downregulating *MYC* expression at the level of transcript initiation. One compound in particular, JQ1, affects MYC through interfering with bromodomain and extra-terminal (BET) containing proteins (201). Principle

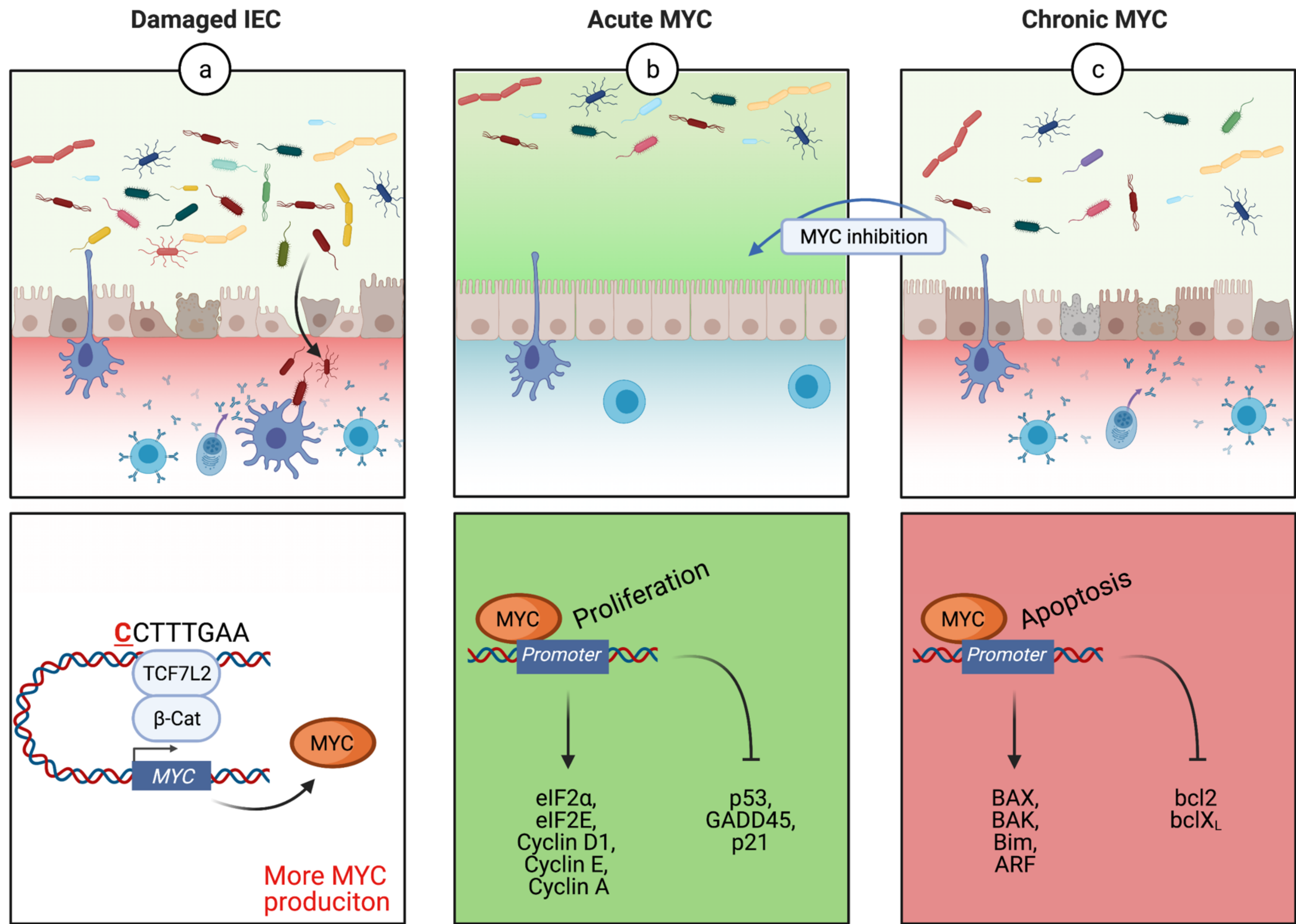

Damaged IEC
a
Acute MYC
b
Chronic MYC
c
MYC inhibition
CCTTTGAA
TCF7L2
β-Cat
MYC
MYC
More MYC
produciton
MYC
Proliferation
Promoter
eIF2α,
eIF2E,
Cyclin D1,
Cyclin E,
Cyclin A
p53,
GADD45,
p21
MYC
Apoptosis
Promoter
BAX,
BAK,
Bim,
ARF
bcl2
bclX$_L$

Figure 4-3. Model of structural changes in intestinal mucosa structure with exposure to acute and chronic MYC overexpression. A; top. In a flare of inflammation due to CD, the IEC barrier is damaged allowing microorganisms and luminal components to infiltrate the underlying tissue. This solicits an immune response wherein the adaptive immune system is activated and immune cells quell invading microorganisms. A; bottom. Within IESC, Wnt signaling is activated in response to the inflammation and damage, stabilizing β-cat and activating expression of Wnt target genes. Germline mutation for the disease-associated allele of rs6651252 (red C) facilitates stronger binding of TCF7L2, which results in elevated *MYC* production. B; top. Proliferative cells expedite restitution, and after all pathogens are removed, immune cells leave allowing for normal homeostasis. B; bottom. A result of elevated MYC, pro-proliferative target genes are activated (65, 72), while cells harboring rs6651252 experience faster repair and restitution. C. Elevated MYC promotes apoptosis through normal mechanisms of gene expression modification(234, 235) to avoid oncogenesis. This results in mediated cell death and IEC permeability. The permeability and cell death results in an immune response and inflammation. Inhibition of MYC in these patients could reduce MYC to normal levels and promote or maintain homeostasis. Figure created with BioRender.com.

members of the BET family are bromodomain (BRD) proteins 2, 3 and 4 (BRD2, BRD3, BRD4). They possess two tandem bromodomains (BD1 and BD2) that serve as acetyl-lysine recognition domains facilitating their functions as coregulators or scaffolding proteins in hyper-acetylated chromatin (236, 237). In fact, BRD4 can recruit transcriptional machinery like pTEFb to activate expression of target genes (238). While the BET family members have overlapping functions, murine knockout studies of specific family members indicate distinct function. Work in several labs have reported that *Brd2* (239-241) or *Brd4* (242) germline knockouts are embryonic lethal, whereas heterozygous knockout of *Brd2* (241) or *Brd4* (242) mice resulted in strong phenotypic defects, suggesting haploinsufficiency. An important discovery through an RNAi screen is apparent in acute myeloid leukemia, which identified Brd4 as a specific target of interest required for maintenance of these cells (243). In fact, suppression of Brd4 by JQ1 was capable of eliminating leukemic stem cells, leading to strong anti-leukemic effects (243).

JQ1 is a cell permeable pan-BRD inhibitor that disrupts anchoring to acetylated proteins through competitive binding to the bromodomain (201). However, JQ1's pan-inhibition includes BRD4 (201), which results in transcriptional regulation of *MYC* and subsequent inhibition of growth of multiple myeloma models (200). Using a dox-inducible shRNA targeting Brd4, Bolden et al. determined strong silencing of Brd4 resulted in depletion of Paneth, goblet and $LGR5^+$ cells after 8 days (244). Importantly, the use of JQ1 did not result in these robust phenotypic changes in mice, yet had strong influence on organoid cultures

(244). As reported elsewhere, this is likely due to poor pharmacokinetic properties of JQ1, which results in rapid clearance from the body (245). These data support prior germline KO models, arguing that complete ablation of BRD4 is dangerous, yet limited knockdown has benefits in repressing tumorigenesis.

Inhibition of BET family members may play an important role in managing inflammatory diseases as well. Critically, BRD4 binds to the p65 subunit of NF-κB, and transcriptionally regulates NF-κB inflammatory targets (246). Two separate BET inhibitors, JQ1 (247) and RVX-297 (248), have shown efficacy in reducing inflammation in murine models of arthritis. Furthermore, these results are attributed to suppression of Th17 differentiation and Th17 cytokine profiles, specifically IL-17, IL-21 and granulocyte macrophage-colony stimulating factor (GCSF) (247, 249). While these findings are important for rheumatoid arthritis, Th17 T-cells are highly prevalent in CD (116), suggesting that treatment may also benefit patients with CD beyond reduction of *MYC* expression. Together these results argue that as the next generation of BET inhibitors is developed, these compounds can be used to treat multiple diseases through diverse mechanisms. Work in this field is promising, and several compounds have been designed with changes from JQ1 structure since its development in 2010, in attempts to improve its function (e.g., (250-252). While second generation BRD4 inhibitors are being developed with improved efficacy (253), bioavailability and greater selectivity (254).

In summary, our findings that patients with the disease-associated allele for rs6651252 may have elevated MYC expression suggest they may benefit from

therapies that target MYC or MYC-regulated genes. These strategies can repress MYC to normal levels, and protect the patient from negative effects of prolonged MYC overexpression. Prior work has shown that *MYC* expression is sensitive to the pan-inhibitors targeting BET family of proteins, like JQ1 (200, 201). Our findings presented here support the idea that as better BET inhibitors are developed, and the role of MYC in IBD becomes more clear, pre-clinical mouse models can provide useful augmenting therapies to improve quality of life, and alter MYC expression to protect against damage caused by chronic IBD.

References

1. Greenwood-Van Meerveld B, Johnson AC, Grundy D. Gastrointestinal Physiology and Function. Handb Exp Pharmacol. 2017;239:1-16.

