The Roles of KCNQ1 (Potassium Voltage-Gated Channel, KQT-like Subfamily, Member 1) and CFTR (Cystic Fibrosis Transmembrane Conductance Regulator) in Mouse and Human GI Cancers

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL OF THE UNIVERSITY OF MINNESOTA

By

Bich Le Ngoc Than

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Thesis Advisors: Robert Cormier, Ph.D.

and Patricia Scott, Ph.D.

September, 2013

Acknowledgements

I would like to thank and acknowledge my thesis advisors, Drs. Robert Cormier and Patricia Scott, for their mentorship on this thesis research project. They have been instrumental in my scientific development, from experimental design to writing. They have also provided support and guidance during the course of my graduate journey, for which I greatly appreciate.

I am thankful for the service of my committee members: Drs. Kendall Wallace, Lester Drewes, David Largaespada and Fekadu Kassie. I greatly appreciate the helpful comments and advice they provided during my graduate training experience.

Dedication

I would like to dedicate this dissertation to my family. Without their trust and constant encouragement, I would not be the way I am today.

I would like to thank my mother, Kim-Chung Le, for her unyielding support. She has been with me every step of the way and given me the inspiration to keep moving forward. I thank my brother, Thanh Than, for his encouragement and confidence in me to follow all of my aspirations. I thank my sister, Phuong-Cac Than, for her support to pursue and complete this degree.

I would also like to dedicate this work to my late father who passed away during the first year of my graduate training. He instilled in me a sense of hard work and dedication toward my goals.

For my family, I have propelled my career towards the field of medicine, knowing, step by step, my research may help thousands of cancer patients out of the place where those more fortunate have never ventured.

Abstract

The ion-channel genes *Kcnq1* and *Cftr* were identified as gastrointestinal (GI) tract cancer susceptibility genes in multiple *Sleeping Beauty* DNA transposon-based forward genetic screens in mice. *Kcnq1* encodes for the pore-forming alpha subunit of a voltage-gated potassium channel and *Cftr* encodes for the chloride conductance channel. These ion channels act together to maintain ion homeostasis in the cellular and extracellular environment. To confirm that *Kcnq1* and *Cftr* have a functional role in GI tract cancer, mouse models in which targeted mutant alleles of *Kcnq1* and *Cftr* were intogressed into the intestinal tumor susceptible Apc^{Min} strain of mice. Results demonstrated that *Kcnq1* mutant mice developed significantly more intestinal tumors, especially in the proximal small intestine and colon, with some of these tumors in the proximal small intestine progressing to adenocarcinomas. Gross tissue abnormalities and neoplasia were also observed in the rectum, pancreas and stomach. Similarly, *Cftr* mutant mice developed significantly more intestinal tumors, both in the colon and the entire small intestine. Colon organoid formation was significantly increased in organoids created from *Kcnq1* mutant and *Cftr* mutant mice compared with wildtype littermate controls, suggesting a role for *Kcnq1* and *Cftr* in regulation of the intestinal crypt stem cell compartment. To identify gene expression changes due to loss of *Kcnq1* and *Cftr*, we carried out microarray studies in the colon and proximal small intestine. We identified an overlapping set of altered genes involved in innate immune responses, goblet and Paneth cell function, ion channels, intestinal stem cells, EGFR and other growth regulatory signaling pathways. We also found genes implicated in inflammation and in cellular detoxification. Pathway analysis using Ingenuity Pathway Analysis (IPA) and gene set enrichment analysis (GSEA) confirmed the importance of these gene clusters and further identified significant overlap with genes regulated by *MUC2*, another important regulator of intestinal homeostasis. To investigate the role of *KCNQ1* in human colorectal cancer (CRC) we measured protein levels of KCNQ1 by immunohistochemistry in tissue microarrays containing samples from CRC patients with liver metastases who had undergone hepatic resection. Results showed that low expression of KCNQ1 expression was significantly associated with poor overall survival (OS). Our results indicate that both *KCNQ1* and *CFTR* are potent tumor suppressor genes in GI cancer. Defining the mechanisms of action of *KCNQ1* and *CFTR*, and elucidating the nature of their interactions in GI cancer can lead to their use as prognostic biomarkers and potential therapeutic targets for human cancers.

Table of Contents

List of Tables

Chapter 2

Chapter 3

Appendices

List of Figures

Chapter 1

Chapter 2

Chapter 3

Appendices

Abbreviations

ABC	ATP Binding Cassette
APC	Adenomatous polyposis coli
Aqp	Aquaporin
Ascl2	ASCL2 achaete-scute complex homolog 2 Cdkn1c
BMP	Bone morphogenetic proteins
BWS	Beckwith-Wiedemann Syndrome
CBC	Crypt Base Columnar
CF	Cystic Fibrosis
CFTR	Cystic Fibrosis Transmembrane Conductance Regulator
CIN	Chromosomal instability
CIS	Common Insertion Site
CRC	Colorectal Cancer
CRCLM	CRC Liver Metastasis
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DDT	Dichloro-diphenyl-trichloroethane
DIOS	Distal Intestinal Obstruction Syndrome
E2	Estrogen
EMT	Epithelial-Mesenchymal Transition
ER	Endoplasmic Reticulum
ER	Estrogen Receptor
FAP	Familial Adenomatous Polyposis

FFPE	Formalin-Fixed and Paraffin-Embedded
FMFP	Fragments per kilobase of exon per Million Fragments mapped
GEO	Gene Expression Omnibus
GERD	Gastroesophageal Reflux Disease
GI	Gastrointestinal
GSEA	Gene Set Enrichment Analysis
HNPCC	Hereditary Non-Polyposis Colorectal Cancer
HRR	Hazard Rate Ratio
IBS	Inflammatory Bowel Disease
IPA	Ingenuity Pathway Analysis
IR	Inverted Repeats
JLNS	*Jervell and Lange-Nielson Syndrome*
KCNE	Potassium Voltage-Gated Channel, Isk-Related Family
KCNQ1	Potassium voltage-gated channel, KQT-like subfamily, member 1
Kcnq1ot1	KCNQ1 overlapping transcript 1
KO	Knockout
LM-PCR	Ligation-Mediated PCR
LOH	Loss of Heterozygosity
LQTS	*Long QT Syndrome*
LRP5	Low-density lipoprotein Receptor-related Protein 5
LSL	Lox-STOP-Lox
LTR	Long Terminal Repeat
MI	Meconium Ileus

Min	Multiple Intestinal Neoplasia
MMR	Mis-Match Repair
MOH	Maintenance of Heterozygosity
MSCV	Murine Stem Cell Virus
MSI	Microsatellite instability
MTT	(3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, a yellow tetrazole
NBD1	Nucleotide Binding Domain 1
NES	Normalized Enrichment Score
NICD	Notch's IntraCellular Domain
NOTCH	Notch homolog 1, translocation-associated (Drosophila)
OS	Overall Survival
PA	Physical Activity
P-gp	P-glycoprotein
PI	Pancreatic Insuffiency
PI3K	Phosphatidylinositol 3-Kinase
PS	Pancreatic Sufficiency
SB	Sleeping Beauty
SD	Splice Donor
SIBO	Small Intestinal Bacterial Overgrowth
TA	Transit-Amplifying
TALEN	Transcription activator-like effector nucleases
TCGA	The Cancer Genome Atlas
TLESR	Transient Lower Esophageal Sphincter Relaxation
TLR	Toll-Like Receptor

UPR	Unfolded Protein Response
WNT	Wingless-type MMTV integration site family
WT	Wild-type

A distorted version of our normal selves.

Harold E. Varmus,
describing a cancer cell.

CHAPTER ONE

Introduction

Battle against gastrointestinal cancer and the importance of targeted therapies

The impact of CRC on society

Excluding skin cancers, colorectal cancer (CRC) is the third most common cancer diagnosed in the United States (ACS 2012), with the lifetime risk of developing CRC at about 1 in 20 people (5.1%). CRC is currently the second leading cause of cancer-related deaths in the United States when both men and women are combined, and the third leading cause when each sex is considered separately. The incidence and mortality from CRC increase with age as 90% of new cases and 94% of deaths occur in individuals age 50 and older. The incidence of CRC is more than 15 times higher in adults age 50 and older than in those age 20 to 49. CRC is expected to cause > 50,000 deaths in the United States during 2013.

Screening and detection methods such as colonoscopy allow more CRCs to be found early when the disease is preventable. Despite a recent annual decline of 3% in the CRC death rate, advanced metastatic colonization of distant organs remains untreatable, reflected by a high mortality rate of ~50% (Siegel *et al.* 2013, Lubbe *et al.* 2009).This is because 30% of CRC patients present with either regional or diststant metastasis at the time of their diagnosis. For patients diagnosed at an advanced stage, only 12% remain alive after five years. Currently, only about half of people age 50 or older, for whom screening is recommended, report having received CRC testing consistent with current

guidelines. These grim statistics highlight the significant impact of CRC on our society and thus the importance of new diagnostic and therapeutic tools for this cancer.

Risk factors contributing to CRC

1) Genetic factors

Many genetic risk factors have been defined for CRC. Inherited dominant predisposition syndromes are estimated to account for approximately 5% of CRC while up to one third of CRC cases demonstrate familial susceptibility. Of the hereditary CRC syndrome, Familial Adenomatous Polyposis (FAP), a highly penetrant, autosomal dominant syndrome, accounts for ~ 0.5- 1% of cases. Patients with FAP develop hundreds of colonic polyps early in life and their lifetime risk of developing CRC is almost 100%. FAP patients carry germline mutations in the *APC* gene – a crucial negative regulator of Wnt signaling. Another hereditary CRC syndrome is Hereditary Nonpolyposis Colorectal Cancer (HNPCC or Lynch Syndrome), which accounts for 2-4% of all hereditary CRC. Germline mutations in the mis-match repair (MMR) genes *MSH2* and *MLH1* account for >60% of the known mutations present in HNPCC patients.

In contrast to hereditary CRC, sporadic CRC is estimated to account for > ~95% of all cases (although up to one third of sporadic CRC is linked to familial susceptibility). Rate limiting mutations in sporadic CRC are similar to the hereditary CRC syndrome with ~85% of sporadic CRC arising from initiating mutations in the *APC* gene and most of the remaining cases arising from genetic or epigenetic alterations in MMR genes. In addition to rate-limiting genetic and epigenetic alterations, every CRC exhibits as many as a hundred or more mutations and epimutations that are critical for CRC development by dysregulating important biological processes and a large number of key cellular signaling pathways such as Wnt/β-catenin, Notch, TGF-β and Hedgehog (Shi and Massague 2003, Saif and Chu 2010, Vanuytsel *et al.* 2012, Ueo *et al.* 2012). See a more extensive discussion of CRC mutations below.

2) Environmental factors

In addition to intrinsic genetic factors, there is growing evidence that multiple extrinsic factors (aspects of our lifestyle and the environment we live in) can contribute to the risk and development of CRC, such as diet and exercise. A role for diet in CRC risk and development is supported by the long-standing observation that poor diet contributes to high CRC incidence. High fat diet, as a notable example, contributes to proinflammatory immune signals that are elevated in excess adipose tissues that provide pro-growth paracrine signals to nearby epithelial cells (Bardou *et al.* 2013). High-fat diet can lead to obesity that subsequently affects the outcomes in CRC patients as it was shown to be associated with a significant increase in overall mortality among women with stage II-III colon cancer. The presence of visceral obesity strongly affects the occurrence of metabolic syndrome which describes a state of metabolic dysregulation characterized by insulin resistance with increased fasting glucose, predisposition to type 2 diabetes, premature atherosclerosis and other disorders. Among biological factors that may support a link between obesity and CRC are inflammation, bile acids and gut microbiota. Inflammation is associated with oxidative stress which subsequently affects the regulation of genes encoding for factors important in colorectal carcinogenesis such as p53, and DNA mismatch repair and base-excision DNA repair proteins. Bile acids are composed of cholic acid and chenodeoxycholic acid, and produced in the liver by the metabolism of cholesterol. Primary bile acids are then converted to secondary bile acids by colonic microbiota, and act to promote processing of dietary fat. The release of bile acids can affect cancer development through tumor-promoting activities such as inducing DNA damage and cellular proliferation. The third factor, gut microbiota, represents an emerging field of interest. The commensal bacteria population in the intestine is composed of at least 400 species. Although the microbiota is essential and beneficial to the human host, dietary factors or inflammation may affect the microbial composition and lead to formation of dysfunctional microbiota. It has been suggested that microbiota plays a role in the development of obesity-dependent low-grade inflammation. In animal

models, it has also been suggested that bile acids regulate the composition of the gut microbiota. These factors that link high-fat diet to CRC are interchangeable.

In addition to high-fat diet, previous studies have also indicated a link between chemicals in meat cooked at high temperatures and cancer risk, as well as a potential link between vitamin D deficiency and the risk of inflammatory bowel disease (IBD) - a predisposition to CRC. In a European cohort study, intake of heterocyclic aromatic amines which arise from cooking meat and fish at high temperatures was found to increase the risk of colorectal adenomas (Rohrmann, Hermann and Linseisen 2009). In a study on the role of vitamin D receptor (VDR) in intestinal mucosal integrity, it was found that $VDR^{+/+}$ mice were mostly resistant to 2.5% DSS while $VDR^{-/-}$ mice developed severe colitis that subsequently led to death (Kong *et al.* 2008).

Exercise is another influencing factor on CRC development. High levels of physical activity have been associated with a decreased risk of colon cancer, observed in both Western and Asian populations (Hall and Crowe 2011, Pham *et al.* 2012, Yun *et al.* 2013). Overall, the prevalence of low-risk and high-risk colorectal adenomas was correlated with increasing body mass index (BMI) in a dose-response manner. It was reported that the incidence rate ratio for finding colonic adenomas among African-American women was 0.94 for groups with low physical activity (PA) versus 0.72 for those with high PA ($P_{trend} = 0.01$) (Watson and Collins 2011). A meta-analysis study of Korean patients with CRC showed that both prediagnosis and postdiagnosis PA were associated with reduced CRC-specific mortality and all-cause mortality (Je *et al.* 2013).

The biology of the GI tract

The role of GI tract is to digest food for energy and to absorb nutrients and excrete waste as food passes through the body (Figure 1, Essential Clinical Anatomy 2011). While intestinal absorption occurs mainly in the duodenum, intestinal secretion occurs throughout the GI tract and each segment specializes in different secretory

functions. Starting in the stomach, acid secretion acidifies the gastric lumen to promote an environment for bacteria and protein degradation. In order to ensure a balanced pH condition for the function of digestive enzymes in the small intestine, bicarbonate dominantly secreted by the pancreas and fractionally secreted from the duodenum is released into the duodenum to neutralize the gastric acid. Therefore, absorptive and secretory processes are critical for the normal function of the GI tract.

Along the longitudinal axis of the intestinal tract, the small intestine, ~20-25 feet long, breaks down food and absorbs most of the nutrients. In the large intestine, a muscular tube of about 5 feet long, water is absorbed and waste matter is temporarily stored. The waste matter then moves from the colon into the rectum and passes out of the body.

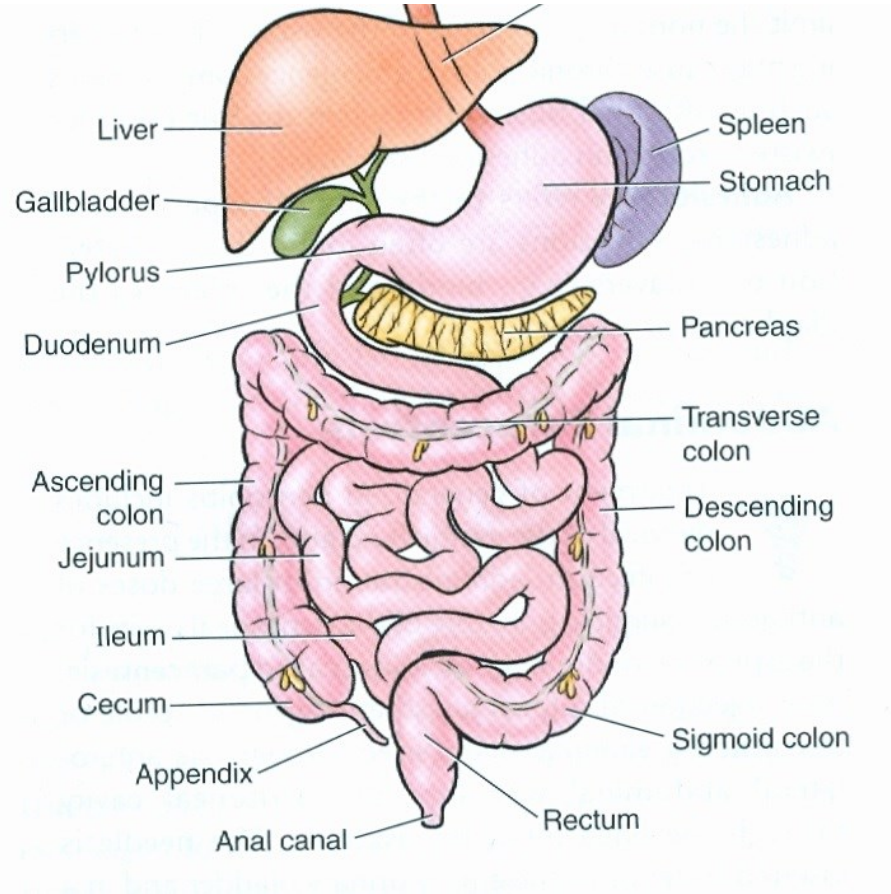


Figure 1. **The Anatomy of the Gastrointestinal Tract.** (Essential Clinical Anatomy 2011).

The dorsoventral axis of the intestine demonstrates radial asymmetry and consists of four layers (Figure 2, Heath JK 2010). Starting from the inside of the intestine or the lumen, the mucosa is where dietary nutrients are absorbed. The muscularis mucosa presents a thin layer of smooth muscle that supports the mucosa and provides it with the ability to move and fold. Rich in vessels like arteries, veins lymphatics and nerves, the second layer - the submucosa - provides the intestine with blood supply. The third layer is the muscularis propria that surrounds the submucosa and forms the basis of peristalsis. The fourth and outermost layer is the serosa which consists of layers of connective tissue and acts as a covering of the digestive tube.

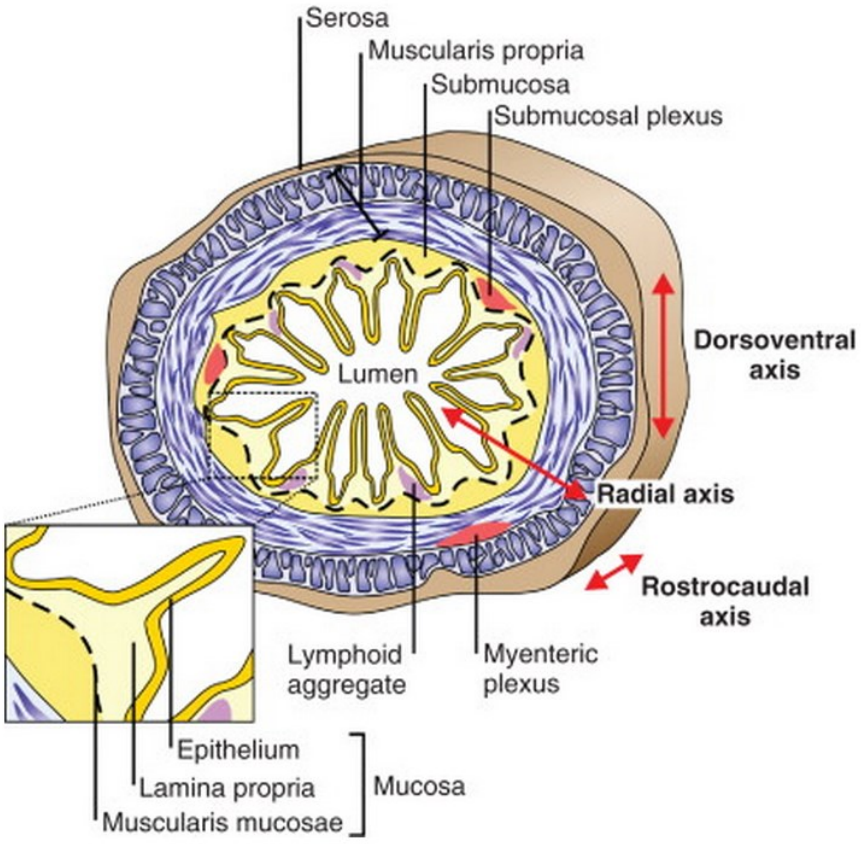


Figure 2. The multi-layered organization of the mature GI tract. (Heath JK 2010). The GI tract is a tube exhibiting rostrocaudal, dorsoventral, and radial asymmetry. It comprises four layers: the mucosa, the submucosa, the muscularis propria (comprising an inner layer of circular muscle fibers and an outer layer of longitudinal muscle fibers), and the serosa. The mucosa (inset) provides the inner lining of the tube and consists of an elaborately folded intestinal epithelium, the lamina propria, a supporting connective tissue containing lymphoid aggregates, and the muscularis mucosae, comprising a thin layer of muscle fibers. Clusters of enteric neurons are found in the submucosa (submucosal plexi) and between the two muscle layers (myenteric neurons).

