Modulation Of Transcription In The Murine Gut Epithelium In Development, Differentiation, And Disease

Julia Elise Kieckhaefer

University of Pennsylvania, kju@mail.med.upenn.edu

Follow this and additional works at: https://repository.upenn.edu/edissertations

Part of the Cell Biology Commons, Genetics Commons, and the Molecular Biology Commons

Recommended Citation

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/edissertations/2388
For more information, please contact repository@pobox.upenn.edu.
Modulation Of Transcription In The Murine Gut Epithelium In Development, Differentiation, And Disease

Abstract
The intestinal epithelium is a continuously self-renewing tissue that must balance its function as the site of nutrient absorption with its role as a barrier between the body and the external environment. With its high cellular turnover and continual exposure to lumenal antigens, the intestinal epithelium is highly susceptible to developing diseases, such as inflammatory bowel disease (IBD) and cancer. Understanding the factors that control intestinal epithelial development are key to discovering disease etiology and to engineering new treatments. In mammals, studies have pointed to dysregulation of RNA Polymerase III as playing a role in oncogenic transformation. A deletion in the gene encoding Polr3b, a Pol III subunit, impairs development of the intestinal epithelial progenitor cells in zebrafish. Thus, I hypothesized that POLR3B also has an important role in the mammalian intestinal epithelium. I characterized mice expressing an intestinal epithelium-specific hypomorphic Polr3b mutation. Neonatal Polr3b mutants have severely reduced survival and growth, reduced intestinal epithelial proliferation, and lack normal crypt development. Thus, Pol III is essential for the establishment of the intestinal epithelium in mice. Given that normal Pol III is required for adequate protein production in order to sustain the high rate of proliferation in mammalian gut, this enzyme is a questionable target for cancer drugs. It was reported that Mucin2 (Muc2), the main component of intestinal mucus, is a direct transcriptional target of FOXA1 and FOXA2 in the mouse intestinal epithelium. I hypothesized that in addition to Muc2, the FOXA factors transcriptionally regulate other genes that are required for the maintenance of the intestinal mucosa. The Foxa1/a2-deficient colonic epithelium in mice of advanced age (>7 months) displays increased susceptibility to spontaneous colitis. I performed FOXA1 and FOXA2 chromatin immunoprecipitation assays followed by ultra-high throughput sequencing (ChIP-Seq) analysis on wild type mouse colonic crypts, as well as RNA-seq on Foxa1/a2-deficient and control aged colon. Of the FOXA1/A2-bound, differentially expressed genes, glycosyltransferases were overrepresented, including the IBD risk gene Fut2, which was downregulated in mutant colon.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Cell & Molecular Biology

First Advisor
Klaus H. Kaestner

Second Advisor
Jonathan P. Katz

Keywords
ChIP-seq, FOXA, gene regulation, inflammatory bowel disease, intestinal epithelium, RNA Polymerase III

Subject Categories
Cell Biology | Genetics | Molecular Biology

This dissertation is available at ScholarlyCommons: https://repository.upenn.edu/edissertations/2388
MODULATION OF TRANSCRIPTION IN THE MURINE GUT EPITHELIUM IN DEVELOPMENT, DIFFERENTIATION, AND DISEASE

Julia Kieckhaefer

A DISSERTATION

in

Cell and Molecular Biology

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2017

Supervisor of Dissertation

_____________________
Klaus H. Kaestner, Ph.D.
Thomas and Evelyn Suor Butterworth Professor in Genetics

Graduate Group Chairperson

_____________________
Daniel S. Kessler, Ph.D., Associate Professor of Cell and Developmental Biology

Dissertation Committee

Jonathan P. Katz, M.D., Associate Professor of Medicine
Christopher D. Brown, Ph.D., Assistant Professor of Genetics
Christopher J. Lengner, Ph.D., Assistant Professor of Cell and Developmental Biology
John P. Lynch, M.D., Ph.D., Associate Professor of Medicine
ACKNOWLEDGMENT

I would like to thank everyone who provided me with the emotional support, guidance, and opportunities that allowed me to get to this point in my education. This includes my parents, who always encouraged hard work and an interest in learning, and my sisters, who were always a phone call away when I needed them. My grandmother, who has been practicing law for almost 70 years, Alice Traeger Koempel, has been a role model for what hard work can achieve. I would also especially like to thank Dr. Kan Cao, who first gave me the chance to show that I could do real research, and whose mentorship opened the door to my future research opportunities. I would not have been able to even enter graduate school without her.

I didn't know what I was doing when I first joined the Kaestner lab, but it might be one of the best naïve decisions I've ever made. As a member of the Kaestner lab, in addition to learning cutting edge science, I also formed relationships with a group of wonderful people. The mentorship of Dr. Klaus Kaestner has been invaluable in shaping how I think about and communicate science. Dr. Karyn Hummel taught me not just about science, but also that it is possible to persevere through setbacks, and that sometimes it can be fun to dance like a crazy person. Dr. Ellen Elliott was one of the best roommates I have ever had. Dr. Adam Zahm, you are a genius, a mentor, and a friend. To the aforementioned people and to everyone else (Mark, Ben, Vasu, Jia, Amanda, Kirk, Aryel, Kristy, Julia (Yue—a.k.a. Julia 2), Amber), thank you for all the fun and the memories that have made the past six years of my life some of the best!
ABSTRACT

MODULATION OF TRANSCRIPTION IN THE MURINE GUT EPITHELIUM IN DEVELOPMENT, DIFFERENTIATION, AND DISEASE

Julia Kieckhaefer
Klaus H. Kaestner

The intestinal epithelium is a continuously self-renewing tissue that must balance its function as the site of nutrient absorption with its role as a barrier between the body and the external environment. With its high cellular turnover and continual exposure to lumenal antigens, the intestinal epithelium is highly susceptible to developing diseases, such as inflammatory bowel disease (IBD) and cancer. Understanding the factors that control intestinal epithelial development are key to discovering disease etiology and to engineering new treatments. In mammals, studies have pointed to dysregulation of RNA Polymerase III as playing a role in oncogenic transformation. A deletion in the gene encoding Polr3b, a Pol III subunit, impairs development of the intestinal epithelial progenitor cells in zebrafish. Thus, I hypothesized that POLR3B also has an important role in the mammalian intestinal epithelium. I characterized mice expressing an intestinal epithelium-specific hypomorphic Polr3b mutation. Neonatal Polr3b mutants have severely reduced survival and growth, reduced intestinal epithelial proliferation, and lack normal crypt development. Thus, Pol III is essential for the establishment of the intestinal epithelium in mice. Given that normal Pol III is required for adequate protein production in order to sustain the high rate of proliferation in mammalian gut, this enzyme is a questionable target for cancer drugs. It was reported that Mucin2 (Muc2), the main component of intestinal mucus, is a direct transcriptional target of FOXA1 and FOXA2 in the mouse intestinal epithelium. I hypothesized that in addition to Muc2, the FOXA factors transcriptionally regulate other genes that are required for the maintenance of the intestinal mucosa. The Foxa1/a2-deficient colonic epithelium in mice of advanced age (>7 months) displays increased susceptibility to spontaneous colitis. I performed
FOXA1 and FOXA2 chromatin immunoprecipitation assays followed by ultra-high throughput sequencing (ChIP-Seq) analysis on wild type mouse colonic crypts, as well as RNA-seq on Foxa1/a2-deficient and control aged colon. Of the FOXA1/A2-bound, differentially expressed genes, glycosyltransferases were overrepresented, including the IBD risk gene Fut2, which was downregulated in mutant colon.
# TABLE OF CONTENTS

ABSTRACT .................................................................................................................................................. iii

LIST OF TABLES ........................................................................................................................................ vi

LIST OF ILLUSTRATIONS ....................................................................................................................... vii

INTRODUCTION ......................................................................................................................................... 1
  Inflammatory bowel disease ..................................................................................................................... 2
  Gut mucus and glycosylation .................................................................................................................. 7
  Signaling pathways: maintenance of stemness and differentiation ...................................................... 13
  The FOXA transcription factor family .................................................................................................. 20

MATERIALS AND METHODS ................................................................................................................ 27
  POLR3B MATERIALS AND METHODS: ............................................................................................. 28
  FOXA1/A2 MATERIALS AND METHODS ............................................................................................. 31

The RNA Polymerase III Subunit Polr3b is required for the maintenance of small intestinal crypts in mice............................................................................................................................................. 36
  ABSTRACT ............................................................................................................................................. 37
  INTRODUCTION ...................................................................................................................................... 38
  RESULTS ............................................................................................................................................... 40
  DISCUSSION ......................................................................................................................................... 45
  FIGURES ................................................................................................................................................ 48

The FOXA Transcription Factors Sustain Murine Intestinal Health by Regulating Crohn's Disease Risk Loci....................................................................................................................................................... 60
  ABSTRACT ............................................................................................................................................. 61
  INTRODUCTION ...................................................................................................................................... 62
  RESULTS ............................................................................................................................................... 64
  DISCUSSION ......................................................................................................................................... 69
  FIGURES ................................................................................................................................................ 72

Conclusions and Future Directions ........................................................................................................ 86

BIBLIOGRAPHY ....................................................................................................................................... 94
LIST OF TABLES

Page number

Chapter 4: The FOXA Transcription Factors Sustain Murine Intestinal Health by Regulating Crohn’s Disease Risk Loci

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 4.1</td>
<td>78</td>
</tr>
<tr>
<td>Table 4.2</td>
<td>80</td>
</tr>
<tr>
<td>Table 4.3</td>
<td>85</td>
</tr>
</tbody>
</table>
# LIST OF ILLUSTRATIONS

<table>
<thead>
<tr>
<th>Chapter 1: Introduction</th>
<th>Page number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>26</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 3: The RNA Polymerase III Subunit Polr3b is required for the maintenance of small intestinal crypts in mice.</th>
<th>Page number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 3.1</td>
<td>48</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>50</td>
</tr>
<tr>
<td>Figure 3.3</td>
<td>51</td>
</tr>
<tr>
<td>Figure 3.4</td>
<td>53</td>
</tr>
<tr>
<td>Figure 3.5</td>
<td>55</td>
</tr>
<tr>
<td>Figure 3.6</td>
<td>57</td>
</tr>
<tr>
<td>Figure 3.7</td>
<td>59</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 4: The FOXA Transcription Factors Sustain Murine Intestinal Health by Regulating Crohn's Disease Risk Loci</th>
<th>Page number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 4.1</td>
<td>72</td>
</tr>
<tr>
<td>Figure 4.2</td>
<td>74</td>
</tr>
<tr>
<td>Figure 4.3</td>
<td>75</td>
</tr>
<tr>
<td>Figure 4.4</td>
<td>76</td>
</tr>
<tr>
<td>Figure 4.5</td>
<td>82</td>
</tr>
<tr>
<td>Figure 4.6</td>
<td>83</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION
Inflammatory bowel disease

Due to the similarity in clinical manifestation of ulcerative colitis (UC) and Crohn’s disease, the two conditions are commonly referred to under the umbrella term ‘inflammatory bowel disease’ (IBD) (Danese and Fiocchi, 2011). Both are chronic conditions with periods of remission and relapse that present clinically with abdominal pain, chronic, bloody diarrhea, fever, and weight loss. Histopathologically, common features of UC and Crohn’s disease in the colon are altered crypt morphology, goblet cell loss, and ulcerations. Complications of IBD include intestinal bleeding and anemia, peripheral arthritis, ocular and skin conditions, and especially after prolonged disease, dysplasia and colorectal cancer (Danese and Fiocchi, 2011; Torres et al., 2016).

Despite their similarities, UC and Crohn’s disease are distinct diseases that affect different regions of the gut. Inflammation in ulcerative colitis is usually restricted to the colon and stays within the mucosal layer, which may be explained by autoimmune targeting of colonocytes. Circulating IgG1 antibodies against an epithelial antigen that is shared by the colon, eyes, skin, joints, and biliary epithelium may result in autoimmune damage to the colon while also explaining the extraintestinal features of the disease (Danese and Fiocchi, 2011). However, all segments of the gastrointestinal tract can be affected in Crohn’s disease, though the most commonly inflamed areas are the terminal ileum and colon. In Crohn’s disease, inflammation occurs transmurally and in segments, forming a pattern that is often called “cobblestoning” (Baumgart and Sandborn, 2012; Torres et al., 2016).

Despite these differences, there is strong overlap in the approach to treating and managing the two diseases. Common medical treatments for IBD are anti-inflammatory drugs such as sulfasalazine and 5-aminosalicylate, glucocorticoids, and infliximab, a monoclonal antibody directed against TNF-α. However, though antibiotics are a successful treatment for luminal Crohn’s, they have not been shown to be an effective treatment for UC (Torres et al., 2016). Failure of medical treatments may indicate colectomy, which can be curative for UC, but not for Crohn’s disease (Baumgart and Sandborn, 2012; Danese and Fiocchi, 2011; Torres et al., 2016). Nevertheless, surgery is often required to treat intestinal complications from Crohn’s
disease, such as strictures, abscesses, and fistulas, which have a high rate of postoperative recurrence (Baumgart and Sandborn, 2012). Thus, IBD is a lifelong condition that requires constant management.

**Disease epidemiology and environmental factors**

Since the mid-20th century, the incidence of IBD has been increasing in developed countries, with the highest prevalence seen in Northern Europe and North America (Ananthakrishnan, 2015; Danese and Fiocchi, 2011). As of 2015, an estimated 1.3% of adults (3.1 million people) had been diagnosed with IBD in the United States. Americans with the highest rate of IBD are aged 65 or older (1.7%). IBD is also more prevalent in urban populations, among people who did not complete high school, and in those that fall below the poverty line (Dahlhamer et al., 2016). The differences in disease incidence between socioeconomic groups point to lifestyle characteristics as having a major impact on disease risk.

Numerous environmental factors associated with a westernized lifestyle are thought to increase the risk of developing IBD, including a diet high in fat and sugar and low in dietary fiber, smoking, medication use, and stress. Environmental factors during childhood may also have an impact on disease development, such as breastfeeding and exposure to animals, which may be protective, and antibiotic use, which may increase disease risk (Ananthakrishnan, 2015; Danese and Fiocchi, 2011; Torres et al., 2016). In support of the role the environment and lifestyle play in disease risk, the incidence of IBD has been rising in Asia in parallel with its economic development, which coincides with a shift towards a westernized lifestyle (Ananthakrishnan, 2015; Torres et al., 2016). Further evidence for this notion comes from the fact that first-generation and second-generation Indian migrants to the UK are diagnosed with IBD at rates that are higher than the IBD incidence in India and more similar to the rate of diagnosis among native residents of the UK. Concordantly, Africa and South America still have low rates of IBD, though there is some evidence that the incidence may be increasing (Ananthakrishnan, 2015).

**Genetic determinants of IBD risk**
Although environmental factors play a prominent role in disease risk, development of disease cannot be explained by these factors alone. There is ample evidence that genetic components make a significant contribution to disease etiology. One of best pieces of evidence for IBD heritability is the significantly higher concordance rate among monozygotic than dizygotic twins for IBD (Orholm et al., 2000). Another well-known substantiator of the genetic basis of IBD is the high rate of IBD in Ashkenazi Jews, which is not observed within other ethnic groups (Roth et al., 1989). Finally, there are numerous family studies that have shown that first-degree relatives of an affected proband are more likely to develop the disease themselves (Russell and Satsangi, 2004). With mounting evidence for a genetic contribution to disease, studies continue to be undertaken to identify new disease risk loci.

IBD is a complex disease with evidence pointing to numerous causal genes. One of the earliest Crohn’s susceptibility loci to be identified was associated with the pattern recognition receptor \textit{NOD2/CARD15}, an activator of the transcriptional regulator NFκB, which reinforced the hypothesis that Crohn’s disease is caused by an altered response to intestinal pathogens (Hugot et al., 2001). Significantly, the cytokine gene \textit{IL10}, which when ablated in mice serves as a model for IBD (Berg et al., 1996), is a disease risk locus for both Crohn’s and UC (Franke et al., 2008, 2010). A genome-wide association study (GWAS) of Crohn’s disease identified a strong association of the \textit{FUT2} null allele \textit{W143X} with the development of Crohn’s disease. The association of \textit{W143X}, which is protective against norovirus infection but increases the risk of Crohn’s disease, strengthens the hypothesis that altered host-microbe interactions in the gut play a role in disease etiology (Franke et al., 2010; McGovern et al., 2010).

To date, GWAS has identified over 200 IBD risk loci (Franke et al., 2010; Jostins et al., 2012; Liu et al., 2015). Of the 163 loci identified by a meta-analysis of 15 Crohn’s disease and ulcerative colitis GWA studies (Jostins et al., 2012), 110 are associated with both diseases, indicating that many mechanisms of pathogenesis in one disease have some kind of role in the other. However, risk alleles at two Crohn’s disease loci, \textit{PTPN22} and \textit{NOD2}, are protective in UC, strong evidence for differences in etiology between the two diseases. However, the autoimmune diseases ankylosing spondylitis and psoriasis have a high enrichment of overlap with IBD loci.
Genes that are associated with the overlapping loci are largely involved in T cell regulation, supporting the idea that IBD is an autoimmune condition. Gene Ontology enrichment analysis of the genes associated with the 163 loci revealed that the most significantly enriched terms were “regulation of cytokine production” and “lymphocyte activation”, though “response to molecules of bacterial origin” was also significantly enriched (Jostins et al., 2012).

**Mouse models of IBD**

Much of what is known about the molecular pathogenesis of IBD comes from the study of mouse models, which has led to a better understanding of the complex role that modulation of the immune response has in disease etiology. Two of the oldest animal models of colitis chemically induce mucosal irritation by the administration of either dextran sulfate sodium (DSS) (Okayasu et al., 1990) or 2,4,6-trinitrobenzenesulfonic acid (TNBS) (Morris et al., 1989). In both the DSS and TNBS models, there is evidence that hyperresponsiveness to antigens in the mucosal microflora plays a critical role in the development of colitis. In DSS-treated animals with a deficiency for trefoil factor 3 (Tff3) (formerly intestinal trefoil factor, Itf), a protein important for the maintenance and healing of the epithelium, colitis is more severe than in normal mice (Mashimo et al., 1996). Furthermore, changes in barrier function in DSS-treated mice have been observed before the onset of DSS-induced inflammation, suggesting that a breakdown in the epithelial barrier is exposing mucosal antigens (Strober et al., 2002). Whereas T cells are normally unresponsive to the resident microbial flora, in mice with TNBS-colitis, tolerance to microbial antigens of commensal bacteria is lost (Duchmann et al., 1999). In addition to determining the importance of gut microbes for the initiation of enterocolitis, the TNBS model also helped to establish the requirement of cytokine dysregulation in disease progression. Notably, TNBS-mediated inflammation could not be induced in TNF-α-deficient mice, and administration of anti-IL-12 antibody prevented TNBS-induced colitis (Strober et al., 2002). Although cytokine dysregulation in the DSS and TNBS models is required for the induction of colitis, resident gut microbes are likely to be an important factor as well.

Gene targeting models have allowed for the dissection of immune regulatory pathways by
enabling the ablation of individual cytokines and immune cell receptors. The T cell receptor (TCR) α and β (Mombaerts, 1993) and the Interleukin-2 (IL-2) (Sadlack et al., 1993) deficient models both support the hypothesis that the immune system plays a primary role in the etiology of UC. The T cell receptor is responsible for recognizing fragments of antigens bound to major histocompatibility complex molecules. IL-2 is a cytokine produced by active T cells that plays key roles in T cell proliferation and differentiation. Both TCR and IL-2 mutant mice developed diarrhea, weight loss, a thickened colonic wall with altered crypt structure, a depletion of goblet cells, and an inflammatory cell infiltrate in the lamina propria (Mombaerts, 1993; Sadlack et al., 1993). Importantly, antibodies directed against colonic antigens were detected in IL-2 mutant mice (Sadlack et al., 1993). Additionally, lesions were not present in the small intestine of either model (Mombaerts, 1993; Sadlack et al., 1993). These data support the hypothesis that UC results from an autoimmune attack on the colonic epithelium, possibly as a result of the high microbial load of the colon creating a highly antigenic environment. Thus, immune-mediated destruction of the colonic epithelium would lead to the production of autoantibodies, creating a cycle of epithelial targeting and damage (Mombaerts, 1993). Despite the role of dysregulated T cells in disease development, the presence of gut microbes was still required for the initiation of disease. Neither TCR-α−/− mice nor IL-2 deficient mice raised in a germ-free environment developed any symptoms of colitis (Dianda et al., 1997; Sadlack et al., 1993).