2. Bleier JIS, Wilkins KB. Colonic Physiology. In: Steele SR, Hull TL, Read TE, Saclarides TJ, Senagore AJ, Whitlow CB, editors. The ASCRS Textbook of Colon and Rectal Surgery. Cham: Springer International Publishing; 2016. p. 27-35.

3. Barker N. Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration. Nat Rev Mol Cell Biol. 2014;15(1):19-33.

4. Helander HF, Fandriks L. Surface area of the digestive tract - revisited. Scand J Gastroenterol. 2014;49(6):681-9.

5. Snoeck V, Goddeeris B, Cox E. The role of enterocytes in the intestinal barrier function and antigen uptake. Microbes Infect. 2005;7(7-8):997-1004.

6. Miron N, Cristea V. Enterocytes: active cells in tolerance to food and microbial antigens in the gut. Clin Exp Immunol. 2012;167(3):405-12.

7. Farin HF, Jordens I, Mosa MH, Basak O, Korving J, Tauriello DV, de Punder K, Angers S, Peters PJ, Maurice MM, Clevers H. Visualization of a short-range Wnt gradient in the intestinal stem-cell niche. Nature. 2016;530(7590):340-3.

8. Allaire JM, Morampudi V, Crowley SM, Stahl M, Yu H, Bhullar K, Knodler LA, Bressler B, Jacobson K, Vallance BA. Frontline defenders: goblet cell mediators dictate host-microbe interactions in the intestinal tract during health and disease. Am J Physiol Gastrointest Liver Physiol. 2018;314(3):G360-g77.

9. Gerbe F, Legraverend C, Jay P. The intestinal epithelium tuft cells: specification and function. Cell Mol Life Sci. 2012;69(17):2907-17.

10. Nevo S, Kadouri N, Abramson J. Tuft cells: From the mucosa to the thymus. Immunol Lett. 2019;210:1-9.

11. Jean-Pierre K, Neutra MR. Epithelial M cells: Differentiation and function. Annual Review of Cell and Developmental Biology. 2000;16.

12. Chia LA, Kuo CJ. The intestinal stem cell. Prog Mol Biol Transl Sci. 2010;96:157-73.

13. Nusse R, Clevers H. Wnt/beta-Catenin Signaling, Disease, and Emerging Therapeutic Modalities. Cell. 2017;169(6):985-99.

14. Cheng H, Leblond CP. Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. V. Unitarian Theory of the origin of the four epithelial cell types. Am J Anat. 1974;141(4):537-61.

15. Barker N, van Es JH, Kuipers J, Kujala P, van den Born M, Cozijnsen M, Haegebarth A, Korving J, Begthel H, Peters PJ, Clevers H. Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature. 2007;449(7165):1003-7.

16. Marshman E, Booth C, Potten CS. The intestinal epithelial stem cell. Bioessays. 2002;24(1):91-8.

17. Niehrs C, Acebron SP. Mitotic and mitogenic Wnt signalling. Embo j. 2012;31(12):2705-13.

18. Nusse R, van Ooyen A, Cox D, Fung YK, Varmus H. Mode of proviral activation of a putative mammary oncogene (int-1) on mouse chromosome 15. Nature. 1984;307(5947):131-6.

19. van Ooyen A, Nusse R. Structure and nucleotide sequence of the putative mammary oncogene int-1; proviral insertions leave the protein-encoding domain intact. Cell. 1984;39(1):233-40.

20. Rijsewijk F, Schuermann M, Wagenaar E, Parren P, Weigel D, Nusse R. The Drosophila homolog of the mouse mammary oncogene int-1 is identical to the segment polarity gene wingless. Cell. 1987;50(4):649-57.

21. Sharma RP. Wingless, a new mutant in Drosophila melanogaster. Drosophila Information Service. 1973;50:134.

22. Nusse R, Brown A, Papkoff J, Scambler P, Shackleford G, McMahon A, Moon R, Varmus H. A new nomenclature for int-1 and related genes: the Wnt gene family. Cell. United States1991. p. 231.

23. McMahon AP, Joyner AL, Bradley A, McMahon JA. The midbrain-hindbrain phenotype of Wnt-1-/Wnt-1- mice results from stepwise deletion of engrailed-expressing cells by 9.5 days postcoitum. Cell. 1992;69(4):581-95.

24. Ikeya M, Lee SM, Johnson JE, McMahon AP, Takada S. Wnt signalling required for expansion of neural crest and CNS progenitors. Nature. 1997;389(6654):966-70.

25. Stark K, Vainio S, Vassileva G, McMahon AP. Epithelial transformation of metanephric mesenchyme in the developing kidney regulated by Wnt-4. Nature. 1994;372(6507):679-83.

26. Nüsslein-Volhard C, Wieschaus E. Mutations affecting segment number and polarity in Drosophila. Nature. 1980;287(5785):795-801.

27. Cadigan KM, Waterman ML. TCF/LEFs and Wnt signaling in the nucleus. Cold Spring Harb Perspect Biol. 2012;4(11).

28. Kohn AD, Moon RT. Wnt and calcium signaling: beta-catenin-independent pathways. Cell Calcium. 2005;38(3-4):439-46.

29. Dijksterhuis JP, Baljinnyam B, Stanger K, Sercan HO, Ji Y, Andres O, Rubin JS, Hannoush RN, Schulte G. Systematic mapping of WNT-FZD protein interactions reveals functional selectivity by distinct WNT-FZD pairs. J Biol Chem. 2015;290(11):6789-98.

30. van de Wetering M, Oosterwegel M, Dooijes D, Clevers H. Identification and cloning of TCF-1, a T lymphocyte-specific transcription factor containing a sequence-specific HMG box. Embo j. 1991;10(1):123-32.

31. Giese K, Amsterdam A, Grosschedl R. DNA-binding properties of the HMG domain of the lymphoid-specific transcriptional regulator LEF-1. Genes Dev. 1991;5(12b):2567-78.

32. Behrens J, von Kries JP, Kühl M, Bruhn L, Wedlich D, Grosschedl R, Birchmeier W. Functional interaction of beta-catenin with the transcription factor LEF-1. Nature. 1996;382(6592):638-42.

33. von Kries JP, Winbeck G, Asbrand C, Schwarz-Romond T, Sochnikova N, Dell'Oro A, Behrens J, Birchmeier W. Hot spots in beta-catenin for interactions with LEF-1, conductin and APC. Nat Struct Biol. 2000;7(9):800-7.

34. Liu C, Li Y, Semenov M, Han C, Baeg GH, Tan Y, Zhang Z, Lin X, He X. Control of beta-catenin phosphorylation/degradation by a dual-kinase mechanism. Cell. 2002;108(6):837-47.

35. Aberle H, Bauer A, Stappert J, Kispert A, Kemler R. beta-catenin is a target for the ubiquitin-proteasome pathway. Embo j. 1997;16(13):3797-804.

36. Liu C, Kato Y, Zhang Z, Do VM, Yankner BA, He X. beta-Trcp couples beta-catenin phosphorylation-degradation and regulates Xenopus axis formation. Proc Natl Acad Sci U S A. 1999;96(11):6273-8.

37. MacDonald BT, Tamai K, He X. Wnt/beta-catenin signaling: components, mechanisms, and diseases. Dev Cell. 2009;17(1):9-26.

38. Mosimann C, Hausmann G, Basler K. Beta-catenin hits chromatin: regulation of Wnt target gene activation. Nat Rev Mol Cell Biol. 2009;10(4):276-86.

39. van Beest M, Dooijes D, van De Wetering M, Kjaerulff S, Bonvin A, Nielsen O, Clevers H. Sequence-specific high mobility group box factors recognize 10-12-base pair minor groove motifs. J Biol Chem. 2000;275(35):27266-73.

40. Van de Wetering M, Castrop J, Korinek V, Clevers H. Extensive alternative splicing and dual promoter usage generate Tcf-1 protein isoforms with differential transcription control properties. Mol Cell Biol. 1996;16(3):745-52.

41. Molenaar M, van de Wetering M, Oosterwegel M, Peterson-Maduro J, Godsave S, Korinek V, Roose J, Destrée O, Clevers H. XTcf-3 transcription factor mediates beta-catenin-induced axis formation in Xenopus embryos. Cell. 1996;86(3):391-9.

42. Roose J, Huls G, van Beest M, Moerer P, van der Horn K, Goldschmeding R, Logtenberg T, Clevers H. Synergy between tumor suppressor APC and the beta-catenin-Tcf4 target Tcf1. Science. 1999;285(5435):1923-6.

43. Travis A, Amsterdam A, Belanger C, Grosschedl R. LEF-1, a gene encoding a lymphoid-specific protein with an HMG domain, regulates T-cell receptor alpha enhancer function [corrected]. Genes Dev. 1991;5(5):880-94.

44. Hoppler S, Kavanagh CL. Wnt signalling: variety at the core. J Cell Sci. 2007;120(Pt 3):385-93.

45. Korinek V, Barker N, Moerer P, van Donselaar E, Huls G, Peters PJ, Clevers H. Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4. Nat Genet. 1998;19(4):379-83.

46. van Es JH, Haegebarth A, Kujala P, Itzkovitz S, Koo BK, Boj SF, Korving J, van den Born M, van Oudenaarden A, Robine S, Clevers H. A critical role for the Wnt effector Tcf4 in adult intestinal homeostatic self-renewal. Mol Cell Biol. 2012;32(10):1918-27.