Within the mucosal layer, two major subpopulations of cells exist: epithelial cells in the epithelium (more details to be discussed in the next section) and stromal cells in the lamina propria. Stromal cells make up the mesenchymal compartment and contain multiple populations of cell types that interact and regulate the various cellular functions in the epithelium. Notable members of stromal cell population are myofibroblasts, fibroblasts, pericytes and smooth muscle cells. In addition, both intestinal epithelial and stromal cells are influenced by multiple immune cell-derived cytokines. Epithelial cells interact with the intestinal microflora and dietary antigens, but the epithelial barrier is not sufficient to prevent these antigens from entering the lamina propria. Therefore, professional and nonprofessional immune cells, which reside close to the epithelial layer, represent a second line of immune defense to escaped luminal invaders. Once reaching the lamina propria, microbial products and foreign invaders may activate the innate immune response through pathogen-associated toll-like receptors (TLRs) present on antigen presenting cells such as macrophages and dendritic cells, which then regulates the efficiency of antigen presentation to T cells. Of note, $CD90^+$ stromal cells, myofibroblasts and fibroblasts, comprise up to 30% of the lamina propria mononuclear cell population and also carry TLRs. Thus, the activation of TLRs on myofibroblasts and fibroblasts can lead to the upregulation in expression of a variety of proinflammatory mediators (such as interleukin [IL]-1α, IL-1β, and IL-8) and adhesion molecules (such as ICAM-1). This event subsequently causes the migration of immune cells to the site of local gastrointestinal mucosal inflammation. Besides innate immune response, adaptive

immunity is also an important regulator of the epithelium-stromal environment. This includes gastrointestinal peripheral CD4+ T cells existing as resting lymphocytes that can be quickly activated by presented antigens. In response to microbial products and to immune mediators, intestinal fibroblasts and myofibroblasts then produce a number of soluble cytokines, chemokines, and growth factors that are key contributors to inflammatory responses, potentially leading to chronic inflammatory gastrointestinal diseases such as Crohn's disease, ulcerative colitis, and celiac disease. In short, via normal immune response, the epithelial lining of the GI tract must detect and quickly respond to bacteria and their byproducts, food antigens and drugs. When the normal regulation is lost, the immune response can wreak havoc, with the overbearing immune response from epithelial, dendritic cells and macrophages leading to intestinal inflammation. Chronic inflammation resulting from the disruption of the epithelial-stromal environment can subsequently predispose patients to intestinal cancer.

The biology of intestinal cancer

Almost all intestinal cancers arise in the intestinal mucosa, the innermost lining of the large intestine, usually as benign polyps (Figure 3, Johns Hopkins Medicine CRC). Some polyps have a high chance of becoming cancerous if not removed, including adenomatous polyps which have villous characteristics. Most adenomas tend to grow slowly over years, which provides a window of time for surveillance and prevention of malignant adenomas that can be readily removed during colonscopy. Adenomas that become malignant or cancerous are termed adenocarcinomas. As cancer cells grow, they can invade nearby layers (i.e. the muscularis mucosa) and eventually grow into and beyond the wall of the intestine. Once the cancer becomes advanced, the tumors may metastasize via shedding cells into the circulatory system and hence spreading the cancer to other organs such as the liver and the lung.

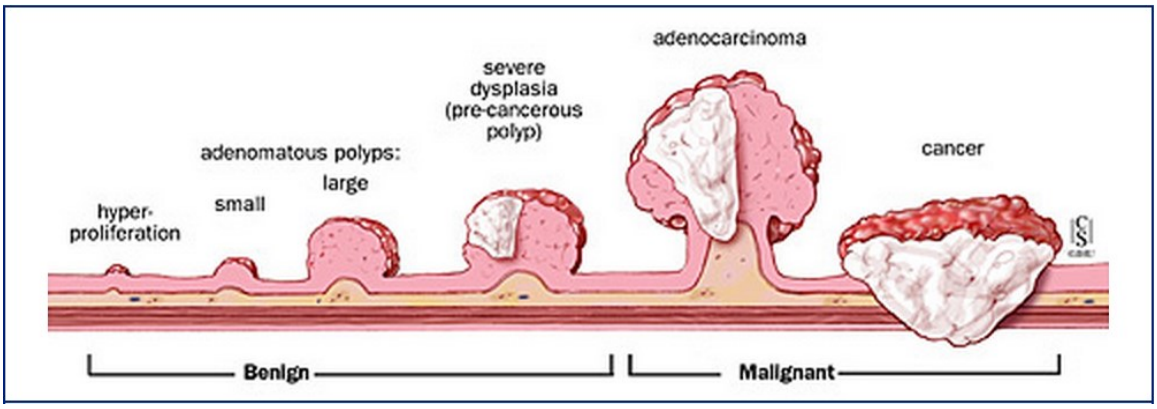


Figure 3. Progression from Polyps to Cancer. (Johns Hopkins Medicine CRC)

Almost all intestinal cancers develop from adenomas that originate in the mucosa. The mucosal surface is lined by an absorptive and secretory epithelium which consists of a monolayer of columnar cells with millions of "crypts of Lieberkuhn" that project deep into the underlying submucosa. Like the skin, the intestinal epithelium has a remarkable self-renewal rate to maintain an active barrier to harmful invaders such as bacteria and toxic waste. The process occurs as the epithelial cells migrate from the site of their division cycle in the base of the crypt to the point of extrusion in the intestinal lumen. Under normal conditions, the turnover of epithelial cell lineages occurs every 2-7 days but this process is faster during wound healing or tissue repair. The intense proliferation which fuels this renewal process originates at the base of the crypt where long-lived stem cells produce progeny that include secretory Paneth cells (in the small intestine) and transit-amplifying (TA) cells (Figure 4, Medema and Vermeulen, 2011). Paneth cells migrate downward to reside at the bottom of the crypt and live for 6-8 weeks. TA cells migrate up toward the crypt/villus border. During this upward migration, TA cells begin to differentiate into the enterocytes, goblet cells, enteroendocrine cells and tuft cells. The enterocytes, the most abundant cell type in the intestinal epithelium, are absorptive cells while the others are all secretory. Each cell type has a different function in the intestinal tract. Populating the base of the crypt, Paneth cells secret digestive enzymes, growth factors and especially diverse host defense (antimicrobial) proteins such as α-defensins,

secretory phospholipase A2 and various proinflammatory cytokines (e.g. TNF-α, IL-1β) to protect the constantly renewing crypt from harmful invaders. No less important, enterocytes absorb water, vitamins and other nutrients important for the overall function of the intestinal tract. Enteroendocrine cells secret peptide hormones that regulate numerous important biological processes such as the control of glucose levels and regulation of food intake. Goblet cells produce different mucins that make up the protective intestinal mucus barrier to harmful foreign invaders and bacteria. The role of tuft cells in the intestinal tract is still largely unknown. When the migration of these differentiated cells reaches the villus tip (small intestine) or intercrypt table (colon), they undergo apoptosis and are shed into the intestinal lumen. The composition of cell types in the large intestine or the colon is similar to that of the small intestine, except for the absence of villi and Paneth cells. The epithelial layer of the colon is made up of a single sheet of columnar epithelial cells that form crypt invaginations.

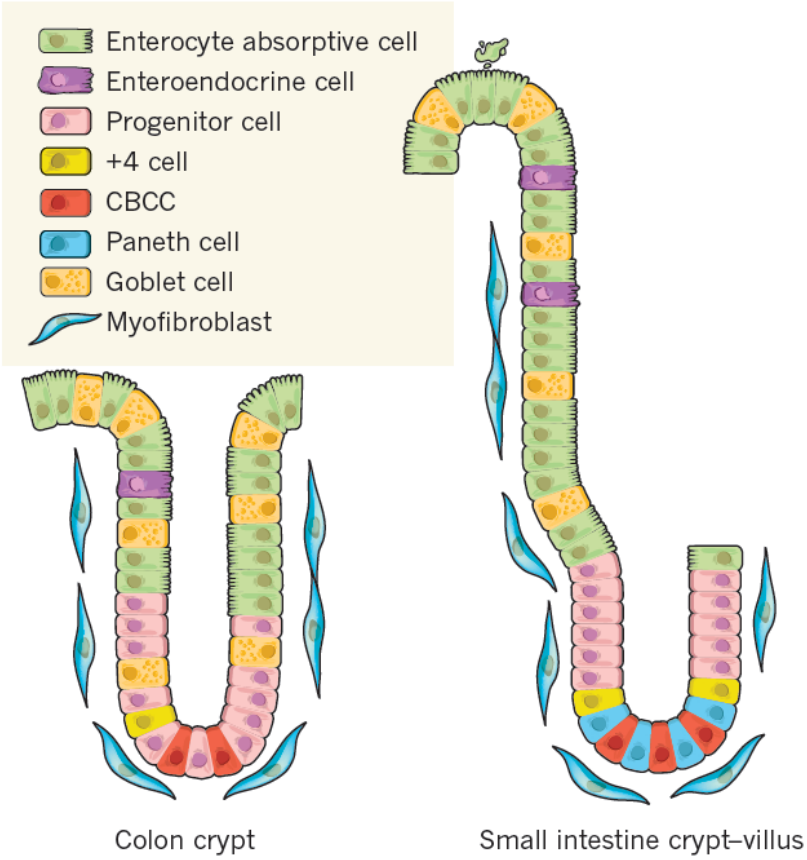


Figure 4. The organization of the small intestinal crypt-villus and the colon crypt. (Medema and Vermeulen, 2011). Both the colon crypt (left) and the small intestinal crypt (right) contain a stem-cell compartment at the crypt bottom. CBCCs and the +4 stem cell have been indicated to be present between and just above the Paneth cells, respectively. Of note, Paneth cells are not detected in the colon, yet a Paneth-like cell has been suggested to be present at the crypt bottom. All four lineages (three in the colon) - enterocytes, Paneth cells, goblet cells and enteroendocrine cells — appear in different, but set ratios. Paneth cells move down to the bottom and are long-lived, whereas other lineages move up and are shed (a few days later) into the lumen while undergoing apoptosis. Rare cell types reported to exist in crypts, such as tuft cells, are not shown.

There are two hypothetical models of stem cells in the intestine (Barker 2012, Li 2012, Clevers 2013). The "+4 model" refers to a common intestinal stem cell origin at position 4-5, directly above the Paneth cell compartment. Potten and colleagues showed that these radiation-sensitive, label-retaining +4 cells were actively proliferating with 24hr cell cycles. The other model termed "the stem cell zone model" originates from the finding by Cheng and Leblond that the crypt base is not solely populated by postmitotic Paneth cells. Instead, using electron microscopy, they found a population of slender crypt-base-columnar (CBC) cells wedged between the Paneth cells. They called them intestinal stem cells. These CBC cells, also radiation-sensitive, are located at position 1-5. But unlike the +4 cells, the CBS cells did not retain DNA labels.

Although the identity of intestinal stem cells is still controversial, it is agreed that intestinal stem cells reside in a niche that provides them with essential signals from important morphogenetic pathways such as Wnt, Notch, and BMP (Crosnier *et al.* 2006, Hardwick *et al.* 2008, Reya and Clevers 2005). (Figure 5, Heath JK 2010).

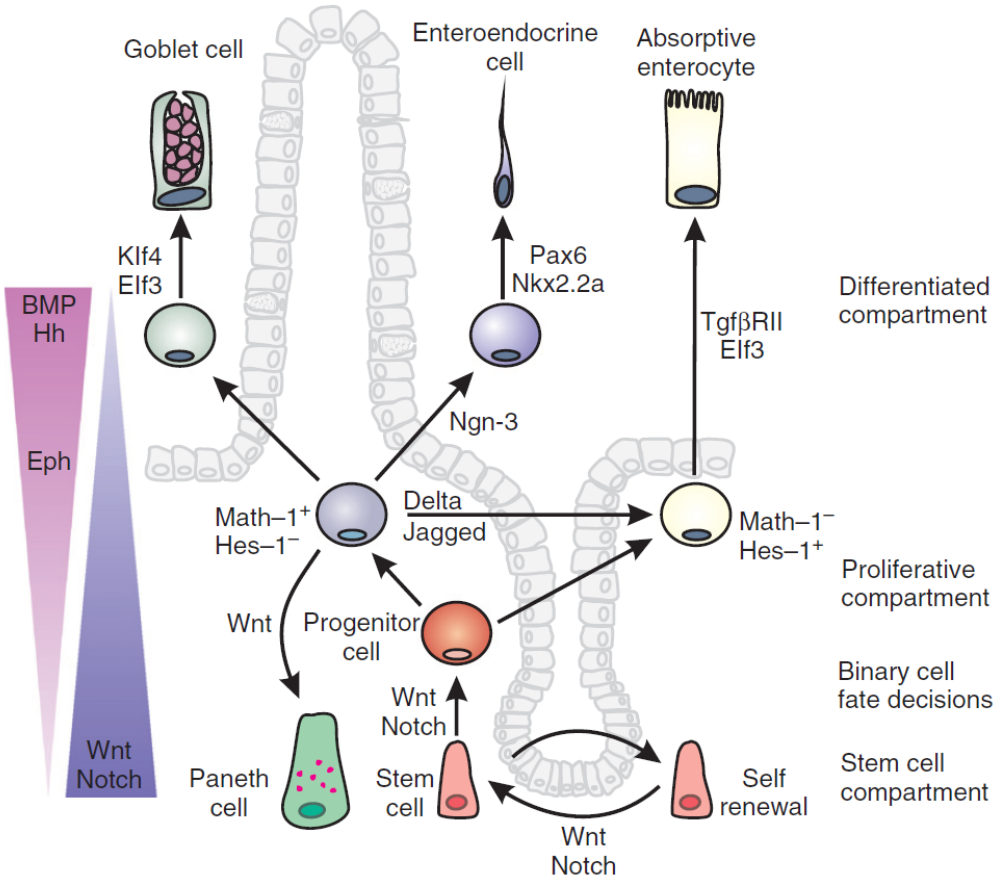


Figure 5. Organization of the intestinal epithelium and the crypts of Lieberkühn and cell lineage determination in the intestinal epithelium. (Heath JK 2010)**.** Stem cells reside at or close to the base of the crypts, either between or just above the Paneth cells. Above the stem cells are the rapidly dividing progenitor/transit amplifying cells. Four different cell lineages (three secretory and one absorptive) arise from a common multipotential progenitor cell through a Notch-mediated binary cell fate decision. The three classes of secretory cells are Paneth cells, which secrete antimicrobial proteins, goblet cells containing thece full of mucus, and the much rarer enteroendocrine cells, which secrete various gut hormones. The absorptive cells have an elaborate brush border on their apical surface. Progenitor cells that receive a Notch signal (Delta/Jagged) from their neighbor(s) turn on the expression of the Hairy and enhancer of split-1 (Hes-1) TF. In turn, Hes-1 antagonizes Math-1, also a TF, which is required for commitment toward secretory lineages. Thus the Notch signal-receiving cell (Hes-1/Math-1) differentiates into an absorptive cell, rather than a secretory cell. In contrast, cells expressing Notch ligands (Hes-$1^{-/}$Math-1^{+}) escape Notch activation and differentiate into secretory cells. Once this secretory versus absorptive cell fate decision has been made, further TFs are required to promote terminal differentiation of the four cell lineages, some examples of which are shown. Neurogenin-3 (also a TF), Nkx2.2a, and Pax6 all

promote enteroendocrine differentiation while the Kruppel-like TF, Klf-4, is required for goblet cell differentiation. An Ets-family TF, Elf-3, activates Tgfβ expression and thereby promotes terminal differentiation of goblet cells and absorptive cells. Epithelial turnover is maintained by a gradient of Wnt/β-catenin signaling, which is highest at the base of the crypts and tapers off toward the top of the crypts and villi. Other prominent signaling pathways (Bmp and Hedgehog) are also established across the crypt/villus axis to oppose the pro-proliferative force of Wnt signaling and permit cell differentiation. Eph–ephrin signaling prevents intermingling of the proliferating and differentiating cells.

Wnt signaling is crucial in controlling cell proliferation/regeneration in the intestinal epithelium and the maintenance of the intestinal stem cells. The Wnt proteins are extracellular glycoproteins that bind Frizzled (a transmembrane receptor protein) and low density lipoprotein receptor-related protein 5 (Lrp5) receptors on the surface of intestinal epithelial cells. In canonical Wnt signaling, when Wnt binds to its receptors, this inactivates the β-catenin destruction complex composed of the tumor suppressor protein APC, Axin, CK1 and GSK3, which results in the accumulation of β-catenin in the cytoplasm. β-catenin is then translocated into the nucleus to bind with Tcf/Lef transcription factors, leading to transcriptional activation of important target genes involved in cell proliferation and oncogenesis such as *Cyclin D1* and *c-Myc*. In the absence of Wnt-receptor binding, the APC/Axin/CK1/GSK3 destruction complex binds to newly synthesized β-catenin and directs degradation of β-catenin via the ubiquitin-proteosome pathway. Wnt/β-catenin signaling is crucial for intestinal crypt regulation due to its powerful control of important downstream targets. For example, in *Tcf4*$^{-/-}$ mice, the proliferating compartment between the villi was entirely absent, suggesting an essential role of Wnt for maintenance of crypt progenitors (Korinek *et al.* 1998). It has been shown that Paneth cells are critical for sustaining Wnt signaling activation.

Also important for intestinal homeostasis and proliferation is Notch signaling. Components of Notch signaling are expressed in the intestinal crypts but only weakly in the intestinal villi. Notch and its ligands of the Delta and Jagged subfamilies are transmembrane proteins that mediate the communication of cells with each other when in

contact. Both human and mice possess four Notch receptors: Notch 1 to 4. In canonical Notch signaling, when Notch receptors bind to their ligands expressed on neighboring cells, this initiates proteolytic cleavage of the receptors by α-secretase and γ-secretase to release Notch receptors' intracellular domain (NICD). The cleaved receptors then translocate into the nucleus and form complexes with RBP-jκ and induce transcriptional activation of Notch target genes such as transcription factor hairy/enhancer of split (Hes1) important for Notch-induced repression of differentiation. A conditional deletion of *Hes1* from intestinal tumor cells in *Apc*-mutant mice was reported to reduce tumor cell proliferation, while promoting differentiation toward epithelial lineages (Ueo *et al.* 2012). Thus, the Notch pathway is a major regulator of intestinal stem cell maintenance, cell fate specification and possibly cell maturation. Importantly, Notch signaling does not act alone but interacts with other key signaling pathways to promote functional signaling networks required for development and tissue homeostasis. Studies have shown that Wnt signaling drives expression of the Notch pathway. In addition, there may be a cross-talk between the Notch and Ras signaling pathways, in which *Ras* activating mutations can activate Notch signaling while Notch activation is required for Ras-mediated transformation (Saif and Chu 2010).

Another important pathway for maintaining intestinal crypt integrity is BMP signaling. Originally identified as factors inducing cartilage and bone formation, they are now known as general regulators of cell development similar to Wnt and Notch (Shi and Massague 2003). BMPs belong to a complex family of ligands consisting of BMP and transforming growth factor-β (TGF- β) members, which share intracellular signaling through the SMAD proteins. The action of BMP is mediated by type 1 and type 2 serine/threonine kinase receptors. The binding of BMP to its receptor BMP Receptor Type 1 (BMPRI) results in activation of BMPRI, leading to phosphorylation of SMAD proteins (SMAD1, SMAD5 and SMAD8). The phosphorylated BMP-SMADs form a complex with a co-SMAD (SMAD4) and translocate into the nucleus. Once in the nucleus, the SMAD complex interacts with co-activators or co-repressors of transcription, in turn activating or repressing expression of target genes. BMP signaling within the

intestine is quite complex as BMPs and their receptors are expressed in both epithelial and mesenchymal compartments. (BMPs are found in the stroma of the villi while phosphorylated SMAD1, 5 and 8 are found in the nuclei of differentiated epithelial cells). It has been suggested that BMP2 and BMP4, expressed by the mesenchymal cells, can halt proliferation at the crypt-villus border. BMPs are active at the top of the crypt where differentiation occurs. Though also produced at the crypt base, they are kept in check by BMP inhibitors such as Noggin. Similarly to Notch signaling, BMPs also have dual functions and they are involved in lineage fate decisions of secretory cells. BMPs are known to antagonize Wnt signaling along the crypt-villus axis by inhibiting stem cell expansion and promoting epithelial differentiation.