The hypothesis that commensal bacteria may play a role in the initiation of disease is supported further by the timing of disease onset in mice with a targeted mutation in IL-10 (IL-10−/−). IL-10−/− mice, when kept under conventional housing conditions, spontaneously developed chronic and highly lethal enterocolitis mainly affecting the duodenum, jejunum, and colon. When maintained in a specific pathogen-free (SPF) environment, IL-10 mutants developed colitis that was far less severe and only present in the proximal colon (Kuhn et al., 1993). At three weeks of age, which is when the normal gut microbiota first begin to stabilize (Schloss et al., 2012), inflammation was undetected in almost a third of SPF-housed IL-10−/− mice. However, by three months of age, intestinal lesions were observed in all mutants, often with transmural inflammation, hyperplasia, and crypt abscesses. Treatment of three-week-old mutants with
injections of IL-10 was able to completely prevent the development of colitis, but treatment of adults was unable to cure the disease (Berg et al., 1996). As IL-10 is a suppressor of macrophage activation, disease in these mice may be caused by the inappropriate release of inflammatory cytokines by unregulated macrophages. In that case, in young IL-10<sup>−/−</sup> mice, commensal gut microbes would initiate an uncontrolled immune response, resulting in the generation of pathogenic effector cells that produce inflammatory cytokines. However, in young IL-10<sup>−/−</sup> mice lacking pathogenic microbiota, there would be no inappropriate immune response and administration of IL-10 would prevent the development of disease. In older mice, administration of IL-10 may not have been able to reverse the progression of inflammation because the dysregulation of immune cells had already been established (Berg et al., 1996; Kuhn et al., 1993). Together, these models all demonstrate the requirement for a balance between protection against and tolerance to commensal bacteria by immune effector cells.

**Gut mucus and glycosylation**

The monosaccharides covalently added to specific amino acids within target proteins are called glycans. The addition in the Golgi apparatus of a glycan to the hydroxyl (-OH) group of serine or threonine residues, termed O-linked glycosylation, is a major form of posttranslational modification. In the gastrointestinal tract, key components of the glycoprotein layer that protects the lumenal cell membranes, termed the glycocalyx, are secreted and membrane-associated mucins (Corfield et al., 2001). Mucins are a family of high molecular weight O-linked glycoproteins that contain domains enriched with serine, threonine, and proline, termed PTS domains. The first step of O-linked glycosylation is the addition of an N-acetylgalactosamine (GalNAc) to the Ser/Thr residues of the PTS domains to form the Tn antigen. The Tn antigen is the foundation for further modifications to form four main core structures, designated cores 1 through 4 (Figure 1.1). Additional glycosyltransferases then go on to modify O-glycans, adding sugars, such as fucose and sialic acid. Most mucin-type O-glycans are derived from core 1 and core 3 structures. Core 3-derived O-glycan expression is restricted primarily to the salivary glands and the intestinal tract (Bergstrom and Xia, 2013; Honke and Taniguchi, 2002; Iwai et al., 2002).
Mucus covers the entire gastrointestinal tract from the mouth to the anus. In the small intestine, a single layer of mucus is secreted from the tops of the crypts and then travels toward the tips of the villi (Johansson et al., 2011). In the colon, there are two well-defined mucus layers: an outer, loose layer that is inhabited by commensal bacteria, and an inner viscous layer that is anchored to the epithelium and is sterile. As bacteria do not normally penetrate the inner mucus layer, they are unable to interact with epithelial cells directly. The main component of mucus in the small intestine and colon is the gel-forming mucin, Mucin-2 (MUC2) (Johansson et al., 2008). MUC2 is heavily modified by glycans, which make up 80% of the mass of the molecule (Bergstrom and Xia, 2013). In addition to their gel-forming function, glycans serve as both an attachment site and an energy source for bacteria. As bacteria degrade glycans and nondigestible food, they release short fatty acids that are taken up by the colonic epithelium and used as an energy source (Johansson et al., 2011).

**MUC2**

Due to its extensive O-glycan modifications, MUC2 has a high capacity to bind water, thus contributing to its gel-forming capability. When MUC2 is stored in mucin granules in goblet cells, it is in a condensed form, but when the granules are released, they expand substantially in volume (Johansson et al., 2008). The numerous chains of carbohydrates extending outward from the protein core of MUC2 are what give the molecule what is commonly termed a “bottle brush” appearance. In humans, the primary O-glycans found on MUC2 are core 3-derived structures, while murine MUC2 is comprised mostly of core-1 derived O-glycans, with core 3-derived structures making up only 1% of the modifications (Bergstrom and Xia, 2013). Muc2-deficient mice (Muc2−/−) develop intestinal inflammation and colorectal cancer (Van der Sluis et al., 2006; Velcich, 2002).

*The fucosylation pathway*

L-fucose is a common component of glycoproteins in mammalian cells. Glycans are fucosylated using the substrate guanosine diphosphate (GDP)-fucose, which is synthesized by
two different pathways in mammals: a *de novo* pathway and a free fucose-dependent salvage pathway. The *de novo* pathway generates fucose through the reactions of two enzymes: GDP-mannose 4,6-dehydratase (GMDS) and GDP-keto-6-deoxymannose 3,5-epimerase, 4-reductase (FX). The majority of fucose is produced through this pathway. The enzyme α(1,2)-fucosyltransferase (FUT2), which modifies salivary and other epithelial secretions, is also called the Secretor (Se) transferase (Kelly et al., 1995). FUT2 synthesizes the H blood group antigen, which may go on to be modified to form the A or B antigens, while unmodified H antigen is expressed in blood-type O individuals. Although A and B antigens are commonly associated with blood type, these cell surface antigens also play a role in host-microbe interactions in the gut (Becker and Lowe, 2003).

*Glycosylation perturbations in IBD*

It has long been known that the composition of glycoconjugates is perturbed in the colonic mucosa of patients during malignant transformation (Boland et al., 1982), as well as in patients with inflammatory bowel disease (Podolsky and Fournier, 1988). Consequently, mouse models have been derived in order to study the effects of changes in glycosylation on the colonic mucosa. Two well-studied examples are the *B3gnt6*-deficient (An et al., 2007) and the *Fx*-deficient mouse (Wang et al., 2016).

Mice homozygous for the *B3gnt6*-null allele, and thus unable to produce core 3-derived O-glycans, produced significantly less MUC2 protein. Despite the difference in core 3-derived O-glycan composition between mice and humans, *B3gnt6<sup>−/−</sup>* mice present with IBD-like features, including increased inflammatory cell infiltration, an increased epithelial proliferation rate, and had an enhanced susceptibility to DSS-induced colitis. Although *B3gnt6* is expressed throughout the GI tract, the inflammatory phenotype was restricted to the colon (An et al., 2007), which is reminiscent of human UC. *Fx<sup>−/−</sup>* mice, which have a deficiency of GDP-fucose, spontaneously develop colitis, with increased epithelial permeability and dysplasia that progresses to adenocarcinoma. The features of colitis in *Fx<sup>−/−</sup>* mice are reversible by dietary fucose supplementation (Wang et al., 2016). The fact that small changes in glycosylation in both of these
models resulted in colitis underlines the importance in maintaining a healthy glycocalyx capable of resisting infiltration by microbes.

**The intestinal epithelium**

The intestinal epithelium, a single cell layer that lines both the small intestine and colon, functions both as the site of nutrient absorption and as a barrier between the underlying lamina propria and lumenal antigens from food and commensal bacteria. In the small intestine, the epithelium forms finger-like projections called villi that extend toward the lumen, while in both the small intestine and the colon, the epithelium forms invaginations into the underlying mesenchyme called crypts. In mice, villi begin to form shortly before birth when outgrowths of the underlying mesenchyme push the pseudostratified epithelium toward the lumen. By birth, the epithelium is a single layer of columnar cells, with depressions between the budding villi called intervillus regions (Mathan et al., 1976). In neonates, intervillus regions are composed of polyclonal progenitor cells. However, as mature crypts form, the polyclonal cells are displaced by a single predominant progenitor at the crypt base. By postnatal day (P) 14, the villi are composed of bands of monoclonal cells arising from each of their surrounding crypts (Schmidt et al., 1988).

**Differentiated cell types of the epithelium**

The single stem cell pool at the crypt base gives rise to all seven of the epithelial cell types: absorptive, enteroendocrine, Paneth, goblet, microfold (M), cup, and tuft. All are found in the small intestine, while the colon lacks Paneth cells. Except for Paneth cells, all cells in the intestinal epithelium undergo renewal every 4-7 days, while Paneth cells are completely replaced after two months. Additionally, Paneth cells remain at the crypt base, while enteroendocrine, goblet, and absorptive cells, which are joined by junctional complexes, are forced as a unit toward the villus tip by new cells that are dividing at the crypt base. The migrating cells are eventually shed from the villus tip into the intestinal lumen (Cheng and Leblond, 1974; Gerbe et al., 2011).

Of the seven types of intestinal epithelial cells, only cup cells, which are distinguished from absorptive cells by their unique apical cup-like morphology and lighter cytoplasmic staining,
lack a clear function (Madara, 1982). The absorptive enterocytes, which are the most abundant villus cell type, are responsible for taking up nutrients, while the other cell types function either in a signaling, barrier forming, or an immunoregulatory capacity. Though enteroendocrine cells make up only approximately 1% of the intestinal epithelial cells, the gut nevertheless constitutes the largest endocrine organ in the body. Enteroendocrine cells modulate nutrient uptake and gut motility by sensing the contents of the lumen and releasing hormones to initiate appropriate responses. For example, in response to sensing lumenal amino acids, the subtype of enteroendocrine cells called G cells release of gastrin, which then enters circulation to stimulate the production of stomach acid (Sternini et al., 2008). Goblet cells, which make up a significant portion of the epithelium, form the first line of defense against gut bacteria by secreting a protective mucus layer that covers the epithelium. By secreting antimicrobial peptides Paneth cells, which are found at the base of every small intestinal crypt, modulate the population of small intestinal microbiota. Instead of preventing bacteria from breaching of the epithelial barrier, two of the least abundant cell types, M cells and tuft cells, may actually have roles as an interface between the gut microbiota and the immune cells that reside in the mesenchyme and the gut-associated lymphoid tissue (GALT) (Gerbe et al., 2012).

Recently, tuft cells were defined as being a unique cell type distinct from enteroendocrine cells (Gerbe et al., 2011). Named for their long apical microvilli that extend into the lumen, tuft cells have been thought to be involved in numerous functions, including chemoreception, smooth muscle contraction, and regulation of ion transport. However, new evidence points to a role of tuft cells in the intestinal response to epithelial parasites (Howitt et al., 2016). The first indication of a tuft cell-specific response was observed in specific pathogen-free (SPF) mice colonized by intestinal parasites when the abundance of tuft cells in the epithelium increased from the 0.4% reported in conventionally housed mice (Gerbe et al., 2012) to as much as 8% in the parasite-colonized mice. Goblet cell hyperplasia and eosinophilia were also observed in the lamina propria of the infected mice. In mice lacking expression of Trpm5, a gene required for taste-chemosensory signal transduction that intestinally is primarily expressed in tuft cells, the numbers of tuft cells, goblet cells, and eosinophils remained at the levels observed in conventionally-
housed animals. Thus, tuft cells may detect parasites, and then in response stimulate changes in epithelial cell differentiation, such as the protective increase in mucus production brought about by goblet cell hyperplasia, as well as the recruitment of mucosal immune cells (Howitt et al., 2016).

M cells comprise the follicle-associated epithelium, which is positioned directly over gut-associated lymphoid follicles. The apical surface of M cells is described as having “microfolds”, while the basal surface forms a “pocket” that contains B and T lymphocytes and macrophages. This interaction between the epithelial M cell and lymphocytes is thought to facilitate the delivery of foreign materials by way of transepithelial vesicular transport to antigen presenting cells. The influence of commensal microbiota on the composition of differentiated epithelial cells is highlighted by the fact that germ-free mice have a reduced number of M cells, but when they are exposed to pathogens, the number of cells differentiating into M cells increases (Neutra, 1998). Together, the functions of tuft cells and M cells illustrate how the intestinal epithelium and the gut microbiota interact in order to modulate both bacterial and differentiated epithelial cell populations.

Commensal bacteria in normal gut development

The presence of commensal bacteria in the gut is mutually beneficial, as the host provides nutrients for the bacteria and in turn, the bacteria are able to break down indigestible materials into nutrients that can then be used by the host. Before birth, the developing mammalian gut is sterile, but at birth mammals acquire gut microflora as they pass through the birth canal. The guts of mice raised in a germ-free environment have several phenotypic differences from conventionally raised mice that are reversible upon bacterial colonization, such as composition of the intestinal glycocalyx. In conventionally raised mice, during weaning, glycoproteins undergo a shift from being mostly terminally sialylated to being mostly terminally fucosylated. However, without stimulation by gut bacteria, the shift in glycosyltransferase activity does not take place. The fucose that modifies glycans is used as an energy source by commensal bacteria, therefore, the regulation of the production of fucose may allow the host to influence the composition of the
Another pathway by which the gut modulates the microbial population is through the secretion antimicrobial peptides, such as angiopoietin-4 (ANG4), a Paneth cell granule protein. During weaning in conventionally raised mice, a dramatic increase in ANG4 is observed that does not occur in germ-free animals (Hooper et al., 2003). Additionally, the increase in ANG4 and other angiogenin family members may stimulate villus capillary network development. In mature germ-free mice, capillary networks do not fully develop. However, the density of villus capillaries rapidly increases following the introduction of normal gut microbes (Stappenbeck et al., 2002). Thus, bacterial colonization may be required to stimulate some aspects of mammalian gut development.

**Signaling pathways: maintenance of stemness and differentiation**

The maintenance of the intestinal epithelium relies on a balance between the proliferative and the differentiated compartments, which is maintained through the coordination of a complex network of signaling pathways.

**Wnt signaling**

Wnt proteins comprise a large family of conserved growth factors that initiate signaling cascades to induce gene expression in multiple developmental processes throughout the body. The canonical process of Wnt signal transduction begins when a secreted Wnt ligand binds to the receptor complex formed by a seven transmembrane protein Frizzled (Fz) and a single-pass low density lipoprotein receptor-related protein (LRP). The Fz-LRP transmembrane complex then recruits and inhibits the intracellular β-catenin destruction complex, which consists of Axin, adenomatous polyposis coli (APC), and glycogen synthase kinase 3 (GSK-3). In the absence of Wnt signaling, this complex tags β-catenin for destruction by phosphorylating the protein, which is then ubiquitylated and degraded by the proteasome. However, following the sequestration of the Axin/APC/GSK-3 complex to the cell membrane following Wnt binding to its receptor, β-catenin accumulates in the cytoplasm and is shuttled to the nucleus. Once in the nucleus, β-catenin displaces the transcriptional repressor Groucho and partners with a complex that includes the T
cell-specific transcription factor/lymphoid enhancer-binding factor 1 (TCF/LEF) family of transcription factors to initiate transcription of Wnt target genes (Gordon and Nusse, 2006).

Canonical Wnt signaling is required for maintenance of epithelial proliferation in both the small intestine and the colon. In mice with disrupted Tcf7l2, which encodes the TCF family member transcription factor 4 (Tcf-4), cycling intestinal epithelial cells were absent by embryonic day 14.5, resulting in the neonatal epithelium being composed completely of differentiated cells (Korinek, 1998). In adult mice, overexpression of the secreted protein Dickkopf-1 (DKK1), which inhibits canonical Wnt signaling by causing the internalization of the LRP coreceptor, led to the breakdown of epithelial architecture in both the small intestine and colon (Kuhnert et al., 2004). Conversely, constitutive activation of the Wnt pathway causes hyperproliferation of the intestinal epithelium and the development of intestinal polyps. For example, in humans that have the disorder familial adenomatous polyposis (FAP), as well as in the ApcMin/+ mouse model of FAP, a mutation in the β-catenin destruction complex gene APC prevents the phosphorylation of β-catenin. The subsequent uncontrolled transcription of Wnt target genes gives rise to the in the development of multiple intestinal neoplasias (Min) (Su et al., 1992). Thus, the activation of Wnt target genes modulates intestinal epithelial proliferation during development and in adulthood.

Although 19 Wnt genes are encoded in the mammalian genome, only a subset of Wnts and their receptors are expressed in the intestine during development and in the mature gastrointestinal tract (Gordon and Nusse, 2006). In the fetal small intestine, Wnt5a has been detected in the mesenchyme at the tips of growing villi, while the receptors Fz5-7 are found in the epithelium. During embryogenesis, epithelial Wnt signaling may be required for the clustering of mesenchymal cells at the start of villus formation at E15.5, as in embryos with a conditional deletion of the β-catenin gene Ctnn1b intestine, these clusters did not form (Chin et al., 2016). However, interpretation of these results is complicated by the fact that β-catenin also functions as a part of epithelial adherens junctions. In the adult intestine, the ligands Wnt3 and Wnt9b are expressed in Paneth cells, while Wnt6 is expressed in the epithelia of small intestinal and colonic crypts. Expression of the Fz and LRP receptors, as well as the TCF proteins, is also restricted to the epithelium. In the mature mesenchyme, Wnt2b and Wnt5a have been detected in the villus,
whereas Wnt5a and the related Wnt5b are both expressed in the mature colon mesenchyme. Notably, Wnt5a is also highly upregulated in the stromal cells of Apc<sup>Min</sup> polyps (Gregorieff et al., 2005).

Due to the multiple Wnts expressed both in the epithelium and mesenchyme, until recently, the primary source of Wnt signaling in the intestine has been unclear. However, new evidence suggests that mesenchymal Wnts are required both during fetal development and in adults. It has recently been shown that in embryos, mesenchymal deletion of Wntless, which is required for the secretion of all Wnts, resulted in a significant decrease in epithelial proliferation near the time of villus formation, suggesting that mesenchymal Wnt secretion may be required for epithelial proliferation once villus formation begins (Chin et al., 2016).

Furthermore, the epithelial stem cell niche is likely comprised of Wnt2b- and Wnt5a-secreting mesenchymal cells labeled by the forkhead box family member FOXL1, which are distributed at the interface of the mesenchyme and the epithelium. Diphtheria toxin (DT)-mediated ablation of FOXL1-expressing cells resulted in the loss of proliferating epithelial cells followed by a loss of crypt architecture as soon as three days following DT administration (Aoki et al., 2016). Together, these data show that the maintenance of proliferation in the intestinal epithelium requires specific Wnt-secreting cells in the mesenchyme.

**Inhibition of Wnt signaling by Hedgehog**

In addition to promoting epithelial growth, epithelial-mesenchymal communication is required to limit proliferation. Inappropriate expansion of the proliferative compartment is prevented by bone morphogenetic protein (BMP) signals in the differentiated cell compartment. In mice with a conditional deletion of the gene for BMP receptor protein BMPR1A targeted to the intestinal epithelium, Wnt is de-repressed and the mice develop intestinal polyposis as a result of the expansion of the stem and progenitor cell populations (He et al., 2004).

Further evidence of the BMP suppression of Wnt signaling was found by investigating the upstream regulation of BMP by hedgehog (Hh) signaling. The Hh ligands sonic hedgehog (Shh) and Indian hedgehog (Ihh) are expressed in the epithelium, while their receptors are expressed in
the mesenchyme. Thus, Hh signaling occurs in a paracrine manner (Kosinski et al., 2010). Neonatal mice epithelially overexpressing the pan-Hh inhibitor, Hhip, which inhibits both Shh and Ihh, had hyperproliferative intestinal epithelia, while the villi failed to fully form. Consistent with the function of BMP as a Wnt inhibitor, the expression of Wnt target genes was upregulated in Hhip mice (Madison et al., 2005).

Despite previous studies showing that Bmp4 expression is regulated by Shh, deletion of Shh targeted to the intestinal epithelium did not phenocopy the Hhip mutants. Rather, ablation of Shh in intestinal epithelium resulted in a decrease in the number of goblet and Paneth cells, impaired fucosylation of secreted glycoproteins, and ER stress resulting from impaired autophagy (Gagné-Sansfaçon et al., 2014). Conversely, deletion of Ihh in the intestinal epithelium disrupted the epithelial-mesenchymal interactions necessary for the regulation of proliferation. The epithelium of Ihh-mutant intestine was hyperproliferative, with expanded crypts and differentiation skewed toward a secretory fate. Moreover, whole intestine gene expression analysis by microarray showed that Ihh mutants had increased expression of Wnt target genes, such as Sox9, cMyc, and Mmp7, increased expression of the intestinal stem cell marker Lgr5, and a significant decrease in the expression of BMP signaling genes (Kosinski et al., 2010). Together, these findings strongly suggest that epithelial Ihh mediates epithelial proliferation by regulating the mesenchymal expression of BMP ligands. BMP signaling then inhibits Wnt signaling, resulting in reduced epithelial proliferation.

**Cdx1 and Cdx2**

During development, the transcription factor caudal-type homeobox 2 (Cdx2) is highly expressed throughout the intestinal epithelium from the duodenum to the distal colon. While expression of Cdx2 is primarily restricted to differentiated cells, Cdx1 is expressed throughout the entire intestinal epithelium, with expression primarily restricted to the crypts, overlapping slightly with Cdx2 (Silberg et al., 2000).

Perturbation of Cdx2 expression during embryogenesis results in disrupted regionalization of the gastrointestinal tract. For example, ectopic expression of CDX2 in the
murine stomach was sufficient to induce intestinal metaplasia, as was shown by the inappropriate
 gastric expression of intestinal epithelial cell markers, including positive staining for intestinal
 mucins (Silberg et al., 2002). Notably, in human intestinal metaplastic pyloric mucosa, expression
 of gastric transcription factor SOX2 is significantly decreased and expression of CDX1, CDX2,
 and the intestinal mucin MUC2 are all increased (Tsukamoto et al., 2004). Thus, CDX2 is capable
 of activating the intestinal program in foregut endodermally derived tissue.