47. Fevr T, Robine S, Louvard D, Huelsken J. Wnt/beta-catenin is essential for intestinal homeostasis and maintenance of intestinal stem cells. Mol Cell Biol. 2007;27(21):7551-9.

48. Pinto D, Gregorieff A, Begthel H, Clevers H. Canonical Wnt signals are essential for homeostasis of the intestinal epithelium. Genes Dev. 2003;17(14):1709-13.

49. Sato T, Vries RG, Snippert HJ, van de Wetering M, Barker N, Stange DE, van Es JH, Abo A, Kujala P, Peters PJ, Clevers H. Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. Nature. 2009;459(7244):262-5.

50. Sato T, Clevers H. Growing self-organizing mini-guts from a single intestinal stem cell: mechanism and applications. Science. 2013;340(6137):1190-4.

51. Farin HF, Van Es JH, Clevers H. Redundant sources of Wnt regulate intestinal stem cells and promote formation of Paneth cells. Gastroenterology. 2012;143(6):1518-29.e7.

52. Gregorieff A, Pinto D, Begthel H, Destrée O, Kielman M, Clevers H. Expression pattern of Wnt signaling components in the adult intestine. Gastroenterology. 2005;129(2):626-38.

53. Kabiri Z, Greicius G, Madan B, Biechele S, Zhong Z, Zaribafzadeh H, Edison, Aliyev J, Wu Y, Bunte R, Williams BO, Rossant J, Virshup DM. Stroma provides an intestinal stem cell niche in the absence of epithelial Wnts. Development. 2014;141(11):2206-15.

54. Tetteh PW, Basak O, Farin HF, Wiebrands K, Kretzschmar K, Begthel H, van den Born M, Korving J, de Sauvage F, van Es JH, van Oudenaarden A, Clevers H. Replacement of Lost Lgr5-Positive Stem Cells through Plasticity of Their Enterocyte-Lineage Daughters. Cell Stem Cell. 2016;18(2):203-13.

55. Tian H, Biehs B, Warming S, Leong KG, Rangell L, Klein OD, de Sauvage FJ. A reserve stem cell population in small intestine renders Lgr5-positive cells dispensable. Nature. 2011;478(7368):255-9.

56. Reya T, Okamura R, Grosschedl R. Control of lymphocyte differentiation by the LEF-1/TCF family of transcription factors. Cold Spring Harb Symp Quant Biol. 1999;64:133-40.

57. Merrill BJ, Pasolli HA, Polak L, Rendl M, García-García MJ, Anderson KV, Fuchs E. Tcf3: a transcriptional regulator of axis induction in the early embryo. Development. 2004;131(2):263-74.

58. Galceran J, Fariñas I, Depew MJ, Clevers H, Grosschedl R. Wnt3a-/--like phenotype and limb deficiency in Lef1(-/-)Tcf1(-/-) mice. Genes Dev. 1999;13(6):709-17. Epub 1999/03/25. doi: 10.1101/gad.13.6.709.

59. Gregorieff A, Grosschedl R, Clevers H. Hindgut defects and transformation of the gastro-intestinal tract in Tcf4(-/-)/Tcf1(-/-) embryos. Embo j. 2004;23(8):1825-33.

60. Nishisho I, Nakamura Y, Miyoshi Y, Miki Y, Ando H, Horii A, Koyama K, Utsunomiya J, Baba S, Hedge P. Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. Science. 1991;253(5020):665-9.

61. Kinzler KW, Nilbert MC, Vogelstein B, Bryan TM, Levy DB, Smith KJ, Preisinger AC, Hamilton SR, Hedge P, Markham A, et al. Identification of a gene located at chromosome 5q21 that is mutated in colorectal cancers. Science. 1991;251(4999):1366-70.

62. Powell SM, Zilz N, Beazer-Barclay Y, Bryan TM, Hamilton SR, Thibodeau SN, Vogelstein B, Kinzler KW. APC mutations occur early during colorectal tumorigenesis. Nature. 1992;359(6392):235-7.

63. Miyoshi Y, Nagase H, Ando H, Horii A, Ichii S, Nakatsuru S, Aoki T, Miki Y, Mori T, Nakamura Y. Somatic mutations of the APC gene in colorectal tumors: mutation cluster region in the APC gene. Hum Mol Genet. 1992;1(4):229-33.

64. He TC, Sparks AB, Rago C, Hermeking H, Zawel L, da Costa LT, Morin PJ, Vogelstein B, Kinzler KW. Identification of c-MYC as a target of the APC pathway. Science. 1998;281(5382):1509-12.

65. Eilers M, Eisenman RN. Myc's broad reach. Genes Dev. 2008;22(20):2755-66.

66. Elbadawy M, Usui T, Yamawaki H, Sasaki K. Emerging Roles of C-Myc in Cancer Stem Cell-Related Signaling and Resistance to Cancer Chemotherapy: A Potential Therapeutic Target Against Colorectal Cancer. Int J Mol Sci. 2019;20(9).

67. Blackwood EM, Eisenman RN. Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. Science. 1991;251(4998):1211-7.

68. Blackwell TK, Kretzner L, Blackwood EM, Eisenman RN, Weintraub H. Sequence-specific DNA binding by the c-Myc protein. Science. 1990;250(4984):1149-51.

69. Kato GJ, Barrett J, Villa-Garcia M, Dang CV. An amino-terminal c-myc domain required for neoplastic transformation activates transcription. Mol Cell Biol. 1990;10(11):5914-20.

70. McMahon SB, Wood MA, Cole MD. The essential cofactor TRRAP recruits the histone acetyltransferase hGCN5 to c-Myc. Mol Cell Biol. 2000;20(2):556-62.

71. Frank SR, Schroeder M, Fernandez P, Taubert S, Amati B. Binding of c-Myc to chromatin mediates mitogen-induced acetylation of histone H4 and gene activation. Genes Dev. 2001;15(16):2069-82.

72. Dang CV, O'Donnell KA, Zeller KI, Nguyen T, Osthus RC, Li F. The c-Myc target gene network. Semin Cancer Biol. 2006;16(4):253-64.

73. Peukert K, Staller P, Schneider A, Carmichael G, Hänel F, Eilers M. An alternative pathway for gene regulation by Myc. Embo j. 1997;16(18):5672-86.

74. Staller P, Peukert K, Kiermaier A, Seoane J, Lukas J, Karsunky H, Möröy T, Bartek J, Massagué J, Hänel F, Eilers M. Repression of p15INK4b expression by Myc through association with Miz-1. Nat Cell Biol. 2001;3(4):392-9.

75. Mateyak MK, Obaya AJ, Adachi S, Sedivy JM. Phenotypes of c-Myc-deficient rat fibroblasts isolated by targeted homologous recombination. Cell Growth Differ. 1997;8(10):1039-48.

76. Tansey WP. Mammalian MYC Proteins and Cancer. New Journal of Science. 2014.

77. Dalla-Favera R, Bregni M, Erikson J, Patterson D, Gallo RC, Croce CM. Human c-myc onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. Proc Natl Acad Sci U S A. 1982;79(24):7824-7.

78. Dave SS, Fu K, Wright GW, Lam LT, Kluin P, Boerma EJ, Greiner TC, Weisenburger DD, Rosenwald A, Ott G, Müller-Hermelink HK, Gascoyne RD, Delabie J, Rimsza LM, Braziel RM, Grogan TM, Campo E, Jaffe ES, Dave BJ, Sanger W, et al. Molecular diagnosis of Burkitt's lymphoma. N Engl J Med. 2006;354(23):2431-42.

79. Taub R, Kirsch I, Morton C, Lenoir G, Swan D, Tronick S, Aaronson S, Leder P. Translocation of the c-myc gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cells. Proc Natl Acad Sci U S A. 1982;79(24):7837-41.

80. Vita M, Henriksson M. The Myc oncoprotein as a therapeutic target for human cancer. Semin Cancer Biol. 2006;16(4):318-30.

81. Vogelstein B, Fearon ER, Hamilton SR, Kern SE, Preisinger AC, Leppert M, Nakamura Y, White R, Smits AM, Bos JL. Genetic alterations during colorectal-tumor development. N Engl J Med. 1988;319(9):525-32.

82. Comprehensive molecular characterization of human colon and rectal cancer. Nature. 2012;487(7407):330-7.

83. Sierra J, Yoshida T, Joazeiro CA, Jones KA. The APC tumor suppressor counteracts beta-catenin activation and H3K4 methylation at Wnt target genes. Genes Dev. 2006;20(5):586-600.

84. Yochum GS, McWeeney S, Rajaraman V, Cleland R, Peters S, Goodman RH. Serial analysis of chromatin occupancy identifies beta-catenin target genes in colorectal carcinoma cells. Proc Natl Acad Sci U S A. 2007;104(9):3324-9.

85. Yochum GS, Cleland R, Goodman RH. A genome-wide screen for beta-catenin binding sites identifies a downstream enhancer element that controls c-Myc gene expression. Mol Cell Biol. 2008;28(24):7368-79.

86. Mautner J, Joos S, Werner T, Eick D, Bornkamm GW, Polack A. Identification of two enhancer elements downstream of the human c-myc gene. Nucleic Acids Res. 1995;23(1):72-80.

87. Yochum GS, Sherrick CM, Macpartlin M, Goodman RH. A beta-catenin/TCF-coordinated chromatin loop at MYC integrates 5' and 3' Wnt responsive enhancers. Proc Natl Acad Sci U S A. 2010;107(1):145-50.