Surrounding the crypt base are stromal (mesenchymal cells) which are largely of myofibroblast lineage and close to the basal lamina – the extracellular matrix (ECM) secreted by the epithelial cells. While normal cells require contact with the ECM where they receive essential signals for growth and survival (otherwise they undergo cell anokis – a form of programmed cell death induced by anchorage-dependent cells detaching from the surrounding ECM), cancer cells do not and therefore can survive outside their normal environment. The ECM in the tumor microenvironment may serve as a barrier against tumor invasion, but it is also a reservoir of cell binding proteins and growth factors that affect tumor cell behavior. The ECM is substantially modified by proteases produced by tumor and stromal cells. As a result of the activity of these proteases, cell-cell and cell-ECM interactions are altered, new biologically active ECM molecules are generated, and the bioavailability and activity of many growth factors, growth factor receptors, and cytokines are modified.

The epithelial-mesenchymal interaction is important in both initiation and progression of tumorigenesis. Regulation of normal and tumor epithelial cell growth, migration and invasion is strongly influenced by the microenvironmental signals coming from the ECM and various stromal cells (myofibroblasts, fibroblasts, smooth muscle cells, pericytes, neutrophils, lymphocytes and macrophages). It has been suggested that the pericryptal myofibroblasts produce Wnt ligands which binds to the Frizzled receptors

on the basal epithelial stem cells (Humphries and Wright, 2008). In addition, by producing factors such as BMP antagonists gremlin 1 and gremlin 2 as well as Notch ligands and mediating the effects of Wnt signaling on Ephrin B1 and receptors EPHB2 and EPHB3, myofibroblasts can contribute to the maintenance of the stem cell compartment, cell migration and differentiation. A recent study found that BMP signaling is not active within the stem cell zone because myofibroblasts and muscle cells located at the base of the crypts secret BMP inhibitory molecules such as gremlin 1 and gremlin 2 (Kosinski *et al.* 2007). Therefore, these stromal cells create an intestinal stem cell environment where Wnt signaling can thrive. Mutations in epithelial cells can cause activation of fibroblasts into a-SMA+ expressing myofibroblasts which are similar to fibroblasts associated with wound healing. These reactive stromal myofibroblasts can stimulate proliferation and malignant transformation of epithelial cells. This has been demonstrated in a model where co-injection of MCF-7 breast cancer cells with reactive stromal, tumor-associated myofibroblasts into nude mice results in enhanced tumor growth and tumor size. In addition, human myofibroblasts treated with PGE2 released elevated levels of amphiregulin, HGF, and VEGF, which modified epithelial proliferation, epithelial migration, and neoangiogenesis (Powell *et al.* 2005, Shao *et al.* 2006).

Lamina propria mesenchymal cells are critical to the instruction of intestinal morphogenesis and regulation of epithelial proliferation and differentiation in the stem cell niche. These processes are subverted by cancer to enhance carcinogenesis, tumorigenesis, and metastasis. Because of the tight interaction between the intestinal epithelium and the stroma, a process termed “Epithelial-mesenchymal transition” (EMT) can arise, where epithelial cells undergo a phenotypic switch to form mesenchymal cells with similar appearance to fibroblasts. This EMT transition allows epithelial cells to acquire oncogenic traits such as enhanced migration capacity, elevated resistance to apoptosis and significantly increased production of the ECM components (Figure 6, Kalluri and Weinberg 2009). The aberrant activation of EMT has been shown to promote tumor metastasis by repressing cell adhesion factors such as E-cadherin. The process of

EMT is completed with the degradation of the underlying basement membrane and the formation of new mesenchymal cells that migrate away from the epithelial layer where they originate.

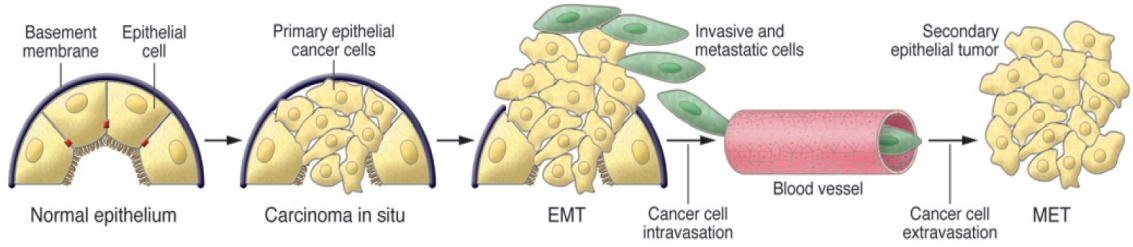


Figure 6. Contribution of EMT to cancer progression. (Kalluri and Weinberg 2009). Progression from normal epithelium to invasive carcinoma goes through several stages. The invasive carcinoma stage involves epithelial cells losing their polarity and detaching from the basement membrane. The composition of the basement membrane also changes, altering cell-ECM interactions and signaling networks. The next step involves EMT and an angiogenic switch, facilitating the malignant phase of tumor growth. Progression from this stage to metastatic cancer also involves EMTs, enabling cancer cells to enter the circulation and exit the blood stream at a remote site, where they may form micro- and macrometastases, which may involve METs and thus a reversion to an epithelial phenotype.

Analysis of CRC mutational spectra

CRC arises through the acquisition of multiple mutations in genes important for normal cellular development and a wide range of normal cellular processes in particular, cell proliferation, metabolism and apoptosis (Figure 7, Johns Hopkins Medicine CRC). It is estimated that for a normal cell to undergo malignant transformation, about 7-15 somatic mutations are required. On the basis of work by Vogelstein and colleagues, in the most common model of adenoma to carcinoma sequence, the first hit that initiates normal tissue to become a dysplastic lesion is through loss of APC function or other means of WNT activation. But the progression from dysplastic or pre-cancerous lesion to adenomas and carcinomas can depend on additional activating mutations in the *RAS* pathway and inactivating mutations of *P53*, *SMAD4* and *PTEN*. As discussed below,

many additional mutations that confer a selective advantage for tumor growth are required for the development of a CRC.

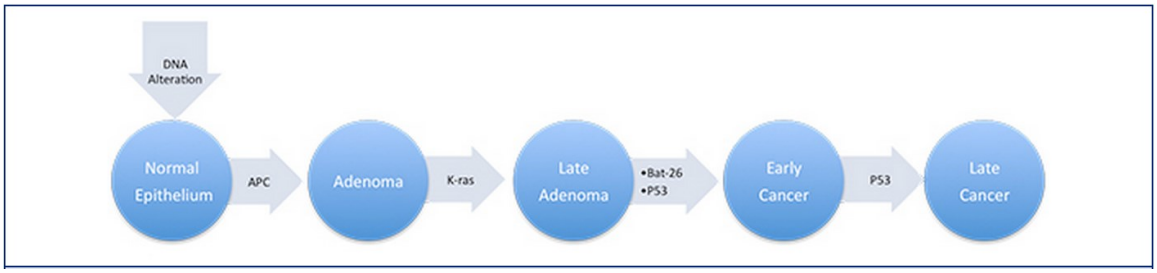


Figure 7. Vogelgram. (Johns Hopkins Medicine CRC). The molecular alterations in multiple genes cause Colorectal Cancer to develop in well-defined stages within an individual cell.

Identification and characterization of genes that drive cancer processes are critical for the development of cancer diagnostic and therapeutic modalities. Advances in DNA sequencing have permitted the discovery of intragenic mutations across multiple cancer genomes. The first large-scale mutational study of CRC was conducted by Bardelli and colleagues, which identified 7 candidate cancer genes in a screen of the tyrosine phosphatase gene family members from 182 CRCs (Bardelli *et al*. 2003). A year later, resequencing of the phosphatidylinositol 3-kinase (*PI3K*) gene family discovered *PIK3CA*, a gene that was frequently mutated in tumors of colon, breast, brain and lung, with most mutations clustering in the catalytic domain (Samuels and Velculescu 2004).

To identify a more complete list of genes dysregulated in CRC, large-scale exome and genome sequencing studies of human CRCs have been conducted (Cancer Genome Atlas Network 2012, Han SW *et al.* 2013, Schee K *et al.* 2013). A prominent example is the early work by the Vogelstein group at John Hopkins Kimmel Cancer Center (Wood *et al*. 2007). In the 2007 study, they identified a small population of genes that were frequently mutated in CRCs. These were called “mountains”, including genes such as *APC*, *KRAS*, *P53*. Surprisingly, their study identified a much larger population of genes mutated at a lower frequency, hence the term “hills”. Their findings suggested that tumor

growth in individual human CRCs may be driven by some genes that are mutated at a low frequency. (Figure 8, Wood *et al*. 2007)

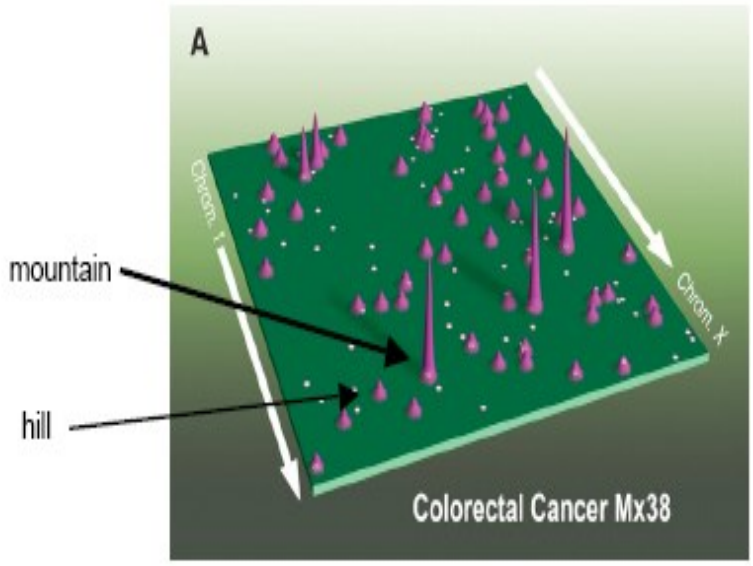


Figure 8. Human Colorectal Cancer genome landscape. (Wood *et al*. 2007). Nonsilent somatic mutations are plotted in two-dimensional space representing chromosomal positions of RefSeq genes. The telomere of the short arm of chromosome 1 is represented in the rear left corner of the green plane and ascending chromosomal positions continue in the direction of the arrow. Chromosomal positions that follow the front edge of the plane are continued at the back edge of the plane of the adjacent row, and chromosomes are appended end to end. Peaks indicate the 60 highest-ranking candidate cancer (CAN)-genes for each tumor type, with peak heights reflecting cancer mutation prevalence (CaMP) scores. The dots represent genes that were somatically mutated in the individual colorectal tumor displayed. The dots corresponding to mutated genes that coincided with hills or mountains are black with white rims; the remaining dots are white with red rims.

These low frequency genes seem to cooperate with high frequency genes to regulate stepwise development of CRC. These low frequency genes can regulate a wide range of biological processes such as metabolic pathways, inflammation and stress responses. However, a remaining challenge was to identify which low frequency genes

act as cancer "driver" genes in this process and which genes only represent as "passenger" mutations. Therefore, identification of low frequency candidate CRC driver genes and their cooperative networks has been a key research objective of many groups of researchers including ours.

Sleeping Beauty mutagenesis

A powerful approach for cancer gene discovery is the use of forward genetic screens in the mouse. Insertional mutagenesis allows for relatively unbiased, genome-wide discovery of novel candidate cancer genes and new cooperative gene networks. Though chemical mutagenesis using mutagens such as ENU presents a highly efficient way to induce tumors in mice, mutant genes are difficult to identify. This has led to the use of retroviruses and DNA transposons as insertional mutagens since each retrovirus and transposon serves as a molecular tag for detection of its location in the genome. While retroviral mutagenesis is a powerful tool for tumor induction in mice, the use of retroviruses poses several disadvantages, in particular, a bias towards insertion in the 5' region of genes and utility primarily limited to hematopoietic cancers.

A different approach using DNA transposon-based mutagenesis systems has emerged as a powerful tool for studying cancer development. Transposons are DNA elements that can mobilize within the host genome. It is estimated that approximately 45% of the human genome is transposon-derived (Ivics *et al.* 2004). While transposons are active in plants and invertebrates, all DNA elements identified in vertebrates to date are nonfunctional (Mattison *et al.* 2009). The cancer gene discovery method described in this dissertation is the *Sleeping Beauty* (SB) system of somatic intestinal mutagenesis. In its natural form, the SB element, which was reconstituted from the consensus sequence of dormant genes identified in Salmonoid fish species, is about 1600bp long consisting of a transposase coding sequence flanked by two inverted repeats (IRs) (Ammar *et al.* 2012). Since the IRs with proper transposase binding sites are sufficient for mobilization, the transposase can be separated from the IRs. Thus, the transposase can be expressed as a

tissue-specific transgene and the IRs can be engineered to flank cargo such as oncogenic elements.

Our SB system presents a combination of three transgenes. The first transgene is an active SB11 transposase that carries a Lox-STOP-Lox (LSL) cassette in the 5' promoter region, permitting tissue specific expression upon expression of a Cre recombinase. The second transgene is an SB transposon called "T2/Onc2" which consists of two IRs of about 230 bp each that flanks a cargo sequence (Dupuy *et al.* 2005). T2/Onc2 transposon contains a murine stem cell virus (MSCV) long terminal repeat (LTR) and a splice donor (SD), which can drive gene expression when integrated upstream or within a gene. The T2/Onc also contains two splice acceptors (SA) and a bi-directional poly(A), which can terminate transcription of a gene when integrated in either of its orientation. Activation and disruption of a candidate gene by SB mutational insertion forms the basic model of activation of an oncogene and inactivation of a tumor suppressor gene, respectively (Figure 9, Copeland and Jenkins 2010). SB transposition is through a "cut and paste" mechanism where the SB transposon is removed by the SB transposase from its native locus and then integrated into another genomic site containing a TA dinucleotide. Upon the removal of the SB transposon from its original locus, the excision site is repaired via non-homologous end joining which leaves a 3bp footprint.

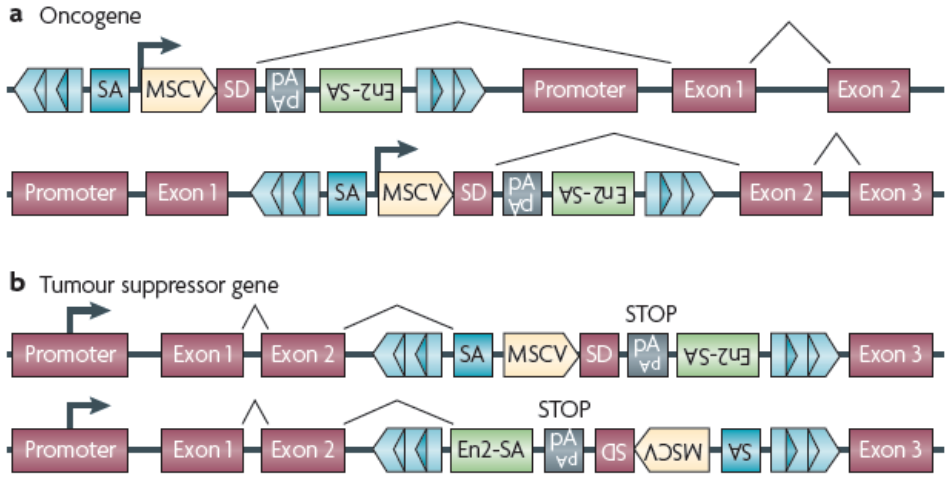


Figure 9. SB Transposon (T2/Onc2) can deregulate the expression of an oncogene or inactivate expression of a tumor suppressor gene. (Copeland and Jenkins 2010). **a)** T2/Onc2 contains a murine stem cell virus (MSCV) 5′ long terminal repeat (LTR) derived from the MSCVneo vector and a splice donor (SD) site derived from exon 1 of the mouse Foxf2 gene50. T2/Onc2 can thus promote the expression of an oncogene when integrated upstream of or within the gene in the same transcriptional orientation. T2/Onc2 is flanked by optimized SB transposase binding sites (light and dark blue triangles) that are located within the transposon inverted terminal repeats (light blue arrows), which increase the frequency of SB transposition. **b)** T2/Onc2 also contains two splice acceptors and a bi-directional polyA (pA) and can thus prematurely terminate transcription of a tumor suppressor gene when integrated within the tumour supressor gene in either orientation. One splice acceptor is derived from exon 2 of mouse engrailed 2 (En2‐SA) and the other from the carp β‐actin gene (SA).

The third transgene is Villin-Cre. Villin is one of the most widely used promoters to target intestinal epithelial gene expression. Villin's basic function is to modulate the structure of actin filaments and support reorganization of actin cytoskeleton under stress (Pott and Hornef 2012). In the intestine, villin is a major component of the brush border of epithelial absorptive cells. Villin is also weakly expressed in exocrine pancreas and liver. By generating a Cre-inducible LSL transposase allele, SB transposition is specific to the epithelial cells of the GI tract when combined with expression of Villin-Cre. (Figure 10, Ammar, Izsvák and Ivics 2012).

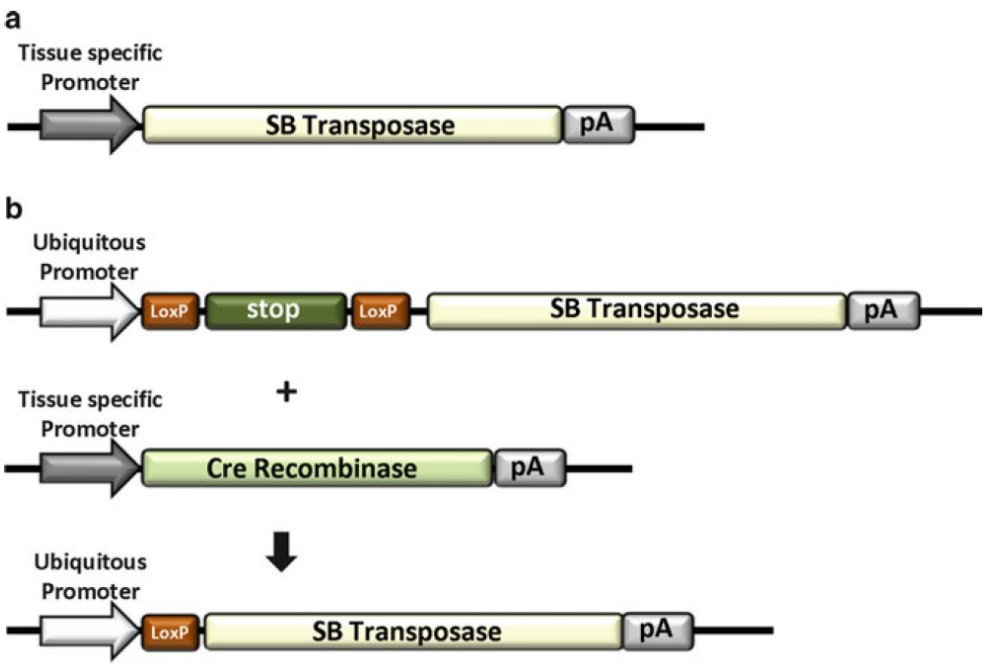


Figure 10. Tissue-specific expression of the SB transposase. (Ammar, Izsvák and Ivics 2012). **a)** A tissue specifi c promoter (gray arrow) regulates the expression of the SB transposase gene (light yellow box). **b)** Tissue-specifi c expression of the SB transposase is achieved by inserting a lox-stop-lox cassette (orange and green boxes) between a ubiquitously active promoter (white arrow) and the SB transposase gene. The lox-stop-lox cassette can be removed by a tissue-specifi cally expressed Cre -recombinase (light green box). The Cre -recombinase excises the lox-stop-lox cassette in the tissue of choice and thus enables tissue-specific expression of the SB transposase gene.

Intercrosses between transgenic founder strains of mice, all on the C57BL/6 genetic background, generated triple transgenic mice in which transposase activity and transposon mobilization were restricted to the GI tract. This system can be further modified by using tumor sensitive strains of mice, such as Apc^{Min} and $LSLp53^{R27OH}$, as a fourth genetic alteration.