Conversely, mice with Cdx2 ablation during early endoderm differentiation developed
 gastrointestinal tracts that terminated without a colon and had no opening at the distal end,
 resulting in complete intestinal obstruction. Histological examination of the Cdx2-null mutants
 revealed that villus development was impaired and that enterocytes failed to develop a brush
 border. Thus, though the gut tube retained expression of the pan-endoderm marker FOXA1, the
 epithelial cells in the mutant intestine did not become mature differentiated intestinal cells (Gao et
 al., 2009).

As opposed to the intestinal metaplasia observed in mice ectopically expressing Cdx2 in
 the proximal endoderm, Cdx2-deficient distal endoderm gave rise to tissue that was reminiscent
 of mouse esophagus. Not only was the anterior foregut endoderm marker Sox2 inappropriately
 expressed in the distal gut, but gene expression profiling of mutant and control E18.5 GI tissue
 also showed that the transcriptome of the mutant ileum was more similar to the control
 esophagus than to the control ileum. Furthermore, in the mutant ileum, intestine-specific
 transcription factor genes, including Hnf4α and Cdx1, were significantly downregulated. Additionally, the factors Shh, Ihh, were downregulated, while expression of multiple Wnt ligands,
 Tcf factors, and Wnt target genes, including cyclin D1 and Sox9, were increased in the mutant
 intestine. Together, the altered gene expression indicates that the foregut endoderm
 differentiation program was ectopically activated in the distal Cdx2 mutant intestine (Gao et al.,
 2009). Thus, CDX2 is the master regulator of the intestinal differentiation program.

In order to better understand the mechanism by which CDX2 regulates gene expression
 in the colon, CDX2 chromatin binding was analyzed in Caco2 cells, a line of human epithelial
 colorectal adenocarcinoma cells that proliferate rapidly until they reach confluence, at which point
they develop characteristics of mature enterocytes. These analyses revealed that CDX2 interacts with multiple different binding partners depending upon the differentiation state of the cells. At CDX2 binding sites both in proliferating and confluent (i.e. differentiated) Caco2 cells, the FOXA motif was enriched. However, in pre-confluent (i.e. undifferentiated) Caco2 dividing cells, CDX2 binding sites were enriched for the GATA binding motif and the Wnt-responsive TCF/LEF motifs, while in differentiated cells, the HNF4A motif was enriched (Verzi et al., 2010). The enrichment of HNF4A/CDX2 co-occupied sites is consistent with the previous finding that HNF4A controls a large number of genes required for differentiation of colonic epithelium (Garrison et al., 2006).

Like Cdx2, Hnf4α is also expressed throughout the intestinal epithelium during development (Garrison et al., 2006). Additionally, chromatin regions that are either co-occupied or individually occupied by CDX2 and HNF4A are enriched with nucleosomes modified by histone H3 dimethylated on lysine 4 (H3K4me2), which label active enhancers. While the binding of HNF4A is dependent on CDX2 binding at normally co-occupied sites, CDX2 binding is unaffected by Hnf4α ablation. Therefore, although the two transcription factors do not directly interact, HNF4A likely binds CDX2-occupied sites as a result of CDX2 maintaining a permissive chromatin structure. CDX1 likely has some functional redundancy with CDX2, as intestine lacking expression of both Cdx1 and Cdx2 lost more H3K4me2-marked enhancers than were lost in CDX2-deficient intestine alone. Therefore, CDX factors regulate gene expression by binding at active enhancers and maintaining an open chromatin conformation (Verzi et al., 2012). Taken together, these data demonstrate the requirement of CDX-mediated chromatin accessibility for intestine-specific transcription factors to initiate gene expression and establishes CDX2 as the master intestinal transcription factor.

**Notch signaling**

Notch signaling is required in the intestinal epithelium for the fate specification of transit amplifying cells. Notch signal transduction, which was first studied in *Drosophila*, is highly conserved throughout evolution. Mammals have four notch receptors (Notch1-4) and five canonical ligands: Dll1, Dll3, and Dll4, Jagged1 and Jagged2 (Kopan and Ilagan, 2009). Both
Notch receptors and their ligands, which are from the DSL (Delta, Serrate, and Lag 2) protein family, are single-pass transmembrane proteins. Thus, activation of Notch signaling can only take place between directly apposed cells (Kopan and Ilagan, 2009; Mumm and Kopan, 2000). This form of cell-cell communication, termed “lateral signaling”, has been suggested to function as an inhibitor of default cell differentiation, thus allowing neighboring cells to follow secondary fate pathways.

Upon ligand-mediated Notch activation, the extracellular domain of the Notch receptor, which binds ligand and inhibits Notch activity in the absence of ligand, is cleaved. Next, the Notch intracellular domain (NICD) is released from the transmembrane domain by the enzyme γ-secretase and transported to the nucleus. In the nucleus, NICD binds the DNA binding protein RBPJ, converting RBPJ from a repressor to a transcriptional activator. This cascade results in the transcription of Notch target genes, such as Hes1 (Mumm and Kopan, 2000).

In the intestine, Notch target genes control the switch between an absorptive and a secretory cell fate. Inhibition of Notch by either the ablation of RBPJ, or by the administration of γ-secretase inhibitor, drives cells to a goblet cell fate (van Es et al., 2005). Additionally, deletion of the transcription factor HES1 in mouse embryos caused a significant increase in the abundance of cells entering the secretory cell fate (Jensen et al., 2000). Conversely, the constitutive activation of Notch1 in the intestinal epithelium results in a complete loss of secretory cell differentiation, while causing the proliferative crypt compartment to expand. This is attributed to the increased transcriptional repression of the transcription factor MATH1 by HES1 (Fre et al., 2005).

The requirement of MATH1 for secretory cell differentiation is illustrated by Math1-deficient mouse embryos, which completely lack secretory cells (Yang et al., 2001). Lineage tracing indicates that the differentiated cells expressing MATH1, including goblet, Paneth, and enteroendocrine, are all from the secretory lineage. Tissue-specific ablation of Math1 in the intestinal epithelium results in the differentiation of all epithelial cells to absorptive enterocytes. Furthermore, consistent with the role of secretory cells as initiators of the primary response to epithelial damage, MATH1-deficient crypts do not increase in proliferation or depth in response to
small bowel resection. In summary, Notch signaling modulates the abundance of secretory cells by activating transcription of \textit{Hes1}, an inhibitor MATH1, a transcription factor required for secretory cell fate determination (Shroyer et al., 2007).

\textbf{The FOXA transcription factor family}

\textit{The forkhead gene family}

The winged helix/forkhead box (Fox) transcription factors are a family of proteins that bind DNA as a monomer and are characterized by a highly conserved 101 amino acid DNA-binding domain, termed the “winged helix” motif (Clark et al., 1993). Fox genes have been identified in eukaryotes from yeast (Zhu et al., 1993), nematodes (Kalb et al., 1998), sea anemones (Fritzenwanker et al., 2004), and zebrafish (Norton et al., 2005) to mice and humans. The mouse genome encodes a total of 44 Fox genes and the human genome encodes 50 Fox genes, which have been classified into 19 subfamilies (FOXA-FOX5) (Jackson et al., 2010; Kaestner et al., 2000).

\textit{Fork head (fkh)}, the first member of the Fox family, was identified in \textit{Drosophila} as a homeotic regulator of proximal and distal gut tube formation (Weigel et al., 1989). At approximately the same time, the three mammalian \textit{FOXA} genes (formerly hepatocyte nuclear factor 3, \textit{HNF3}), \textit{FOXa1}, \textit{FOXa2}, and \textit{FOXa3}, were identified as important transcriptional regulators in rat liver because of their enriched binding to hepatocyte-specific promoters (Lai et al., 1990). Based on its high degree of amino acid identity (approximately 75\%) to the mammalian \textit{FOXA} and \textit{fkh} winged helix domains (Azzaria et al., 1996; Weigel and Jäckle, 1990), in addition to similar tissue expression to \textit{fkh}, the \textit{C. elegans} gene \textit{Pha-4} was subsequently identified as an additional homolog (Kalb et al., 1998). Like \textit{fkh} in \textit{Drosophila}, \textit{Pha-4} is required for the development of its digestive organ, the pharynx. Notably, the PHA-4 binding motif TRTTKRY (Gaudet and Mango, 2002) is nearly identical to the \textit{FOXa2} binding motif identified in mouse liver nuclear lysates (Tuteja et al., 2009). Together, these data point to a highly conserved role for FOXA homologs throughout evolution in the specification of the gut tube.

\textit{Chromatin remodeling}
To investigate the mechanism by which FOXAs regulate gene expression, the DNA-binding domain of rat FOXA3 and the promoter sequence of the liver gene transthyretin (TTR) were cocrystallized, thus revealing the butterfly-like “winged helix” three-dimensional structure of FOX proteins. Additionally, the crystallographic structure showed that there is a high degree of structural similarity between FOXA3 and linker histone H1 (Clark et al., 1993). A recent comparison of linker histone H1 deposition in wild type and FOXA1/A2-null (mutant) liver shows that H1 binding is increased at FOXA1/A2 target regions in mutant liver, while in controls, higher FOXA2 occupancy is correlated with H1 depletion. Furthermore, binding of liver-specific transcription factors, such as HNF4A, at FOXA2-bound enhancers was reduced in mutant hepatocytes (Iwafuchi-Doi et al., 2016). Moreover, it was shown in vitro that by C-terminal domain-mediated binding of histones H3 and H4, mouse FOXA1 is capable of remodeling chromatin (Cirillo et al., 2002). Thus, FOXA proteins may maintain an accessible chromatin conformation for tissue-specific transcription factors by directly interacting with histone proteins.

**Tissue expression**

Analysis of expression patterns in adult mouse tissues revealed that all three Foxa genes are highly expressed in the adult liver, stomach, and intestines. However, Foxa1 and Foxa2 are expressed in the lungs, while Foxa3 is absent. Conversely, Foxa3 alone is highly expressed in the ovary and testes, heart, thymus, and adipose tissue. Although all three factors are expressed in the liver, stomach, and intestines, Foxa3 is the most highly expressed (Kaestner et al., 1994). Despite the similarity in coding sequence between the Foxa genes, with 95% sequence identity between the Foxa1 and Foxa3 forkhead domains (Clark et al., 1993), differences in regulation and tissue expression may be explained by the differing sequences of their promoters (Kaestner et al., 1994). Thus, in some tissues, but not others, ablation of one factor may be compensated for by upregulation of the one or both of the other factors.

Still, the mechanism by which transcription of the three Foxas is regulated remains unclear. A FOXA binding site was found upstream of the Foxa1 gene that bound all three FOXAs in liver nuclear extracts, suggesting autoregulation by FOXA1 and crossregulation by FOXA2 and
FOXA3. However, in Fox1-LacZ homozygous animals (Fox1<sup>−/−</sup>), in which the fragment of Fox1 encoding the DNA binding domain was replaced by LacZ, the expression level of the Fox1-LacZ allele was the same as for the wild type Fox1, suggesting FOXA1 does not autoregulate its expression (Kaestner et al., 1999).

Regulation of tissue specification and organogenesis by FOXA1 and FOXA2
FOXA1 and FOXA2 have both separate and overlapping roles during embryonic development and in mature tissues. Early in embryogenesis, FOXA2 is required in the visceral endoderm for proper primitive streak morphogenesis (Dufort et al., 1998), notochord development (Weinstein et al., 1994) and node formation (Ang and Rossant, 1994). Later, it is required for specification of the endodermal foregut (Dufort et al., 1998), though not the midgut or hindgut (McKnight et al., 2010). Organs derived from foregut and midgut endoderm, including the lungs, liver, pancreas, and intestines, require both FOXA1 and FOXA2 during patterning and in the mature animal.

Lungs: In the embryonic mouse respiratory epithelium, absence of FOXA1 or FOXA2 alone does not inhibit branching morphogenesis or cell differentiation. However, Foxa1/Foxa2-null lungs, despite having a normal pulmonary vascular structure, are smaller than control and single mutant lungs and are lined by immature cuboidal epithelial cells. Furthermore, there is an increase in Foxa2 mRNA in the Foxa1-mutant respiratory epithelium, while in Foxa2-mutants there is an increase in Foxa1 mRNA, suggesting that the factors are repressing each other's expression in normal lung development. In tissue mosaic for the deletion of Foxa1/Foxa2, absence of differentiation markers is restricted to the mutant regions. Thus, FOXA1 and FOXA2 have cell autonomous, overlapping roles in lung morphogenesis (Wan et al., 2005).

Despite the redundancy of FOXA1 and FOXA2 during embryogenesis, postnatal morphogenesis is significantly altered in Foxa2-null lungs. In addition to increased airspace enlargement, the Foxa2-mutant respiratory epithelium exhibits goblet cell hyperplasia and increased macrophage and neutrophil infiltration. Goblet cell hyperplasia in mutant lungs may result from the de-repression of Mucin5ac (Muc5ac) transcription by FOXA2, as in vitro it was
shown that FOXA2 directly inhibits MUC5AC expression. Notably, in lung tissue from human subjects with chronic lung disease, regions of goblet cell hyperplasia lacked FOXA2. Thus, FOXA2 may be required in the lung for the inhibition of goblet cell differentiation (Wan et al., 2004).

Liver: The absence of Foxa1, Foxa2, or Foxa3 individually in the endoderm is not sufficient to prevent liver specification. However, when both Foxa1 and Foxa2 are deleted in the foregut endoderm of mouse embryos at E8.5, progenitor cells in the foregut endoderm fail to form a liver bud and the hepatoblast marker alpha-fetoprotein (Afp) is not expressed (Lee et al., 2005). Ablation of Foxa1 and Foxa2 following liver specification results in hyperproliferation of the biliary tree from the increased cycling of cholangiocytes as a result of increased signaling by IL-6, an activator of cholangiocyte proliferation. In control liver, the IL-6 promoter is co-occupied by its negative regulator, glucocorticoid receptor (GR), and FOXA1/FOXA2. Following deletion of Foxa1/Foxa2 in the liver, binding at the IL-6 promoter by GR was significantly decreased. Although the proliferation rate of hepatocytes is unaffected, by three months of age, the bile ducts of mutant livers become fibrotic. Thus, the FOXA proteins limit bile duct expansion through regulation of IL-6 expression (Li et al., 2009).

Pancreas: Although newborn Foxa1−/− mice are the same weight at birth as their littermates, their growth lags and they die of severe hypoglycemia by P14. Plasma levels of both insulin and glucagon, as well as preproglucagon mRNA levels, are reduced in mutants, thus they are unable to modulate their blood glucose (Kaestner et al., 1999). Like Foxa1−/− mice, when deletion of Foxa2 is targeted to insulin-secreting β-cells by Cre driven by the insulin promoter (InsCre), Foxa2 mutants die before weaning of hyperinsulinemic hypoglycemia. Unlike Foxa1−/− islets, which have significantly reduced glucose-stimulated insulin secretion (Vatamaniuk et al., 2006), Foxa2loxP/loxP,InsCre islets are hypersensitive to glucose-stimulation, resulting in a significantly increased ratio of plasma insulin to glucagon, despite the decreased levels of both hormones compared to controls. An explanation for the abnormal increase in insulin secretion may be that
FOXA2 targets genes that regulate glucose sensing, as mRNA levels of Kcnj11 and Abcc8, the genes that encode both subunits of the K\textsubscript{ATP} channel, are significantly downregulated in Foxa2-null islets (Sund et al., 2001).

Both FOXA1 and FOXA2 bind a critical enhancer for the transcription factor pancreatic and duodenal homeobox 1 (Pdx1), which is required for pancreatic development (Lee et al., 2002). In Foxa1\textsuperscript{loxP/loxP}; Foxa2\textsuperscript{loxP/loxP}; Pdx1Cre mice, conditional deletion of both Foxa1 and Foxa2 results in mutant mice having severe pancreatic hypoplasia and lacking all Pdx1 expression (Gao et al., 2008). Pdx1CreERT2-mediated ablation of FOXA1 and FOXA2 in mature islets results in mice developing severe hypoglycemia. Gene expression array analysis comparing mature Foxa1\textsuperscript{loxP/loxP}; Foxa2\textsuperscript{loxP/loxP}; Pdx1CreER and control islets revealed that genes involved in neural differentiation are upregulated in mutant islets, while factors involved in secretory pathways are downregulated, such as insulin receptor-related receptor (Insrr) and MLX interacting protein like (Mlxipl). Significantly, transcription of Mlxipl, which encodes the carbohydrate response element-binding protein (ChREBP), is directly regulated by FOXA1 and FOXA2, both during development and in mature islets (Gao et al., 2010). Additionally, in vitro, all three FOXAs are able to transactivate the preproglucagon promoter (Kaestner et al., 1999). Thus, FOXA proteins are important factor for the control of glucose homeostasis through the direct regulation of genes that control glucose sensing and hormone secretion.

**Intestines:** While Foxa1 is expressed uniformly throughout the epithelium of the developing and the mature gut, Foxa2 expression is restricted to the intervillus region of perinatal intestines and to the crypts and goblets cells of the adult intestine. In the murine rectal cancer cell line CMT-93, which expresses Muc2, Foxa1, and Foxa2, downregulation of Foxa1 and Foxa2 by siRNA resulted in a dramatic decrease in Muc2 expression. ChIP-PCR confirmed that FOXA1 and FOXA2 bind to the Muc2 promoter in vivo. Thus, Muc2 is a direct target of FOXA1 and FOXA2 (van der Sluis et al., 2008).

In order to further investigate the role of FOXA1 and FOXA2 in intestinal differentiation, deletion of Foxa1 and Foxa2 was targeted to the intestinal epithelium by Cre driven by the Villin
promoter. Although the anterior/posterior patterning of the intestine was normal Foxa1^{loxP/loxP}; Foxa2^{loxP/loxP};VillinCre mice, the number of goblet cells was significantly decreased and mucin granule density was reduced compared to controls. Notably, Muc2 and Tff3 mRNA levels were significantly decreased in the colon, but not in the small intestine of mutant mice. However, the number of chromogranin A-positive enteroendocrine cells was significantly down in the mutant small intestine (Ye and Kaestner, 2009). Thus, FOXA1 and FOXA2 are not only required for regulating Muc2 expression, they are also vital for the normal differentiation of intestinal secretory cells.

Regulation of tissue specification by FOXA3

Although there was initially little evidence for a significant regulatory role for FOXA3, new functions have recently been discovered. Foxa3-null mouse embryos are morphologically normal and mice are born healthy and fertile at normal Mendelian ratios. Although Foxa1 and Foxa2 transcript levels are unchanged in mutant stomach and colon, their levels are increased in Foxa3-null liver, signifying a possible compensatory increase in transcription. In the mutant liver, transcription rates for FOXA target genes, including phosphoenolpyruvate carboxykinase (Pck1), a part of the gluconeogenesis pathway, are reduced, indicating that FOXA3 activates expression of liver genes (Kaestner et al., 1998). Additionally, FOXA3 may be a significant factor for promoting liver regeneration following toxic liver injury, as a competitive overexpression screen of liver genes found that FOXA3 significantly promotes hepatocyte repopulation, while FOXA1 and FOXA2 do not (Wangensteen et al., 2015). The importance of FOXA3 as a hepatic transcription factor is underscored by the finding that lentiviral expression of FOXA3, along with HNF1A and HNF4A, transdifferentiates human fibroblasts to cells that express hepatic gene programs (Huang et al., 2014).

With their highly conserved overlapping and unique functions, FOXA1, FOXA2, and FOXA3 are essential for the development and maintenance of tissues derived from the foregut and midgut endoderm.
**Figure 1.1**: Synthesis of the four main mucin type O-glycans. The foundation of mucin protein O-linked glycosylation is the Tn antigen, which is comprised of N-acetylgalactosamine (GalNAc) bound to the -OH group of a serine (Ser) or a threonine (Thr) residue. The Tn antigen is modified by the addition of galactose and/or N-acetylglucosamine (GlcNAc) to form the four common core O-glycan structures. Addition of GlcNAc and galactose is carried out by the enzymes listed above the black arrows. The core structures may then be branched and extended by the addition of more sugars.
Parts of this chapter were adapted from Kieckhaefer et al. 2016. The RNA Polymerase III Subunit Polr3b is required for the maintenance of small intestinal crypts in mice. Cell Mol Gastroenterol Hepatol. 2: 783–795.
**POLR3B MATERIALS AND METHODS:**

**Derivation of Polr3b\textsubscript{loxP/loxP};VillinCre mice**

The Polr3b conditional mutant allele was derived in mouse embryonic stem cells via homologous recombination using the targeting construct schematized in Figure 1A. LoxP sites flanking exon 10 allow for ablation of exon 10 in Polr3b\textsubscript{loxP/loxP} upon Cre activation. The neomycin cassette was excised by crossing the F1 offspring with Fip1 deleter mice (Rodríguez et al., 2000). In order to ablate Polr3b conditionally in the intestinal epithelium, Polr3b\textsubscript{loxP/+};VillinCre mice were crossed with VillinCre mice, kindly provided to us by Deborah Gumucio (Madison et al., 2002), to obtain Polr3b\textsubscript{loxP/+};VillinCre and Polr3b\textsubscript{loxP/loxP};VillinCre mice. Genotyping was performed by polymerase chain reaction (PCR) analysis of genomic tail DNA (data not shown). The following primers were used to detect the Polr3b wild type and mutant allele, yielding a 222 bp product and a 519 bp product, respectively: 5’-AGCCTCATTTCTTCGTGGTG-3’ and 5’-TGGCCTAGGAAGGGAAGAAA-3’. Polr3b\textsubscript{loxP/+};VillinCre and Polr3b\textsubscript{loxP/loxP} mice show no abnormalities and are viable and fertile. The Cre transgene was detected using: 5’-GCGGCATGGTGCAAGTTGAA-3’ and 5’-CGTTCACCGGCATCAACGTT-3’ primers (product size of 232 bp). To perform lineage tracing experiments, Polr3b\textsubscript{loxP/+};VillinCre and Polr3b\textsubscript{loxP/loxP} mice were crossed with R26-stop-EYFP mice (B6.129X1-Gt(ROSA)26Sor\textsuperscript{tm1(EYFP)Cos}J/J) obtained from The Jackson Laboratory (Bar Harbor, ME). All procedures involving mice were conducted in accordance with approved Institutional Animal Care and Use Committee (IACUC) protocols.