88. Pomerantz MM, Ahmadiyeh N, Jia L, Herman P, Verzi MP, Doddapaneni H, Beckwith CA, Chan JA, Hills A, Davis M, Yao K, Kehoe SM, Lenz HJ, Haiman CA, Yan C, Henderson BE, Frenkel B, Barretina J, Bass A, Tabernero J, Baselga

J, Regan MM, Manak JR, Shivdasani R, Coetzee GA, Freedman ML. The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. Nat Genet. 2009;41(8):882-4.

89. Ahmadiyeh N, Pomerantz MM, Grisanzio C, Herman P, Jia L, Almendro V, He HH, Brown M, Liu XS, Davis M, Caswell JL, Beckwith CA, Hills A, Macconaill L, Coetzee GA, Regan MM, Freedman ML. 8q24 prostate, breast, and colon cancer risk loci show tissue-specific long-range interaction with MYC. Proc Natl Acad Sci U S A. 2010;107(21):9742-6.

90. Sotelo J, Esposito D, Duhagon MA, Banfield K, Mehalko J, Liao H, Stephens RM, Harris TJ, Munroe DJ, Wu X. Long-range enhancers on 8q24 regulate c-Myc. Proc Natl Acad Sci U S A. 2010;107(7):3001-5.

91. Bottomly D, Kyler SL, McWeeney SK, Yochum GS. Identification of {beta}-catenin binding regions in colon cancer cells using ChIP-Seq. Nucleic Acids Res. 2010;38(17):5735-45.

92. Hnisz D, Schuijers J, Lin CY, Weintraub AS, Abraham BJ, Lee TI, Bradner JE, Young RA. Convergence of developmental and oncogenic signaling pathways at transcriptional super-enhancers. Mol Cell. 2015;58(2):362-70.

93. Yochum GS. Multiple Wnt/ss-catenin responsive enhancers align with the MYC promoter through long-range chromatin loops. PLoS One. 2011;6(4):e18966.

94. Murphy DJ, Junttila MR, Pouyet L, Karnezis A, Shchors K, Bui DA, Brown-Swigart L, Johnson L, Evan GI. Distinct thresholds govern Myc's biological output in vivo. Cancer Cell. 2008;14(6):447-57.

95. Raup-Konsavage WM, Cooper TK, Yochum GS. A Role for MYC in Lithium-Stimulated Repair of the Colonic Epithelium After DSS-Induced Damage in Mice. Dig Dis Sci. 2016;61(2):410-22.

96. Xing Y, Chen X, Cao Y, Huang J, Xie X, Wei Y. Expression of Wnt and Notch signaling pathways in inflammatory bowel disease treated with mesenchymal stem cell transplantation: evaluation in a rat model. Stem Cell Res Ther. 2015;6(1):101.

97. Bandukwala HS, Gagnon J, Togher S, Greenbaum JA, Lamperti ED, Parr NJ, Molesworth AM, Smithers N, Lee K, Witherington J, Tough DF, Prinjha RK, Peters B, Rao A. Selective inhibition of CD4+ T-cell cytokine production and autoimmunity by BET protein and c-Myc inhibitors. Proc Natl Acad Sci U S A. 2012;109(36):14532-7.

98. Macpherson AJ, Chester KA, Robson L, Bjarnason I, Malcolm AD, Peters TJ. Increased expression of c-myc proto-oncogene in biopsies of ulcerative colitis and Crohn's colitis. Gut. 1992;33(5):651-6.

99. Alexander RJ, Panja A, Kaplan-Liss E, Mayer L, Raicht RF. Expression of protooncogene-encoded mRNA by colonic epithelial cells in inflammatory bowel disease. Dig Dis Sci. 1996;41(4):660-9.

100. Kaser A, Zeissig S, Blumberg RS. Inflammatory bowel disease. Annu Rev Immunol. 2010;28:573-621. Epub 2010/03/03. doi: 10.1146/annurev-immunol-030409-101225.

101. Philp MM, Ross HM. IBD Diagnosis and Evaluation. In: Steele SR, Hull TL, Read TE, Saclarides TJ, Senagore AJ, Whitlow CB, editors. The ASCRS Textbook of Colon and Rectal Surgery. Cham: Springer International Publishing; 2016. p. 771-86.

102. Holubar SD, Soop M. Medical Management of Chronic Ulcerative Colitis. In: Steele SR, Hull TL, Read TE, Saclarides TJ, Senagore AJ, Whitlow CB, editors. The ASCRS Textbook of Colon and Rectal Surgery. Cham: Springer International Publishing; 2016. p. 787-806.

103. Connelly TM, Koltun WA. Inflammatory Bowel Disease: Pathobiology. In: Steele SR, Hull TL, Read TE, Saclarides TJ, Senagore AJ, Whitlow CB, editors. The ASCRS Textbook of Colon and Rectal Surgery. Cham: Springer International Publishing; 2016. p. 749-69.

104. Strong SA. Medical Management of Crohn's Disease. In: Steele SR, Hull TL, Read TE, Saclarides TJ, Senagore AJ, Whitlow CB, editors. The ASCRS Textbook of Colon and Rectal Surgery. Cham: Springer International Publishing; 2016. p. 807-17.

105. Silverberg MS, Satsangi J, Ahmad T, Arnott ID, Bernstein CN, Brant SR, Caprilli R, Colombel JF, Gasche C, Geboes K, Jewell DP, Karban A, Loftus EV, Jr., Peña AS, Riddell RH, Sachar DB, Schreiber S, Steinhart AH, Targan SR, Vermeire S, Warren BF. Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. Can J Gastroenterol. 2005;19 Suppl A:5a-36a.

106. Feagan BG, Macdonald JK. Oral 5-aminosalicylic acid for induction of remission in ulcerative colitis. Cochrane Database Syst Rev. 2012;10:Cd000543.

107. Sands BE, Tremaine WJ, Sandborn WJ, Rutgeerts PJ, Hanauer SB, Mayer L, Targan SR, Podolsky DK. Infliximab in the treatment of severe, steroid-refractory ulcerative colitis: a pilot study. Inflamm Bowel Dis. 2001;7(2):83-8.

108. Gustavsson A, Järnerot G, Hertervig E, Friis-Liby I, Blomquist L, Karlén P, Grännö C, Vilien M, Ström M, Verbaan H, Hellström PM, Magnuson A, Halfvarson J, Tysk C. Clinical trial: colectomy after rescue therapy in ulcerative colitis - 3-year follow-up of the Swedish-Danish controlled infliximab study. Aliment Pharmacol Ther. 2010;32(8):984-9.

109. Singh S, Garg SK, Pardi DS, Wang Z, Murad MH, Loftus EV, Jr. Comparative efficacy of biologic therapy in biologic-naïve patients with Crohn disease: a systematic review and network meta-analysis. Mayo Clin Proc. 2014;89(12):1621-35.

110. Samuel S, Ingle SB, Dhillon S, Yadav S, Harmsen WS, Zinsmeister AR, Tremaine WJ, Sandborn WJ, Loftus EV, Jr. Cumulative incidence and risk factors for hospitalization and surgery in a population-based cohort of ulcerative colitis. Inflamm Bowel Dis. 2013;19(9):1858-66.

111. Frolkis AD, Dykeman J, Negrón ME, Debruyn J, Jette N, Fiest KM, Frolkis T, Barkema HW, Rioux KP, Panaccione R, Ghosh S, Wiebe S, Kaplan GG. Risk of surgery for inflammatory bowel diseases has decreased over time: a systematic review and meta-analysis of population-based studies. Gastroenterology. 2013;145(5):996-1006.

112. Kühn R, Löhler J, Rennick D, Rajewsky K, Müller W. Interleukin-10-deficient mice develop chronic enterocolitis. Cell. 1993;75(2):263-74.

113. Sadlack B, Merz H, Schorle H, Schimpl A, Feller AC, Horak I. Ulcerative colitis-like disease in mice with a disrupted interleukin-2 gene. Cell. 1993;75(2):253-61.

114. Mombaerts P, Mizoguchi E, Grusby MJ, Glimcher LH, Bhan AK, Tonegawa S. Spontaneous development of inflammatory bowel disease in T cell receptor mutant mice. Cell. 1993;75(2):274-82.

115. Chen ML, Sundrud MS. Cytokine Networks and T-Cell Subsets in Inflammatory Bowel Diseases. Inflamm Bowel Dis. 2016;22(5):1157-67.

116. Fuss IJ, Neurath M, Boirivant M, Klein JS, de la Motte C, Strong SA, Fiocchi C, Strober W. Disparate CD4+ lamina propria (LP) lymphokine secretion profiles in inflammatory bowel disease. Crohn's disease LP cells manifest increased secretion of IFN-gamma, whereas ulcerative colitis LP cells manifest increased secretion of IL-5. J Immunol. 1996;157(3):1261-70.

117. Sellon RK, Tonkonogy S, Schultz M, Dieleman LA, Grenther W, Balish E, Rennick DM, Sartor RB. Resident enteric bacteria are necessary for development of spontaneous colitis and immune system activation in interleukin-10-deficient mice. Infect Immun. 1998;66(11):5224-31.

118. Taurog JD, Richardson JA, Croft JT, Simmons WA, Zhou M, Fernández-Sueiro JL, Balish E, Hammer RE. The germfree state prevents development of gut and joint inflammatory disease in HLA-B27 transgenic rats. J Exp Med. 1994;180(6):2359-64.

119. Cong Y, Brandwein SL, McCabe RP, Lazenby A, Birkenmeier EH, Sundberg JP, Elson CO. CD4+ T cells reactive to enteric bacterial antigens in spontaneously

colitic C3H/HeJBir mice: increased T helper cell type 1 response and ability to transfer disease. J Exp Med. 1998;187(6):855-64.