Figure 11 depicts a general strategy for the identification of candidate CRC driver genes using SB mutagenesis. Tumors are harvested, flash frozen and DNA isolated. The next step used ligation-mediated PCR (LM-PCR) to amplify and sequence SB

transposition sites, which confirms the reintegration of excised SB transposons into the mouse genome. Using ABI Prism and Roche GSFlex 454 sequenching, thousands of transposon-genomic junctions were sequenced. Based on the number of mapped insertions and distribution of TA dinucleotides per chromosome, Monte Carlo simulations were performed to determine the window size for any given number of insertions where this did not occur by random chance. Our groups then identified the common insertion sites (CIS) that represented areas of the genome that had significantly more insertions than one would expect by chance. The presence of a CIS indicated that transposon insertional mutations at that locus provided a selective advantage and thus likely contributed to tumor development. We assigned a candidate gene to each CIS if the majority of the insertions were in or near a single gene. If transposon insertions in all the tumors in a single CIS landed in the same intron and oriented in the same direction of transcription, we would predict transposon insertions caused a gain-of-function mutation (or activation of candidate oncogene). If the transposon insertions distributed randomly in both orientations in all the tumors of a CIS, we would predict transposon insertions caused a loss-of-function mutation (or inactivation of candidate tumor suppressor gene). Validation and characterization of these candidate cancer driver genes using both *in-vitro* and *in-vivo* analyses help us better understand their roles and mechanisms in GI cancer.

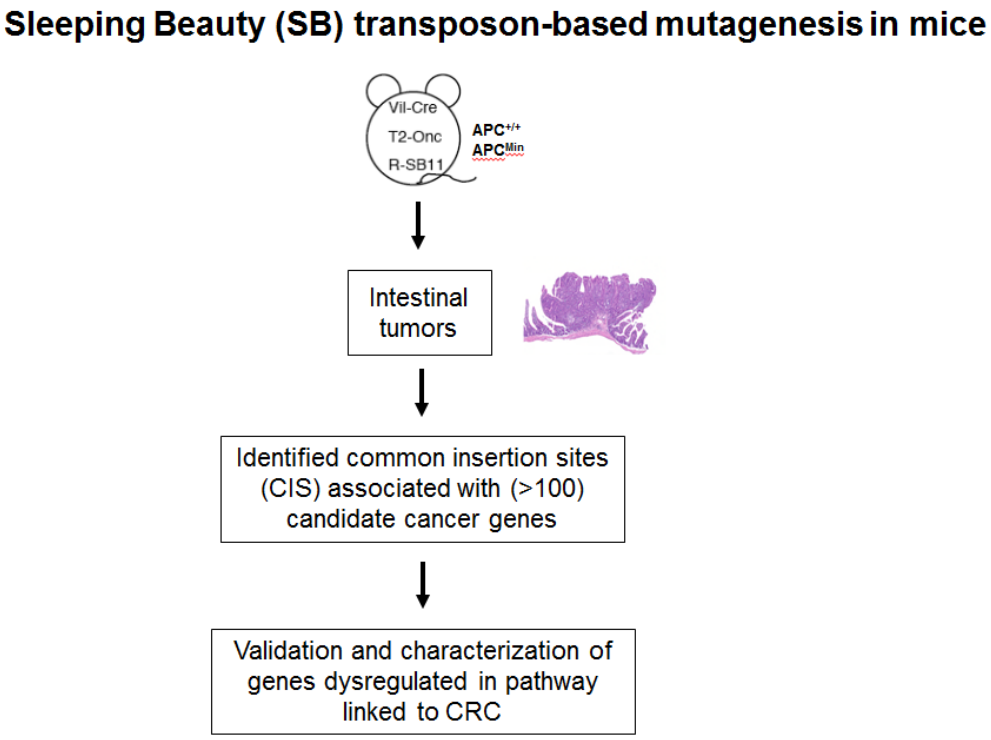


Figure 11. Scheme for validation of intestinal candidate cancer genes.

SB-transposon-mediated mutagenesis screens in the GI tract of mice represent random, unbiased, whole genome scans for somatic gene mutations that can act as drivers of intestinal tumorigenesis. While the role of mutations in rate-limiting genes such as *APC* or *P53* is well known, there is little information on a larger number of low frequency candidate genes that are also required for tumorigenesis. By examining sets of SB-induced mutant candidate genes in hundreds of individual tumors and on different host genetic backgrounds, we can develop a much fuller picture of cancer gene cooperative networks that underlie CRC.

For example, SB mutagenesis in *Apc* wild-type mice identified 77 CIS that are associated with 74 known genes. More recently, SB mutagenesis in Apc^{Min} mice identified 30 genes that cooperated with a germline deficiency in the *Apc* gene. While some of these genes, such as *Apc*, have previously been reported to be tumor suppressors

or oncogenes in human CRC, many of the CIS are novel. Of 104 genes associated with the CIS from the SB screens of *Apc*$^{+/+}$ and *Apc*^{Min} *mice*, 2 CIS candidate cancer genes in this dissertation have been shown to play a crucial role in ion signaling.

Ion channels

Ion channel proteins are found in every cell type and located at the cell membranes to allow regulated transfer of ions. Ion channels can be broadly classified into voltage-gated ion channels, ligand-gated ion channels and mechanosensitive ion channels according to their primary gating modality. Most ion channel gating is regulated by more than one of these mechanisms. Since an indispensable role of the GI tract lies in its absorptive and secretory functions, the interplay among electrolyte balance and water secretion, regulated by ion channels, is essential to a normal GI condition.

The intestinal epithelium is an electrolyte transporting layer that moves large quantities of salt and water from the mucosal side to the blood side and vice versa. Under physiological conditions, normal tuning of salt excretion in the stool requires colonic absorption of about 1.5 liter of electrolyte-rich fluid per day (Mall *et al.* 2002). The tightly coordinated transport of major ions, K^+, Cl^-, $HCO3^-$ and Na^+ are crucial for the bulk-fluid movement. For example, Cl^- and Na^+ absorption is achieved by the regulation of the apical membrane Na^+/H^+ exchangers (mostly NHE2 and NHE3) and anion exchangers (SLC26A3, SLC26A6 and SLC4A9) (Venkatasubramanian, Ao and Rao 2010). On the other hand, Cl^- secretion requires the actions of basolateral Na^+-K^+-$2Cl^-$ transporters (NKCC), basolateral K^+ channels and apical Cl^- channels.

This dissertation focuses on the two ion-channel encoding genes, *KCNQ1* and *CFTR*, that regulate the secretion of potassium and chloride ions and potentially mediate cellular homeostasis via mediating the process of Cl^- secretion. An increase in basolateral K conductance is essential for inward K^+ recycling by the Na/K pump and to maintain a driving force for anion exit across the apical membrane of the cell, in turn allowing the

balance of water transport across the cell membrane. Interestingly, activation of apical CFTR-mediated anion secretion by cAMP/PKA is correlated with phosphorylation and activation of a basolateral membrane K channel, KCNQ1 (Frizzell and Hanrahan 2012). Studies that employed barium showed that this generalized K channel blocker could inhibit apical anion secretion (ie. Cl^-) by depolarizing the membrane potential. Importantly, Cl^- and anion exchangers are also required for HCO_3^- secretion especially in the duodenum where balance of luminal pH is required. This is because the duodenum is constantly exposed to strongly acidic juice released from the stomach, where subsequently low pH signals complex neuroendocrine and possibly hormonal pathways to promote the secretion of duodenal HCO_3- against acid juice damage. Recent studies support the concept of CFTR as one of the apical HCO_3- exit pathway during the stimulation of HCO_3- secretion (Xiao *et al.* 2012, Yang *et al.* 2013). As potassium and chloride are among the most abundant intracellular ions in the human body, disruption of this ion homeostasis can lead to a variety of disturbances in intestinal cellular function.

With high relevance to human physiological processes and cell development, ion channels, such as KCNQ1 and CFTR, and their encoding genes present important targets in toxicological studies and cancer genetics. In the area of toxicology, factors such as natural toxins, environmental pollutants, and chemical compounds have been reported to target and disrupt many important ion channels. One example of natural toxins is chlorotoxin, a 36 amino acid peptide found in the venom of the Israeli desert scorpion *Leiurus quinquestriatus* (Restrepo-Angulo, De Vizcaya-Ruiz and Camacho 2010). This toxin inhibits chloride channels in a reversible fashion. Chlorotoxin was reported to block outwardly rectifying chloride current and alter cell proliferation in astrocytomas. Another example of natural toxins is dendrotoxin from the venom of the snake *Dendroaspis augusticeps* (green mamba) which blocks potassium channels at nanomolar concentrations. Humans bitten by the green mamba develop symptoms including abdominal pain, sweating, and respiratory weakness. Environmental pollutants are another extrinsic source that targets ion channels. Dichloro-diphenyl-trichloroethane (DDT), a type of organochlorine pesticide, slows the rate of activation of voltage-gated

sodium channels causing long sodium currents in mammals. In addition to natural toxins and environmental pollutants, chemical molecules like estrogen (E2) have recently emerged as an important cellular regulator in the intestinal tract and a potential mediator of several ion channels. It has been long known that E2 plays an important role in a wide range of physiological and pathological processes. E2 mimickers such as phytoestrogens (i.e. soy supplementation) and synthetic chemicals can also exert major effects on estrogen receptor (ER) signaling, which, in turn, may alter the expression of genes important for cell development and ultimately oncogenesis. Specifically in GI malignancy, E2 has been shown to confer a protective effect (Hogan *et al.* 2009, Giroux, Betnatchez and Carrier 2011). Interestingly, a few studies have reported the link between E2 and KCNQ1, and E2 and CFTR. Treatment with E2 results in colonic Kcnq1 internalization (Rapetti-Mauss *et al.* 2013) and increased *Cftr* mRNA and protein expression in rats (Gholami, Muniandy and Salleh 2012). However, the link between these two ion channels and E2 has not been fully explored in the GI tract. Thus, further tissue-specific investigations will allow a better understanding of how this E2-KCNQ1/CFTR network regulates the normal development versus oncogenesis of the GI tract. In short, the evidence linking ion channels and toxicology strongly suggests them as excellent candidates to investigate the toxic effects of environmental factors and drug compounds, and as potential early indicators of life-threatening conditions such as cancer.

Ion channels and transporters have only been the focus in cancer research the past 15 years (Pedersen and Stock, 2013). Studies have shown that potassium channels may be involved in physiological and pathological cell proliferation. For example, the voltage-gated potassium channel Kv1.3 was reported to be important for T-lymphocyte proliferation (Wonderlin and Strobl, 1996). Ether a go-go 1 (EAG1) potassium channels were found to be overexpressed in crypt cells in a CRC mouse model produced by administration of dimethylhydrazine and N-methyl-N-nitrosourea (Ousingsawat *et al*. 2007). Importantly, early activity of plasma membrane ion channels seems to be a critical step during apoptosis, and mitochondrial ion channels are key players in the apoptotic

mitochondrial pathway. Due to the convenient location on the plasma membrane and the highly regulated nature of the opening and closing of most ion channels, they are considered as attractive drug targets. There are over 400 ion channel genes representing over 1.5% of the human genome, yet less than 10 ion channels are currently the target of available drugs; thus presenting a potential therapeutic opportunity.

KCNQ1 (Kv7.1)

The *KCNQ1* gene on chromosome 11 encodes for the pore-forming alpha subunit of a voltage-gated potassium channel which enables a K^+ current after electrical depolarization of the cell membrane. KCNQ1 is part of a family of five potassium channels: KCNQ1 to KCNQ5. All KCNQ members have six transmembrane domains and can form functional homomeric channels (Dedek and Waldegger, 2001). While KCNQ3 can form heterodimers with KCNQ2, KCNQ4 or KCNQ5, KCNQ1 does not associate with any other KCNQ subunits. Instead it associates with β-subunits of the KCNE family which have a single transmembrane domain and can markedly affect the gating of KCNQ1. A domain with residues 589-620 located near the COOH terminal of KCNQ1 is accountable for assembly specificity (Figure 12: Jespersen, Grunnet and Olesen 2005). Deletion of a part of this assembly domain leads to impaired channel complexes and mistrafficking. Similar to the KCNQ family, the KCNE gene family consists of five known members: KCNE1 to KCNE5. All KCNE subunits consist of a small one-transmembrane protein and an extracellular NH_2 terminal. Interestingly, several KCNE proteins not only interact with KCNQ1 but also with other a-subunits such as HERG and Kv3.4. It is believed that assembly of four KCNQ1 α-subunit proteins is required to form a functional channel.

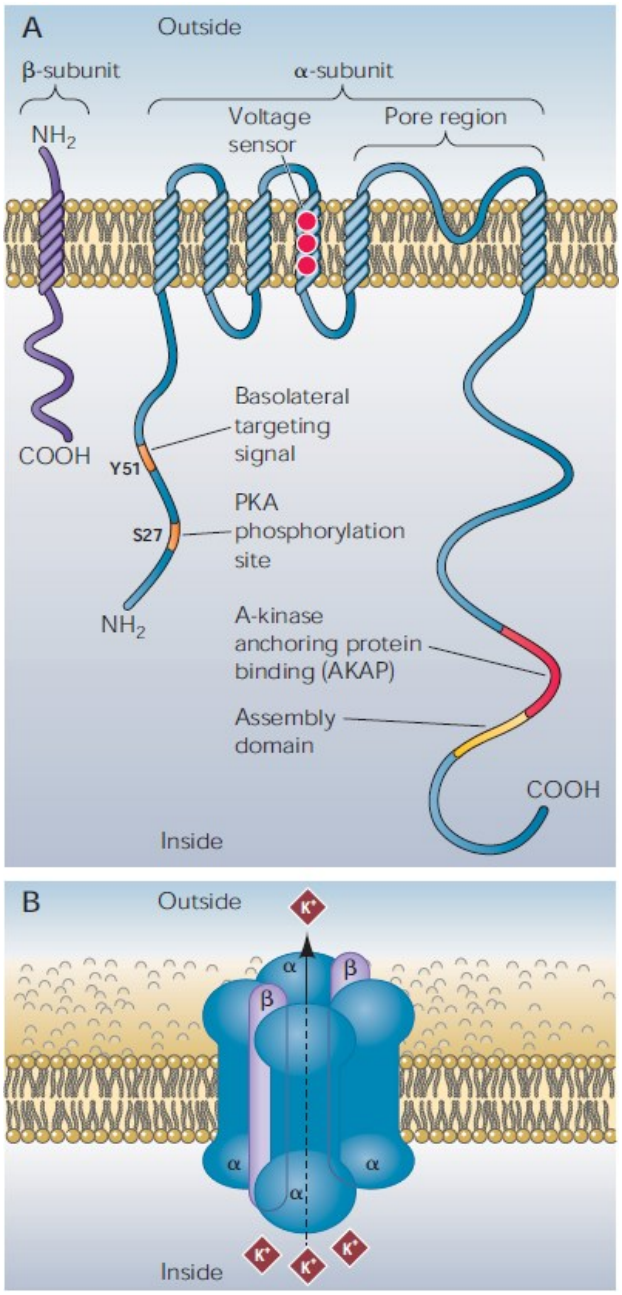


Figure 12. Structures of KCNE and KCNQ1 proteins. (Jespersen, Grunnet and Olesen 2005). **A)** Structure of the KCNE proteins as well as the delayed rectifier potassium channel KCNQ1 with indications of some of the domains important for regulation of the channel. All five KCNE proteins hold one transmembrane domain but have varying lengths of NH_2 and COOH termini. **B)** KCNQ1/KCNE channel architecture. Four KCNQ1 α-subunits assemble to form the basic channel. Coassembly with a number of KCNE β-subunits leads to a β-subunit-specific change in the current characteristics.

Assembling with different KCNEs causes activation and deactivation of KCNQ1. Homomeric KCNQ1 and KCNQ1/KCNE1 heteromers are almost completely closed at voltages more negative than –50 mV and opened at potentials positive to –20 mV. KCNE2, prominently expressed in gastric glands of the stomach (Dedek and Waldegger 2001), renders the KCNQ1 channel constitutively open. Less studied and not well-defined, KCNE5, identified as a KCNE1-like gene, assembles with KCNQ1 and only activates the channels at potentials >40 mV, indicating that these complexes are closed under physiological conditions (Dedek and Waldegger 2001, Palmer Br *et al.* 2012). KCNE4 β-subunit inhibits the KCNQ1 current in both *Xenopus* oocytes and mammalian cells at physiologically relevant potentials, whereas potentials >50 mV reveal currents with very slow activation kinetics (Teng *et al.* 2003, Grunnet *et al.* 2005). Unlike all other KCNQ1-KCNE assembles, association with KCNE3 accelerates gating kinetics and eliminates voltage dependence; thus KCNQ1/KCNE3 heteromers are constitutively open at the negative membrane voltages of intestinal epithelial cells and can be further activated by the action of cAMP. Overall, as presented, although the mechanisms by which KCNE proteins modulate KCNQ1 channels have been well studied, it is still not well understood how different KCNE proteins exert considerably different effects on KCNQ1 channels.

KCNQ1 is prominently expressed in myocardium, inner ear, airway, stomach, pancreas and intestine, tissues in which *KCNQ1* expression is critical for ion homeostasis. KCNQ1 expression is also found in the proximal and distal tubule of the nephron, though there remains controversy about the function of KCNQ1 channels in the kidney (Vallon *et al.* 2001, Neal *et al.* 2011).

In the heart, KCNQ1 plays a crucial role in repolarization of the cardiac tissue following an action potential. In cardiac myocytes, KCNQ1 forms heteromeric channels with KCNE1, which activate very slowly upon depolarization. The KCNQ1/KCNE current is mediated by β-adrenergic receptor activation, causing an increased level of

cAMP and thus PKA stimulation (Jespersen, Grunnet and Olesen 2005). Upon PKA activation, residual S27 on the NH2 terminal of KCNQ1 is phosphorylated. Inherited mutations of *KCNQ1* underlie congenital *long QT syndrome* (LQTS), a disorder caused by abnormal ventricular repolarization that increases the risk of sudden death from cardiac arrhythmias. In fact, the transmembrane region and 3'-end of *KCNQ1* was cloned by Wang and colleagues in 1996 using linkage analysis of LQTS1 patients (Wang *et al.* 1996). Dominant negative mutations in either *KCNQ1* or *KCNE1* causes prolonged cardiac action potentials in the Romano-Ward form of LQTS. Mutations in *KCNQ1* are also associated with *Jervell and Lange-Nielson syndrome* (JLNS), a rare autosomal disorder characterized by deafness in addition to LQTS. Recent reports indicate that JLNS patients can also be susceptible to gastrointestinal defects including iron-deficiency anemia, gastric and duodenal hyperplasia, elevated gastrin levels, and more rarely, gastric adenocarcinoma (Wimbo *et al.* 2012, Rice *et al.* 2011).

Another interesting function of KCNQ1 is seen in epithelial cells of the inner ear. In the cochlea, KCNQ1 is co-expressed with KCNE1 in the apical membrane of marginal cells of the stria vascularis (Neyroud *et al.* 1997). The role of the KCNQ1/KCNE1 channel here is to secrete potassium to the endolymph of the scala media, which is essential for proper hearing. Hearing loss and inner ear abnormalities have been reported in both a mouse model with a dysfunctional KCNQ1/KNCE1 channel complex and human samples with a *KCNE1-p.85N* (susceptibility) variant (Rivas and Francis 2005, Van Laer *et al.* 2006).

In the airway, KCNQ1 mediates a basolateral K^+ conductance that exerts an important role in maintaining cAMP-dependent Cl^- secretion. KCNQ1 assembles with KCNE1 and the channels have also been suggested to regulate cell volume (Mall *et al.* 2000, Lock and Valverde 2000). However, a different study by Grahammer *et al* reported that KCNE1 was not detected in mouse tracheal epithelial cells and that instead a complex consisting of KCNQ1 and KCNE3 formed a basolateral K^+ channel (Grahammer *et al.* 2001). In 2010, Preston and colleages disrupted KCNE3 in a mouse model, examined the effect on KCNQ1 activity and concluded that KCNE3 is important

for ion transport across tracheal epithelia although Ca^{2+}-activated K^+ channels could substitute partially for KCNQ1/KCNE3 (Preston *et al.* 2010). Regardless of the dispute over KCNE3's importance, all studies agree that the KCNQ1 channel seems to be the dominant basolateral K^+ conductance mediator in airway epithelia.

Though less well studied, KCNQ1 was found to co-express with KCNE1 in the proximal tubule and distal tubule of nephrons (Vallon *et al.* 2001). *Kcnq1* knockout mice suffer from hypokalemia, urinary and fecal salt wasting as well as volume depletion indicating an important role of the KCNQ1 complex in renal function. KCNQ1 in the pancreas presents another interesting role in cell regulation. KCNQ1 is found in the insulin-secreting beta cells. Inhibiting KCNQ1 channels increased insulin secretion while promoting KCNQ1 impaired insulin secretion by causing premature repolarization of the action potential (Ullrich *et al.* 2005).