**Laser Capture Microdissection (LCM) and subsequent PCR analysis of wild type and mutant allele**

Paraffin-fixed intestinal tissue was sectioned (5 µm) on nuclease and nucleic acid free frame slides for laser microdissection (ASEE, FS-LMD-M-50r). Unstained and dehydrated tissue was air-dried and processed immediately for LCM. Crypt regions were dissected from Polr3b\textsubscript{loxP/loxP};VillinCre mice and their wild type littermates into PCR sample tubes for laser microdissection (ASSEE, ST-LMD-M-500), subsequent overnight proteinase K digestion, and
PCR analysis for the presence of the wild type and mutant Polr3b allele using the following primers: 5’-GACGAATTCCTGCAGCCCAA-3’ (Polr3bF) with 5’-CCCTGGAATGAACAGGGGAAA-3’ (Polr3bR1), which yield a wild type product of 930 bp and a mutant product of 337 bp. Additional PCR analysis was performed using the same Polr3bF primer and the Polr3bR2 primer: 5’-GTGGCCGCTTACTGAGATGAA-3’ in order to detect only the wild type product of 530 bp.

Antibodies

The following antibodies were used for immunofluorescent or immunohistochemical analysis: rabbit anti-Ki67 (1:1000; Vector Laboratories; Burlingame, CA; VP-K451), rabbit anti-lysozyme (1:1000; Dako; Glostrup Municipality, Denmark; A0099), mouse anti-E-cadherin (1:500; BD Transduction Laboratories; San Jose, CA; 610181), rabbit anti-Epcam (1:100; Abcam; Cambridge, United Kingdom; ab71916), chicken anti-GFP (1:2000; Aves Labs; Tigard, OR; GFP-1020), rabbit anti-Sox9 (1:250; Millipore; Billerica, MA; AB5535), rabbit anti-cyclin D1 (1:100; Biocare Medical; Concord, CA; CRM 307), rabbit anti-cleaved caspase-3 (1:200; Cell Signaling Technology; Danvers, MA; 9661), and rabbit anti-mucin 2 (1:1000; Santa Cruz Biotechnology; Dallas, TX; sc-15334). Cy2- and Cy3-conjugated fluorescent secondary antibodies were purchased from The Jackson Laboratory and used at 1:200 dilution. Biotinylated secondary antibodies for immunohistochemical assays (Vector Laboratories) were used at a 1:200 dilution along with the ABC detection system (Vector Laboratories; PK-6100).

Histology, immunohistochemistry and immunofluorescence

Standard Hematoxylin and eosin (H&E) and Alcian blue stains were performed by the Morphology Core of the Penn Center for Molecular Studies in Digestive and Liver Diseases. For quantification of Alcian blue stained sections, images of age-matched control and Polr3b-mutant intestines were analyzed. Positive cells, as well as the total number of epithelial cells, within five villi per section (from the middle of the crypt base to the middle of the villus tip) were manually counted in blinded fashion from four control and mutant intestinal sections each. In addition, five
regions that displayed branching or hyperplasia were counted in P21 mice for total epithelial cells and Alcian blue-positive cells. The numbers of Alcian blue-positive cells were normalized to total number of epithelial cells.

**TUNEL assay**

In order to test the apoptotic cells in the intestinal epithelium, we performed the TUNEL assay using an in situ cell death detection kit according to the manufacturer’s instructions (Roche Applied Science; Penzburg, Germany). Cell death was quantified by counting the number of TUNEL-positive cells per crypt in jejunum sections from P21 mice (control, n = 4 and mutant, n = 4). An average of 20 crypts were counted per animal.

**Quantitative PCR analysis**

For the analysis of stem cell markers, total RNA was extracted from the jejunal tissue using the TRIzol reagent method (Invitrogen; Carlsbad, CA; 15596-026). cDNA synthesis and quantitative PCR analysis were performed as described (Gao et al., 2007). Gene expression was normalized to mRNA levels of Gapdh. Primer sequences used for the quantitative PCR analysis have been reported (Foureau et al., 2010; Inoue et al., 2008; Sakamori et al., 2012; Salzman et al., 2010).

To measure the expression of Pol III targets, total RNA was extracted from the jejunal tissue of embryonic day 18.5 (E18.5) mice using the mirVana miRNA Isolation Kit (Life Technologies, AM1560). cDNA synthesis and quantitative PCR analysis were performed using Custom TaqMan Small RNA Assays (Life Technologies, 4440418) for tRNA-Arg, tRNA-Asn, tRNA-Cys, tRNA-His, and 5S rRNA. TaqMan MicroRNA Assays (Life Technologies, 4427975) were used for cDNA synthesis and quantitative PCR analysis of U6 snRNA, snoRNA 202, and snoRNA 234. Gene expression was normalized to the mean level of snoRNA 202 and snoRNA 234. Primer sequences available on request.

**Intestinal crypt/enteroid culture**
Small intestine was harvested from P1 Polr3b-mutant and control mice. After a 30-minute EDTA (5 mM) incubation, the intestinal crypts were filtered through a 70 μm cell strainer. Crypts were plated suspended in Matrigel (BD Biosciences, 356231) and cultured in Advanced DMEM/F-12 (Thermo Fisher Scientific; Waltham, MA; 12634010) containing the Wnt agonist R-spondin1 (Protein Expression Facility, The Wistar Institute; Philadelphia, PA), murine epidermal growth factor (mEGF) (Life Technologies; Carlsbad, CA; PMG8043), murine Noggin (mNoggin) (Peprotech; Rocky Hill, NJ; 250-38), and the Gsk3 inhibitor (CHIR99021) (BioVision; Milpitas, CA; 1677-5) as published (de Lau et al., 2011; Sato et al., 2009). On days 2, 4, and 6, pooled enteroids were genotyped by qPCR for the presence of the deleted allele using primers Polr3bF and Polr3bR1, as described above, and undeleted allele using primers Polr3bF2: 5’-GCTTGATATCGAATTCCGAAG-3’ and Polr3bR: 5’-CCTCTCTGGAACTCCAACCAA-3’. The quantity of genomic DNA present in qPCR reactions was normalized to genomic β-actin. In addition, proliferation and apoptosis were assessed in developing enteroids by immunofluorescent analysis with Ki67 and cleaved caspase-3, respectively, as described above.

**Statistical Analysis**

Values represent means ± standard error of the mean for the indicated number of mice per group. Differences between groups were determined using an unpaired Student's t test (2-tailed), with a significance cutoff of p<0.05, unless indicated otherwise. Percentage undeleted allele quantified in enteroids by qPCR at three different time points (2, 4, and 6 days in culture) was compared using the Kruskal-Wallis test, followed by Dunn’s Multiple Comparison Test. Analysis was performed with GraphPad Prism 5 software (GraphPad Software Inc.; La Jolla, CA).

**FOX1/A2 MATERIALS AND METHODS**

**Derivation of Foxa1^LoxP/LoxP^; Foxa2^LoxP/LoxP^; VillinCre Mice**

The derivation of the Foxa1^Lox^, Foxa2^Lox^, and VillinCre mice has been described previously (Ye, 2009). No difference was observed between Foxa1^LoxP/+^; Foxa2^LoxP/LoxP^ and
F*oxa*1*LoxP/LoxP; Foxa2*LoxP/LoxP* mice, hence both genotypes were used as controls and genotyping was performed as described previously (Ye, 2009).

**RNA preparation and quantitative Real-Time PCR (qRT-PCR)**

Large intestines were dissected and divided into proximal and distal colon. Total RNA was extracted using TRIzol (Thermo Fisher Scientific; Waltham, MA) according to the manufacturer’s protocol. RNA was reverse-transcribed using SuperScriptTM II Reverse Transcriptase (Thermo Fisher Scientific; Waltham, MA) and quantitative real-time PCR was performed as described before (control n=3, mutant n=3). (Gupta, 2009).

**Immunohistochemistry**

Bromodeoxyuridine (BrdU) labeling was performed by intraperitoneal injection of 1ml/100g body weight BrdU solution (catalog no. 00-0103; Thermo Fisher Scientific; Waltham, MA) 90 minutes prior to dissection. Large intestines were dissected, fixed in 4% PFA overnight, washed in 1xPBS, and embedded in paraffin. Hematoxylin and eosin staining was performed on sections of the proximal and distal colon. The following primary antibodies and dilutions were used for immunostaining: rabbit anti-CD3 1:200 (catalog no. RM-9107; Lab vision; Fremont, CA), rat anti-CD45R 1:1000 (catalog no. 550786; Pharmingen; San Diego, CA), rat anti-BrdU 1:1000 (catalog no. OBT 0030; Accurate Chemical and Scientific Corp.; Westborg, NY), rabbit anti-mRELM-β 1:1000 (provided by Dr. Gary Wu—see He et al., 2003). The antibodies were diluted in PBT (1X PBS with 2% bovine serum albumin and 0.1% Triton X-100). The slides were incubated with the primary antibody overnight at 4°C, washed in 1X PBS, incubated with biotinylated anti-goat, anti-rabbit or anti-rat antibodies, as appropriate, (1:200; Vector Laboratories; Burlingame, CA) for 30 minutes at 37°C, and then washed in 1X PBS. The slides were incubated with ABC reagent (VECTASTAIN ABC Kit; Vector Laboratories; Burlingame, CA) for 30 minutes at 37°C, washed with 1X PBS, and signals developed using the DAB Peroxidase (HRP) Substrate Kit (with Nickel), 3,3’-diaminobenzidine (Vector Laboratories; Burlingame, CA), before counterstaining of the sections with hematoxylin. Prism software (GraphPad Software; San Diego, CA, USA) was used for all statistical analyses of histological experiments.
Chromatin immunoprecipitation (ChIP), ChIP-sequencing, and analysis

Chromatin preparation and chromatin immunoprecipitation were performed as described previously (Ye, 2009). Briefly, chromatin was prepared from the isolated colonic crypts of four wild type C57BL/6 male mice. From each chromatin preparation, two libraries were constructed using 10µg of chromatin and 2µg of anti-rabbit FOXA1 antibody (ab23738; Abcam; Cambridge, United Kingdom) or 2µg FOXA2 antibody (sc-6554; Santa Cruz Biotechnology; Dallas, TX), respectively. The immunoprecipitated DNA was modified according to the Illumina protocol (http://www.illumina.com), subjected to 100-bp sequencing analysis on an Illumina HiSeq 2500 System, and aligned to the mouse genome (mm9).

HOMER software (Heinz, 2010) in ‘factor mode’ was used to call Foxa1 and Foxa2 peaks. Peaks were selected at a false discovery rate (FDR) of 1%, cumulative Poisson p-value of 0.0001, and at least 4-fold greater coverage of reads in the target experiment compared with the background input reads. A union set of FOXA1 and FOXA2 peaks was constructed using Bioconductor software’s DiffBind package (Ross-Innes, 2012; Stark, 2011). Regions that were called in at least two of the eight ChIP samples were selected into the union set.

Genes associated with peaks were identified using GREAT (version 3.0.0) with default settings (Mclean, 2010). To assess whether DNA-binding sites of other DNA-binding proteins were enriched in the union set of FOXA1/A2-bound chromatin regions, motif analysis was performed using MEME-ChIP (version 4.11.1) (Machanick, 2011; Bailey, 2015). The file used for motif analysis was formatted using the MEME-ChIP workflow on Galaxy (usegalaxy.org) (Blankenburg, 2010; Giardine, 2005; Goecks, 2010).

The Bioconductor software package ChIPseeker (Yu, 2015) was used to plot the frequency of Foxa1/A2 target loci within ±3000kb of transcription start sites. To determine the enrichment of genomic regions bound by FOXA1/A2, ChIPseeker was used to annotate the union set of FOXA1/A2-bound loci and the background input loci.

RNA-Seq library preparation and analysis
Total RNA was purified as described above. Libraries were prepared from proximal whole colons of control (n=3, 14-23 months) and mutant (n=3, 7-18 months) mice. Poly(A) selection was carried out using the NEBNext Poly(A) mRNA Magnetic Isolation Module (New England Biolabs; Ipswich, MA). The NEBNext Ultra RNA Library Prep Kit for Illumina was used to complete library preparation. The libraries were multiplexed and subjected to 100-bp sequencing analysis on an Illumina HiSeq 2500 System.

Raw sequenced reads were filtered to retain only high quality reads. Ribosomal reads and repeats were eliminated. Remaining reads were processed with RNA-Seq Unified Mapper (RUM) (Grant, 2011), which aligns reads to the set of known transcripts included in RefSeq, UCSC known genes, VEGA transcripts, and the genome, and then outputs feature-level quantitation (transcript, exon, and intron). To analyze global gene expression profiles, the number of uniquely aligning read counts to mRNA transcripts in RefSeq and UCSC genes were extracted from the RUM output. Pairwise comparisons between groups were carried using a custom script that implemented the Bioconductor software package *edgeR* (Robinson, 2010) to compute a p-value and fold-change for each transcript. The p-values were corrected for multiple testing using the Benjamini & Hochberg mode of the R function *p.adjust* to compute a false discovery rate (FDR). The data were summarized for individual genes by selecting a representative transcript with the highest read counts.

In order to identify physiologically relevant targets of FOXA1/A2, the FOXA1/A2 ChIP-seq and RNA-seq data were integrated using R (R Core Team, 2016). The probability of the overlap between differentially expressed genes identified by RNA-seq and peak-associated genes identified by GREAT was calculated using the hypergeometric distribution function *phyper* in R. The disease and functional annotation and prediction upstream regulators of differentially expressed genes with an FDR < 5% that were also identified as peak-associated genes by GREAT (above) was performed using QIAGEN's Ingenuity Pathway Analysis (IPA®, QIAGEN Redwood City, CA; https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/) (Krämer et al., 2014).
The FOXA1/A2 ChIP-seq data were compared to publicly available colonic epithelial HNF4A ChIP-seq data (GEO series GSM1266727) (Chahar, 2014). HNF4A peak-associated genes were identified by GREAT and genes targeted by both FOXA1/A2 and HNF4A were identified using online Bioinformatics and Research Computing software from the Whitehead Institute (http://jura.wi.mit.edu/bio/). The probability of the overlap between FOXA1/A2 and HNF4A target genes using the hypergeometric distribution function \textit{phyper} in R.
CHAPTER 3

The RNA Polymerase III Subunit Polr3b is required for the maintenance of small intestinal crypts in mice.

Parts of this chapter were adapted from Kieckhaefer et al. 2016. The RNA Polymerase III Subunit Polr3b is required for the maintenance of small intestinal crypts in mice. Cell Mol Gastroenterol Hepatol. 2: 783–795.
ABSTRACT

Background & aims: The continuously self-renewing mammalian intestinal epithelium, with high cellular turnover, depends on adequate protein synthesis for its proliferative capacity. RNA polymerase III activity is closely related to cellular growth and proliferation. Here, we studied the role of Polr3b, a large RNA polymerase III subunit, in the mammalian intestinal epithelium.

Methods: We derived mice with an intestinal epithelium-specific hypomorphic mutation of the Polr3b gene, using VillinCre-mediated gene ablation. Phenotypic consequences of the Polr3b mutation on the intestinal epithelium in mice were assessed using histological and molecular methodologies.

Results: The Polr3b mutation severely reduced survival and growth in mice during the first postnatal week, the period when the expansion of the intestinal epithelium, and thus the requirement for protein synthesis, are highest. The neonatal intestinal epithelium of Polr3b<sup>loxP/loxP;VillinCre</sup> mice was characterized by areas with reduced proliferation, abnormal epithelial architecture, lack of crypts, and impaired Paneth cell differentiation; all accompanied by reduced Wnt signaling activity in these regions. In addition, enteroids cultured from Polr3b<sup>loxP/loxP;VillinCre</sup> mice show reduced proliferative activity and increased apoptosis.

Conclusions: We provide evidence for an essential role of the Pol III subunit Polr3b in orchestrating the establishment of the intestinal epithelium during early postnatal development in mice.
INTRODUCTION

The small intestinal epithelium is the fastest self-renewing tissue in mammals, and thus highly dependent on adequate protein synthesis for proliferation, especially during early postnatal development, when the expansion of the epithelial layer takes place. In neonatal mice, intervillus regions of the intestinal epithelium reorganize to give rise to highly proliferative crypt structures, which increase in number and size during this transition from neonatal to adult function (De Santa Barbara et al., 2003).

The RNA polymerase III (Pol III) complex transcribes small noncoding RNAs (sncRNAs), such as transfer RNAs (tRNAs), U6 snRNAs, and 5S ribosomal RNAs (rRNAs) (White, 2008), which all are involved directly or indirectly in protein synthesis. Studies in Saccharomyces cerevisiae have shown many factors and mechanisms that regulate Pol III transcription (Dumay-Odelot et al., 2010). These studies clearly show that high Pol III activity is required during conditions of rapid growth, whereas reduction in Pol III activity is observed during nutrient deprivation (Roberts et al., 2003). Although the Pol III complex is highly conserved from yeast to human beings, unlike in yeast, Pol III targets are repressed differentially in response to nutrient deprivation in human fibroblasts, indicating that mechanisms of transcriptional regulation of Pol III genes are more complex in mammals (Orioli et al., 2016).

Although Pol III traditionally has been regarded as functioning in a housekeeping capacity, studies in metazoans have shown sophisticated mechanisms of transcriptional regulation of Pol III targets, including cell type–specific transcription of different tRNA genes (White, 2011). In mammalian cells, numerous studies have pointed to Pol III dysregulation as playing a role in oncogenic transformation (Grewal, 2014; White, 2011). In support of this finding, overexpression of Pol III transcripts has been described in multiple cancers and in transformed cell lines (Chen et al., 1997; Pavon-Eternod et al., 2009; Singh et al., 1985; Winter et al., 2000). The proto-oncogene proteins mechanistic target of rapamycin complex 1 (mTORC1), myelocytomatosis oncogene (MYC), p53, and retinoblastoma all directly regulate transcription by Pol III in human cells (Chesnokov et al., 1996; Gomez-Roman et al., 2003; Michels et al., 2010;
White et al., 1996). Thus, comprehending the function and regulation of Pol III is vital to the understanding of tissue growth in health and disease.

Despite the importance of Pol III, little is known about its requirement during development. In zebrafish, a large-scale mutagenesis screen recovered a mutation, termed *slim jim* (*slj*), that causes an in-frame deletion in the gene encoding the second-largest Pol III subunit, Polr3b (Yee et al., 2007). The mutation has a profound effect on the proliferation and growth of intestinal epithelial progenitor cells, consistent with a high demand for protein synthesis in the rapidly proliferating intestinal epithelium.

The continuously self-renewing mammalian intestinal epithelium, with high cellular turnover, provides an attractive system to study the role of the Pol III complex in mammalian cell proliferation. Based on data from previous studies in the *slj* zebrafish, we hypothesized that Polr3b might play an important role in the growth and development of the mammalian gut. To study this hypothesis, we generated a conditional hypomorphic mutation of *Polr3b* specifically in the mouse intestinal epithelium.

Here, we show that the disruption of the Pol III complex in the intestinal epithelium by *Polr3b* mutation leads to a severe developmental phenotype, mainly characterized by increased mortality, accompanied by reduced proliferative and differentiation capacity, impaired crypt maturation, and reduced Wnt signaling in the intestinal epithelium. In summary, our data show the important role of RNA polymerase III in regulating the establishment of the intestinal epithelium in mammals.
RESULTS

Polr3b$^{loxP/loxP}$, VillinCre Mice Have Reduced Survival and Growth

To study the contribution of the RNA polymerase III complex in the intestinal epithelium of mice, we derived mice with a conditional hypomorphic mutation of the Polr3b gene, which encodes Polr3b, a major subunit of Pol III (Figure 3.1A). The Polr3b mutation resulted in major mortality, with only 60% of Polr3b$^{loxP/loxP}$, VillinCre mice (hereafter referred to as Polr3b-mutants) surviving the first postnatal week (Figure 1B). Surviving mutant mice lagged behind their control littermates in growth until the third postnatal week (Figure 3.1B and C). Though there was no difference in intestinal length when normalized to body weight (data not shown), early postnatal mutant mice were significantly smaller and had shorter intestines than their littermate controls (Figure 1C–E). As expected, before birth (E18.5), the intestinal tissue of Polr3b$^{loxP/loxP}$, VillinCre mice already showed a significant reduction in the expression levels of the 5S and U6 RNAs, key components of the ribosome and splicing machinery, respectively, compared with control (Polr3b$^{loxP/loxP}$) mice (Figure 3.1F). Milk was present in the stomach of Polr3b-mutant mice from the first postnatal day until 3 weeks of age (data not shown), suggesting that Polr3b-mutant mice fed normally.