120. Iqbal N, Oliver JR, Wagner FH, Lazenby AS, Elson CO, Weaver CT. T helper 1 and T helper 2 cells are pathogenic in an antigen-specific model of colitis. J Exp Med. 2002;195(1):71-84.

121. Rath HC, Herfarth HH, Ikeda JS, Grenther WB, Hamm TE, Jr., Balish E, Taurog JD, Hammer RE, Wilson KH, Sartor RB. Normal luminal bacteria, especially Bacteroides species, mediate chronic colitis, gastritis, and arthritis in HLA-B27/human beta2 microglobulin transgenic rats. J Clin Invest. 1996;98(4):945-53.

122. Strachan DP. Hay fever, hygiene, and household size. Bmj. 1989;299(6710):1259-60.

123. Olszak T, An D, Zeissig S, Vera MP, Richter J, Franke A, Glickman JN, Siebert R, Baron RM, Kasper DL, Blumberg RS. Microbial exposure during early life has persistent effects on natural killer T cell function. Science. 2012;336(6080):489-93.

124. Nishida A, Inoue R, Inatomi O, Bamba S, Naito Y, Andoh A. Gut microbiota in the pathogenesis of inflammatory bowel disease. Clin J Gastroenterol. 2018;11(1):1-10.

125. Tamboli CP, Neut C, Desreumaux P, Colombel JF. Dysbiosis in inflammatory bowel disease. Gut. 2004;53(1):1-4.

126. Frank DN, St Amand AL, Feldman RA, Boedeker EC, Harpaz N, Pace NR. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. Proc Natl Acad Sci U S A. 2007;104(34):13780-5.

127. Damman CJ, Miller SI, Surawicz CM, Zisman TL. The microbiome and inflammatory bowel disease: is there a therapeutic role for fecal microbiota transplantation? Am J Gastroenterol. 2012;107(10):1452-9.

128. Khoruts A, Weingarden AR. Emergence of fecal microbiota transplantation as an approach to repair disrupted microbial gut ecology. Immunol Lett. 2014;162(2 Pt A):77-81.

129. Brandt LJ, Reddy SS. Fecal microbiota transplantation for recurrent clostridium difficile infection. J Clin Gastroenterol. 2011;45 Suppl:S159-67.

130. Capaldo CT, Powell DN, Kalman D. Layered defense: how mucus and tight junctions seal the intestinal barrier. J Mol Med (Berl). 2017;95(9):927-34.

131. Vetrano S, Rescigno M, Cera MR, Correale C, Rumio C, Doni A, Fantini M, Sturm A, Borroni E, Repici A, Locati M, Malesci A, Dejana E, Danese S. Unique

role of junctional adhesion molecule-a in maintaining mucosal homeostasis in inflammatory bowel disease. Gastroenterology. 2008;135(1):173-84.

132. Oshima T, Miwa H, Joh T. Changes in the expression of claudins in active ulcerative colitis. J Gastroenterol Hepatol. 2008;23 Suppl 2:S146-50.

133. Laukoetter MG, Nava P, Lee WY, Severson EA, Capaldo CT, Babbin BA, Williams IR, Koval M, Peatman E, Campbell JA, Dermody TS, Nusrat A, Parkos CA. JAM-A regulates permeability and inflammation in the intestine in vivo. J Exp Med. 2007;204(13):3067-76.

134. Dorofeyev AE, Vasilenko IV, Rassokhina OA, Kondratiuk RB. Mucosal barrier in ulcerative colitis and Crohn's disease. Gastroenterol Res Pract. 2013;2013:431231.

135. Wenzel UA, Magnusson MK, Rydström A, Jonstrand C, Hengst J, Johansson ME, Velcich A, Öhman L, Strid H, Sjövall H, Hansson GC, Wick MJ. Spontaneous colitis in Muc2-deficient mice reflects clinical and cellular features of active ulcerative colitis. PLoS One. 2014;9(6):e100217.

136. Lu P, Burger-van Paassen N, van der Sluis M, Witte-Bouma J, Kerckaert JP, van Goudoever JB, Van Seuningen I, Renes IB. Colonic gene expression patterns of mucin Muc2 knockout mice reveal various phases in colitis development. Inflamm Bowel Dis. 2011;17(10):2047-57.

137. Eshelman MA, Jeganathan NA, Schieffer KM, Kline BP, Mendenhall M, Deiling S, Harris L, Koltun WA, Yochum GS. Elevated Colonic Mucin Expression Correlates with Extended Time to Surgery for Ulcerative Colitis Patients. J Gastrointestin Liver Dis. 2019;28(4):405-13.

138. Rehman A, Rausch P, Wang J, Skieceviciene J, Kiudelis G, Bhagalia K, Amarapurkar D, Kupcinskas L, Schreiber S, Rosenstiel P, Baines JF, Ott S. Geographical patterns of the standing and active human gut microbiome in health and IBD. Gut. 2016;65(2):238-48.

139. Montgomery SM, Morris DL, Pounder RE, Wakefield AJ. Asian ethnic origin and the risk of inflammatory bowel disease. Eur J Gastroenterol Hepatol. 1999;11(5):543-6.

140. Soon IS, Molodecky NA, Rabi DM, Ghali WA, Barkema HW, Kaplan GG. The relationship between urban environment and the inflammatory bowel diseases: a systematic review and meta-analysis. BMC Gastroenterol. 2012;12:51.

141. Kaplan GG, Ng SC. Understanding and Preventing the Global Increase of Inflammatory Bowel Disease. Gastroenterology. 2017;152(2):313-21.e2.

142. Mahmud N, Weir DG. The urban diet and Crohn's disease: is there a relationship? Eur J Gastroenterol Hepatol. 2001;13(2):93-5.

143. Felder JB, Korelitz BI, Rajapakse R, Schwarz S, Horatagis AP, Gleim G. Effects of nonsteroidal antiinflammatory drugs on inflammatory bowel disease: a case-control study. Am J Gastroenterol. 2000;95(8):1949-54.

144. McGovern DP, Kugathasan S, Cho JH. Genetics of Inflammatory Bowel Diseases. Gastroenterology. 2015;149(5):1163-76.e2..

145. Ek WE, D'Amato M, Halfvarson J. The history of genetics in inflammatory bowel disease. Ann Gastroenterol. 2014;27(4):294-303.

146. Bennett RA, Rubin PH, Present DH. Frequency of inflammatory bowel disease in offspring of couples both presenting with inflammatory bowel disease. Gastroenterology. 1991;100(6):1638-43.

147. Laharie D, Debeugny S, Peeters M, Van Gossum A, Gower-Rousseau C, Bélaïche J, Fiasse R, Dupas JL, Lerebours E, Piotte S, Cortot A, Vermeire S, Grandbastien B, Colombel JF. Inflammatory bowel disease in spouses and their offspring. Gastroenterology. 2001;120(4):816-9.

148. Peeters M, Nevens H, Baert F, Hiele M, de Meyer AM, Vlietinck R, Rutgeerts P. Familial aggregation in Crohn's disease: increased age-adjusted risk and concordance in clinical characteristics. Gastroenterology. 1996;111(3):597-603.

149. Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. Science. 2008;322(5903):881-8.

150. Roth MP, Petersen GM, McElree C, Feldman E, Rotter JI. Geographic origins of Jewish patients with inflammatory bowel disease. Gastroenterology. 1989;97(4):900-4.

151. Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Pasternak S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, Hu W, Li C, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature. 2007;449(7164):851-61.

152. The International HapMap Project. Nature. 2003;426(6968):789-96.

153. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, et al. Initial sequencing and analysis of the human genome. Nature. 2001;409(6822):860-921.

154. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, et al. The sequence of the human genome. Science. 2001;291(5507):1304-51.

155. Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, Henning AK, SanGiovanni JP, Mane SM, Mayne ST, Bracken MB, Ferris FL, Ott J, Barnstable C, Hoh J. Complement factor H polymorphism in age-related macular degeneration. Science. 2005;308(5720):385-9.

156. Consortium WTCC. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007;447(7145):661-78.

157. Bodmer W, Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. Nat Genet. 2008;40(6):695-701.

158. Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. Benefits and limitations of genome-wide association studies. Nat Rev Genet. 2019;20(8):467-84.

159. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll SA, Visscher PM. Finding the missing heritability of complex diseases. Nature. 2009;461(7265):747-53.

160. Anderson CA, Boucher G, Lees CW, Franke A, D'Amato M, Taylor KD, Lee JC, Goyette P, Imielinski M, Latiano A, Lagacé C, Scott R, Amininejad L, Bumpstead S, Baidoo L, Baldassano RN, Barclay M, Bayless TM, Brand S, Büning C, et al. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. Nat Genet. 2011;43(3):246-52.

161. Liu JZ, van Sommeren S, Huang H, Ng SC, Alberts R, Takahashi A, Ripke S, Lee JC, Jostins L, Shah T, Abedian S, Cheon JH, Cho J, Dayani NE, Franke L, Fuyuno Y, Hart A, Juyal RC, Juyal G, Kim WH, et al. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. Nat Genet. 2015;47(9):979-86.

162. Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, Lee JC, Schumm LP, Sharma Y, Anderson CA, Essers J, Mitrovic M, Ning K, Cleynen I, Theatre E, Spain SL, Raychaudhuri S, Goyette P, Wei Z, Abraham C, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature. 2012;491(7422):119-24.