In the stomach, KCNQ1 predominently assembles with KCNE2 in gastric glands and is coexpressed with an H^+/K^+-ATPase in acid-secreting parietal cells (Dedek and Waldegger 2001, Roepke *et al.* 2010). KCNQ1-KCNE2 was found to give rise to a time independent current with nearly linear voltage dependence. Both KCNQ1 and KCNE2 are expressed at or near the parietal cell apical membrane. It is suggested that KCNQ1-KCNE2 channels may be necessary for gastric secretion by recycling potassium transported by the K^+/H^+ ATPase across the apical membrane. *Kcnq1*$^{-/-}$ mice developed metaplasia, dysplasia and premalignant adenomatous hyperplasia in the stomach independent of infection; and *Kcne2*$^{-/-}$ mice exhibited gastritis cystic profunda, increased proliferative markers in gastric mucosa and progressive gastric hyperplasia (Roepke *et al.* 2010). The same study also reported reduced parietal cell KCNE2 expression in human gastric cancer tissue.

In the small intestine and colon, KCNQ1's major role is suggestedly to regulate water and salt transport in epithelial tissues via recycling potassium that is transported into the cell by basolateral NaK2Cl cotransporters and Na/K-ATPases. KCNQ1 forms constitutively open heteromultimeric channels with KCNE3 (Schroeder *et al.* 2000,

Preston *et al.* 2010, Nakajo *et al.* 2011). It was found that KCNQ1 and KCNE3 colocalize at the basolateral membrane of crypt cells of both the small intestine and the colon (Dedek and Waldegger 2001). In mouse small intestine, *Kcnq1* expression is strongest in the duodenum and proximal jejunum and decreases towards the ileum. In the large intestine, *Kcnq1* expression is strongest in distal colon. The full spectrum of functions of KCNQ1 in the GI tract remains to be explored. It has been suggested that similar to CFTR, KCNQ1 channels express in the colonic crypts (Rapetti-Mauss *et al.* 2013) although this co-localization and mechanism of action is not definitively known. A few studies have suggested that KCNQ1 promotes basolateral potassium secretion which in turn acts as a driving force for apical chloride secretion by CFTR (O'Mahony *et al.* 2009, Bajwa *et al.* 2009), which is important for the hydration and protective function of the intestinal mucus layer. Disruption of this process potentially could lead to inflammation and alterations in other important biological processes, a predisposition to CRC. A recent study has proposed that basolateral intermediate conductance K^+ channels may instead coexist with low conductance KCNQ1 channels to drive Cl^- secretion (Linley *et al.* 2013).

Six mutant mouse alleles of *Kcnq1* have been created that partially model human disorders arising from inherited mutations in *KCNQ1*, especially JLNS. These mutant alleles, generated by traditional gene targeting, chemical mutagenesis and identification of spontaneous mutations, share similar but not identical phenotypes, that include bilateral deafness from birth, gastric hyperplasia, gastric achlorhydria, elevated serum gastrin, headbobbing, bi-directional circling, ataxia, body tremor, and hyperactivity with a consequent significant decrease in body weight. Conflicting results were reported regarding a cardiac phenotype in these models.

It should be noted that *Kcnq1* is an imprinted gene during early embryogenesis. It is only expressed from the maternal allele but becomes biallelic during fetal heart development. This transition coincides with conformational changes involving the interaction between the *Kcnq1* promoter and heart-specific enhancers. The human *KCNQ1* gene is located in a region of chromosome 11 that contains a large number of

contiguous genes that are abnormally imprinted in cancer (i.e. *HRAS* in bladder cancer). The telomeric region of *Kcnq1* contains at least eight maternally expressed genes that include *Ascl2* (an essential placenta-specific transcription factor and also a transcription factor essential for the intestinal stem cell compartment), and *Cdkn1c* (a cyclin-dependent kinase inhibitor). The extent of imprinting of these genes is highly variable and not fully understood. *Ascl2* was found to be imprinted only in the placenta while *Cdkn1* is expressed from the maternal allele in all expressing tissues and at all stages of development (Guillemot *et al.* 1995, Gould and Pfeifer 1998).

The distal region of the *Kcnq1* locus also includes *Kcnq1ot1* which is an imprinted non-coding RNA transcribed from the paternal allele (of the ~1MB *Kcnq1*-imprinted domain), emerging from intron 11 of the *Kcnq1* gene in an antisense direction (Mancini-Dinardo *et al.* 2006, Lee and Bartolomei 2013). In contrast to most long ncRNA, *Kcnq1ot1* is highly expressed and detectable. Its transcription has a role in repressing several neighboring genes in *cis*. For example, silencing of placentally imprinted genes involves the function of several histone methyltransferase systems including G9a and both polycomb complexes. It is suggested that *Kcnq1ot1*'s imprinting mechanism occurs through the production of the transcript itself or its elongation past the engineered terminational signal 1.5kb downstream of the transcription start site. While the ncRNA is not responsible for imprinted expression of *Kcnq1* in the heart during heart development, it is reported to have a role in silencing *Kcnq1* in tissues other than heart.

Overall, ~ 300 *KCNQ1* mutations have been identified in human patients with various diseases. *KCNQ1* is a frequent target of germline chromosomal rearrangements in Beckwith-Wiedemann syndrome (BWS) which causes prenatal overgrowth and predisposition to childhood cancers (Netchine *et al.* 2013, Kaltenback *et al.* 2013). Variants in *KCNQ1* are also associated with enhanced risk for type-2 diabetes (Unoki *et al*. 2008, Yasuda *et al*. 2008, Wang *et al*. 2013). Mutations in *KCNE* family also contribute to several human abnormalities. Variants in *KCNE2* may cause human cardiac arrhythmia by altering HERG currents (Abbott 1999). In addition, *Kcne2*$^{-/-}$ mice showed reduced cardiac currents (Roepke *et al*. 2008). Polymorphism rs697829 in *KCNE5* was

identified with QT interval and survival in acute coronary syndromes (Palmer *et al.* 2012). Notably until the work described in chapter 2, relatively little was known about the role of *KCNQ1* in intestinal cancer, but a potential role in GI tract cancers was indicated by its identification as a high frequency CIS locus (Figure 13) in *Sleeping Beauty* (*SB*) DNA transposon-based forward mutagenesis genetic screens for intestinal (Starr *et al*. 2009, Starr *et al.* 2011, March *et al.* 2011) and pancreatic cancers (Perez-Mancera *et al.* 2012, Mann *et al.* 2012).

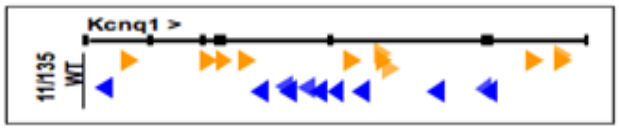


Figure 13. CIS map of *Kcnq1* in *Apc*^{wt} screen. Yellow and blue markers represent sites of transposon insertions on forward and reserve strands in GI tumors.

CFTR

The human *CFTR* gene covers ~250 kb on chromosome 7. The *CFTR* gene encodes for the protein CFTR, a member of the ATP Binding Cassette (ABC) family which includes several clinically important proteins such as P-glycoprotein (P-gp), multidrug resistance associated protein and the TAP transporters. ABC family members perform several tasks, including energy-dependent transmembrane transport of large molecules, regulation of other transmembrane transporters, and the transport of ions across membranes. Unlike other members of the ABC family, CFTR is not involved in active transport. It is an anion channel that uses the energy of ATP hydrolysis to transit through conducting and non-conducting conformations. Ion channels are equally capable of allowing ions to flow either into or out of cells, enabling the CFTR anion channel to play crucial roles in both absorption and secretion.

CFTR channels are found primarily at the apical surfaces of epithelial cells in multiple organs including airways, pancreas, stomach, intestine, kidney and testis (Strong, Boehm and Collins 1994, Hodges *et al.* 2008, deHostos *et al.* 2011) where they regulate chloride secretion, and in the sweat glands where they modulate the reabsorption of chloride from sweat. CFTR's role is to maintain transepithelial fluid homeostasis via its control of Cl^- flow and water movement from the cells through osmotic driving force. CFTR is not the sole pathway for apical Cl^- exit but it is the predominant factor for Cl^- secretion upon being stimulated by cAMP-PKA dependent agonists. In the intestine, CFTR is localized to the crypt epithelium where it regulates both electroneutral and electrogenic absorption of electrolytes in intestinal epithelium. The secretion and absorption of electrolytes may facilitate the transport of mucus out of the crypts and maintain hydration of mucus. In a recent study using intact human colonic crypts, Cl^- secretion was reported to originate from the dominant population of colonocytes expressing apical CFTR (Linley *et al.* 2013). Specifically in the duodenum, CFTR is postulated as an important pH regulator. Duodenal pH is controlled by bicarbonate-rich pancreatic secretions and intestinal bicarbonate secretion which neutralize acid from the stomach. Thus, CFTR seems to also function as a bicarbonate channel, even at high concentrations under critical conditions.

The CFTR protein consists of two, six span membrane bound regions each connected to a nucleotide-binding factor which binds ATP (Figure 14, source: OxfordU Med). Between these two units is an R (regulatory)-domain which consists of many charged amino acids. The R-domain is a unique feature of CFTR within the ABC superfamily (Ostedgaard *et al.* 2011, Sebastian *et al.* 2013). The R-domain contains several strong consensus sites for phosphorylation by cAMP-dependent PKA or PKC. The activity of CFTR as an ion channel depends upon phosphorylation of the R domain and binding of ATP to the nuclear binding domains. (When PKA phosphorylates the R region, this enables the gating of CFTR.) The N terminal portion of the R domain is highly conserved between species while there is a lower degree of conservation between the rest of the domain. About 4 % of the CFTR protein is found in the extracellular loops

while 19% of the CFTR protein makes up the twelve transmembrane domains. Sequences within the intracellular loops (ICL1 - 4) have been shown to be important for the processing of CFTR and correct localization to the cell membrane.

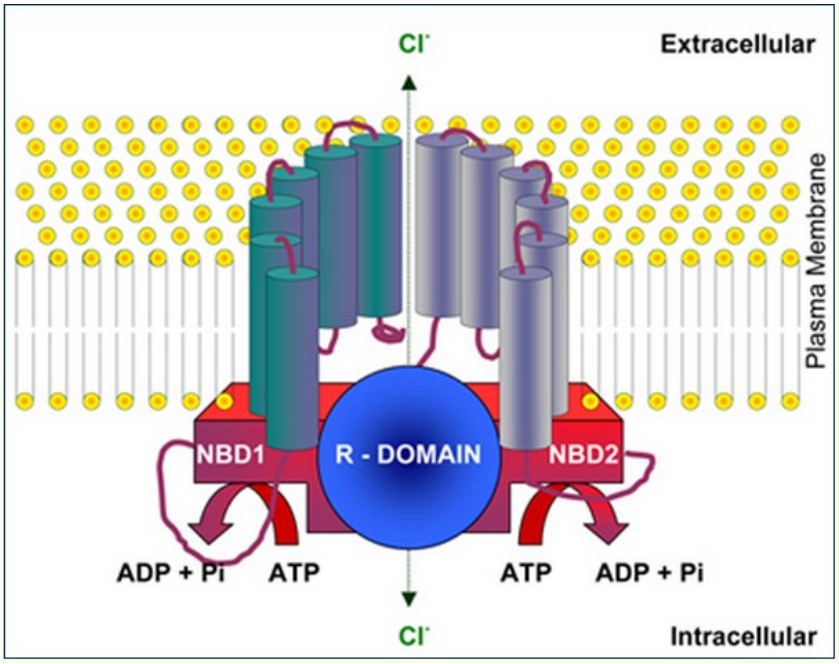


Figure 14. Proposed model of CFTR structure in the cell membrane. (OxfordU Med)

In 1989, CFTR was identified (Riordan *et al*. 1989) as the mutation which causes cystic fibrosis (CF) - an autosomal recessive disorder that occurs in ~1 in every 3000 Causasian live births (Hodges *et al.* 2008). Many of the mutations identified in CF occur in the first nucleotide-binding domain (NBD1). Although there are >400 known *CFTR* mutations, a single mutation - a deletion of the phenylalanine at position 508 (F508Δ) in exon 10 in NBD1 - accounts for about 70% of all CF chromosomes worldwide. In the United States, 48% of patients are homozygous for F508Δ and 40% are heterozygous for F508Δ (Cystic Fibrosis Foundation 2011). This class 2 *CFTR* gene mutation (further explained in the pancreas section) produces an abnormal protein which

does not escape the endoplasmic reticulum for further processing and trafficking of CFTR channels to the cell membrane. Having two copies of this mutation (one inherited from each parent) is the leading cause of CF. The basic defect in CF is a decrease in chloride conductance across apical membranes accompanied by an increase in the uptake of sodium ions (When active, the NBD1 region of CFTR has an inhibitory effect on the epithelial sodium channel).

Currently, several CF clinical trials are being conducted to target the basic defect of CFTR channels via three mechanisms: inducing a premature stop codon readthrough, correcting membrane trafficking of CFTR (corrector) and amplifying CFTR conductance (potentiator). For example, the potentiator VX-770 (Ivacaftor) has successfully passed phase 3 clinical trial for treating CF patients with a CFTR-G551D mutation. Another trial in phase 2 is testing a combination therapy of VX-770 and the corrector VX-809 for patients with a CFTR-F508del mutation (Dekkers *et al.* 2013).

Due to CFTR's crucial role in fluid homeostasis of diverse tissues, clinical features of CF range from meconium ileus to pancreatic exocrine insufficiency, liver disease, chronic lung infection, GI complications and shortened lifespan (Hostos *et al.* 2011). In fact, CF was first reported in 1938 by a New York pathologist, Dorothy Andersen, who described autopsy findings of 49 children who died of pancreatic and airway destruction instead of previously presumed celiac disease and vitamin A deficiency (Andersen D. 1938). Since then, additional reported symptoms have included blockage and eventual degeneration of the vas deferens in males, dehydrated cervical mucus and a failure of the mucus to show appropriate hydration during ovulation in females, and greatly elevated concentrations of salt in the sweat.

Approximately 10% of children with CF experience meconium ileus, an intestinal blockage caused by impaction of the meconium (fetal form of feces). This is manifested shortly after birth and can be fatal unless corrected. The cause seems to be inadequate secretion of fluid by CFTR-expressing crypt cells in the intestine. The failure of

chloride-mediated fluid then causes mucus secretion to be abnormally viscous, thus retarding the normal movement of the contents of the intestines.

Disruption of CFTR in CF also affects pancreatic function. CFTR is present in the duct cells of the exocrine pancreas where it couples with anion exchangers SLC26A3 and SLC26A6 to generate $NaHCO_3$ secretion to neutralize the duodenal lumen. This also allows the delivery of pancreatic digestive enzymes to assist the food breakdown process in the duodenum. ~ 85% of CF patients at the US Cystic Fibrosis Foundation centers take exogenous pancreatic supplements on a regular basis. Disruption of pancreatic function presents a powerful biomarker tool to determine the severity of CF. Patients can be subcategorized into two groups: one with pancreatic sufficiency (PS) and one with pancreatic insufficiency (PI). Patients with PI tend to have class 1-3 "severe" mutations which result in absent or nonfunctional CFTR at the cell surface, whereas patients with PS are likely to have class 4-5 mutations which result in partially functional CFTR (Gelfond and and Borowitz 2012). Reduced fluid and anion secretion is seen in the pancreas of both PS and PI patients. Currently, an enzyme-linked immunosorbent assay for fecal elastase has been used as a good indicator of PI with a sensitivity of 98-100% and a specificity of 93-100%. Concurrently, a serum test for trypsinogen is also used to diagnose PI in CF patients older than 8. While the secretory defect is more pronounced in the PI group, PS patients have a high risk of developing pancreatitis. The common pancreatic phenotype in PS patients is decreased ductal flow that over time leads to scattered ductal obstruction, enzymatic breakdown with trypsin activation and local inflammation. When the condition worsens, local inflammation can lead to systemic inflammation that may cause multisystem organ failure.

Recently reported by the Cystic Fibrosis Registry, liver-related mortality in CF patients increased 28% from 2008 to 2010, presenting liver disease as an important clinical feature of CF. Various liver disorders in CF patients include cholelithiasis, microgallbladder, hepatic steatosis, nodular hyperplasia and focal biliary cirrhosis. It seems that liver disease in the CF population occurs mostly during the first 10 years of life and declines in the second decade. In addition to the *CFTR* mutation, envonromental

factors and genetic modifiers can greatly affect the phenotype and onset of liver disease. For example, a study on genetic modifiers in the subset of 124 CF patients with severe liver disease reported the *SERPINA1 Z* allele to be associated with the greatest risk (Bartlett *et al.* 2009). Most biomarkers for liver disease in CF have not been validated to guide clinical decisions and liver transplantation remains the definitive approach to the management of CF patients with terminal-stage liver disease.

In the airway epithelia, CFTR plays an important role in mucus and bacterial clearance, thus the most prominent CF phenotype is airway obstruction. Interestingly, CF patients that underwent lung transplantation had significantly smaller airways than patients with chronic obstructive pulmonary disease, and the density of small airways in these patients even decreased with age (Tiddens *et al.* 2013). In a normal condition, anion secretions provides the airways with a thin airway surface liquid layer adjacent to the apical membrane and a high-viscosity gel layer where secreted mucins trap debris and microorganisms. In CF, defective or absent CFTR leads to a dysfunctional anion secretion which affects the formation and function of these protective layers, subsequently leading to the accumulation of mucus in the airways. Disruption in the protective mucus layer can potentially cause bacterial colonization in the lung, such as *Pseudomonas aeruginosa* and *Staphylococcus* (Tiddens *et al.* 2013, Cohen-Cymberknoh *et al.* 2013). This is further worsened by the loss of innate immune factors present in the anion secretion (Frizzel and Hanrahan 2012) and the increased production of proinflammatory mediators in response to pathogens. Ultimately CF-related dysfunction of the lung leads to an exaggerated airway inflammation while the protective immune system fails to eradicate invading pathogens. Once advanced, chronic bacterial infection can progressively destroy the lung and ultimately ends with respiratory failure.

Though less common, adult CF patients also experience GI complications. In the mid-to-upper GI tract region, gastroesphophageal reflux disease (GERD) and reflux symptoms are seen in ~30% of adult CF patients. Some phenotypes such as heartburn, acid brash and dysphagia are experienced on a daily basis by 24% of patients. The primary mechanism by which GERD occurs is inappropriate transient lower esophageal

sphincter relaxation (TLESR) that causes gastric contents to move up to the esophagus and thus elicits discomfort and pain. Interestingly, there is a well-recognized connection between GERD and chronic obstructive pulmonary disease (Rabinovich and MacNee 2011), suggesting a likely important role of ion channels (ie. CFTR) in these tissues. Abnormalities are also observed in the lower GI tract. In normal intestine, the secretion of salt and water is essential for the maintenance of a functional luminal environment which supports enzyme activities, nutrient absorption and toxic waste clearance. A major intestinal CF phenotype caused by disruption of CFTR is that less fluid gets secreted than normal, potentially leading to intestinal blockage by improperly hydrated stools. More specifically, CF patients of all ages experience distal intestinal obstruction syndrome (DIOS) with European studies, reporting that up to 7 times more adult CF patients develop DIOS than pediatric patients. DIOS is characterized by an incomplete or complete fecal obstruction of the ileocecum, due to the accumulation of fecal materials that strongly attach to the villi and crypts of the intestinal mucosa. This explains chronic constipation experienced by CF adult humans. In addition to intestinal blockade, about 30-55% of CF patients also have small intestinal bacterial overgrowth caused by a combination of factors such as thick mucus secretion and the use of acid-suppressing agents. Once overgrowth of bacteria occurs, toxic byproducts are produced, which subsequently can lead to enterocyte damage and malabsorption. Interestingly, the incidence of inflammatory bowel disease (IBD) is also markedly higher in CF patients.

The first mouse model of CFTR deficiency was created in 1992 (Snouwaert *et al.* 1992), where inactivation of *CFTR* in murine embryonic stem cells was achieved by inserting an in frame stop codon in the coding sequence of exon 10. This designated \$489X mutation of CFTR created phenotypes similar to those with the F508Δ mutation in human patients. Since then, there have been several other mouse models of CFTR deficiency created, including an improved model of FABP-h*CFTR* to correct the lethal intestinal abnormalities (Zhou *et al.* 1994). Most recently, researchers at Case Western Reverse University generated mice with a conditional null allele of *Cftr*, $Cftr^{fl10}$, in which loxP sites were inserted around exon 10 of *Cftr* (Hodges *et al.* 2008). (These mice were

used for the studies in Chapter 3). In this conditional homozygous CF knockout mouse model, a more severe intestinal phenotype is exhibited resulting in peri-natal lethality. Inspection of the intestines of dead CF mice indicates the presence of a mucus-fecal mixture and considerable inflammation. These effects can be prevented or reversed with administration of an oral osmotic laxative such as Colyte.