Newborn Polr3b-Mutant Mice Show Reduced Proliferation and Impaired Crypt Development

Given the slower growth rate of mutant animals and the significant reduction in the mutant epithelium of RNAs required for translation, we hypothesized that there were defects in proliferation in the mutant intestine. Therefore, we performed immunostaining for Ki67, a protein that marks the transit-amplifying cells of the intestinal epithelium. Polr3b$^{loxP/loxP}$, VillinCre intestine had clear regions lacking Ki67-positive cells at P7, P14, and P21 (Figure 3.2D, F, and H), which was indicative of reduced proliferative capacity in the intestinal epithelium of young Polr3b-mutant mice. In addition, P14 mutant intestinal epithelial areas with reduced numbers of Ki67-positive cells were lacking fully developed crypts, which normally are present by this stage in control mice (Figure 2E vs F). In regions lacking Ki67, crypts still were not present by P21 in Polr3b-mutant mice (Figure 2H). At P21, the areas with reduced numbers of Ki67-positive cells were surrounded
by hyperproliferative regions that had increased numbers of Ki67-positive cells (Figure 3.2H), suggesting regenerative mechanisms were activated in regions with surviving crypts.

**Mutant Mice Develop Abnormal Intestinal Epithelial Architecture**

Morphologically, the intestinal epithelium of young Polr3b-mutant mice was affected severely, showing abnormal crypt and villus architecture (Figure 3.3A). At earlier developmental stages (E17.5 and P1), no obvious morphologic defects were detected in mutants (data not shown and Figure 3.3A, respectively). By P7, epithelial defects were present in all regions of the mutant small intestine and were characterized by an expanded mesenchymal layer, lack of developing crypts, and branched villus-like structures (Figure 3.3A). At weaning (P21), the intestinal epithelium of Polr3b-mutant mice continued to show similar morphologic changes to those seen at P7. In addition, the P21 epithelium contained areas with dramatic hyperplasia (Figure 3.3A). In contrast to younger animals, a large fraction of the epithelium of adult mutant mice had normal morphology. However, regions of disorganized epithelium with branched villi and crypt expansion still were detected in adult Polr3b-mutants (Figure 3.3A).

**Undeleted Cre Escaper Cells Proliferate in Hyperplastic Regions of Mutant Intestinal Epithelium**

The regenerative response, indicated by the presence of hyperplastic regions and increased proliferation of the intestinal epithelium seen in P21 Polr3b-mutant mice (Figure 3.3A), could be explained by stem cells that had escaped Cre-mediated gene ablation. Such incomplete gene deletion has been reported previously in other studies using the Cre system in the mouse intestinal epithelium (de Lau et al., 2011). Epithelial stem and transit-amplifying cells that escape Cre-mediated mutation, and thus have retained normal Polr3b function, would be expected to have a significant growth advantage over mutant cells.

To assess the presence of undeleted Cre escaper cells in Polr3b-mutant mice in vivo, we performed laser capture microdissection to dissect crypts from control mice and from hyperplastic regions in Polr3b-mutant mice (Figure 3.3B). Subsequently, we PCR-amplified the region of the Polr3b locus encoding the loxP sites to detect undeleted and deleted Polr3b-mutant alleles.
Indeed, the presence of the undeleted \textit{Polr3b} allele in the hyperplastic areas of the intestinal epithelium of \textit{Polr3b}\textsuperscript{loxP/loxP}; \textit{VillinCre} mice was confirmed by PCR assay using 2 sets of primers (Figure 3.3C and D). These data indicate the presence of wild-type escaper cells with the undeleted allele in the \textit{Polr3b}-mutant intestinal epithelium, which likely were driving the regenerative response described earlier.

To determine the contribution of apoptosis to the loss of developing crypts in the \textit{Pol3b}-mutant intestinal epithelium, we performed TUNEL assays on P3 (Figure 3.4A and B) and P21 (Figure 3.4C and D) intestines. Because of the mosaicism in Cre activity observed in the mutant mouse, to distinguish between escaper regions and recombined (mutant) regions in the epithelium, we crossed our \textit{Polr3b}\textsuperscript{loxP/loxP}; \textit{VillinCre} mice with \textit{R26-stop-EYFP} mice. Thus, in regions where Cre was efficient, the epithelium expressed EYFP as detected by immunostaining (Figure 3.4A and B). In the control epithelium at P3, no apoptotic cells were detected (Figure 3.4A), while in the mutant epithelium, apoptotic cells were present in EYFP-positive crypts, but not in EYFP-negative cells (Figure 3.4B). At P21, apoptotic cells were detected in the mesenchyme, but not the epithelium, of control mice (Figure 3.4C), but apoptotic crypts with TUNEL-positive cells were detected in the mutant epithelium (Figure 3.4D). A blinded count of the TUNEL-positive cells per crypt in P21 intestine showed that there were significantly more apoptotic cells in the mutant intestinal epithelium than in the control epithelium (control, 0.2 TUNEL-positive cells/crypt, \(N = 4\); mutant, 1.2 TUNEL-positive cells/crypt, \(N = 4\); \(P = .0012\)) (Figure 3.4E).

\textit{Undeleted Cre Escaper Cell Populations Expand Over Time In Vitro}

We hypothesized that if stem cells had escaped Cre-mediated gene ablation, a progressive increase would result in the proportion of cells carrying the undeleted allele as the epithelium expands during development. Therefore, to directly measure changes in the abundance of epithelial cells expressing undeleted \textit{Polr3b}\textsuperscript{loxP/loxP} over time, and without the potential contamination of mesenchymal cells, we harvested epithelial tissue from newborn control and \textit{Polr3b}-mutant mice and cultured it in vitro into enteroids (Figure 3.5; experimental
scheme shown in Figure 3.5A). Indeed, the fraction of the undeleted allele increased significantly between days 2 and 6 in culture (P < .05), initially comprising 18% on day 2 in culture, and increasing to 46% on day 6, as shown by qPCR-based quantification of the undeleted allele shown in Figure 3.5B.

To determine whether the defects in proliferation observed in vivo in mutant epithelium persisted when the epithelium was cultured in vitro, we immunostained enteroids for Ki67. Although all of the control enteroids formed Ki67-positive crypt-like domains (Figure 3.5C), indicating the presence of stem cells and transit-amplifying cells, mutant enteroids developed into 2 distinctly different structures: enteroids that resembled controls (Figure 3.5D), and enteroids that formed dilated, cyst-like structures (Figure 3.5E). Although the cyst-like enteroids expressed Ki67, they failed to form crypt-like domains.

To assess cell death in vitro, after 6 days in culture we stained enteroids for activated caspase-3 (Figure 3.5F and G). Control enteroids showed no apoptotic cells within the epithelial layer, caspase-3 staining was present only in dead cells shed into the enteroid lumen (Figure 3.5F). In contrast, as we also observed by TUNEL staining in vivo, Polr3b-deficient enteroids showed a high frequency of apoptotic cells within the enteroid epithelium itself (Figure 3.5G). This is consistent with the hypothesis that the intestinal epithelium is being repopulated by cells expressing the undeleted allele in response to the increased rate of apoptosis in mutant cells.

**Over Time, Polr3b^-loxP/loxP;VillinCre Crypts Are Replaced by Highly Proliferative Cre-Escaper Cells**

**In Vivo**

To confirm that the regions of disrupted proliferation observed in Polr3b-mutant intestine were in tissue where Cre was active, while regions of hyperproliferation were derived from Cre-escapers, we performed Ki67 immunostaining on epithelia from Rosa26-lox-stop-lox-EYFP;VillinCre (control) and Polr3b^-loxP/loxP;Rosa26-lox-stop-lox-EYFP;VillinCre (mutant) P3 mice (described earlier). As expected, controls uniformly expressed Ki67 at the base of each intervillus region (Figure 3.6C). In contrast, in regions of mutant epithelium where Cre was active, as indicated by EYFP staining, very little Ki67 was detected, whereas in EYFP-negative Cre-escaper
regions, intervillus regions were expanded and contained dense clusters of Ki67-positive cells (Figure 3.6D, yellow bracket).

The nonproliferative, Ki67-negative regions interspersed with hyperproliferative crypts observed in the Polr3b-mutant small intestine resemble small intestinal epithelium that regenerates after exposure to ionizing radiation. The regenerative response in injured intestine is driven by Wnt signaling (Clevers et al., 2014). Indeed, even in irradiated crypts in which Lgr5+ stem cells have been ablated, constitutive activation of Wnt by ablation of Apc is sufficient to drive crypt hyperplasia (Metcalfe et al., 2014). Therefore, to determine whether Wnt pathway activation contributed to proliferation in escaper regions of Polr3b-deficient gut, we performed immunostaining for the Wnt target cyclin D1.

At P3, immunohistochemical staining showed that in controls, cyclin D1 was expressed uniformly in the intervillus regions (Figure 3.6E), whereas mutant regions (EYFP-positive regions) of the Polr3b-mutant epithelium lacked cyclin D1-positive cells (Figure 3.6F). In contrast, EYFP-negative regions of the P3 Polr3b-mutant epithelium (highlighted by a yellow bracket), showed dense clusters of cyclin D1-positive cells (Figure 3.6F). Lack of active Wnt signaling in the Polr3b-deficient small intestine was confirmed by staining for b-catenin. As shown in Figure 3.6G, in control tissue, intervillus regions show strong cytoplasmic and nuclear b-catenin staining, indicative of Wnt pathway activation. In striking contrast, most of the P3 Polr3b-mutant epithelium is b-catenin negative, with the exception of an area of regenerative epithelium bracketed in yellow (Figure 3.6H).
DISCUSSION

Previous studies in yeast and zebrafish have shown the importance of the RNA Pol III complex in cell growth and tissue development. More specifically, the zebrafish mutant slj, with a mutation in the second largest Pol III subunit Polr3b, causes defects in digestive tract development.\textsuperscript{16} Here, we present a model to study the role of Pol III in mammalian intestinal epithelial development and proliferation. We show that decreased activity of the Pol III complex in the mouse intestinal epithelium leads to increased mortality of Polr3b-mutant mice in the first postnatal week. Surviving Polr3b-mutant mice show reduced proliferation, impaired crypt development, and loss of active Wnt signaling. This initial phenotype is followed by a regenerative response around weaning, which is driven by cells that escape Cre-mediated mutation (Figure 3.7A). Because of the diminished Pol III activity caused by the Polr3b mutation, and resulting low levels of 5S rRNA and U6 snRNA, the high demand for protein synthesis during early intestinal epithelial tissue expansion cannot be met. These events result in impaired crypt development and increased apoptosis in crypts. These data show the important role of Polr3b and the Pol III complex in the establishment of the intestinal epithelium during early postnatal life in mammals (Figure 3.7B). The conditional hypomorphic Polr3b-mutant mice described earlier will allow a detailed investigation of the contribution of RNA polymerase III in all tissue and cell types for which a Cre line is available.

The epithelium of the small intestine is the most rapidly self-renewing mammalian tissue, and thus it is required to be capable of modulating protein synthesis in response to environmental stress or during tissue expansion/growth. During periods of starvation, the gut responds by decreasing villus size, crypt size, and the rate of cell cycling (Altmann, 1972). Protein malnutrition alone also results in histologic abnormalities, such as villi fused at the base accompanied by villus blunting, or even the absence of villus architecture (Tandon et al., 1968). The Pol III complex plays an essential role in protein synthesis in all eukaryotic cells (White, 2004). Therefore, sufficient Pol III activity is predicted to be essential for intestinal epithelial function and survival.

The simplest explanation of the disrupted development of the small intestine of Polr3b mutants is that the reduced Pol III activity impairs cell proliferation because the high metabolic
demands of the epithelium cannot be met. Indeed, the mutant intestine shows several similarities
to both fasted and protein-starved small intestines: a decrease in cycling cells, abnormal
branched villus architecture, and increased apoptosis in epithelial cells (Altmann, 1972; Boza et
al., 1999; Tandon et al., 1968; White, 2004). The Polr3b mutant intestine thus mimics the
malnourished state because the levels of 5S rRNA and U6 snRNA, which are required for protein
synthesis, are decreased as a result of low Pol III activity. Defects in protein production
machinery thus have a phenotype reminiscent of amino acid deficiency, because in both cases
the high demand for protein synthesis during early intestinal epithelial tissue expansion cannot be
met.

Yeast and zebrafish studies have suggested that the structural effects of Polr3b
mutations resulting in impaired transcription by Pol III are caused by the destabilization of the
interaction between the Polr3b protein and another Pol III subunit, Polr3k (Landrieux et al., 2006;
Yee et al., 2007). The yeast homolog of Polr3k is required for Pol III to efficiently perform multiple
rounds of transcription on the same gene, which determines the overall output of the complex
(Flores et al., 1999). Thus, the Polr3b mutation in mice results in a reduction in the overall
abundance of Pol III transcripts, but without a complete loss of transcription by Pol III, as also was
evident by the reduction, but not elimination, of Pol III transcripts in the mutant intestine we
documented in Figure 1.

Not surprisingly, rapidly dividing cancer cells have an extremely high demand for protein
synthesis, and tRNA abundance is increased in many cancers such as multiple myeloma and
breast cancer (Pavon-Eternod et al., 2009; Zhou et al., 2009). Multiple oncogenic (extracellular
signal-regulated kinase (ERK), TORC1, and MYC) and tumor suppressor (p53, MAF1,
retinoblastoma) pathways regulate RNA polymerase III activity and thus tRNA abundance
(reviewed by (White, 2008) and (Grewal, 2014)). Given the development of small-molecule
inhibitors of RNA polymerase III, reducing tRNA abundance has been suggested as a novel
targeted approach for cancer therapy. Indeed, in a recent report, Yee et al. used the RNA
polymerase III inhibitor ML-60218 developed by Wu et al., 2003 in combination with the histone
deacetylase inhibitor suberoylanilide hydroxamic acid to reduce proliferation in two human
pancreatic adenocarcinoma cell lines (Yee et al., 2012). As noted by the investigators, the safety of ML-60218 in human beings has not yet been evaluated. Given the striking dependency of the mammalian intestinal epithelium on normal rates of RNA polymerase III we showed earlier, it is questionable if globally acting, small-molecule inhibitors of this enzyme can be developed into successful drugs to treat cancer.
**Figure 3.1**

The Polr3b mutation in mice leads to reduced survival, delayed growth, and decreased intestinal epithelial transcription of snRNAs. (A) Targeting strategy employed for conditional ablation of exon 10 (E10) of the Polr3b gene. After germ line transmission, the FRT-flanked neomycin resistance gene (PGK-Neo cassette) was excised by crossing to Flp deleter mice (Flp). Polr3b<sup>loxP/loxP</sup>;VillinCre mice were derived by intercrossing Polr3b<sup>loxP/+</sup> with VillinCre mice. (B) Kaplan-Meier survival curve of control and mutant mice during the first eight postnatal days, showing a significantly lower survival rate in Polr3b-mutant mice compared to littermate
controls (Log-rank (Mantel-Cox) and Gehan-Breslow-Wilcoxon Test; p<0.0001, n=12-28). (C) Significantly reduced body weight during the pre-weaning period in Polr3b-mutant mice (mean ± SEM; *p<0.05; n=3-47). During weaning, when the regenerative response is observed, Polr3b-mutant mice start to catch up in body weight with their littermate controls. (D) Polr3b-mutant mice (right) are visibly smaller than their control littermates (left), at P3. (E) P1 control and mutant stomach, large intestine, and colon. (F) Real-time qPCR analysis of expression of tRNA-Arg, tRNA-Asn, tRNA-Cys, tRNA-His, 5S rRNA, and U6 snRNA in E18.5 control (n=7) and mutant (n=10) mice (mean ± SEM; *p= 0.0496, Student’s t-test, **p=0.002, Mann Whitney test). Data were normalized to the mean of snoRNA 202 and snoRNA 234.
**Figure 3.2**

Proliferation is reduced in neonatal Polr3b<sup>loxP/loxP</sup>;VillinCre small intestine, followed by sporadic hyperplasia and enhanced proliferative activity during weaning. Ki67 immunohistochemistry of control and Polr3b-mutant mice at P1 (A and B), P7 (C and D), P14 (E and F), and P21 (G and H). At all stages, Polr3b-mutant mice show jejunal regions lacking Ki67-positive cells (indicated by arrows). At P21, the epithelium of Polr3b-mutant mice displays hyperproliferative regions with expanded crypts and an increased number of Ki67-positive cells.
Figure 3.3: Hyperplastic regions of the Polr3b-mutant intestinal epithelium at P21 contain wild type ‘Cre-escaper’ cells with the undeleted Polr3b allele. (A) H&E staining of the intestinal epithelium showing the morphology of the jejunal epithelium of control and mutant mice at P1 (20X magnification), P7 (10X magnification), P21 (20X magnification), and adult stage (20X magnification). Areas with expanded mesenchyme, villus branching, and lack of developing crypts are present in P7 Polr3b-mutant epithelium (yellow box). At P21, hyperplastic areas are clearly visible in mutant mice (yellow box). In adults, the intestinal epithelium of Polr3b-mutant mice shows improved morphology; however, areas with villus branching are still detected (yellow box). (B) Representative pictures before and after laser capture microdissection (LCM) of normal
crypts dissected from control and hyperplastic regions from P21 mutant mice. (C) Scheme of primer design and PCR product size for amplification using different primer sets for LCM-PCR analysis. Primers F and R1 were designed surrounding the loxP sites flanking exon 10 of the Polr3b gene resulting in the amplification of the longer, undeleted Polr3b PCR product (1,018 bp) or the shorter, mutant Polr3b gene PCR product (337 bp). An additional reverse primer (R2) was designed within the loxP-flanked region of the Polr3b gene, which when used with the forward (F) primer, results in the amplification of the undeleted Pol3b PCR product only (530 bp). (D) PCR results for LCM-dissected regions from control crypts (control) and Polr3b-mutant crypts from hyperplastic regions (mutant HR), showing the presence of the undeleted Polr3b allele in control crypts, and both deleted as well as undeleted alleles in Polr3b-mutant crypts from hyperplastic regions using both primer combinations (F+R1 and F+R2).
**Figure 3.4:** The Polr3b mutation is associated with crypt cell death. Apoptotic cells (red) in the crypts were analyzed by performing the TUNEL assay on intestinal epithelial tissue at P3 (A and B) and P21 (weaning) (C and D). (A and B) Green staining represents YFP, which is visible only in Cre-expressing tissue. Red = TUNEL, white = E-cadherin, and blue = DAPI. For increased clarity, the inset in B, which is a 3x enlargement of the dying crypt indicated by the yellow arrow, shows only YFP and TUNEL staining. (C and D) Green = E-cadherin, red = TUNEL, and blue = DAPI. During weaning (P21), TUNEL-positive crypts were still not detected in control tissue (C), but were seen in mutant animals (D). The inset in D is a 2x enlargement of the dying crypt indicated by the yellow arrow. No TUNEL-positive cells were detected in crypts of control mice (A and C), though apoptotic cells were detected in the mesenchyme. Crypt epithelial cells positive
for TUNEL staining were present in Polr3b-mutant intestinal epithelium in neonates and during weaning (B and D). TUNEL-positive epithelial cells were only present in the tissue where cre was active in mutant animals (B). (E) Apoptosis is significantly increased in mutant epithelium at P21. Cell death was quantified by counting the number of TUNEL-positive cells per crypt in jejunal sections from P21 mice (control, n = 4; mutant, n = 4; *p = 0.0012, Student's t-test).
**Figure 3.5**: Increased cell death is accompanied by expansion of ‘Cre-escaper’ cells in mutant epithelium *in vitro*. (A and B) Cleaved caspase 3 (red) staining of enteroids from P1 gut on day 6 in culture (blue = DAPI). Increased cell death was detected in Polr3b-mutant enteroids compared to controls. In control enteroids (A), cell death was detected only in cells shed into the lumen, while in mutant enteroids (B), cell death was also seen in peripheral cells (yellow arrows). Images are at 40x magnification. Scale bar = 500 μM. (C, D, and F) Ki67 staining of enteroids from P1 gut on day 4 in culture. Mutant enteroids developed into two distinctly different structures: enteroids
with Ki67 expression at the tips of crypt-like buds (F), resembling controls (C), and dilated, cyst-like structures with patches of Ki67 expression surrounded by flattened Ki67-negative cells (D).