163. Barrett JC, Lee JC, Lees CW, Prescott NJ, Anderson CA, Phillips A, Wesley E, Parnell K, Zhang H, Drummond H, Nimmo ER, Massey D, Blaszczyk K, Elliott T, Cotterill L, Dallal H, Lobo AJ, Mowat C, Sanderson JD, Jewell DP, et al. Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the HNF4A region. Nat Genet. 2009;41(12):1330-4.

164. Franke A, McGovern DP, Barrett JC, Wang K, Radford-Smith GL, Ahmad T, Lees CW, Balschun T, Lee J, Roberts R, Anderson CA, Bis JC, Bumpstead S, Ellinghaus D, Festen EM, Georges M, Green T, Haritunians T, Jostins L, Latiano

A, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat Genet. 2010;42(12):1118-25.

165. Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, Rioux JD, Brant SR, Silverberg MS, Taylor KD, Barmada MM, Bitton A, Dassopoulos T, Datta LW, Green T, Griffiths AM, Kistner EO, Murtha MT, Regueiro MD, Rotter JI, Schumm LP, et al. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. Nat Genet. 2008;40(8):955-62.

166. Momozawa Y, Dmitrieva J, Theatre E, Deffontaine V, Rahmouni S, Charloteaux B, Crins F, Docampo E, Elansary M, Gori AS, Lecut C, Mariman R, Mni M, Oury C, Altukhov I, Alexeev D, Aulchenko Y, Amininejad L, Bouma G, Hoentjen F, et al. IBD risk loci are enriched in multigenic regulatory modules encompassing putative causative genes. Nat Commun. 2018;9(1):2427.

167. Huang H, Fang M, Jostins L, Umićević Mirkov M, Boucher G, Anderson CA, Andersen V, Cleynen I, Cortes A, Crins F, D'Amato M, Deffontaine V, Dmitrieva J, Docampo E, Elansary M, Farh KK, Franke A, Gori AS, Goyette P, Halfvarson et al. Fine-mapping inflammatory bowel disease loci to single-variant resolution. Nature. 2017;547(7662):173-8.

168. Mokry M, Middendorp S, Wiegerinck CL, Witte M, Teunissen H, Meddens CA, Cuppen E, Clevers H, Nieuwenhuis EE. Many inflammatory bowel disease risk loci include regions that regulate gene expression in immune cells and the intestinal epithelium. Gastroenterology. 2014;146(4):1040-7.

169. Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, Reynolds AP, Sandstrom R, Qu H, Brody J, Shafer A, Neri F, Lee K, Kutyavin T, Stehling-Sun S, Johnson AK, Canfield TK, Giste E, Diegel M, et al. Systematic localization of common disease-associated variation in regulatory DNA. Science. 2012;337(6099):1190-5.

170. McGovern DP, Gardet A, Törkvist L, Goyette P, Essers J, Taylor KD, Neale BM, Ong RT, Lagacé C, Li C, Green T, Stevens CR, Beauchamp C, Fleshner PR, Carlson M, D'Amato M, Halfvarson J, Hibberd ML, Lördal M, Padyukov L, et al. Genome-wide association identifies multiple ulcerative colitis susceptibility loci. Nat Genet. 2010;42(4):332-7.

171. Meddens CA, Harakalova M, van den Dungen NA, Foroughi Asl H, Hijma HJ, Cuppen EP, Bjorkegren JL, Asselbergs FW, Nieuwenhuis EE, Mokry M. Systematic analysis of chromatin interactions at disease associated loci links novel candidate genes to inflammatory bowel disease. Genome Biol. 2016;17(1):247.

172. Yochum GS, Cleland R, McWeeney S, Goodman RH. An antisense transcript induced by Wnt/beta-catenin signaling decreases E2F4. J Biol Chem. 2007;282(2):871-8.

173. Rennoll SA, Eshelman MA, Raup-Konsavage WM, Kawasawa YI, Yochum GS. The MYC 3' Wnt-Responsive Element Drives Oncogenic MYC Expression in Human Colorectal Cancer Cells. Cancers (Basel). 2016;8(5).

174. Konsavage WM, Jr., Kyler SL, Rennoll SA, Jin G, Yochum GS. Wnt/beta-catenin signaling regulates Yes-associated protein (YAP) gene expression in colorectal carcinoma cells. J Biol Chem. 2012;287(15):11730-9.

175. Schieffer KM, Choi CS, Emrich S, Harris L, Deiling S, Karamchandani DM, Salzberg A, Kawasawa YI, Yochum GS, Koltun WA. RNA-seq implicates deregulation of the immune system in the pathogenesis of diverticulitis. Am J Physiol Gastrointest Liver Physiol. 2017;313(3):G277-g84.

176. Sehgal R, Berg A, Polinski JI, Hegarty JP, Lin Z, McKenna KJ, Stewart DB, Poritz LS, Koltun WA. Mutations in IRGM are associated with more frequent need for surgery in patients with ileocolonic Crohn's disease. Dis Colon Rectum. 2012;55(2):115-21.

177. Khor B, Gardet A, Xavier RJ. Genetics and pathogenesis of inflammatory bowel disease. Nature. 2011;474(7351):307-17.

178. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489(7414):57-74.

179. Mokry M, Harakalova M, Asselbergs FW, de Bakker PI, Nieuwenhuis EE. Extensive Association of Common Disease Variants with Regulatory Sequence. PLoS One. 2016;11(11):e0165893.

180. Korinek V, Barker N, Morin PJ, van Wichen D, de Weger R, Kinzler KW, Vogelstein B, Clevers H. Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC-/- colon carcinoma. Science. 1997;275(5307):1784-7.

181. Rennoll S, Yochum G. Regulation of MYC gene expression by aberrant Wnt/beta-catenin signaling in colorectal cancer. World J Biol Chem. 2015;6(4):290-300.

182. Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA, Jaenisch R. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci U S A. 2010;107(50):21931-6.

183. Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. A unique chromatin signature uncovers early developmental enhancers in humans. Nature. 2011;470(7333):279-83.

184. Komor AC, Badran AH, Liu DR. CRISPR-Based Technologies for the Manipulation of Eukaryotic Genomes. Cell. 2017;169(3):559.

185. Sanyal A, Lajoie BR, Jain G, Dekker J. The long-range interaction landscape of gene promoters. Nature. 2012;489(7414):109-13.

186. Huppi K, Pitt JJ, Wahlberg BM, Caplen NJ. The 8q24 gene desert: an oasis of non-coding transcriptional activity. Front Genet. 2012;3:69.

187. Tuupanen S, Turunen M, Lehtonen R, Hallikas O, Vanharanta S, Kivioja T, Bjorklund M, Wei G, Yan J, Niittymaki I, Mecklin JP, Jarvinen H, Ristimaki A, Di-Bernardo M, East P, Carvajal-Carmona L, Houlston RS, Tomlinson I, Palin K, Ukkonen E, Karhu A, Taipale J, Aaltonen LA. The common colorectal cancer predisposition SNP rs6983267 at chromosome 8q24 confers potential to enhanced Wnt signaling. Nat Genet. 2009;41(8):885-90.

188. Wright JB, Brown SJ, Cole MD. Upregulation of c-MYC in cis through a large chromatin loop linked to a cancer risk-associated single-nucleotide polymorphism in colorectal cancer cells. Mol Cell Biol. 2010;30(6):1411-20.

189. Grisanzio C, Freedman ML. Chromosome 8q24-Associated Cancers and MYC. Genes Cancer. 2010;1(6):555-9.

190. Sipos F, Firneisz G, Muzes G. Therapeutic aspects of c-MYC signaling in inflammatory and cancerous colonic diseases. World J Gastroenterol. 2016;22(35):7938-50.

191. Konsavage WM, Jr., Jin G, Yochum GS. The Myc 3' Wnt-responsive element regulates homeostasis and regeneration in the mouse intestinal tract. Mol Cell Biol. 2012;32(19):3891-902.

192. Konsavage WM, Jr., Roper JN, Ishmael FT, Yochum GS. The Myc 3' Wnt responsive element regulates neutrophil recruitment after acute colonic injury in mice. Dig Dis Sci. 2013;58(10):2858-67.

193. Frey MR. Regenerating Reputations: Are Wnt and Myc the Good Guys After All? Dig Dis Sci. United States 2016. p. 327-9.

194. Ciclitira PJ, Macartney JC, Evan G. Expression of c-myc in non-malignant and pre-malignant gastrointestinal disorders. J Pathol. 1987;151(4):293-6.

195. Robles AI, Traverso G, Zhang M, Roberts NJ, Khan MA, Joseph C, Lauwers GY, Selaru FM, Popoli M, Pittman ME, Ke X, Hruban RH, Meltzer SJ, Kinzler KW, Vogelstein B, Harris CC, Papadopoulos N. Whole-Exome Sequencing Analyses of Inflammatory Bowel Disease-Associated Colorectal Cancers. Gastroenterology. 2016;150(4):931-43.

196. Yaeger R, Shah MA, Miller VA, Kelsen JR, Wang K, Heins ZJ, Ross JS, He Y, Sanford E, Yantiss RK, Balasubramanian S, Stephens PJ, Schultz N, Oren M, Tang L, Kelsen D. Genomic Alterations Observed in Colitis-Associated Cancers

Are Distinct From Those Found in Sporadic Colorectal Cancers and Vary by Type of Inflammatory Bowel Disease. Gastroenterology. 2016;151(2):278-87.e6.