Newly explored is the link between CF and GI cancer. Due to its important role in the process of ion and acid-base transport required for normal GI physiology, dysregulation of CFTR may potentially lead to GI cancer. For example, in the proximal small intestine, CFTR's role is to regulate transepithelial export of Cl^- and $HCO3^-$, important for the maintenance of proper pH levels. When this process is disrupted, the proximal small intestine is now flooded with highly acidic material released from the stomach, subsequently creating a thriving environment for harmful bacteria. Chronic colonization of bacteria can cause intestinal inflammation, a predisposition to intestinal cancer. Another potential connection between CFTR deficiency and cancer development lies in its role of regulating water balance and intestinal mucus clearance. In the large intestine, disruption of CFTR cause water dehydration and intestinal blockade, which can lead to severe sepsis - a high risk for developing intestinal cancer. These hypotheses of how CFTR deficiency may lead to cancer development are supported by several lines of epidemiological evidence. A recent epidemiological study reported CF patients to be at higher risk for esophageal, gastric, hepatobiliary gall bladder and intestinal cancers (Maisonneuve *et al.* 2013). It is suggested that other factors such as defective mucus function, excessive acid exposure and deficiencies, can increase the likelihood of GI tumors in CF patients. Though there is yet a national consensus on a definitive causation of GI cancer by CF, physicians are advised to be alert of this possibility in older CF patients. Therefore, it is critical to investigate the direct role of *CFTR* in GI tract cancers. Evidence for a potential role of CFTR in GI cancers was its identification as a CIS gene (Figure 15) in multiple SB mutagenesis screens (Starr *et al.* 2009, March *et al.* 2011).

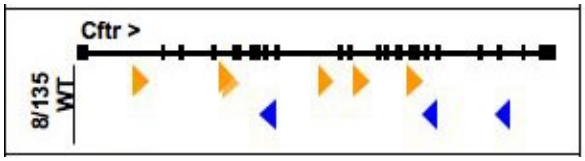


Figure 15. CIS map of *Cftr* in *Apc*^{wt} screen. Yellow and blue markers represent sites of transposon insertions on forward and reserve strands in GI tumors.

CHAPTER TWO

The role of *KCNQ1* in mouse and human gastrointestinal cancers

BLN Than[1,2], JACM Goos[3], AL Sarver[4], MG O'Sullivan[5], A Rod[1], TK Starr[6,7],RJA Fijneman[3], GA Meijer[3], L Zhao[1], Y Zhang[8], DA Largaespada[7], PM Scott[1] and RT Cormier[1]

[1]Department of Biomedical Sciences, University of Minnesota Medical School, Duluth, MN, USA; [2]Toxicology Graduate Program, University of Minnesota, Duluth, MN, USA; [3]Department of Pathology, VU University Medical Center, Amsterdam, The Netherlands; [4]Department of Biostatistics and Informatics, Masonic Cancer Center, University of Minnesota, Minneapolis, MN, USA; [5]College of Veterinary Medicine, University of Minnesota, St Paul, MN, USA; [6]Department of Genetics, Cell Biology and Development, Center for Genome Engineering, Masonic Cancer Center, University of Minnesota, Minneapolis, MN, USA; [7]Department of Obstetrics, Gynecology and Women's Health, Masonic Cancer Center, University of Minnesota Medical School, Minneapolis, MN, USA and [8]University of Minnesota Supercomputing Institute, Minneapolis, MN, USA.

Than BL, Goos JA, Sarver AL, O'Sullivan MG, Rod A, Starr TK, Fijneman RJ, Meijer GA, Zhao L, Zhang Y, Largaespada DA, Scott PM, Cormier RT. The role of KCNQ1 in mouse and human gastrointestinal cancers. *Oncogene*. 2013 Aug 26. doi: 10.1038/onc.2013.350.

Abstract. *Kcnq1*, which encodes for the pore-forming a-subunit of a voltage-gated potassium channel, was identified as a gastrointestinal (GI) tract cancer susceptibility gene in multiple Sleeping Beauty DNA transposon-based forward genetic screens in mice. To confirm that *Kcnq1* has a functional role in GI tract cancer, we created Apc^{Min}

mice that carried a targeted deletion mutation in *Kcnq1*. Results demonstrated that *Kcnq1* is a tumor suppressor gene as *Kcnq1* mutant mice developed significantly more intestinal tumors, especially in the proximal small intestine and colon, and some of these tumors progressed to become aggressive adenocarcinomas. Gross tissue abnormalities were also observed in the rectum, pancreas and stomach. Colon organoid formation was significantly increased in organoids created from *Kcnq1* mutant mice compared with wild-type littermate controls, suggesting a role for *Kcnq1* in the regulation of the intestinal crypt stem cell compartment. To identify gene expression changes due to loss of *Kcnq1*, we carried out microarray studies in the colon and proximal small intestine. We identified altered genes involved in innate immune responses, goblet and Paneth cell function, ion channels, intestinal stem cells, epidermal growth factor receptor and other growth regulatory signaling pathways. We also found genes implicated in inflammation and in cellular detoxification. Pathway analysis using Ingenuity Pathway Analysis and Gene Set Enrichment Analysis confirmed the importance of these gene clusters and further identified significant overlap with genes regulated by MUC2 and CFTR, two important regulators of intestinal homeostasis. To investigate the role of KCNQ1 in human colorectal cancer (CRC), we measured protein levels of KCNQ1 by immunohistochemistry in tissue microarrays containing samples from CRC patients with liver metastases who had undergone hepatic resection. Results showed that low expression of KCNQ1 expression was significantly associated with poor overall survival.

INTRODUCTION

The *KCNQ1* gene encodes for the pore-forming a-subunit of a voltage-gated potassium channel that enables a K^+ current after electrical depolarization of the cell membrane. KCNQ1 is predominantly expressed in the myocardium, inner ear, stomach, intestine and pancreas, tissues in which KCNQ1 expression is critical for ion homeostasis (Peroz *et al.* 2008). Inherited mutations in *KCNQ1* underlie several human disease syndromes. For example, mutations in *KCNQ1* are associated with congenital long QT syndrome, a disorder caused by abnormal ventricular repolarization that increases the risk

of sudden death from cardiac arrhythmias. Jervell and Lange–Nielson syndrome is a rare autosomal disorder characterized by deafness in addition to long QT syndrome. Recent reports indicate that Jervell and Lange–Nielson syndrome patients can also be susceptible to GI defects, including iron-deficiency anemia, gastric and duodenal hyperplasia, elevated gastrin levels and, more rarely, gastric adenocarcinoma (Winbo *et al.* 2012, Grahammer *et al.* 2001, Rice *et al.* 2011, Nikou *et al.* 2011). Variants in *KCNQ1* are also associated withenhanced risk for type 2 diabetes (Unoki *et al.* 2008, Yasuda *et al.* 2008).

Relatively little is known about the role of *KCNQ1* in cancer, but a potential role in GI tract cancers was indicated by its identification as a high-frequency common insertion site locus in Sleeping Beauty DNA transposon-based forward mutagenesis genetic screens for intestinal (Starr *et al.* 2009, March *et al.* 2011) and pancreatic cancer (Mann *et al*. 2011, Perez-Mancera *et al*. 2012). In our forward genetic screen in *Apc* wild-type mice, we found *Kcnq1* mutations in 18 intestinal tumors (13%) (Starr *et al.* 2009) ranking it among the top 10 common insertion site genes. In this study, while transposon insertions in *Kcnq1* were found in tumors from all regions of the intestine, approximately two-thirds (64%) were located in the duodenum and jejunum, with 26% in the ileum and 10% in the colon. Furthermore, transposon insertions in *Kcnq1* were observed equally in both the forward and reverse strand orientation, consistent with a model whereby the transposon is acting to disrupt gene function. We performed a second forward genetic screen in Apc^{Min} mice and found *Kcnq1* transposon insertions in 13 intestinal tumors (14%) (Starr *et al.* 2011). Another screen in Apc^{Min} mice found *Kcnq1* transposon mutations in 120 intestinal tumors (27%), ranking it no. 15 out of 641 common insertion site genes (120 kb window) (March *et al.* 2011). We have also conducted a screen using p53R270H mutant mice and identified insertions in *Kcnq1* in 15 intestinal tumors (23%) (TK Starr, personal communication). Six mutant mouse alleles of *Kcnq1* have been created that partially model human disorders arising from inherited mutations in *KCNQ1*, especially Jervell and Lange–Nielson syndrome (Lee *et al.* 2000, Casimiro *et al.* 2001, Casimiro *et al.* 2004, Takagi *et al.* 2007, Elso *et al.* 2004). These mutant alleles share similar but not identical phenotypes, including bilateral deafness

from birth, gastric hyperplasia, gastric achlorhydria, elevated serum gastrin, headbobbing, bidirectional circling, ataxia, body tremor and hyperactivity with a consequent significant decrease in body weight. Conflicting results were reported regarding a cardiac phenotype in these models. In mouse small intestine, *Kcnq1* expression is strongest in the duodenum and proximal jejunum and decreases towards the ileum (Demolombe *et al.* 2001, Dedek *et al.* 2001). In the large intestine, *Kcnq1* expression is strongest in the distal colon (Vallon *et al*. 2005, Warth *et al.* 2002). To further investigate the role of *Kcnq1* in GI tract cancers, we measured GI cancer phenotypes following introgression of a targeted knockout (KO) allele of *Kcnq1* (Casimiro *et al.* 2001) into the Apc^{Min} model of intestinal tumorigenesis. We then investigated the role of KCNQ1 in progression of human colorectal cancer (CRC) by measuring KCNQ1 protein levels in tissue microarrays containing samples from more than 500 CRC patients with liver metastases. Results from both of these studies indicate that *KCNQ1* is a tumor suppressor.

RESULTS

Loss of *Kcnq1* enhances tumor multiplicity in Apc^{Min} mice.

Haploinsufficiency for *Kcnq1* significantly enhanced Apc^{Min} intestinal tumor multiplicity overall, in both males and females, by ~40% (***Table 1***), with the phenotype strongest in the proximal small intestine and colon, where tumors increased 2-fold. $Kcnq1^{-/-}$ mice in our study manifested a range of previously reported phenotypes: rapid bidirectional circling, ataxia, tremor, head bobbing and they were smaller ($\downarrow$ ~ 25%, body weight) and leaner than littermate *Kcnq1* wild-type and heterozygous mice. Less than half of the expected number of Apc^{Min} $Kcnq1^{-/-}$ mice survived until weaning (23 vs ~60). Survivor Apc^{Min} $Kcnq1^{-/-}$ mice demonstrated a very strong increase in tumors in the proximal half of the small intestine (B3-fold), especially in the proximal quarter of the small intestine containing the duodenum and proximal jejunum (***Table 2***), the region exhibiting strongest expression of *Kcnq1* in the intestine. Apc^{Min} $Kcnq1^{-/-}$ mice developed B2-fold increase in colon tumors but no increase in the distal small intestine of males, and a decrease in females. No significant differences between *Kcnq1* genotypes in tumor sizes were observed.

Table 1. Loss of *Kcnq1* enhances tumor multiplicity in Apc^{Min} mice

	N	*Colon*	*Proximal SI*	*Distal SI*
Males				
$Kcnq1^{+/+}$	30	1.8 ± 1.7	11 ± 8	56 ± 41
$Kcnq1^{+/-}$	34	3.8 ± 2.3*	18 ± 13*	76 ± 38*
$Kcnq1^{-/-}$	10	2.6 ± 2.8	36 ± 16*	52 ± 43
Females				
$Kcnq1^{+/+}$	26	1.2 ± 1.7	14 ± 9	58 ± 25*
$Kcnq1^{+/-}$	46	2.5 ± 2.2*	20 ± 12	84 ± 37*
$Kcnq1^{-/-}$	13	2.1 ± 2.0	22 ± 11*	30 ± 16

Abbreviation: SI, small intestine. Apc^{Min} $Kcnq1^{+/+}$, Apc^{Min} $Kcnq1^{+/-}$ and $Apc^{Min}Kcnq1^{-/-}$ mice were killed at 120 days of age and tumors counted and measured. *P-values: <0.05; two-sided P-values for tumor counts were determined by use of the Wilcoxon rank-sum test comparing gender- and age-matched classes produced in the same genetic crosses.

Table 2. Apc^{Min} $Kcnq1^{-/-}$ tumor phenotype is strongest in the proximal quarter of the small intestine

Males	*Tumors*	*Females*	*Tumors*
$Kcnq1^{+/+}$	4 ± 2	$Kcnq1^{+/+}$	5 ± 3
$Kcnq1^{+/-}$	6 ± 3	$Kcnq1^{+/-}$	8 ± 4
$Kcnq1^{-/-}$	21 ± 11*	$Kcnq1^{-/-}$	15 ± 9*

Mice and tissues are the same as Table 1. *P-values: <0.05; two-sided P-values for tumor counts were determined by use of the Wilcoxon rank-sum test comparing gender- and age-matched classes produced in the same genetic crosses.

Loss of *Kcnq1* promotes tumor progression

Tumors in Apc^{Min} mice are invariably benign adenomas that usually vary in size between 0.3 and ~5 mm. Adenocarcinomas are extremely rare. In Apc^{Min} mice carrying either heterozygous or homozygous mutations in *Kcnq1* development of adenocarcinomas occurred at a frequency of ~10–15%. These adenocarcinomas appeared

as large lesions (some >20 mm), located in the proximal small intestine in a region approximate with the duodenal–jejunal flexure at the head of the pancreas. In almost all cases, there was buckling of the intestinal tube, in two cases there was a pronounced intussusception of the tissue and in two cases the tumors extended across the entire inner surface of the intestinal lumen. Six of these large tumors were analyzed by histopathology and all were found to be invasive adenocarcinomas (***Figure 16***). ***Figure 16a*** depicts a lesion of > 2 cm in length from the duodenum of an Apc^{Min} $Kcnq1^{+/-}$ mouse, which when sectioned regionally revealed adenocarcinoma tissue in each section. No adenocarcinomas were detected in the remainder of the small intestine or in the large intestine, and no putative adenocarcinomas were detected in Apc^{Min} $Kcnq1^{+/+}$ mice.

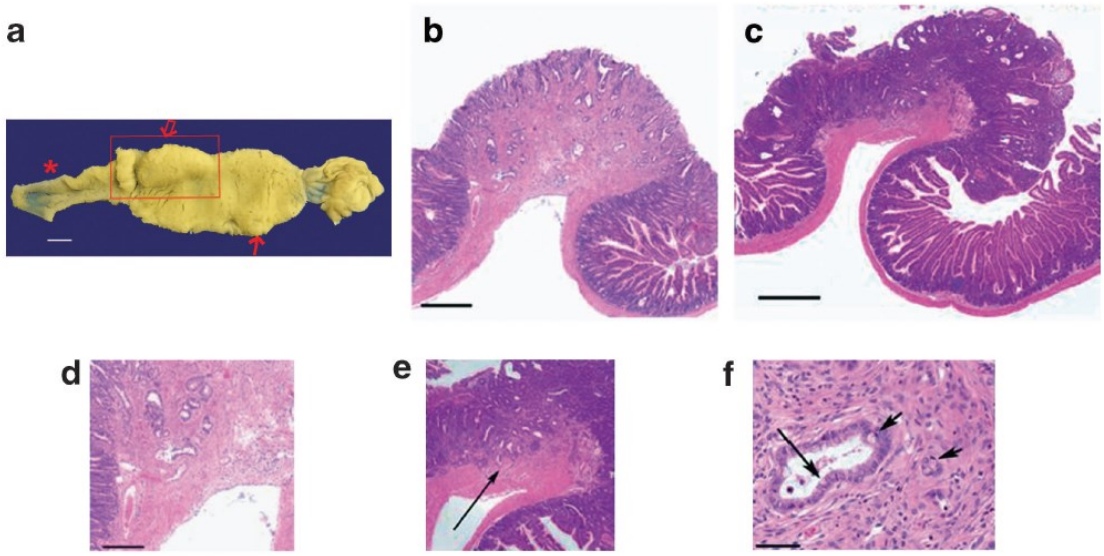


Figure 16. Loss of *Kcnq1* promotes tumor progression. **(a)** Duodenum from Apc^{Min} $Kcnq1^{+/-}$ mouse with stomach at right. Bar = 5 mm. Broad area of hyperplastic tissue with embedded tumors extends from pylorus–dudoenum juncture past the ampulla to a region approximate with the duodenal jejunal flexure. Asterisk indicates area of normal duodenum–proximal jejunum. Box encloses a region of adenocarcinoma tissue of > 2 cm. Sections from this lesion were taken from various regions and all showed evidence of invasion. Two of these adenocarcinoma sections are depicted in **(b–f).** Smaller arrow depicts normal large adenoma of ~5 mm that is embedded in hyperplastic tissue. **(b** and **c)** Bars = 500 µm; invasion into muscularis mucosa, **(d** and **e**) bar = 250 µm; prominent scirrhous response to invading tubules that exhibit differentiated goblet cells (long arrow) and signet ring cells (short arrows) **(f)**, bar = 50 µm.

Loss of *Kcnq1* is associated with other pathologies in the GI tract.

Rectal prolapse was evident in ~10% of Apc^{Min} $Kcnq1^{+/-}$ and Apc^{Min} $Kcnq1^{-/-}$ mice and in several cases the rectal prolapses were associated with enlarged rectums characterized by diffused severe crypt hyperplasia (***Figure 23***). In addition, $Kcnq1^{-/-}$ mice developed severe gastric hyperplasia that was characterized by mucosal hyperplasia, with parietal cells increased in size and crypt dilation with extensive cellular debris (***Figure 24***). Hyperplastic adenomas in the pylorus were also present in Apc^{Min} *Kcnq1* mutant mice. Although Apc^{Min} $Kcnq1^{+/-}$ mice showed an incidence of pyloric tumors of ~30%, the phenotype was much stronger in Apc^{Min} $Kcnq1^{-/-}$ mice where the incidence of pyloric tumors was ~90% (21/23). Furthermore, in Apc^{Min} $Kcnq1^{-/-}$ mice pyloric tumor multiplicity was much larger, with some mice developing as many as a dozen tumors. No pyloric tumors were observed in Apc^{Min} $Kcnq1^{+/+}$ mice. Finally, roughly one-third of Apc^{Min} $Kcnq1^{-/-}$ mice developed grossly and severely enlarged (~2- to 3-fold) pancreases, characterized by extra lobes. Histopathology of a matched pair of samples indicated that zymogen granules were less eosinophilic and less prominent, and that both exocrine cells and islets were smaller (***Figure 25***).

Loss of *Kcnq1* has a strong effect on gene expression

We conducted cDNA microarray expression studies using tissue from distal colon and the proximal quarter of the small intestine of $Apc^{+/+}$ $Kcnq1^{+/+}$ and $Apc^{+/+}$ $Kcnq1^{-/-}$ mice. Overall, more than 400 genes showed a > 1.5-fold change in expression, and > 200 of these genes also had a P-value of < 0.05 (***Tables 3*** and ***4***). Expression of 20 genes was confirmed by reverse transcription–polymerase chain reaction (RT–PCR) (***Figures 17*** and ***18***).