(E) Scheme for the culture and harvesting of P1 intestinal enteroids. (F) Quantification of undeleted Polr3b exon 10 in genomic DNA from enteroids after 2, 4, and 6 days in culture, as shown in (E) (Error bars indicate SEM; *p < 0.05, paired, one-tailed Student’s t-test; control, n=5; mutants, n=6).
Figure 3.6: Hyperproliferative Cre-escaper cells repopulate the Polr3b-mutant intestinal epithelium. *Rosa26-lox-stop-lox-YFP* mice were crossed with *Polr3b<sup>LoxP/LoxP</sup>;*VillinCre* mice to show which sections of the tissue are *Polr3b* mutant (YFP-positive), and which are *Polr3b* wild-type Cre-escaper cells (YFP-negative). (A and B) Red, E-cadherin; green, EYFP; and blue, 4′,6-diamidino-2-phenylindole. (A) *Rosa26-lox-stop-lox-YFP;VillinCre* (control) jejunum. (B) *Polr3b<sup>LoxP/LoxP</sup>; Rosa26-lox-stop-lox-YFP;VillinCre* (mutant) jejunum. The yellow bracket indicates a Cre-escaper region that is negative for YFP. (C and D) Red, Ki67; green,
EYFP; and blue, 4′,6-diamidino-2-phenylindole. (D) Crypts in the Cre-escaper region (yellow bracket) of the mutant intestine have (C) a higher density of Ki67-positive cells than crypts in the control intestine. (D) Almost no Ki67 was visible in Polr3b mutant tissue outside the bracketed area. (E and F) Red, cyclin D1; green, EYFP; and blue, 4′,6-diamidino-2-phenylindole. (F) In Polr3b-mutant mice, expression of the Wnt target protein cyclin D1 is reduced in mutant tissue compared with the (E) control intestine, whereas escaper regions have dramatically increased cyclin D1 staining (yellow bracket). (G and H) Brown, β-catenin; red, nuclear fast red counterstain. (G) Rosa26-lox-stop-lox-YFP; VillinCre (control) jejunum. (H) Polr3bLoxP/LoxP; Rosa26-lox-stop-lox-YFP; VillinCre (mutant) jejunum. The yellow bracket indicates a Cre-escaper region that has active Wnt signaling, as indicated by nuclear β-catenin staining.
Figure 3.7: Summary of the effects of the Polr3b mutation on the mouse intestinal epithelial development and maintenance. (A) Right after birth, large areas lacking proliferative capacity are present (Newborn (P3)). During weaning (P21), regions lacking proliferative capacity are present with very low or absent Wnt-pathway activity, consistent with affected crypt architecture. By P21 the affected regions in the Polr3b-mutant epithelium are surrounded by hyperplastic regions, which are most likely initiated by the wild type ‘escaper’ stem cells that have escaped Cre recombination of the Polr3b gene. These regions are repopulating the intestinal epithelium, resulting in normalization of epithelial morphology in Polr3b-mutant intestinal epithelium in adults. (B) The mutation in Polr3b leads to reduced Pol III activity, which leads to reduced transcription of small noncoding RNAs required for translation results in increased crypt death, reduced Wnt signaling, decreased proliferative activity, impaired crypt maturation, and altered epithelial architecture in mutant tissue. The increased death of mutant cells creates a selective advantage for Cre-escaper cells in the intestinal epithelium of Polr3b-mutant mice.
CHAPTER 4

The FOXA Transcription Factors Sustain Murine Intestinal Health by Regulating Crohn’s Disease Risk Loci
ABSTRACT

While the primary cause of intestinal inflammation in IBD is thought to be an inappropriate immune response to luminal antigens, genome-wide association studies (GWAS) have also implicated multiple genetic loci associated with genes expressed in the intestinal epithelium itself. Defects in the mucus layer, which is required for limiting the exposure of epithelial cells to pathogens, result in inflammation. It was previously shown that mucin expression is altered and Mucin2 (Muc2), the main component of intestinal mucus, is downregulated by deletion of the paralogous transcription factors FOXA1 and FOXA2 in the mouse intestinal epithelium. Mice deficient in Muc2 develop intestinal inflammation and colorectal cancer, suggesting a potential role for the FOXA factors in the maintenance of an intact mucus layer and intestinal epithelial health. To investigate the requirement of Foxa1/a2 in the maintenance of intestinal barrier function, I analyzed immune cell infiltration into the epithelium of the Foxa1/a2-deficient colon in mice of advanced age (>7 months). Immunohistochemical staining showed that the Foxa1/a2-deficient colonic epithelium displays increased susceptibility to spontaneous colitis. To analyze the mechanism of increased susceptibility to colitis in Foxa1/a2 mutants, I performed FOXA1 and FOXA2 chromatin immunoprecipitation assays followed by ultra-high throughput sequencing (ChIP-Seq) analysis on wild type mouse colonic crypts followed by RNA-seq on Foxa1/a2-deficient and control aged colon to identify relevant FOXA targets. Significantly enriched FOXA1/A2 binding sites were identified in close proximity to the IBD risk gene Fut2, which was downregulated in mutant colon. Additionally, FOXA1/A2 binding sites were found upstream of the glycosyltransferase genes B3gnt6 and Gcnt3 and downstream of the gene for the secreted immune effector RELM-β. These findings suggest that FOXA factors regulate modulators of barrier function and that downregulation of these genes may contribute to IBD pathogenesis.
INTRODUCTION

Inflammatory bowel disease (IBD), which includes Crohn’s disease (CD) and ulcerative colitis (UC), currently affects 3.1 million people in the US (Dahlhamer et al., 2016). Genetic causes play an important role in the pathogenesis of IBD (Abraham and Cho, 2009), but inheritance does not follow simple Mendelian laws, suggesting that multiple genetic factors and environmental triggers are involved. Genome-wide association studies (GWAS) have identified single-nucleotide polymorphisms associated with increased susceptibility to IBD (Franke et al., 2010; Jostins et al., 2012; Liu et al., 2015). Among these loci are the epithelially-expressed genes ‘autophagy-related, 16-like’ (ATG16L1), ‘nucleotide-binding oligomerization domain 2’ (NOD2), and ‘immunity-related GTPase’ (IRGM), all part of the autophagosome pathway (Rioux et al., 2007). Autophagy is responsible for the degradation of cellular constituents and the digestion of pathogens (“xenophagy”), including bacteria, viruses, and parasites, and the activation of the innate and adaptive immunity pathways (Levine and Deretic, 2007).

An important barrier between the gut epithelium and lumenal pathogens is the protective mucus layer on the apical surface of the intestinal epithelium. Notably, MUC19, a member of the gel-forming mucin protein family, was identified as a potential IBD risk gene (Franke et al., 2010). This reinforces the notion that in addition to autophagy, mucus, which is produced and secreted by epithelial goblet cells and which limits exposure of epithelial cells to pathogens, is vital for maintaining intestinal health. Consequently, the association of these intestinal epithelial genes with human IBD emphasizes that physiological adaptive and innate immune responses, as well as an intact intestinal epithelial cell and mucus layer, are critical for the manifestation of IBD.

As previously shown, mucus-secreting goblet cell differentiation and Mucin2 (Muc2) transcription are regulated by the winged-helix transcription factors FOXA1 and FOXA2 (Ye and Kaestner, 2009). Deletion of Foxa1 and Foxa2 in the mouse intestinal epithelium impairs secretory cell differentiation, and results in altered mucin expression, including Muc2 downregulation. Interestingly, the Muc2 gene is a direct FOXA1/A2 target (van der Sluis et al., 2008; Ye and Kaestner, 2009) and Muc2-deficient mice (Muc2<sup>−/−</sup>) develop intestinal inflammation.
and colorectal cancer (Van der Sluis et al., 2006; Velcich, 2002), suggesting a potential role for the FOXA factors in the maintenance of an intact mucus layer and intestinal epithelial health.

Here, I show that simultaneous deletion of Foxa1 and Foxa2 in the gastrointestinal epithelium of mice leads to increased susceptibility to spontaneous colitis. FOXA1 and FOXA2 chromatin immunoprecipitation followed by ultra-high throughput sequencing (ChIP-Seq analysis) of the colon revealed FOXA1 and FOXA2 binding sites associated with multiple genes that influence intestinal health, including the ortholog of the human IBD-associated gene ‘fucosyltransferase 2’ (Fut2). The predisposition to colitis in Foxa1/Foxa2-deficient mice and the regulatory role of FOXA factors in human IBD-associated genes suggest that additional FOXA-regulated factors may contribute to human IBD pathogenesis.
RESULTS

Aged Foxa1/a2 mutant mice spontaneously develop colitis-like features.

While examining mice deficient for Foxa1/a2 in the intestinal epithelium (Foxa1<sup>LoxP/LoxP</sup>, Foxa2<sup>LoxP/LoxP</sup> VillinCre) (Ye, 2009), I noticed that aged mutant mice spontaneously develop several features of colitis. Upon dissection, I observed a distended and fragile colon in Foxa1/a2 mutants.

To determine whether the altered morphology of Foxa1/a2-deficient colon was caused by inflammation, I stained for CD45R, a B-lymphocyte marker, and CD3, a T-lymphocyte marker (Figure 4.1B and C, respectively). The dramatically increased number of CD45R<sup>+</sup> lymphoid nodules in aged Foxa1/a2 mutants compared to controls demonstrates that spontaneous inflammation is increased in the aged mutant colonic epithelium. This finding was accompanied by a large increase in the number of T-lymphocytes in the lamina propria.

Using a defined set of criteria (Kim et al., 2012), blinded assessment by a veterinary pathologist of colon sections stained with hematoxylin and eosin confirmed my observations of the spontaneous colitis phenotype. Although 8-week-old mutants did not have significantly increased damage scores, the assessment revealed that the aged mutant colonic epithelium displayed distorted crypt architecture, increased immune cell infiltration, crypt abscesses, and goblet cell depletion, resulting in significantly increased total histological damage scores when compared to aged littermate controls (Figure 4.1A and 1D).

Alcian blue staining of mucopolysaccharides highlighted the altered goblet cell morphology in Foxa1/a2-deficient colons. The goblet cells in aged mutants displayed smaller and fewer mucin granules than age-matched controls (Figure 4.2A), consistent with the previous observations in the small intestine (Ye and Kaestner, 2009) and histological features of IBD (Baumgart and Sandborn, 2012; Danese and Fiocchi, 2011). In addition, by measuring BrdU incorporation into epithelial cells in S-phase of the cell cycle, I observed a significant increase in epithelial proliferation in the Foxa1/a2-deficient colon, which is reminiscent of observations made previously in the Muc2<sup>−/−</sup> colon (Van der Sluis et al., 2006; Velcich, 2002) (Figure 4.2B and 2C).
FOXA1 and FOXA2 bind genes orthologous to IBD risk loci in colonic crypt epithelium.

To elucidate the molecular mechanisms underlying the spontaneous colitis in aged Foxa1/a2 mutants, I performed FOXA1 and FOXA2 ChIP-seq analysis on purified crypts from wild type mouse colon to identify relevant FOXA targets. Due to the high percentage of sequence identity between the two target proteins, especially in the DNA binding domains (Clark, 1993), I first validated that the antibodies I used to immunoprecipitate FOXA1 and FOXA2 were monospecific and performed ChIP-qPCR for the known FOXA liver target Cdx2. At the Cdx2 enhancer in Foxa1-null liver lysates, binding by the antibody against FOXA1 was depleted (Figure 4.3A). Likewise, binding by the antibody against FOXA2 was depleted in Foxa2-null liver lysates (Figure 4.3B). Thus, the two antibodies specifically bind their respective target proteins.

An important question in transcriptional regulation by paralogous DNA binding proteins such as FOXA1 and FOXA2 is whether they occupy the same, overlapping, or distinct binding sites in chromatin. Early work measuring the in vitro in DNA binding affinities had suggested distinct preferences for DNA sequence motifs for FOXA1 and FOXA2, then known as Hnf3α and Hnf3β, respectively (Lai et al., 1990). In addition, ChIP-seq experiments on murine liver had found partially overlapping target sets for FOXA1 and FOXA2 with hundreds of genes uniquely bound by FOXA1 (Bochkis et al., 2012).

Therefore, I compared the two FOXA cistromes by plotting the normalized reads per million values (RPM) for each peak called by HOMER in the union of the FOXA1 and the FOXA2 ChIP-seq datasets (Figure 4.4A). Linear regression of the plotted points in Figure 4.4A indicates the overall binding strength of Foxa2 is marginally stronger than that of Foxa1 (y = 0.27 + 1.1x, r² = 0.828). Remarkably, the plot shows that the two FOXA factors have the same relative binding frequency at thousands of target loci. Thus, in colonic crypts FOXA1 and FOXA2 bind to the same set of target genes in colonic crypts, providing a molecular explanation for the lack of intestinal phenotypes in the single gene conditional gene ablation models (Ye and Kaestner, 2009).

Motif analysis of the union FOXA1/A2 ChIP data by MEME-ChIP identified a de novo binding motif (E = 2.2x10^{-638}) that is nearly identical to both the previously published FOXA1 and
the FOXA2 canonical binding motifs (FOXA1, \( p = 2.32 \times 10^{-9} \); FOXA2, \( p = 1.71 \times 10^{-15} \)) (Figure 4.4B) (Tuteja et al., 2009; Zhang et al., 2008). Therefore, the regions identified by FOXA1/A2 binding in my ChIP-seq data represent cis-regulatory sites in the colonic epithelium.

The binding of FOXA1/A2 in colonic crypts is enriched for sites close to transcriptional start sites (Figures 4.4C and D), which suggests that FOXA factors regulate transcription by binding both enhancers and promoters. Analysis of the genomic binding distribution of FOXA1/A2-bound regions (Figure 4.4D) shows that there is an enrichment of FOXA binding in promoter regions relative to the frequency of promoters in the genome. This is in contrast to the liver, where FOXA binding occurs predominantly at distal enhancers (Bochkis et al., 2012).

In addition to the enriched FOXA motif identified in the union dataset, analysis of sequences within the FOXA-bound regions also identified significant enrichment of an HNF4A-like motif (\( p = 2.45 \times 10^{-6} \)) (Figure 4.4E). Intersection of FOXA1/A2 and HNF4A (Chahar et al., 2014) target genes in the colonic epithelium revealed a significant overlap (Figures 4.4F, 4.5, and 4.6A). Regulatory regions co-occupied by FOXA2 and HNF4A have previously been identified in the fetal and the mature liver (Alder et al., 2014). These data support the idea that FOXA1/A2 and HNF4A also co-regulate a set of target genes in the colon.

**FOXA targets in the colon regulate barrier function and inflammatory response**

To identify physiologically relevant targets of FOXA in the colonic epithelium, I performed transcriptomic analysis of control and Foxa1/a2-deficient colon. Integration of these data with my ChIP-seq results revealed 86 genes that were both differentially expressed (FDR < 0.05) and also bound by FOXA1/A2 (Table 4.1). Of these genes, 67 were also bound by HNF4A, including Muc2, Fut2, B3gnt6, Gcnt3, and Tlr1 (discussed below). In addition to identifying an enrichment of genes associated with different types of epithelial cancer, functional and disease annotation by Ingenuity Pathway Analysis (Krämer et al., 2014) (Table 4.2) of differentially expressed, FOXA1/A2-bound genes identified an enrichment of genes involved in maintaining the morphology of the mucosa, as well as genes associated with the inflammatory response. Within the subset of mucosa-associated genes, I identified the gene for the main component of intestinal
mucus, Muc2. Consistent with the direct binding by FOXA1/A2 to the Muc2 promoter in vitro (van der Sluis et al., 2008), I observed not only that FOXA1/A2 binds the Muc2 promoter in vivo, but also that in the Foxa1/a2-null colon Muc2 steady state transcription levels are significantly decreased (Figure 4.5A and B).

Notably, genes required for post-translational glycosylation of MUC2, including Fut2, ‘glucosaminyl (N-acetyl) transferase 3, mucin type’ (Gcnt3), and the gene for the upstream enzyme of GCNT3, ‘UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 6’ (B3gnt6), are also downregulated FOXA1/A2 targets (Figure 4.5C-D). While B3GNT6 and GCNT3 are required for the enzymatic addition of O-type glycans to mucin proteins (Figure 1.1), which form the core structures of the extensive carbohydrate branches of MUC2 (Honke and Taniguchi, 2002; Iwai et al., 2002), the FUT2 enzyme is responsible for further modifying glycosylated proteins on the mucosa and in bodily secretions (Becker and Lowe, 2003). The heavily glycosylated MUC2 protein constitutes the main barrier in the colon between the epithelium and commensal gut bacteria (Bergstrom and Xia, 2013).

In concert with mucus barrier genes, the transcription of FOXA1/A2 target genes involved in host defense against bacteria is also significantly reduced in Foxa1/a2 mutants. The most downregulated gene is defensin alpha 6 (Defa6) (Table 4.1), which encodes a secreted antimicrobial peptide (Jarczak et al., 2013). Regulators of intestinal inflammation, including Toll-like receptor 1 (Tlr1) and Resistin-like β (Retnlb or Relm-β), are also FOXA1/A2 targets with reduced expression in mutants (Figure 4.6). The pattern recognition receptor TLR1 heterodimerizes with TLR2 to detect bacterial lipopeptides and initiate an immune response (Farhat et al., 2008). RELM-β is secreted by goblet cells in the colon in response to bacterial colonization (He et al., 2003) and is protective against TNBS-induced inflammation (Krimi et al., 2008; Steppan et al., 2001). In addition to significantly decreased levels of Tlr1 and Relm-β mRNA (Figure 4.6A-D), goblet cell RELM-β protein is also considerably depleted in FOXA1/A2 deficient colon of aged mice (Figure 4.6E).

Prediction of control nodes for differentially expressed FOXA1/A2 targets by Ingenuity Pathway Analysis (Krämer et al., 2014) identified IL10RA as an inhibited upstream regulator.
Notably, IL10RA is a subunit of the receptor for the cytokine IL-10, and mice that are IL-10-deficient develop chronic enterocolitis in response to bacterial colonization of the colon (Berg et al., 1996; Kuhn et al., 1993). Thus, Foxa1/a2 ablation and disruption of IL-10 signaling may result in the activation of similar proinflammatory pathways. Together, these data suggest that targets of FOXA1/A2 in the colon regulate barrier function by maintaining the mucus layer and modulating the mucosal inflammatory response to enteric pathogens.
DISCUSSION

Aged Foxa1/a2 mutant colon shares several features with the Muc2−/− colon: a loss of goblet cells, an increase in inflammatory cell infiltration, and altered crypt morphology. However, as Muc2 is downregulated, but not completely absent from the Foxa1/a2-deficient colon, it is not surprising that I did not observe the colorectal cancer phenotype that was seen in the aged Muc2-null colon (Van der Sluis et al., 2006; Velcich, 2002). The Kaestner lab had previously found a decrease in the number of goblet cells and reduced Muc2 mRNA levels in 8-week-old Foxa1/a2-deficient intestine, but did not detect an increase in inflammation or in the number of BrdU+ cells (Ye and Kaestner, 2009). Conversely, as in Muc2−/− mice (Van der Sluis et al., 2006; Velcich, 2002), proliferation was significantly increased in the colonic epithelium of our aged Foxa1/a2 mutants.

There are two possible explanations for the discrepancy between young and aged Foxa1/a2 mutants. First, the increase in inflammation and crypt proliferation seen in the aged Foxa1/a2-deficient colon could be a secondary effect due to the decreased goblet cell number and mucin secretion that begins early in life. The reduction in Muc2 mRNA levels, and therefore decreased lubrication of the intestine, could cause microlesions, possibly resulting in impaired barrier function and the increased inflammation observed in the aged mutant mice. Second, loss of Foxa1/a2 could directly cause dysregulation of different pathways controlling mucosal health, thus resulting in impaired barrier function, which can be compensated for in very young mice, but becomes apparent only at an older age. Several findings, which I will discuss in detail below, support the latter explanation.

Although microarray analysis of wild type and Muc2−/− colonic epithelium by Yang and colleagues revealed that Muc2−/− colon displays a chronic subclinical inflammation phenotype, it did not identify expression changes in glycosylation-related genes (Yang et al., 2008). This suggests that the expression changes in our Foxa1/a2 mutants are not solely due to reduced expression of Muc2. Moreover, expression of RELM-β is actually increased in Muc2−/− colon (Yang et al., 2008), as opposed to the downregulation of RELM-β observed in Foxa1/a2-null
colon (Figure 4.6C-E). Together, these data establish a distinctive role for the *Foxa1/a2* transcription factors in the pathogenesis of colitis.

Remarkably, my findings also establish a regulatory role for the FOXA factors in the transcription of the human IBD susceptibility locus *Fut2* (Franke et al., 2010; Jostins et al., 2012; McGovern et al., 2010). The *FUT2*-associated SNP rs516246 was identified by GWAS as an IBD risk locus (Franke et al., 2010; Jostins et al., 2012). The same SNP was later established as an IBD-associated expression quantitative trait locus (Singh, 2015). FUT2 is necessary for the fucosylation of epithelial glycoproteins, which are the main components of the intestinal mucus layer (Becker and Lowe, 2003; Wang et al., 2016). Optimal interactions between epithelial glycoproteins and the luminal microbiome ensure colonization of the intestine with physiological microflora, and therefore contribute to mucosal defense (Bergstrom and Xia, 2013; Hooper, 2004; Patsos and Corfield, 2009). Interestingly, epithelial fucosylation of glycoproteins is induced in mice treated with DSS (Terahara et al., 2011) as well as with toll-like receptor ligands (Pickard et al., 2014), supporting the importance of fucosylation for maintaining or restoring the mucosal barrier. My ChIP-seq data and the *Fut2*-downregulation in *Foxa1/a2*-mutants (Figure 4.5C and D) demonstrate that FOXA factors control *Fut2* and suggest that decreased fucosylation of intestinal glycoconjugates contributes to impaired host-microbe crosstalk, concomitant with colonization of the intestine by pathogenic flora and promotion of colitis.