197. Hartman DJ, Binion DG, Regueiro MD, Miller C, Herbst C, Pai RK. Distinct Histopathologic and Molecular Alterations in Inflammatory Bowel Disease-Associated Intestinal Adenocarcinoma: c-MYC Amplification is Common and Associated with Mucinous/Signet Ring Cell Differentiation. Inflamm Bowel Dis. 2018;24(8):1780-90.

198. Yamada A, Arakaki R, Saito M, Tsunematsu T, Kudo Y, Ishimaru N. Role of regulatory T cell in the pathogenesis of inflammatory bowel disease. World J Gastroenterol. 2016;22(7):2195-205.

199. van Loosdregt J, Coffer PJ. The Role of WNT Signaling in Mature T Cells: T Cell Factor Is Coming Home. J Immunol. 2018;201(8):2193-200.

200. Delmore JE, Issa GC, Lemieux ME, Rahl PB, Shi J, Jacobs HM, Kastritis E, Gilpatrick T, Paranal RM, Qi J, Chesi M, Schinzel AC, McKeown MR, Heffernan TP, Vakoc CR, Bergsagel PL, Ghobrial IM, Richardson PG, Young RA, Hahn WC, Anderson KC, Kung AL, Bradner JE, Mitsiades CS. BET bromodomain inhibition as a therapeutic strategy to target c-Myc. Cell. 2011;146(6):904-17.

201. Filippakopoulos P, Qi J, Picaud S, Shen Y, Smith WB, Fedorov O, Morse EM, Keates T, Hickman TT, Felletar I, Philpott M, Munro S, McKeown MR, Wang Y, Christie AL, West N, Cameron MJ, Schwartz B, Heightman TD, La Thangue N, French CA, Wiest O, Kung AL, Knapp S, Bradner JE. Selective inhibition of BET bromodomains. Nature. 2010;468(7327):1067-73.

202. Matthews SM, Eshelman MA, Berg AS, Koltun WA, Yochum GS. The Crohn's disease associated SNP rs6651252 impacts MYC gene expression in human colonic epithelial cells. PLoS One. 2019;14(2):e0212850.

203. Mateos B, Palanca-Ballester C, Saez-Gonzalez E, Moret I, Lopez A, Sandoval J. Epigenetics of Inflammatory Bowel Disease: Unraveling Pathogenic Events. Crohn's & Colitis 360. 2019;1(2).

204. Wilson C, Kanhere A. 8q24.21 Locus: A Paradigm to Link Non-Coding RNAs, Genome Polymorphisms and Cancer. Int J Mol Sci. 2021;22(3).

205. Ling H, Spizzo R, Atlasi Y, Nicoloso M, Shimizu M, Redis RS, Nishida N, Gafà R, Song J, Guo Z, Ivan C, Barbarotto E, De Vries I, Zhang X, Ferracin M, Churchman M, van Galen JF, Beverloo BH, Shariati M, Haderk F, et al. CCAT2, a novel noncoding RNA mapping to 8q24, underlies metastatic progression and chromosomal instability in colon cancer. Genome Res. 2013;23(9):1446-61.

206. Saini Y, Chen J, Patial S. The Tristetraprolin Family of RNA-Binding Proteins in Cancer: Progress and Future Prospects. Cancers (Basel). 2020;12(6).

207. Montorsi L, Guizzetti F, Alecci C, Caporali A, Martello A, Atene CG, Parenti S, Pizzini S, Zanovello P, Bortoluzzi S, Ferrari S, Grande A, Zanocco-Marani T. Loss of ZFP36 expression in colorectal cancer correlates to wnt/ ß-catenin activity and enhances epithelial-to-mesenchymal transition through upregulation of ZEB1, SOX9 and MACC1. Oncotarget. 2016;7(37):59144-57.

208. Zanfi ED, Fantini S, Lotti R, Bertesi M, Marconi A, Grande A, Manfredini R, Pincelli C, Zanocco-Marani T. Wnt/CTNNB1 Signal Transduction Pathway Inhibits the Expression of ZFP36 in Squamous Cell Carcinoma, by Inducing Transcriptional Repressors SNAI1, SLUG and TWIST. Int J Mol Sci. 2020;21(16).

209. Eshelman MA, Matthews SM, Schleicher EM, Fleeman RM, Kawasawa YI, Stumpo DJ, Blackshear PJ, Koltun WA, Ishmael FT, Yochum GS. Tristetraprolin targets Nos2 expression in the colonic epithelium. Sci Rep. 2019;9(1):14413.

210. Taylor GA, Carballo E, Lee DM, Lai WS, Thompson MJ, Patel DD, Schenkman DI, Gilkeson GS, Broxmeyer HE, Haynes BF, Blackshear PJ. A pathogenetic role for TNF alpha in the syndrome of cachexia, arthritis, and autoimmunity resulting from tristetraprolin (TTP) deficiency. Immunity. 1996;4(5):445-54.

211. Nakao A, Imamura T, Souchelnytskyi S, Kawabata M, Ishisaki A, Oeda E, Tamaki K, Hanai J, Heldin CH, Miyazono K, ten Dijke P. TGF-beta receptor-mediated signalling through Smad2, Smad3 and Smad4. Embo j. 1997;16(17):5353-62.

212. Zhang Y, Feng X, We R, Derynck R. Receptor-associated Mad homologues synergize as effectors of the TGF-beta response. Nature. 1996;383(6596):168-72..

213. Rivas MA, Beaudoin M, Gardet A, Stevens C, Sharma Y, Zhang CK, Boucher G, Ripke S, Ellinghaus D, Burtt N, Fennell T, Kirby A, Latiano A, Goyette P, Green T, Halfvarson J, Haritunians T, Korn JM, Kuruvilla F, Lagacé C, et al. Deep resequencing of GWAS loci identifies independent rare variants associated with inflammatory bowel disease. Nat Genet. 2011;43(11):1066-73.

214. Miyoshi H, Ajima R, Luo CT, Yamaguchi TP, Stappenbeck TS. Wnt5a potentiates TGF-β signaling to promote colonic crypt regeneration after tissue injury. Science. 2012;338(6103):108-13.

215. Liu L, Rao JN, Zou T, Xiao L, Smith A, Zhuang R, Turner DJ, Wang JY. Activation of Wnt3a signaling stimulates intestinal epithelial repair by promoting c-Myc-regulated gene expression. Am J Physiol Cell Physiol. 2012;302(1):C277-85.

216. Terzić J, Grivennikov S, Karin E, Karin M. Inflammation and colon cancer. Gastroenterology. 2010;138(6):2101-14.e5.

217. Neufert C, Becker C, Neurath MF. An inducible mouse model of colon carcinogenesis for the analysis of sporadic and inflammation-driven tumor progression. Nat Protoc. 2007;2(8):1998-2004.

218. Okayasu I, Ohkusa T, Kajiura K, Kanno J, Sakamoto S. Promotion of colorectal neoplasia in experimental murine ulcerative colitis. Gut. 1996;39(1):87-92.

219. Konsavage WM, Jr., Yochum GS. The myc 3' wnt-responsive element suppresses colonic tumorigenesis. Mol Cell Biol. 2014;34(9):1659-69.

220. Grivennikov SI. Inflammation and colorectal cancer: colitis-associated neoplasia. Semin Immunopathol. 2013;35(2):229-44. Epub 2012/11/20. doi: 10.1007/s00281-012-0352-6. PubMed PMID: 23161445; PMCID: PMC3568220.

221. Tarmin L, Yin J, Harpaz N, Kozam M, Noordzij J, Antonio LB, Jiang HY, Chan O, Cymes K, Meltzer SJ. Adenomatous polyposis coli gene mutations in ulcerative colitis-associated dysplasias and cancers versus sporadic colon neoplasms. Cancer Res. 1995;55(10):2035-8.

222. Cooper HS, Murthy S, Kido K, Yoshitake H, Flanigan A. Dysplasia and cancer in the dextran sulfate sodium mouse colitis model. Relevance to colitis-associated neoplasia in the human: a study of histopathology, B-catenin and p53 expression and the role of inflammation. Carcinogenesis. 2000;21(4):757-68.

223. Claessen MM, Schipper ME, Oldenburg B, Siersema PD, Offerhaus GJ, Vleggaar FP. WNT-pathway activation in IBD-associated colorectal carcinogenesis: potential biomarkers for colonic surveillance. Cell Oncol. 2010;32(4):303-10.

224. van Dekken H, Wink JC, Vissers KJ, Franken PF, Ruud Schouten W, WC JH, Kuipers EJ, Fodde R, Janneke van der Woude C. Wnt pathway-related gene expression during malignant progression in ulcerative colitis. Acta Histochem. 2007;109(4):266-72.

225. Dhir M, Montgomery EA, Glockner SC, Schuebel KE, Hooker CM, Herman JG, Baylin SB, Gearhart SL, Ahuja N. Epigenetic regulation of WNT signaling pathway genes in inflammatory bowel disease (IBD) associated neoplasia. J Gastrointest Surg. 2008;12(10):1745-53.

226. Kaler P, Godasi BN, Augenlicht L, Klampfer L. The NF-κB/AKT-dependent Induction of Wnt Signaling in Colon Cancer Cells by Macrophages and IL-1β. Cancer Microenviron. 2009;2(1):69-80.

227. Lee G, Goretsky T, Managlia E, Dirisina R, Singh AP, Brown JB, May R, Yang GY, Ragheb JW, Evers BM, Weber CR, Turner JR, He XC, Katzman RB, Li L, Barrett TA. Phosphoinositide 3-kinase signaling mediates beta-catenin activation in intestinal epithelial stem and progenitor cells in colitis. Gastroenterology. 2010;139(3):869-81, 81.e1-9.