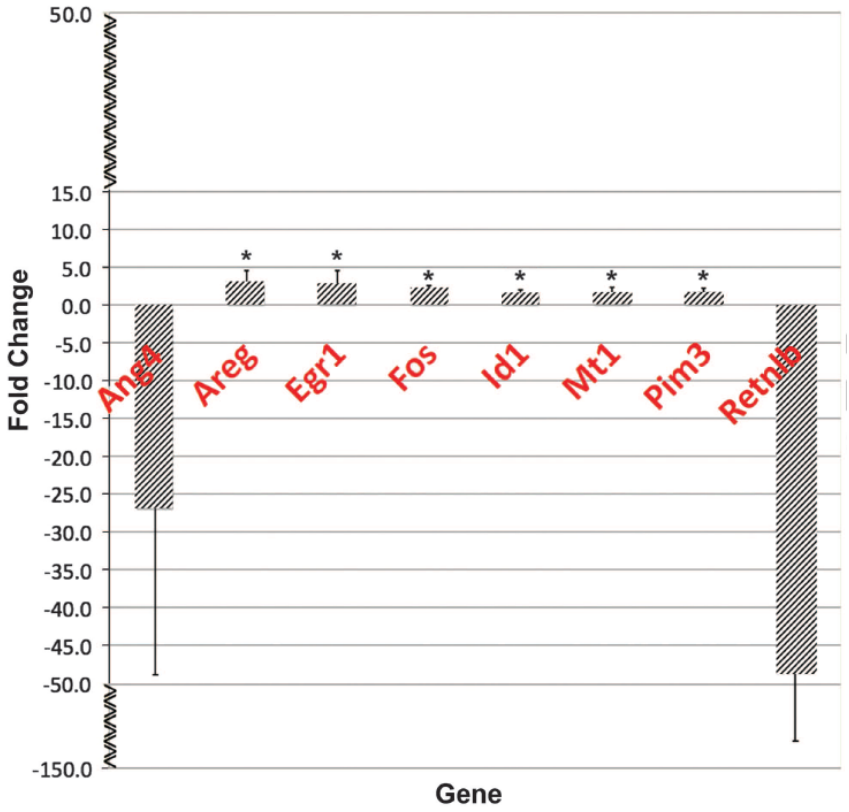


Figure 17. qRT–PCR gene expression analysis of mouse colon. Samples were analyzed in triplicate and normalized to 18S ribosomal RNA. Data are presented as the mean fold change ± s.d. Each bar represents the mean and s.d. of multiple experiments that measured fold differences in the mRNA expression of whole colon tissue isolated from adult (~100 days), littermate and gendermatched pairs of $Apc^{+/+}$ $Kcnq1^{+/+}$ and $Kcnq1^{-/-}$ mice. RNA was isolated from 1 cm sections from the same region of distal colon. At least two matched pairs of mRNAs were tested for each gene with most genes tested in at least three matched pairs of mRNAs. To be included in this figure, genes showed a mean fold difference of at least 1.5. In all cases, the direction in changes in gene expression confirmed microarray data. *P < 0.05.

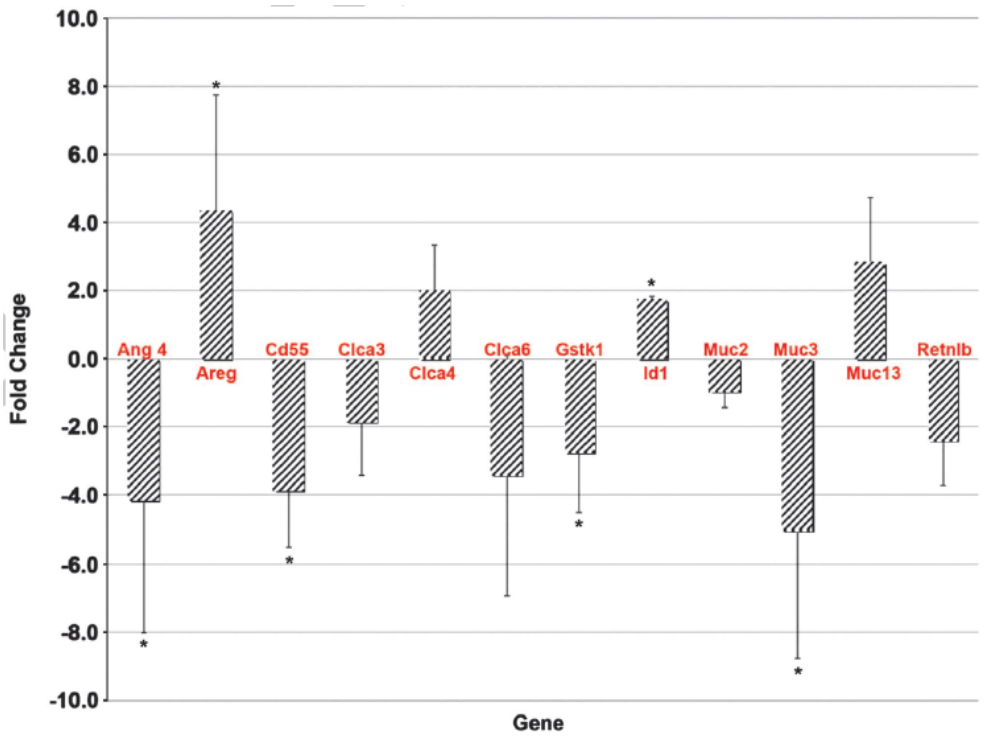


Figure 18. qRT–PCR gene expression analysis of mouse proximal small intestine.

Each gene sample was run in triplicate and gene expression was normalized to the expression of 18S. Data are presented as the mean fold change ± s.d. Each bar represents the mean and s.d. of multiple experiments that measured fold differences in the mRNA expression in proximal small intestine tissue isolated from adult (~100 days), littermate and gendermatched pairs of *Apc*$^{+/+}$ *Kcnq1*$^{+/+}$ and *Kcnq1*$^{-/-}$ mice. mRNAs were isolated from 1 cm sections of proximal small intestine from the same region for all mice. At least two matched pairs of mRNAs were tested for each gene with most genes tested in at least three matched pairs of mRNAs. To be included in this figure, genes showed a mean fold difference of at least 1.5. In all cases, the direction of changes in gene expression confirmed microarray data. *$P < 0.05$.

Target genes were clustered into several functional groups: innate immune response/goblet and Paneth cell function (*Ang4, Retnlb, Clps, Pnliprp2, Clca3, Muc2, Reg3a, Reg3b, Reg3g, Car1*), ion channels (*Trpv6, Slc12a8, Slc30a10, Aqp7, Aqp1, Aqp8, Aqp4, Clca3, Clca6*), mucins (*Muc2, Muc13, Muc4, Muc3*), growth regulatory signaling pathways (*Areg, Egr1, Fos, Ccdn1*), cancer cell migration (*Mmp9, Mmp15, S100a14, Mt1*), apoptosis (*Casp14, Casp2, Bcl6*) inflammation, mediated by mechanisms such as detoxification and stress responses (*Cyp2c55, Gstk1, Sirt1, Sirt3, Aldh1a1,*

Aldh1b1, Aldh2, Mt1, Gstm2), the adaptive immune response and intestinal stem cell-related genes (*Clca4, Aqp4, Aldh1a1, Olfm4*), consistent with a role for *Kcnq1* in the intestinal stem cell compartment (Munoz *et al.* 2012). Interestingly, while loss of *Kcnq1* showed a strong effect on colon secretory cell genes, there was no change in colon tissues in the secretory cell populations or cell proliferation (***Figure 26***).

Kcnq1, Cftr and *Muc2* share common genetic pathways in the GI tract

Ingenuity Pathway Analysis (IPA) identified significant networks and functional groups altered in *Kcnq1* KO mice (see ***Figures 27A*** and ***B***, and ***Table 5*** (colon); and ***Figure 28*** and ***Table 6*** (proximal small intestine). In particular, IPA analysis identified Cftr as a top upstream regulator of *Kcnq1* KO dysregulated genes, indicating that genes differentially expressed in the *Kcnq1* KO mice overlap with targets of *Cftr* (overlap P-value of 5.81E-13, activation Z-score -2.722). Furthermore, the *Cftr* network is predicted to be downregulated in *Kcnq1* KO small intestine, suggesting that *Kcnq1* may be necessary for some aspects of Cftr function.

To identify common pathways and genes disrupted by *Kcnq1* and *Cftr* deficiencies, we used Gene Set Enrichment Analysis (GSEA). The entire ranked expression data set from the *Kcnq1* KO proximal small intestine array was compared with gene sets compiled from a microarray analysis of gene expression in the small intestine of *Cftr* KO mice (Norkina *et al.* 2004). *Cftr* gene sets consisted of all genes up- or downregulated by twofold in the *Cftr* KO mouse. GSEA revealed significant correlation, with overlap of gene identity and direction of regulation (***Figure 19*** and ***Table 7***; normalized enrichment score (NES) for Cftr KO upregulated genes 1.97, $P < 0.001$; NES for *Cftr* KO downregulated genes -1.92, $P < 0.001$). These shared targets included a large subset of genes involved in lipid metabolism and, in particular, downregulation of genes involved in lipid oxidation. In support of the significance of this overlap, IPA analyses of *Kcnq1* KO and *Cftr* KO mice each identified lipid metabolism as a major functional category, with oxidation of lipids as the subcategory showing the most significant down regulation (*Kcnq1*, P-value 9.78E-06, activation Z-score, -2.26; *Cftr*, P-value 9.37E-09, Z-score, -3.62). Common targets included upregulation of *Idi1*,

involved in cholesterol synthesis, and downregulation of *Hsd17β4* and *Acaa1b*, both involved in fatty acid beta oxidation. Inhibition of fatty acid oxidation is associated with a switch to lipogenesis and this switch is characteristic of metabolic remodeling that commonly occurs during oncogenesis (DeBerardinis *et al*. 2008). Also, fatty acid metabolism has recently been shown to have a key role in stem cell maintenance with lipogenesis associated with more highly proliferating stem cells (Knobloch *et al.* 2013). Thus, altered lipid metabolism may contribute to the development of CRC in *Kcnq1* and *Cftr* KO mice. A second group of genes dysregulated in both *Kcnq1* KO mice and *Cftr* KO mice is involved in inflammation. Because the intestinal epithelium is constantly exposed to flora and foreign substances, it must maintain protective and repair mechanisms. Disruption of these protective or homeostatic processes can result in inflammatory changes, which in turn are associated with inflammatory bowel disease and increased risk of CRC (Duerkop *et al.* 2009). Genes involved in inflammation are dysregulated in both *Kcnq1* KO mice and *Cftr* KO mice. These include upregulation of factors normally produced by intestinal epithelial cells to maintain immune homeostasis, including *Il7* (Shalapour *et al.* 2010), *Reg3g30* and *Pglryp1* (Vaishnava *et al.* 2011, Dziarski *et al.* 2010). In addition, a number of protective genes are downregulated including those involved in detoxification, *Gstt1, Cyp3a11, Cyp2b10*, and others involved in protective signaling, such as *Sema7a*, which promotes *Il10* signaling (Kang *et al.* 2012), and *Ace2*, which protects against inflammation-induced epithelial damage (Hashimoto *et al.* 2012). Further evidence of overlap between *KCNQ1* and *CFTR* was provided by analyzing RNA Seq data from 20 CRC liver metastases samples for concordance between expression levels of each gene. Using Pearson's correlation analysis of FPKM (fragments per kilobase pair of exon model per million fragments mapped) values in the 20 samples, it was found that *KCNQ1* and *CFTR* were significantly coexpressed, despite up to a 10-fold sample-to-sample FPKM variability for each gene (see ***Figures 29A*** and ***B***). Finally, both *KCNQ1* and *CFTR* are regulated downstream of the serine threonine kinase gene *SGK1*, which is activated by factors including insulin, via the PI3K pathway, under conditions of osmotic stress (Lang *et al.* 2013). *SGK1* levels in turn appear to be influenced by the activity of both *KCNQ1* and

CFTR, as *Sgk1* expression was upregulated by more than twofold in *Kcnq1* KO colon (***Table 5***) and was also upregulated by > 1.6-fold in *Cftr* KO colon (BLN Than, personal communication). Of further interest, *Sgk1* expression levels were downregulated by 1.5-fold in the proximal small intestine of *Kcnq1* KO mice and downregulated by 5.5-fold in the small intestine of *Cftr* KO mice (Norkina *et al.* 2004).

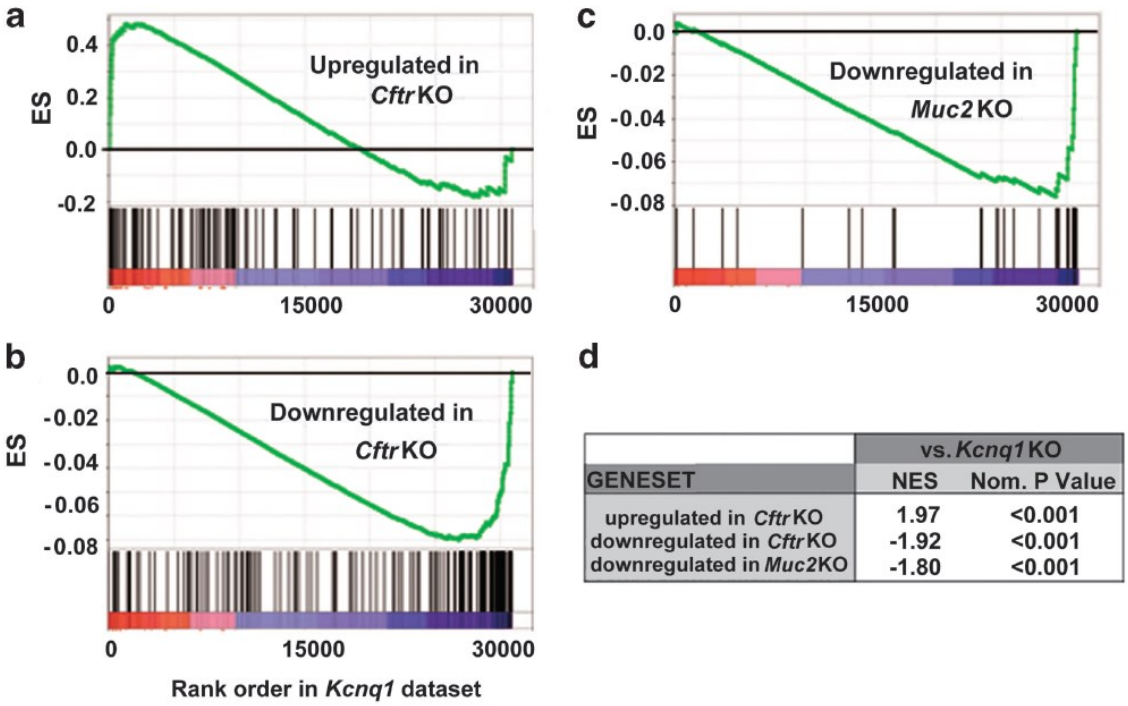


Figure 19. Results of GSEA showing enrichment scores (ES) of *Cftr* and *Muc2* gene sets with respect to the ranked *Kcnq1* expression data set. Shown are enrichment plots for gene sets consisting of **(a)** genes upregulated in the small intestine of *Cftr* KO mice (Norkina *et al.* 2004); **(b)** genes downregulated in the small intestine of *Cftr* KO mice (Norkina *et al.* 2004); and **(c)** genes downregulated in the small intestine of *Muc2* KO mice (Yang *et al.* 2008); **(d)** NES and nominal P-value (nom. P-value) are shown for each comparison.

CFTR export of Cl^- and HCO_3^- maintains extracellular H2O homeostasis, which in turn is necessary for hydration and function of the protective lumenal mucin layer (De Lisle *et al.* 2011). Another gene directly involved in maintenance of the mucin layer is

MUC2, which encodes MUCIN2, one of the major glycoproteins making up this layer. As with *Kcnq1* and *Cftr*, *Muc2* deficiency is implicated in CRC (Yang *et al.* 2008). To test if common pathways were disrupted by loss of *Kcnq1* and *Muc2*, we compared gene expression changes in the *Kcnq1* KO mouse with those reported for *Muc2* KO mice (Yang *et al.* 2008). GSEA identified significant overlap in gene expression, with genes downregulated in *Muc2* KO mice enriched in genes downregulated in *Kcnq1* KO mice (***Figure 19*** and ***Table 7***, NES -1.80, $P < 0.001$). This group included a number of genes involved in detoxification, including cytochrome oxidase P450 enzymes and several glutathione S-transferases, *Gsta2, Gstm2, Gstm4, Gstm6* and *Gstt1*. Moreover, in a separate study (Burger-van Paassen *et al.* 2012), deficiency for *Muc2* in the mouse intestine resulted in significant increases in the innate immune responders *Reg3g* and *Reg3b*, matching our findings in *Kcnq1* KO mice. Finally, we found that *Areg*, an epidermal growth factor receptor pathway gene that was upregulated in both the colon and proximal small intestine of *Kcnq1* KO mice, was also upregulated by >10-fold in the colon of *Muc2* KO mice (RTC and RJAF, unpublished microarray results) and in the colon of *Cftr* KO mice (BLNT, manuscript in preparation) (see ***Table 8***).

Loss of *Kcnq1* promotes colon organoid development.

To investigate a role for *Kcnq1* in the intestinal stem cell compartment, we created intestinal organoids from the colons of a group of $Apc^{+/+}$ $Kcnq1^{+/+}$ and $Kcnq1^{-/-}$ age- and gendermatched littermate mice. At days 2–5, organoid growth was monitored and organoid number counted on day 5. Loss of *Kcnq1* increased the number of colon organoids by greater than threefold, confirming a potential regulatory effect of *Kcnq1* on the intestinal stem cell compartment (Munoz *et al.* 2012) (***Figure 20***).

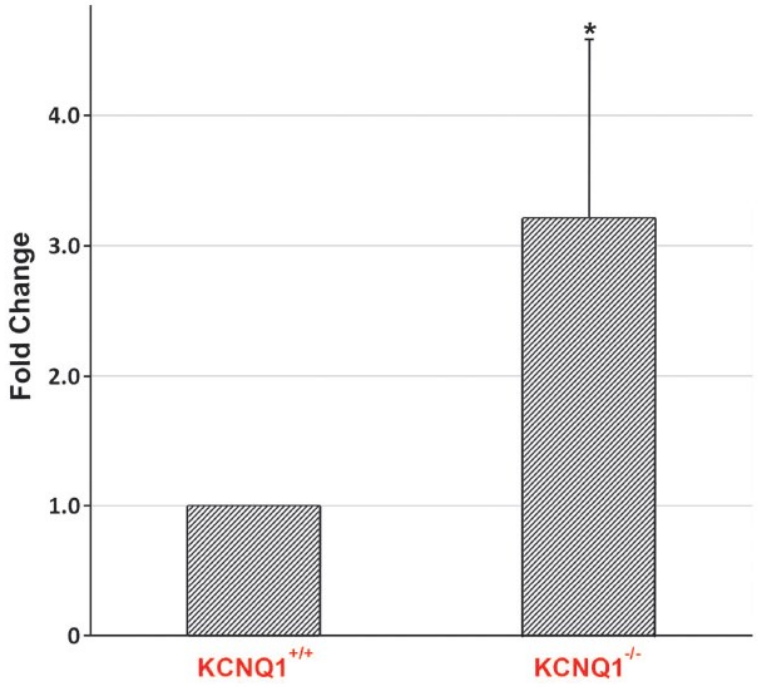


Figure 20. Colon organoids. Colon organoids were created from 500 crypt bottoms and plated in triplicate in 24-well plates. Crypt bottoms were isolated on separate days from 5 of each, age-, gender- and littermate-matched *Apc*$^{+/+}$ *Kcnq1*$^{+/+}$ and *Apc*$^{+/+}$ *Kcnq1*$^{-/-}$ mice (one matched pair per day). Organoids were examined daily and counted at 5 days post plating. *P < 0.05.

KCNQ1 expression is associated with good prognosis in late-stage human CRC.

On the basis of our finding that deficiency for *Kcnq1* drove the progression of intestinal tumors in mice, we tested the hypothesis that low levels of KCNQ1 in human CRC would correlate with a worse prognosis. To this end, we evaluated KCNQ1 protein expression in tissue microarrays containing samples from more than 500 stage IV CRC patients with liver metastases who had undergone hepatic resection. Immunohistochemical staining for KCNQ1 expression was evaluated for 311 patients. Membranes of neoplastic CRC liver metastasis (CRCLM) epithelium were scored for KCNQ1 staining intensity (***Figure 21***). We found that high KCNQ1 expression was related to improved overall survival (OS) (hazard rate ratio (HRR) 0.64; 95% confidence interval (CI): 0.45–0.92; P = 0.02). Median OS for patients with high KCNQ1 expression

was 60 months, whereas patients with low KCNQ1 levels had a median OS of 37 months, indicating a significant 23-month difference in survival (***Figure 22***).

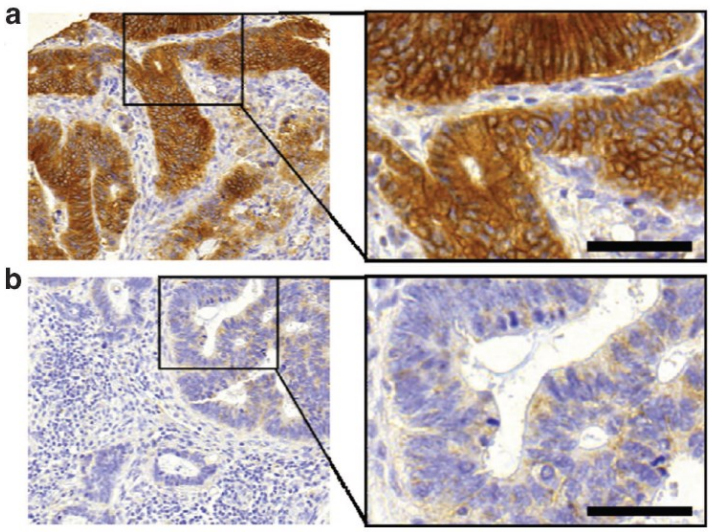


Figure 21. Representative human samples showing the expression pattern of KCNQ1 in the epithelium of CRCLM. Staining intensity was evaluated as **(a)** high or **(b)** low. Scale bars are 50 µm.