In addition to FOXA1/A2 activating *Fut2* expression, my data also suggest that FOXA1/A2 co-regulates thousands of genes in concert with HNF4A. Given the requirement of HNF4A for the development of the colonic epithelium and for the maturation of goblet cells (Garrison et al., 2006), FOXA1/A2 likely cooperates with HNF4A in the regulation of genes required for goblet cell differentiation and mucosal barrier maintenance. Furthermore, *Hnf4α* has also been implicated by GWAS as an IBD risk locus (Barrett et al., 2009), and expression of HNF4A is protective against inflammation (Chahar et al., 2014). Thus, inflammation in the *Foxa1/a2*-null colonic epithelium may result from the loss of cooperative gene regulation by FOXA1/A2 and HNF4A.
In summary, deficiency of Foxa1/a2 in the gastrointestinal epithelium predisposes mice to colitis, with the FOXA factors acting as central regulators of key genes influencing intestinal epithelial barrier function (Figure 4.5). I provide a mechanistic explanation for the predisposition to colitis seen in aged Foxa1/a2-mutant mice and link the observed phenotype to two genes, Fut2 and Hnf4α, for which single-nucleotide polymorphisms have been identified in GWAS for human IBD. This highlights the role of FOXA1 and FOXA2 in intestinal epithelial health, and suggests that the FOXA factors may play a major role in human IBD pathogenesis.
Figure 4.1: Aged Foxa1/a2-null colon develops features of colitis. (A) Histological analysis of aged control and Foxa1/a2 mutants shows spontaneous colitis in the absence of Foxa1/a2. Hematoxylin and eosin (H&E) stained sections show increased crypt distortion, inflammatory cell
infiltration, and goblet cell depletion. (B) CD45R staining for B-lymphocytes of the distal colon (20x) reveals an increased number of lymphoid nodules in the Foxa1/a2 mutants. (C) CD3-staining for T-lymphocytes of the distal colon (4x) shows occasional T-lymphocyte infiltration in the control, but increased intraepithelial T-lymphocytes in the Foxa1/a2 mutants. (D) Blinded histological damage scoring (maximum possible = 11) of H&E sections by a veterinary pathologist reveals that colonic inflammation in aged mutants is significantly more severe than in age matched controls. Each data point represents one mouse. (p = 0.0369).
**Figure 4.2**

**A** Alcian blue staining of mucopolysaccharides in control and Foxa1/a2 mutant colon of mice aged 14 months shows decreased mucin production in aged Foxa1/a2 mutants (20x). **B** BrdU staining of control and Foxa1/a2 mutant colon of mice aged 14 months (20x). **C** BrdU-positive cells per crypt in proximal colon of mice aged 7-14 months. Increased BrdU incorporation in aged Foxa1/a2 mutant epithelium indicates increased proliferation. Each data point represents average BrdU-positive proximal colonic crypts per mouse. (control n = 4, mutant n = 5; *p = 0.0317, Mann-Whitney test).

**Figure 4.2: Foxa1/a2 mutant colonic epithelium exhibits significantly increased cell cycling as well as altered differentiation.** (A) Alcian blue staining of mucopolysaccharides in control and Foxa1/a2 mutant colon of mice aged 14 months shows decreased mucin production in aged Foxa1/a2 mutants (20x) (B) BrdU staining of control and Foxa1/a2 mutant colon of mice aged 14 months (20x). (C) Brdu-positive cells per crypt in proximal colon of mice aged 7-14 months. Increased BrdU incorporation in aged Foxa1/a2 mutant epithelium indicates increased proliferation. Each data point represents average BrdU-positive proximal colonic crypts per mouse. (control n = 4, mutant n = 5; *p = 0.0317, Mann-Whitney test).
The antibodies directed against FOXA1 (ab23738) and FOXA2 (sc-6554), which were used to perform ChIP-seq, do not cross-react. The antibodies were validated by using each of them to perform ChIP in (A) Foxa1-null liver lysates and (B) Foxa2-null liver lysates followed by qPCR for the FOXA liver target Cdx2. Binding of the Cdx2 promoter by the FOXA antibodies was depleted in liver lysates lacking their respective target proteins.
Figure 4.4: FOXA1 and FOXA2 bind overlapping targets in the colonic epithelium. (A) Scatter-plot of average normalized peak height in reads per million (RPM) of 4,702 regions from the union of FOXA1 and FOXA2 ChIP-seq peaks. Diagonal blue line marks y = x. Gray error bars indicate SEM. (B) De novo motif analysis of regions bound by both FOXA1 and FOXA2 identified the
Foxa consensus site as being the most enriched ($p = 2.2x10^{-638}$). The de novo motif had high identity to both the FOXA1 and the FOXA2 canonical motifs (FOXA1, $p = 2.32x10^{-9}$; FOXA2, $p = 1.71x10^{-16}$). (C) Frequency of FOXA1/A2-bound regions binned by distance to TSS. (D) Enrichment of FOXA1/A2 binding at genomic regions compared to the frequency of the region in the genome. (E) De novo motif analysis of Foxa1/A2-bound regions identified an HNF4A-like motif as being significantly enriched ($p = 2.45x10^{-6}$). (F) There is significant overlap between the genes identified as being FOXA targets and the genes identified as being HNF4A targets in the colon, which may indicate co-regulation by the two proteins. ($p = 2.34x10^{-279}$, hypergeometric test).
Table 4.1: Differentially expressed FOXA1/A2 target genes (FDR<0.05) in mutant colonic epithelium

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>UCSC Transcript ID</th>
<th>Log2(Mut. RPM/Cntrl. RPM)</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Defa6</td>
<td>uc009bix.2</td>
<td>-6.88</td>
<td>1.37E-02</td>
</tr>
<tr>
<td>Sycn</td>
<td>uc009fzj.2</td>
<td>-6.30</td>
<td>1.30E-32</td>
</tr>
<tr>
<td>Insr</td>
<td>uc008psy.2</td>
<td>-4.89</td>
<td>1.62E-10</td>
</tr>
<tr>
<td>Rgs17</td>
<td>uc007egm.2</td>
<td>-4.61</td>
<td>4.50E-07</td>
</tr>
<tr>
<td>Pltx2</td>
<td>uc008rhx.1</td>
<td>-4.42</td>
<td>1.15E-02</td>
</tr>
<tr>
<td>Fut9</td>
<td>uc008sei.1</td>
<td>-4.07</td>
<td>1.47E-05</td>
</tr>
<tr>
<td>Sult1c2</td>
<td>uc008cuu.1</td>
<td>-4.01</td>
<td>6.40E-10</td>
</tr>
<tr>
<td>Aqp4</td>
<td>uc008edr.2</td>
<td>-3.97</td>
<td>8.04E-06</td>
</tr>
<tr>
<td>Cubn</td>
<td>uc008ijz.1</td>
<td>-3.76</td>
<td>8.92E-09</td>
</tr>
<tr>
<td>Reg4</td>
<td>uc008aqq.1</td>
<td>-3.50</td>
<td>6.60E-05</td>
</tr>
<tr>
<td>A4gnt</td>
<td>uc009rej.2</td>
<td>-3.47</td>
<td>3.53E-07</td>
</tr>
<tr>
<td>Foxa1</td>
<td>uc007nps.1</td>
<td>-3.46</td>
<td>4.95E-21</td>
</tr>
<tr>
<td>Itih2</td>
<td>uc008ihq.2</td>
<td>-3.44</td>
<td>3.82E-02</td>
</tr>
<tr>
<td>Nov</td>
<td>uc007vm.1</td>
<td>-3.23</td>
<td>1.82E-02</td>
</tr>
<tr>
<td>Retnlib (Relm-β)</td>
<td>uc007zju.1</td>
<td>-3.18</td>
<td>9.53E-05</td>
</tr>
<tr>
<td>Capn8</td>
<td>uc011xwv.1</td>
<td>-3.08</td>
<td>1.75E-03</td>
</tr>
<tr>
<td>Car2</td>
<td>uc008oqt.2</td>
<td>-3.05</td>
<td>2.99E-10</td>
</tr>
<tr>
<td>Chga</td>
<td>uc007oui.1</td>
<td>-2.87</td>
<td>1.56E-12</td>
</tr>
<tr>
<td>Ephx4</td>
<td>uc008yme.1</td>
<td>-2.83</td>
<td>1.72E-03</td>
</tr>
<tr>
<td>2810459M11Rik</td>
<td>uc011wog.1</td>
<td>-2.80</td>
<td>1.30E-12</td>
</tr>
<tr>
<td>Samd5</td>
<td>uc007ew.1</td>
<td>-2.75</td>
<td>9.85E-03</td>
</tr>
<tr>
<td>Gna14</td>
<td>uc008gju.2</td>
<td>-2.61</td>
<td>6.31E-03</td>
</tr>
<tr>
<td>Phd1</td>
<td>uc007akv.1</td>
<td>-2.61</td>
<td>9.23E-04</td>
</tr>
<tr>
<td>Car8</td>
<td>uc008rjuv.1</td>
<td>-2.59</td>
<td>2.09E-10</td>
</tr>
<tr>
<td>Mmp15</td>
<td>uc009myh.1</td>
<td>-2.55</td>
<td>4.45E-09</td>
</tr>
<tr>
<td>Adams15</td>
<td>uc009eqz.1</td>
<td>-2.51</td>
<td>1.13E-03</td>
</tr>
<tr>
<td>Cyp4b1</td>
<td>uc008fc.1</td>
<td>-2.49</td>
<td>2.09E-10</td>
</tr>
<tr>
<td>Slc45a3</td>
<td>uc007cnz.1</td>
<td>-2.45</td>
<td>4.19E-07</td>
</tr>
<tr>
<td>B3gnt6</td>
<td>uc009ikf.1</td>
<td>-2.41</td>
<td>1.46E-03</td>
</tr>
<tr>
<td>Pla2g10</td>
<td>uc007ygi.1</td>
<td>-2.22</td>
<td>1.64E-05</td>
</tr>
<tr>
<td>Pllp</td>
<td>uc009mww.1</td>
<td>-2.20</td>
<td>1.28E-06</td>
</tr>
<tr>
<td>Aldh111</td>
<td>uc009cym.1</td>
<td>-2.10</td>
<td>1.55E-03</td>
</tr>
<tr>
<td>Paqr8</td>
<td>uc007ald.1</td>
<td>-2.10</td>
<td>2.87E-02</td>
</tr>
<tr>
<td>Kcne3</td>
<td>uc009imp.2</td>
<td>-2.06</td>
<td>2.11E-07</td>
</tr>
<tr>
<td>Ak4</td>
<td>uc008tvo.2</td>
<td>-2.01</td>
<td>2.22E-02</td>
</tr>
<tr>
<td>Gstm3</td>
<td>uc008xv.1</td>
<td>-2.00</td>
<td>2.19E-02</td>
</tr>
<tr>
<td>Mbof</td>
<td>uc009ssb.2</td>
<td>-1.94</td>
<td>8.18E-04</td>
</tr>
<tr>
<td>Ccl6</td>
<td>uc007kp.1</td>
<td>-1.93</td>
<td>2.30E-05</td>
</tr>
<tr>
<td>2210407C18Rik</td>
<td>uc007jw.2</td>
<td>-1.76</td>
<td>5.25E-03</td>
</tr>
<tr>
<td>Arhgap20</td>
<td>uc008pl.1</td>
<td>-1.74</td>
<td>3.02E-04</td>
</tr>
<tr>
<td>Ptprn2</td>
<td>uc007phx.2</td>
<td>-1.72</td>
<td>6.99E-03</td>
</tr>
<tr>
<td>Mkx</td>
<td>uc008zep.2</td>
<td>-1.72</td>
<td>3.19E-03</td>
</tr>
<tr>
<td>Mpst</td>
<td>uc007wpf.2</td>
<td>-1.69</td>
<td>2.59E-02</td>
</tr>
<tr>
<td>At2c2</td>
<td>uc009nqg.1</td>
<td>-1.64</td>
<td>2.17E-02</td>
</tr>
<tr>
<td>Ramp1</td>
<td>uc007bzx.2</td>
<td>-1.58</td>
<td>5.43E-03</td>
</tr>
<tr>
<td>Igdc4</td>
<td>uc012gyg.1</td>
<td>-1.57</td>
<td>2.29E-03</td>
</tr>
<tr>
<td>Gcnt3</td>
<td>uc009qnt.1</td>
<td>-1.55</td>
<td>8.26E-04</td>
</tr>
<tr>
<td>Cd177</td>
<td>uc009fl.1</td>
<td>-1.55</td>
<td>1.69E-02</td>
</tr>
<tr>
<td>Bace2</td>
<td>uc008ade.2</td>
<td>-1.54</td>
<td>1.58E-03</td>
</tr>
<tr>
<td>Muc2</td>
<td>uc009km.2</td>
<td>-1.53</td>
<td>1.72E-02</td>
</tr>
<tr>
<td>Rnase4</td>
<td>uc007tmo.2</td>
<td>-1.51</td>
<td>2.19E-03</td>
</tr>
<tr>
<td>Nupr1</td>
<td>uc007jsd.1</td>
<td>-1.51</td>
<td>8.69E-03</td>
</tr>
<tr>
<td>Gene</td>
<td>Accession</td>
<td>Fold Change</td>
<td>p-Value</td>
</tr>
<tr>
<td>--------</td>
<td>-----------</td>
<td>-------------</td>
<td>---------</td>
</tr>
<tr>
<td>Hsd17b2</td>
<td>uc009npf.1</td>
<td>-1.47</td>
<td>3.58E-03</td>
</tr>
<tr>
<td>Gmds</td>
<td>uc007pzq.1</td>
<td>-1.38</td>
<td>7.76E-03</td>
</tr>
<tr>
<td>Pde8b</td>
<td>uc007rmd.2</td>
<td>-1.35</td>
<td>3.33E-03</td>
</tr>
<tr>
<td>Fut2</td>
<td>uc009gwn.1</td>
<td>-1.35</td>
<td>1.75E-03</td>
</tr>
<tr>
<td>Ffar2</td>
<td>uc009ggg.1</td>
<td>-1.34</td>
<td>1.94E-02</td>
</tr>
<tr>
<td>Smepd3</td>
<td>uc009nfx.1</td>
<td>-1.26</td>
<td>7.05E-03</td>
</tr>
<tr>
<td>Wfdc2</td>
<td>uc008nve.1</td>
<td>-1.24</td>
<td>1.97E-02</td>
</tr>
<tr>
<td>Fam83e</td>
<td>uc009gwy.1</td>
<td>-1.22</td>
<td>1.12E-02</td>
</tr>
<tr>
<td>Sgsm3</td>
<td>uc007wwb.2</td>
<td>-1.22</td>
<td>3.41E-02</td>
</tr>
<tr>
<td>Tlr1</td>
<td>uc008xmr.1</td>
<td>-1.22</td>
<td>7.05E-03</td>
</tr>
<tr>
<td>Slco2a1</td>
<td>uc009rgd.1</td>
<td>-1.22</td>
<td>2.66E-02</td>
</tr>
<tr>
<td>Cyp2s1</td>
<td>uc009fua.1</td>
<td>-1.20</td>
<td>5.12E-03</td>
</tr>
<tr>
<td>Acsm3</td>
<td>uc009ijq.1</td>
<td>-1.18</td>
<td>3.82E-02</td>
</tr>
<tr>
<td>Sord</td>
<td>uc008maj.1</td>
<td>-1.13</td>
<td>3.61E-02</td>
</tr>
<tr>
<td>Far1</td>
<td>uc009jhp.1</td>
<td>-1.11</td>
<td>2.82E-02</td>
</tr>
<tr>
<td>Pim3</td>
<td>uc007seq.1</td>
<td>-1.04</td>
<td>4.21E-02</td>
</tr>
<tr>
<td>Cytip</td>
<td>uc008jsi.1</td>
<td>1.28</td>
<td>1.47E-02</td>
</tr>
<tr>
<td>Pou2af1</td>
<td>uc009pig.1</td>
<td>1.78</td>
<td>1.45E-02</td>
</tr>
<tr>
<td>Pdk4</td>
<td>uc009awk.1</td>
<td>1.86</td>
<td>1.78E-07</td>
</tr>
<tr>
<td>Gcnt1</td>
<td>uc008gxn.2</td>
<td>1.93</td>
<td>4.01E-05</td>
</tr>
<tr>
<td>Rdh12</td>
<td>uc011yog.1</td>
<td>2.26</td>
<td>3.82E-02</td>
</tr>
<tr>
<td>Maf</td>
<td>uc009nof.1</td>
<td>2.38</td>
<td>4.11E-06</td>
</tr>
<tr>
<td>Ccl25</td>
<td>uc009kw.2</td>
<td>2.45</td>
<td>5.02E-03</td>
</tr>
<tr>
<td>Ereg</td>
<td>uc008yls.1</td>
<td>2.70</td>
<td>4.59E-07</td>
</tr>
<tr>
<td>Trim29</td>
<td>uc009pbi.2</td>
<td>2.73</td>
<td>1.07E-02</td>
</tr>
<tr>
<td>Pla2g2f</td>
<td>uc008vlf.2</td>
<td>3.09</td>
<td>8.31E-03</td>
</tr>
<tr>
<td>Scl34a2</td>
<td>uc008xbl.2</td>
<td>3.10</td>
<td>6.40E-10</td>
</tr>
<tr>
<td>Cck</td>
<td>uc009sdg.2</td>
<td>3.22</td>
<td>1.45E-06</td>
</tr>
<tr>
<td>Nav3</td>
<td>uc007gzl.1</td>
<td>3.56</td>
<td>1.01E-03</td>
</tr>
<tr>
<td>Fabp7</td>
<td>uc007ict.1</td>
<td>3.77</td>
<td>2.03E-03</td>
</tr>
<tr>
<td>Pcdh17</td>
<td>uc007utp.2</td>
<td>3.94</td>
<td>1.75E-03</td>
</tr>
<tr>
<td>Pnoc</td>
<td>uc007uji.1</td>
<td>6.19</td>
<td>2.29E-06</td>
</tr>
<tr>
<td>Kdm5d</td>
<td>uc009uzb.1</td>
<td>7.68</td>
<td>3.31E-05</td>
</tr>
<tr>
<td>Eif2s3y</td>
<td>uc009uzd.1</td>
<td>9.18</td>
<td>3.37E-06</td>
</tr>
<tr>
<td>Categories</td>
<td>Diseases or Functions Annotation</td>
<td>P-Value</td>
<td>Molecules</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
<td>-------------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Cancer, Organismal Injury and Abnormalities, Reproductive System Disease</td>
<td>uterine serous papillary cancer</td>
<td>8.93E-05</td>
<td>CAR2, MAF, PDE8B, PLLP, RAMP1, RNASE4, WFDC2</td>
</tr>
<tr>
<td>Molecular Transport</td>
<td>transport of carbon dioxide</td>
<td>1.51E-04</td>
<td>AQP4, CAR2</td>
</tr>
<tr>
<td>Lipid Metabolism, Small Molecule Biochemistry</td>
<td>synthesis of fatty acid</td>
<td>1.51E-04</td>
<td>ACSM3, EREG, FAR1, FOXA1, PDK4, PLA2G10, PNOC, SLC45A3</td>
</tr>
<tr>
<td>Tissue Morphology</td>
<td>abnormal morphology of membrane tissue</td>
<td>1.75E-04</td>
<td>A4GNT, CCL25, GCNT1, GCNT3, MUC1, SLC45A2</td>
</tr>
<tr>
<td>Post-Translational Modification</td>
<td>glycosylation of protein</td>
<td>1.93E-04</td>
<td>A4GNT, FUT2, FUT9, GCNT1, GCNT3</td>
</tr>
<tr>
<td>Digestive System Development and Function, Organ Development</td>
<td>function of intestine</td>
<td>2.20E-04</td>
<td>KCNE3, MUC2, SLC34A2</td>
</tr>
<tr>
<td>Developmental Disorder, Hematological Disease, Hereditary Disorder,</td>
<td>hyperkalemic periodic paralysis</td>
<td>2.26E-04</td>
<td>CAR2, KCNE3</td>
</tr>
<tr>
<td>Metabolic Disease, Neurological Disease, Organismal Injury and Abnormalities, Skeletal and Muscular Disorders</td>
<td>female genital tract serous carcinoma</td>
<td>2.36E-04</td>
<td>CA2, MAF, NAV3, PDE8B, PDK4, PLLP, RAMP1, RNASE4, WFDC2</td>
</tr>
<tr>
<td>Lipid Metabolism, Small Molecule Biochemistry</td>
<td>fatty acid metabolism</td>
<td>2.70E-04</td>
<td>ACSM3, EREG, FABP7, FAR1, FOXA1, PDK4, PLA2G10, PNOC, SLC45A3, SLC45A3, SLC20A1, SMPD3</td>
</tr>
<tr>
<td>Auditory Disease, Dermatological Diseases and Conditions, Immunological Disease, Inflammatory Disease, Inflammatory Response, Organismal Injury and Abnormalities</td>
<td>contact dermatitis of ear</td>
<td>3.16E-04</td>
<td>CYTIP, PLA2G2F</td>
</tr>
<tr>
<td>Digestive System Development and Function, Gastrointestinal Disease, Organ Morphology, Organismal Development, Organismal Injury and Abnormalities, Tissue</td>
<td>abnormal morphology of intestinal mucosa</td>
<td>3.20E-04</td>
<td>GCNT1, GCNT3, MUC2</td>
</tr>
<tr>
<td>Category</td>
<td>Term</td>
<td>p-value</td>
<td>Gene Symbols</td>
</tr>
<tr>
<td>-------------------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>----------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Morphology</td>
<td>hyperplasia of stomach tissue</td>
<td>6.72E-04</td>
<td>A4GNT, KCNE3</td>
</tr>
<tr>
<td>Post-Translational Modification</td>
<td>O-glycosylation of protein</td>
<td>7.86E-04</td>
<td>A4GNT, GCNT1, GCNT3</td>
</tr>
<tr>
<td>Organismal Injury and Abnormalities</td>
<td>lithiasis</td>
<td>8.41E-04</td>
<td>CAR2, CUBN, GNA14, SLC34A2</td>
</tr>
<tr>
<td>Organismal Development</td>
<td>development of genitourinary system</td>
<td>9.70E-04</td>
<td>Elf2s3y, EREG, FABP7, GCNT1, GCNT3, HSD17B2, KDM5D, MMP15, MPST, NUPR1, PKHD1, PTPRN2, SMPD3, WFDC2</td>
</tr>
<tr>
<td>Inflammatory Response</td>
<td>inflammatory response</td>
<td>9.88E-04</td>
<td>CCL25, CHGA, GCNT1, KCNE3, MAF, MAOB, NOV, NUPR1, PLA2G10, PNOC, RETNLB, TLR1</td>
</tr>
<tr>
<td>Post-Translational Modification</td>
<td>O-mannosylation of protein</td>
<td>1.35E-03</td>
<td>GCNT1, GCNT3</td>
</tr>
<tr>
<td>Ophthalmic Disease, Organismal Injury and Abnormalities</td>
<td>cataract disease</td>
<td>1.37E-03</td>
<td>MAF, NOV, PITX2, SORD</td>
</tr>
<tr>
<td>Lipid Metabolism, Small Molecule Biochemistry</td>
<td>synthesis of lipid</td>
<td>1.42E-03</td>
<td>ACSM3, EREG, FAR1, FOXA1, HSD17B2, PDE8B, PKD4, PLA2G10, PNOC, SLC45A3, SMPD3</td>
</tr>
<tr>
<td>Amino Acid Metabolism</td>
<td>quantity of amino acids</td>
<td>1.56E-03</td>
<td>AQP4, CUBN, FUT2, PDK4, SMPD3</td>
</tr>
<tr>
<td>Carbohydrate Metabolism</td>
<td>metabolism of carbohydrate</td>
<td>1.62E-03</td>
<td>A4GNT, CCK, CHGA, FUT2, FUT9, GCNT3, GMDS, PLA2G10, RETNLB, SORD</td>
</tr>
<tr>
<td>Cell-To-Cell Signaling and Interaction, Molecular Transport, Small Molecule Biochemistry</td>
<td>quantity of monoamines</td>
<td>1.87E-03</td>
<td>AQP4, CHGA, MAOB, MPST, PTPRN2</td>
</tr>
<tr>
<td>Behavior</td>
<td>anxiety-like behavior</td>
<td>1.96E-03</td>
<td>CCK, PNOC, PTPRN2</td>
</tr>
<tr>
<td>Cancer, Organismal Injury and Abnormalities</td>
<td>serous adenocarcinoma</td>
<td>1.96E-03</td>
<td>CAR2, MAF, NAV3, PDE8B, PKD4, PLLP, RAMP1, RDH12, RNASE4, WFDC2</td>
</tr>
<tr>
<td>Organismal Development</td>
<td>morphology of body cavity</td>
<td>2.16E-03</td>
<td>A4GNT, CCK, CCL25, CHGA, GCNT1, GCNT3, HSD17B2, KCNE3, MMP15, MUC2, NOV, NUPR1, PITX2, PKHD1, RETNLB, SLC02A1, SMPD3</td>
</tr>
</tbody>
</table>
Figure 4.5: Genes involved in mucosal barrier function are bound by FOXA1/A2 and HNF4A. Genome browser images displaying FOXA1/A2 and HNF4A binding (blue), transcription in control colonic epithelium (green), and Foxa1/a2-null colonic epithelium (red) at the genomic loci of (A) Muc2, (C) Fut2, (E) B3gnt6, and (G) Gcnt3. (B, D, F, and H) Validation of RNA-seq transcript levels by RT-qPCR (*p < 0.05; Student’s t-test).
Figure 4.5: Genes involved in mucosal inflammatory response are bound by FOXA1/A2 and downregulated in Foxa1/a2-null colonic epithelium. (A) Genome browser images displaying FOXA1/A2 and HNF4A binding (blue), transcription in control colonic epithelium (green), and Foxa1/a2-null colonic epithelium (red) at the Tlr1 locus. (B) Validation of Tlr1 RNA-seq transcript levels by RT-qPCR. (C) Genome browser images displaying FOXA1/A2 binding (blue), transcription in control colonic epithelium (green), and Foxa1/a2-null colonic epithelium (red) at the Relm-β locus. Relm-β is not bound by HNF4A. (D) Validation of Relm-β RNA-seq transcript
levels by RT-qPCR. (E) RELM-β staining in control and Foxa1/a2 mutant colon of mice aged 14 months shows significant RELM-β depletion in mutant goblet cells (10x).
Table 4.3: Predicted upstream regulators of differentially expressed, FOXA1/A2-bound genes