228. Askew DS, Ashmun RA, Simmons BC, Cleveland JL. Constitutive c-myc expression in an IL-3-dependent myeloid cell line suppresses cell cycle arrest and accelerates apoptosis. Oncogene. 1991;6(10):1915-22.

229. Askew DS, Ihle JN, Cleveland JL. Activation of apoptosis associated with enforced myc expression in myeloid progenitor cells is dominant to the suppression of apoptosis by interleukin-3 or erythropoietin. Blood. 1993;82(7):2079-87.

230. Evan GI, Wyllie AH, Gilbert CS, Littlewood TD, Land H, Brooks M, Waters CM, Penn LZ, Hancock DC. Induction of apoptosis in fibroblasts by c-myc protein. Cell. 1992;69(1):119-28.

231. Neiman PE, Thomas SJ, Loring G. Induction of apoptosis during normal and neoplastic B-cell development in the bursa of Fabricius. Proc Natl Acad Sci U S A. 1991;88(13):5857-61.

232. Albuquerque C, Breukel C, van der Luijt R, Fidalgo P, Lage P, Slors FJ, Leitão CN, Fodde R, Smits R. The 'just-right' signaling model: APC somatic mutations are selected based on a specific level of activation of the beta-catenin signaling cascade. Hum Mol Genet. 2002;11(13):1549-60.

233. Lamlum H, Ilyas M, Rowan A, Clark S, Johnson V, Bell J, Frayling I, Efstathiou J, Pack K, Payne S, Roylance R, Gorman P, Sheer D, Neale K, Phillips R, Talbot I, Bodmer W, Tomlinson I. The type of somatic mutation at APC in familial adenomatous polyposis is determined by the site of the germline mutation: a new facet to Knudson's 'two-hit' hypothesis. Nat Med. 1999;5(9):1071-5.

234. McMahon SB. MYC and the control of apoptosis. Cold Spring Harb Perspect Med. 2014;4(7):a014407.

235. Hoffman B, Liebermann DA. Apoptotic signaling by c-MYC. Oncogene. 2008;27(50):6462-72.

236. Andrieu G, Belkina AC, Denis GV. Clinical trials for BET inhibitors run ahead of the science. Drug Discov Today Technol. 2016;19:45-50.

237. Denis GV, Green MR. A novel, mitogen-activated nuclear kinase is related to a Drosophila developmental regulator. Genes Dev. 1996;10(3):261-71.

238. Yang Z, Yik JH, Chen R, He N, Jang MK, Ozato K, Zhou Q. Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. Mol Cell. 2005;19(4):535-45.

239. Gyuris A, Donovan DJ, Seymour KA, Lovasco LA, Smilowitz NR, Halperin AL, Klysik JE, Freiman RN. The chromatin-targeting protein Brd2 is required for neural tube closure and embryogenesis. Biochim Biophys Acta. 2009;1789(5):413-21.

240. Shang E, Wang X, Wen D, Greenberg DA, Wolgemuth DJ. Double bromodomain-containing gene Brd2 is essential for embryonic development in mouse. Dev Dyn. 2009;238(4):908-17.

241. Wang F, Liu H, Blanton WP, Belkina A, Lebrasseur NK, Denis GV. Brd2 disruption in mice causes severe obesity without Type 2 diabetes. Biochem J. 2009;425(1):71-83.

242. Houzelstein D, Bullock SL, Lynch DE, Grigorieva EF, Wilson VA, Beddington RS. Growth and early postimplantation defects in mice deficient for the bromodomain-containing protein Brd4. Mol Cell Biol. 2002;22(11):3794-802.

243. Zuber J, Shi J, Wang E, Rappaport AR, Herrmann H, Sison EA, Magoon D, Qi J, Blatt K, Wunderlich M, Taylor MJ, Johns C, Chicas A, Mulloy JC, Kogan SC, Brown P, Valent P, Bradner JE, Lowe SW, Vakoc CR. RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. Nature. 2011;478(7370):524-8.

244. Bolden JE, Tasdemir N, Dow LE, van Es JH, Wilkinson JE, Zhao Z, Clevers H, Lowe SW. Inducible in vivo silencing of Brd4 identifies potential toxicities of sustained BET protein inhibition. Cell Rep. 2014;8(6):1919-29.

245. Matzuk MM, McKeown MR, Filippakopoulos P, Li Q, Ma L, Agno JE, Lemieux ME, Picaud S, Yu RN, Qi J, Knapp S, Bradner JE. Small-molecule inhibition of BRDT for male contraception. Cell. 2012;150(4):673-84.

246. Huang B, Yang XD, Zhou MM, Ozato K, Chen LF. Brd4 coactivates transcriptional activation of NF-kappaB via specific binding to acetylated RelA. Mol Cell Biol. 2009;29(5):1375-87.

247. Mele DA, Salmeron A, Ghosh S, Huang HR, Bryant BM, Lora JM. BET bromodomain inhibition suppresses TH17-mediated pathology. J Exp Med. 2013;210(11):2181-90.

248. Jahagirdar R, Attwell S, Marusic S, Bendele A, Shenoy N, McLure KG, Gilham D, Norek K, Hansen HC, Yu R, Tobin J, Wagner GS, Young PR, Wong NCW, Kulikowski E. RVX-297, a BET Bromodomain Inhibitor, Has Therapeutic Effects in Preclinical Models of Acute Inflammation and Autoimmune Disease. Mol Pharmacol. 2017;92(6):694-706.

249. Zhang QG, Qian J, Zhu YC. Targeting bromodomain-containing protein 4 (BRD4) benefits rheumatoid arthritis. Immunol Lett. 2015;166(2):103-8.

250. Zengerle M, Chan KH, Ciulli A. Selective Small Molecule Induced Degradation of the BET Bromodomain Protein BRD4. ACS Chem Biol. 2015;10(8):1770-7..

251. Winter GE, Buckley DL, Paulk J, Roberts JM, Souza A, Dhe-Paganon S, Bradner JE. DRUG DEVELOPMENT. Phthalimide conjugation as a strategy for in vivo target protein degradation. Science. 2015;348(6241):1376-81.

252. Tanaka M, Roberts JM, Seo HS, Souza A, Paulk J, Scott TG, DeAngelo SL, Dhe-Paganon S, Bradner JE. Design and characterization of bivalent BET inhibitors. Nat Chem Biol. 2016;12(12):1089-96.

253. Zhou B, Hu J, Xu F, Chen Z, Bai L, Fernandez-Salas E, Lin M, Liu L, Yang CY, Zhao Y, McEachern D, Przybranowski S, Wen B, Sun D, Wang S. Discovery of a Small-Molecule Degrader of Bromodomain and Extra-Terminal (BET) Proteins with Picomolar Cellular Potencies and Capable of Achieving Tumor Regression. J Med Chem. 2018;61(2):462-81.

254. Liu Z, Chen H, Wang P, Li Y, Wold EA, Leonard PG, Joseph S, Brasier AR, Tian B, Zhou J. Discovery of Orally Bioavailable Chromone Derivatives as Potent and Selective BRD4 Inhibitors: Scaffold Hopping, Optimization, and Pharmacological Evaluation. J Med Chem. 2020;63(10):5242-56.

Vita

EDUCATION

2010-2014 B.S. in Biology, York College of Pennsylvania, York PA

2015-2021 Ph.D. in Biomedical Sciences, laboratory of Gregory S. Yochum, The Pennsylvania State University College of Medicine, Hershey, PA

HONORS AND AWARDS

2018 Graduate Student Travel Grant

2019 Summer Symposium in Molecular Biology at Penn State University travel assistance

PUBLICATIONS

1. **Matthews SM**, Eshelman MA, Berg AS, Koltun WA, Yochum GS. (2019). The Crohn's disease associated SNP rs6651252 impacts MYC gene expression in human colonic epithelial cells. PLoS One 14:e0212850. https://doi.org/10.1371/journal.pone.0212850

2. Eshelman MA, **Matthews SM,** Schleicher EM, Fleeman RM, Kawasawa YI, Stumpo DJ, Blackshear PJ, Koltun WA, Ishmael FT, Yochum GS. (2019). Tristetraprolin targets Nos2 expression in the colonic epithelium. Sci. Rep. 9:14413 https://doi.org/10.1038/s41598-019-50957-9

SELECTED MEETING PRESENTATIONS

Stephen Matthews, Melanie A. Eshelman, Katie Schieffer, Leonard Harris, Sue Deiling, Walter A. Koltun and Gregory S. Yochum, The Crohn's disease associated SNP, rs6651252, impacts MYC gene expression in human colonic epithelial cells. 2019 Summer Symposium in Molecular Biology at Pennsylvania State University, State College, PA, USA. (poster presentation)

Stephen Matthews, Melanie A. Eshelman, Katie Schieffer, Leonard Harris, Sue Deiling, Walter A. Koltun and Gregory S. Yochum, The Crohn's disease associated SNP, rs6651252, impacts MYC gene expression in human colonic epithelial cells. 2018. Work presented at Digestive Disease Week (DDW) conference, Washington, D.C., USA. (poster presentation)

Stephen Matthews, Melanie A. Eshelman, Katie Schieffer, Walter A. Koltun and Gregory S. Yochum. Crohn's Disease-associated SNP rs6651252 demarcates a Wnt-Responsive Element in Intestinal Epithelial Cells. 2017. Inflammatory Bowel Disease Research Summit; Hershey PA, USA. (oral presentation)