<table>
<thead>
<tr>
<th>Upstream Regulator</th>
<th>Molecule Type</th>
<th>Predicted Activation State</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL10RA</td>
<td>transmembrane receptor</td>
<td>Inhibited</td>
<td>5.49E-08</td>
</tr>
<tr>
<td>HNF1A</td>
<td>transcription regulator</td>
<td>Inhibited</td>
<td>4.53E-04</td>
</tr>
<tr>
<td>miR-6742-3p (miRNAs w/seed CCUGGGU)</td>
<td>mature microrna</td>
<td>Activated</td>
<td>5.15E-04</td>
</tr>
<tr>
<td>tretinoin</td>
<td>chemical - endogenous mammalian</td>
<td>Inhibited</td>
<td>2.62E-03</td>
</tr>
<tr>
<td>miR-204-5p (and other miRNAs w/seed UCCUUUU)</td>
<td>mature microrna</td>
<td>Activated</td>
<td>6.29E-03</td>
</tr>
<tr>
<td>calcitriol</td>
<td>chemical drug</td>
<td>Inhibited</td>
<td>6.87E-03</td>
</tr>
<tr>
<td>miR-204-3p (and other miRNAs w/seed CUGGGAA)</td>
<td>mature microrna</td>
<td>Activated</td>
<td>1.15E-02</td>
</tr>
<tr>
<td>miR-92a-1-5p (miRNAs w/seed GGUUGGG)</td>
<td>mature microrna</td>
<td>Activated</td>
<td>2.88E-02</td>
</tr>
<tr>
<td>miR-3083-5p (and other miRNAs w/seed GGCUGGGG)</td>
<td>mature microrna</td>
<td>Activated</td>
<td>3.44E-02</td>
</tr>
</tbody>
</table>
CHAPTER 5

Conclusions and Future Directions
The incidence of IBD in Westernized countries has increased significantly during the last century. Factors contributing to this increase include improved diagnosis due to improved medical technology and access to healthcare, and a spread of the Westernized lifestyle. Notably, IBD incidence has increased Asia, the Middle East, and South America alongside the industrialization of those regions (Kaplan and Ng, 2016). Thus, IBD is a condition with ever-increasing worldwide significance. As IBD is a chronic condition with an increasing prevalence, the global burden of caring for an aging population of IBD patients is expected to increase even further in the coming decade (Kaplan and Ng, 2016). Early treatment of IBD is key to preventing disease progression and the development of complications (Colombel et al., 2016). Therefore, it is important to understand the pathogenesis of IBD in order to improve treatment outcomes and to prevent the devastating sequelae such as colorectal cancer.

The molecular pathogenesis of colitis in the Foxa1/a2-mutant mouse

The study of mouse models of colitis has led to a better understanding of the molecular pathogenesis of IBD and the role that the immune response has in etiology of the disease. Characterization of mice deficient for interleukin 10 (IL-10), a suppressor of macrophage activation (Ren et al., 1991), uncovered altered MUC2 expression in colitis. In response to exposure to bacteria, IL-10-deficient mice develop colitis and display reduced in MUC2 synthesis in comparison to controls, as well as reduced post-translational MUC2 sulfation (Schwerbrock et al., 2004). This discovery stimulated detailed analyses of mice homozygous for a Muc2-null allele in order to gain insight into the role of MUC2 in colonic barrier protection, as well as its possible role in IBD. The importance of MUC2 as a component of the barrier is illustrated by the fact that Muc2-mutant mice develop spontaneous intestinal inflammation that becomes exacerbated with age, and eventually leads to colorectal cancer (Van der Sluis et al., 2006; Velcich, 2002). Analysis of upstream regulators of Muc2 revealed that the Muc2 promoter is bound in vivo by the transcription factors FOXA1 and FOXA2 and that they are capable of activating the Muc2 promoter in vitro (van der Sluis et al., 2008; Ye and Kaestner, 2009).

In my thesis work, through cistromic and transcriptomic analyses, I have established that
in the mature colon, FOXA1 and FOXA2 direct transcription not only of Muc2, but of a network of genes required for the normal differentiation of secretory cells, the maintenance of the mucus layer, and for defense against microbial challenge. Like in the colon, FOXA1/A2 is required for the differentiation and function of other secretory organs, including the liver (Bochkis et al., 2012; Lee et al., 2005), pancreas (Gao et al., 2008; Kaestner et al., 1999), and lungs (Wan et al., 2004, 2005). However, I have also shown that, although FOXA is required in each of these developmentally related organs, FOXA1 and FOXA2 have distinct colon-specific functions. For example, while FOXA2 regulates expression of mucins in both the lungs and the colon, in the lungs FOXA2 directly inhibits transcription of the gel-forming mucin Mucin5ac (Muc5ac) (Wan et al., 2004), but in the colon FOXA2 activates expression of Muc2.

Furthermore, I have shown that genomic binding enrichment of FOXA1 and FOXA2 differs between the colon and other organs. In contrast to the colon, ChIP-seq experiments on adult murine liver revealed that FOXA1 and FOXA2 occupy an overlapping set of targets, but thousands of genes are uniquely bound by either FOXA1 or FOXA2. Additionally, most FOXA1/FOX2 co-occupied regions in the liver are at distal enhancers (Bochkis et al., 2012). Conversely, my ChIP-seq data demonstrate that in the colon, FOXA1 and FOXA2 share nearly identical sets of binding sites. Moreover, as opposed to in the mature liver, I have shown that FOXA binding is enriched at transcriptional start sites in the colon, which suggests that the FOXAs co-regulate transcription by binding both enhancers and promoters.

My data show that, in addition to Muc2, direct targets of FOXA1/A2 in the colon include the O-glycosyltransferase genes B3gnt6 and Gcnt3, and the fucosyltransferase gene Fut2. The Kaestner lab had previously reported that the number of goblet cells is decreased and Muc2 mRNA levels are reduced in the 8-week-old Foxa1/a2-deficient intestine (Ye and Kaestner, 2009). Likewise, I have shown in vivo that ablation of Foxa1/a2 results in a significant decrease of colon Muc2 steady state transcription levels in aged mutants.

Although the requirement of FOXA1/A2 for expression of MUC2 had already been established, I have expanded our understanding of the role of FOXA in the formation of the colonic mucus layer by identifying regulatory targets that are required for the post-translational
modification of mucins. Importantly, I have shown that transcript levels for B3gnt6, which is required to produce core 3-derived O-glycans, and for Gcnt3, which produces core 2- and core 4-derived O-glycans, are also significantly decreased in Foxa1/a2-null colon compared to age-matched controls. Notably, core 3-derived O-glycans are restricted primarily to the salivary glands and the intestinal tract (Bergstrom and Xia, 2013; Honke and Taniguchi, 2002; Iwai et al., 2002) and MUC2 O-glycan modifications are mostly derived from core 1 and core 3 structures (Bergstrom and Xia, 2013). Thus, adequate levels of B3GNT6 are imperative for the post-translational processing of MUC2.

Thus far, although the evidence is persuasive, I have not directly shown that binding by FOXA1/A2 at the promoters of my proposed target genes is required for their transcription. Furthermore, I have yet to show that these binding sites are required for target gene expression in humans. Given the potential for off-target effects in a physiologically artificial system like a luciferase reporter assay, which includes the possibility of increased aberrant promoter binding by FOXA to weak targets as a result of overexpression, I have instead opted to utilize the CRISPR/Cas9 system to specifically delete the FOXA binding motif at potential target genes in human colon cell lines. Given that many enhancers contain evolutionarily conserved sequences (Heintzman et al., 2009), and also that FOXA2 binding data are available in the Caco-2 colorectal adenocarcinoma cell line (Gosalia et al., 2015), I was able to identify FOXA2-bound enhancers in human cells that correspond to the FOXA1/A2 target regions I had mapped in mouse colonic crypts. Through the CRISPR/Cas9-mediated insertion of indels into these regions, I expect that I will be able to disrupt FOXA1/A2 binding, which will then result in the downregulation of target genes if indeed the FOXA proteins regulate these genes by direct binding to their promoters.

**Altered glycoprotein composition in FOXA1/A2-deficient colon**

The extensive O-glycan modifications on MUC2 give it a high capacity to bind water, which contributes substantially to its ability to form the viscoelastic gel bilayer in the colon. While the glycan modifications in the outer, less viscous layer of mucus serve as both an attachment site and an energy source for bacteria (Johansson et al., 2011), the viscous inner layer is
impenetrable to bacteria, and thus is maintained as a sterile environment between lumenal antigens and the epithelial surface (Johansson et al., 2008). Not surprisingly, mice homozygous for the B3gnt6-null allele, and consequently unable to produce core 3-derived O-glycans, also produce significantly less MUC2 protein. As I observed in FOXA1/A2-deficient colon, B3gnt6-mutants present with an inflammatory phenotype that includes increased inflammatory cell infiltration and an increased epithelial proliferation rate (An et al., 2007).

Compounding the potential deleterious effects of Muc2 and B3gnt6 downregulation in the aged Foxa1/a2 mutant colon, I have also shown that steady state transcript levels are significantly decreased for the FOXA binding target Fut2. FUT2 catalyzes the transfer of GDP-fucose, the final step of H blood group antigen synthesis, in salivary and other epithelial secretions (Kelly et al., 1995). Although the H cell surface antigen is associated with blood type, it also plays a role in host-microbe interactions in the gut (Becker and Lowe, 2003). As fucose is used as an energy source by commensal bacteria, the host may regulate fucose production in order to influence the composition of the gut microbiome (Hooper, 2004). In support of this idea, epithelial fucosylation of glycoproteins is induced in mice treated with toll-like receptor ligands (Pickard et al., 2014). Furthermore, mice that have a GDP-fucose deficiency have increased epithelial permeability and spontaneously develop colitis, dysplasia, and, ultimately, adenocarcinoma (Wang et al., 2016).

Importantly, in humans the FUT2-associated SNP rs516246, which was first identified by GWAS as an IBD risk locus (Franke et al., 2010; Jostins et al., 2012; McGovern et al., 2010), has also been established as an IBD-associated expression quantitative trait locus (Singh et al., 2015). Additionally, abnormal glycosylation has been observed in patients with inflammatory bowel disease (Podolsky and Fournier, 1988), in addition to being present in the colonic mucosa of patients during malignant transformation (Boland et al., 1982). Thus, my data suggest that host-microbe crosstalk may be perturbed in the Foxa1/a2-mutant colon as a result of inhibited glycosylation, resulting in colonization of the intestine by pathogenic microflora, thus modeling certain aspects of inflammatory bowel disease.
Although my data provide compelling evidence for disrupted MUC2 glycoconjugation that leads to dysbiosis, I have yet to directly show that mucus composition is altered in Foxa1/a2-mutant colon, or that bacterial interactions with the mucosa are altered in mutants. In healthy colon tissue, the Tn antigen that serves as the foundation for O-type glycosylation is normally masked by additional glycans and is thus virtually undetectable by antibody staining. However, in colon cancer, glycosylation is frequently disrupted and unmodified Tn antigen can be detected (Itzkowitz et al., 1991). Therefore, in the future, I plan to demonstrate the reduction of O-glycan biosynthesis by staining for Tn antigen in mutant and control epithelium.

Despite the significantly increased histological damage scores and the considerably increased number of CD45R⁺ lymphoid nodules in aged Foxa1/a2 mutants compared to age-matched controls, I have not yet directly shown that there is an increase in the permeability of the colonic mucus layer in mutants. Thus, in order to confirm compromised barrier function in Foxa1/a2-mutant colon, I plan to visualize bacterial localization relative to the mucus layer by fluorescence in situ hybridization in tandem with MUC2 staining, as described by Johansson and Hansson (Hill and Button, 2012). MUC2 that lacks glycan branches would not only have an impaired ability to bind water (Johansson et al., 2011), but would also leave the protein core exposed to proteolytic cleavage by bacteria (Van Der Post et al., 2013). Therefore, I expect that the normally thick, sterile mucus layer would be thinner and allow the epithelium to be directly apposed to resident microbiota in Foxa1/a2 mutant colon.

Application of the colonic FOXA1/A2 cistrome to IBD risk locus prediction

In addition to the environmental triggers of disease, GWAS has identified over 200 IBD risk loci, which have implicated genes involved pathways including innate and adaptive immunity, ER stress, autophagy, microbial defense, and intestinal barrier function (Jostins et al., 2012; Kaplan and Ng, 2016). However, numerous IBD risk loci have also been identified in intergenic regions, emphasizing the need to identify enhancer elements in order to understand the effects of disease-associated mutations at these loci on gene regulatory networks.

Many enhancers contain evolutionarily conserved sequences and are frequently marked
by H3K4me1 enrichment, as well as acetylation of H3K27 (Heintzman et al., 2009). Given the requirement of FOXA for the maintenance of chromatin accessibility at enhancer regions in the liver (Iwafuchi-Doi et al., 2016), I expect that by aligning human colonic FOXA ChIP-seq data with maps of H3K4me1 and H3K27ac enrichment in the colonic epithelium, I will be able to identify evolutionarily conserved colon-specific enhancers. At present, there are no FOXA occupancy data available for human colonic epithelium. Therefore, I plan to perform ChIP-seq in colonic human crypts for both FOXA1 and FOXA2.

**Foxa1/a2 in Barrett’s esophagus**

In addition to better understanding IBD etiology, studying the developmental pathways regulated by FOXA1/A2 may also aid in the investigation of Barrett’s esophagus (BE) pathogenesis. BE is characterized by the presence of intestinal columnar epithelial cells, enteroendocrine cells, Paneth cells, and MUC2-expressing goblet cells in the esophagus. Intestinal metaplasia in the esophagus is thought to develop as a result of chronic gastroesophageal reflux, which may activate intestinal transcription factors, such as CDX2 and FOXA2. The risk of developing esophageal adenocarcinoma (EAC) is substantially increased in patients with BE compared to the general population, however EAC pathogenesis is still not well understood (Chen et al., 2011).

FOXA2 is highly expressed in the columnar embryonic esophageal epithelium, but not in the mature squamous esophageal epithelium. However, FOXA2 is expressed in human BE, dysplasia, and EAC. Additionally, though FOXA1 is expressed in the squamous epithelium of the esophagus (Besnard et al., 2004), it is overexpressed in EAC (Lin et al., 2002). Notably, overexpression of FOXA2 in human esophageal squamous epithelial cells induces expression of MUC2 and AGR2, which is a protein found in the ER of intestinal secretory cells and is required for folding MUC2 (Wang et al., 2014).

It has been shown that environmental stimuli, such as tissue damage from acid reflux, may activate developmental signaling pathways resulting in one type of tissue abnormally giving rise to cells of a different, but embryonically related, tissue (Tosh and Slack, 2002). In the liver,
the enhancer of Cdx2, which is transcriptionally silent, is co-occupied by FOXA2 and the repressive transcription factor RFX1. Likewise, in the normal esophageal epithelium, there is high expression of RFX1, but CDX2 is undetectable. Notably, in vitro, Cdx2 transcription is enhanced by the deletion of its associated RFX1 binding motif, suggesting that RFX1 suppresses activity of FOXA2-occupied enhancers. In addition, in the intestinal metaplastic epithelium of humans with the Barrett’s esophagus with dysplasia and EAC, RFX1 expression is significantly downregulated (Watts et al., 2011). Thus, FOXA may initiate the intestinal differentiation program in non-intestinal epithelium by activating transcription of the network of genes that are required for the differentiation of colonic secretory cells.
CHAPTER 6

BIBLIOGRAPHY


Metcalfe, C., Kljavin, N.M., Ybarra, R., and De Sauvage, F.J. (2014). Lgr5+ stem cells are


Sox2, and ectopic expression of intestinal homeobox genes, Cdx1 and Cdx2: Inverse correlation during progression from gastric/intestinal-mixed to complete intestinal metaplasia. J. Cancer Res. Clin. Oncol. 130, 135–145.