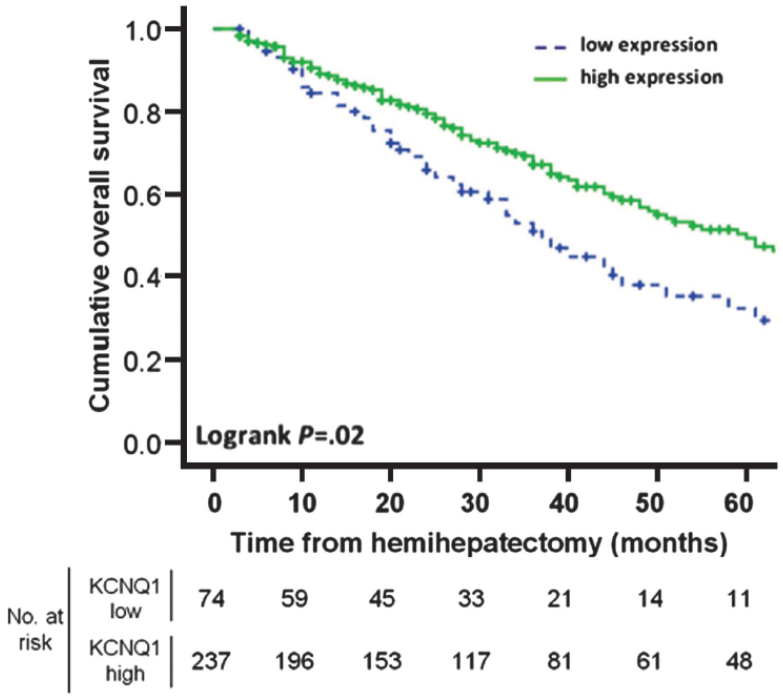


Figure 22. Kaplan–Meier graph depicting OS in months, stratified by intensity of KCNQ1 expression in CRCLMs. P-values were calculated using the log-rank statistic.

Next, we performed a multivariate analysis to evaluate whether the prognostic value of KCNQ1 protein expression was independent of established prognostic clinicopathological variables, that is, primary tumor-to-liver metastasis interval > 12 months, number of liver metastases > 1, maximal tumor diameter > 5.0 cm, lymph node positivity at the time of diagnosis of the primary tumor and serum carcinoembryonic level > 200 ng/ml. Upon stepwise backward Cox regression analysis with OS as the dependent variable, KCNQ1 expression (HRR, 0.49; 95%: 0.32–0.77; P = 0.002), lymph node positivity at the time of diagnosis of the primary tumor (HRR, 1.64; 95% CI: 1.08–2.48; P = 0.02) and maximal tumor diameter > 5.0 cm (HRR, 1.44; 95% CI: 0.96–2.16; P = 0.08) were retained as prognostic variables. Note that the association between maximal tumor diameter and survival was not significant; however, the variable was retained in the model because we only excluded variables when P > 0.1.

DISCUSSION

We first identified *Kcnq1* as a potential CRC driver gene in our forward genetic screens for GI tract cancer genes. Here, we have followed up with studies using a transgenic mouse model and analysis of human CRC tissues that support a tumor suppressor role for *Kcnq1*. We demonstrated the tumor suppressor function of *Kcnq1* in Apc^{Min} mice, where both hetero- and homozygous inactivation of *Kcnq1* resulted in significantly more adenomas and, importantly, progression to adenocarcinoma. This was especially evident in the proximal small intestine. Furthermore, *Kcnq1* deficiency caused rectal adenomatous hyperplasia, pyloric tumorigenesis, gastric hyperplasia and pancreatic abnormalities. Notably, in human CRC that had metastasized to the liver, low KCNQ1 protein expression correlated with significantly lower OS, by almost 2 years, compared to patients with high KCNQ1 expression. These findings suggest that KCNQ1 expression could be a biomarker to help determine eligibility for hepatic resection in CRC patients

with liver metastases. In addition, newly developed small molecules that enhance K^+ ion channel activity (Xiong *et al.* 2008, Wulff *et al.* 2009) could potentially be used to treat CRC patients with low levels of KCNQ1.

By what mechanism does loss of KCNQ1 influence cancer processes in the GI tract? A promising model is via KCNQ1's functional interactions with CFTR and MUC2. In the healthy intestinal epithelium, KCNQ1 is physiologically linked to the CFTR ion channel, with basolateral export of K^+ ions by KCNQ1 providing the electrochemical driving force for apical export of Cl^- ions by CFTR (Preston *et al.* 2010). *Cftr* was identified as a CRC driver gene in two *Sleeping Beauty* GI tract forward genetic screens (Starr *et al.* 2009, March *et al.* 2011), and individuals with cystic fibrosis, caused by inactivating mutations in *CFTR*, are at increased risk for GI cancers including CRC (Maisonneuve *et al.* 2003). To test if common pathways and genes are disrupted by loss of *Kcnq1* and *Cftr* activity, we used GSEA to compare gene sets consisting of genes reported to be up- or downregulated in the small intestine of *Cftr* KO mice to the entire ranked *Kcnq1* expression data set and found significant correlation (***Figure 19***). Genes upregulated in common include several involved in immune responses and inflammation, whereas genes downregulated in common include a subset involved in fatty acid metabolism, in particular lipid oxidation. Further evidence of overlap between *KCNQ1* and *CFTR* was the finding of concordance between *KCNQ1* and *CFTR* expression in RNA Seq analysis of 20 CRC liver metastases (***Figures 29A*** and ***B***).

As with *Kcnq1* and *Cftr*, *Muc2* deficiency is implicated in CRC (Yang *et al.* 2008). To test if common pathways were disrupted by loss of *Kcnq1* and *Muc2*, we compared gene expression changes in the *Kcnq1* KO mice with those reported for *Muc2* KO mice. GSEA identified significant overlap in gene expression with genes downregulated in *Muc2* KO mice enriched in genes downregulated in *Kcnq1* KO mice (***Figure 19***). This group included a number of genes involved in detoxification including cytochrome oxidase P450 enzymes and several glutathione S transferases, and genes involved inflammation.

These results point to common pathways (inflammation, lipid metabolism, detoxification and stress responses) disrupted by *Kcnq1, Cftr* and *Muc2* deficiencies. Therefore, we can hypothesize that KCNQ1, like MUC2 and CFTR, may act through one or more of these pathways in preventing cancer in the GI tract. Moreover, precisely defining mechanisms of action for KCNQ1 can lead to its use as a prognostic predictor for CRC patients and potential therapeutic target.

MATERIALS AND METHODS

Mice

C57BL/6J mice and C57BL/6J-*Apc*^{Min} mice were obtained from the Jackson Laboratory (Bar Harbor, ME, USA). C57BL/6-*Kncq1* knockout mice were obtained from Dr Karl Pfeifer (NIH, Bethesda, MD, USA) (Casimiro *et al.* 2001). Details of mouse husbandry are as described previously (Fijneman *et al.* 2012). The genotype of the *Apc* and *Kcnq1* loci were determined by PCR assays as described previously (Casimiro *et al.* 2001, Fijneman *et al.* 2012).

Tumor analysis

Scoring of GI tumor tissues was performed as described previously (Fijneman *et al.* 2012). Two-sided P-values for tumor counts were determined by use of the Wilcoxon rank-sum test comparing gender- and age-matched classes produced in the same genetic crosses.

Histopathology

Histopathological analysis of tumors and adjoining normal tissue was performed on formalin-fixed paraffin-embedded tissues by an ACVPcertified veterinary pathologist (MGO0S) from the University of Minnesota Masonic Cancer Comparative Pathology Shared Resources facility using protocols as described previously (Fijneman *et al.* 2012).

Organoid culture

Gender- and age-matched littermate C57Bl/6J *Kcnq1*$^{+/+}$ and *Kcnq1*$^{-/-}$ mice were killed between 8 and 12 weeks of age. Colons were removed, cut open and washed in cold phosphate-buffered saline. Colon organoids were then cultured using the protocol of Sato *et al*. (Sato *et al.* 2013) following the plating of 500 crypt bottoms per well in triplicate per sample. Two-sample t-test was used to determine statistical significance.

RNA processing for Illumina bead arrays

Mouse intestinal tissues were removed, opened longitudinally and rinsed in phosphate-buffered saline. Colon tissues were processed in RNAlater (Qiagen, Valencia, CA, USA) as per the manufacturer's protocol. The distal quarter of the proximal quarter of the small intestine was flash frozen in liquid nitrogen. The tissue was then transferred from liquid nitrogen to precooled RNALaterIce (Ambion, Foster City, CA, USA) at -80^{o}C for 30 min, and later transferred to -20^{o}C for at least 48h. Immediately before RNA isolation, tissue in either RNALater or RNALaterIce was placed in 2–4 ml of RLT (guanidine thiocyanate buffer)/14.3 M b-mercaptoethanol buffer. The tissue was then homogenized using an IKA Ultra-Turrax T25 digital homogenizer (Fisher Scientific, Waltham, MA, USA). RNA isolation was then conducted using an RNeasy Mini Kit with an additional DNAse Digestion step (Qiagen). RNA sample concentrations were measured using a Nanodrop-1000 Spectrophotometer (Thermo Scientific, Rochester, NY, USA). RNA was stored at -80^{o}C.

Illumina bead microarray and data analysis

RNA labeling, microarray hybridization and scanning were performed at the University of Minnesota BioMedical Genomics Center using Illumina MouseWG-6 v.2.0 Expression BeadChips (Illumina, San Diego, CA, USA), according to the manufacturer's instructions. Differential expression of genes was quantified by the moderated t-statistic. Data were analyzed using GeneData Expressionist Software (GeneData Inc., San Francisco, CA, USA); genes that showed statistically significant differences in expression between groups were determined using the two-group t-test and analysis of variance.

Gene expression data has been submitted to the Gene Expression Omnibus, and accession number is pending.

qRT–PCR

Quantitative RT–PCR was performed as described previously (Starr *et al.* 2011). Significant differences in expression between groups were determined using the two-group t-test.

List of primers used:

All primers were designed using Primer-BLAST (National Center for Biotechnology Information) and obtained from Integrated DNA Technologies.

Mouse primers

Ang4: forward, 5'-AGCACACAGCTAGACTCGTCCC-3'; reverse, 5'-ACCAGAC CCAGCACGAAGACC-3'; *Areg*: forward, 5'-GGTCTTAGGCTCAGGCCATTA-3'; reverse, 5'-CGCTTATGGTGGAAACCTCTC-3'; *Cd55*: forward, 5'-GGGGCTATG ATCCGTGGGCG-3'; reverse, 5'-TGCCCAAGATTGGCCTGGCA-3'; *Clca3*: forward, 5'-CCACACCAAAACGAGAAGGC-3'; reverse, 5'-TGCTTCGGAGATTG CATCGT-3'; *Clca4*: forward, 5'-ACATGGACCGGCCTTTCTAC-3'; reverse, 5'-CG ACATCTCCTCGACACACA-3'; *Clca6*: forward, 5'-GTTCAATCAGAAAAGGCTT CCA-3'; reverse, 5'-ACTTCCCAAGTGCTTCTGTAATTG-3'; *Gstk1*: forward, 5'-TGGATGCGTGTATGGTCTCG-3'; reverse, 5'-CAGAAAGTGTTGGGCTTGCG-3'; *Egr1*: forward, 5'-CCTCAAGGGGAGCCGAGCG-3'; reverse, 5'-AACCGAGT CGTTTGGCTGGG-3'; *Fos*: forward, 5'-CGGGTTTCAACGCCGACTA-3'; reverse, 5'-TTGGCACTAGAGACGGACAGA-3'; *Id1*: forward, 5'-CTCAGCACCCTGAAC GGCGA-3'; reverse, 5'-CATCTGGTCCCTCAGTGCGCC-3'; *Muc2*: forward, 5'-AACGATGCCTACACCAAGGTC-3'; reverse, 5'-ACTGAACTGTATGCCTTCC TCA-3'; *Muc3*: forward, 5'-CACTACCACCCCAGCACCTA-3'; reverse, 5'-GCCT CCATTCTCGCAGTTGA-3'; *Muc13*: forward, 5'-TTTGGCTACAGCGGGAT GAA-3'; reverse, 5'-TCTTTGACCTCGCAGAGACG-3'; *Mt1*: forward, 5'-AAAC CCTTTGCGCCCGGACT-3'; reverse, 5'-AGCAGGAGCAGTTGGGGTCC-3'; *Pim3*: forward, 5'-GCTCATCGACTTCGGCTCGGG-3'; reverse, 5'-AGTGGCAGACCGC

CCGTGATA-3'; and *Retnlb*: forward, 5'-AAGCCTACACTGTGTTTCCTTTT-3'; reverse, 5'-GCTTCCTTGATCCTTTGATCCAC-3'.

Pathway analysis

GSEA. GSEA v.2.0, http://www.broad.mit.edu/gsea/ (Subramanian *et al.* 2005, Mootha *et al.* 2003). Gene sets analyzed were obtained from published reports and the National Center for Biotechnology Information Gene Expression Omnibus database. GSEA compared the entire ranked *Kcnq1* proximal small intestine KO expression data set to *Cftr* KO (Norkina *et al.* 2004) (twofold change in expression, $t \leq 0.05$) and *Muc2* KO (Yang *et al.* 2008) (1.5-fold change in the duodenum at 3 or 6 months of age).

IPA. Genes showing a 1.5-fold change in expression in whole colon (N = 89) and proximal small intestine (N = 345) microarrays were analyzed using IPA (IPA Ingenuity Systems).

Tissue microarrays

Patient study population. Patients who underwent liver resection with curative intent, sometimes with addition of RFA, in one of the seven Dutch hospitals affiliated with the DeCoDe PET group were identified by crossreferencing surgery and pathology databases using Dutch MeSH terms for 'colon', 'rectum', 'carcinoma', 'adenocarcinoma', 'colorectal neoplasms', 'liver', 'neoplasm metastasis' and '(hemi)hepatectomy'. Clinicopathological data from 507 patients that were operated on between 1990 and 2010 were extracted from these databases (see ***Table 9***; a manuscript describing the patient population in more detail is in preparation). Formalin-fixed paraffin-embedded tissue specimens were collected from one CRCLM sample and an adjacent control liver sample. Only specimens of patients with histologically confirmed CRCLM were included in the study, whereas tissue samples of patients with multiple primary tumors were excluded. Collection, storage and use of clinicopathological data and tissue specimens were performed in compliance with the 'Code for Proper Secondary Use of Human Tissue in The Netherlands', and approved in protocol 2011-03 of the Department of Pathology (Stichting FMWV Rotterdam 2011).

Tissue microarrays. A total of 21 tissue microarrays (TMAs) were generated as described previously (Simon *et al.* 2004, Belt *et al.* 2011). In brief, three tissue core biopsies of 0.6mm in diameter were punched from morphologically representative areas of all formalin-fixed paraffin-embedded donor blocks, and transferred into TMA recipient paraffin blocks using the 3DHISTECH TMA Master (v.1.14, 3DHISTECH Ltd, Budapest, Hungary).

Immunohistochemistry. Four-micrometer sections of TMAs were mounted on glass slides, deparaffinized by xylene and rehydrated with a decreasing alcohol series. Staining for KCNQ1 was performed upon antigen retrieval by microwave heating in citric acid (10mM, pH 6.0) and endogenous peroxidase neutralization in 0.3% hydrogen peroxide in methanol for 25 min. The primary rabbit polyclonal antibody directed against human KCNQ1 (sc-20816; Santa Cruz Biotechnology Inc., Santa Cruz, CA, USA) was incubated overnight at a 1:200 dilution at 4^{o}C, followed by incubation with anti-rabbit secondary antibodies for 30 min at room temperature (Envision Plus; Dako, Heverlee, Belgium). Secondary antibodies were visualized by liquid diaminobenzidine substrate chromogen system. Slides were counterstained with Mayer's hematoxylin. Staining of formalin-fixed paraffin-embedded colon tissue was used as a positive control, and incubation without primary antibody as a negative control.

Evaluation of KCNQ1 protein expression. Immunohistochemical stainings were digitally captured using the Mirax slide scanner system equipped with a x20 objective with a numerical aperture of 0.75 (Carl Zeiss BV, Sliedrecht, The Netherlands) and a Sony DFW-X710 Fire Wire 1/3 in-type progressive SCAN IT CCD (pixel size 4.65 x 4.65 μm^2). Actual scan resolution at x 20 was 0.23 µm. Computer monitors were calibrated using Spyder2PRO software (v.1.0–16; Pantone Colorvision, Regensdorf, Switzerland). TMA core biopsies were scored for intensity of KCNQ1 protein expression on membranes of neoplastic epithelial cells (categories negative, weak, moderate, strong) using dedicated TMA scoring software (v.1.14.25.1; 3DHISTECH Ltd). For facilitating

scoring, a chart with visual analogue scales of staining patterns was used. Intensity scores were dichotomized into low and high KCNQ1 intensities using receiver operating characteristic curve analysis (Zlobec *et al.* 2007).

Statistical analysis of TMAs. OS was defined as the time in months after surgery until death in a follow-up period of 10 years. Patients were excluded from analysis if OS < 2 months (n = 72), if OS or survival status were unknown (n = 36) or if tissue cores could not be evaluated for technical reasons (n = 88). The relation between KCNQ1 expression and survival was studied using Kaplan–Meier curves and tested using log-rank statistics. HRR was calculated using Cox regression. Multivariate analysis was performed by inclusion of KCNQ1 protein expression in a proportional hazards (Cox regression) analysis together with established prognostic clinicopathological variables combined in the clinical risk score as defined by Fong *et al*. (Fong *et al* 1999) primary tumor-to-liver metastasis interval > 12 months, number of liver metastases > 1, maximal tumor diameter > 5.0 cm, lymph node positivity at the time of diagnosis of the primary tumor, serum carcinoembryonic level > 200 ng/ml. Stepwise backward regression was used to exclude variables from the model when P > 0.1. All statistical tests were two-sided and executed using IBM SPSS Statistics 20.0 software (SPSS Inc., Chicago, IL, USA). P-values < 0.05 were considered significant.

Supplemental Information:

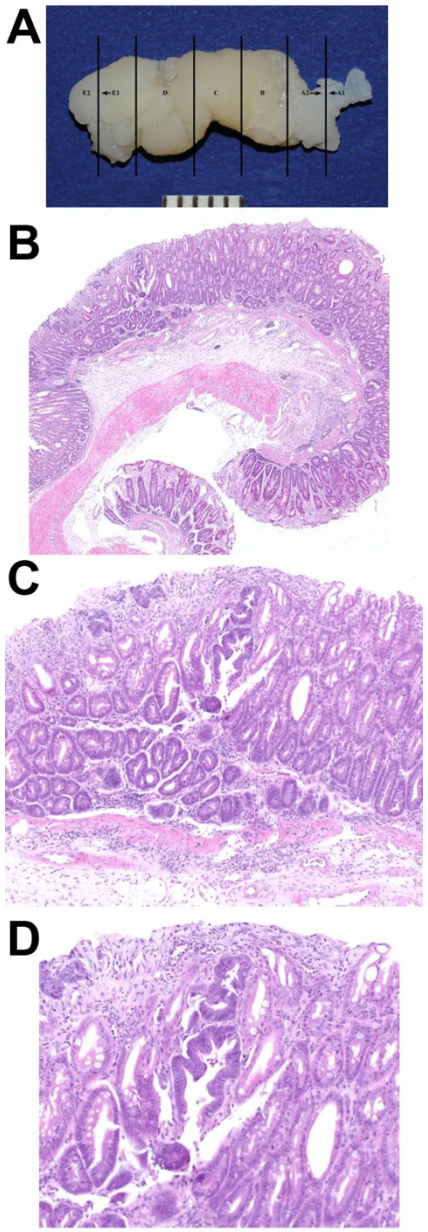


Figure 23. Rectal Adenomatous Hyperplasia. Panel A. Gross image of rectum. Lines represent locations of cross sections taken for histopathological analysis. **Panels B, C & D**, three cross sections of the same tissue showing diffuse crypt hyperplasia. The incidence of abnormal rectums was 10 out of 102 *Apc*^{Min} *Kcnq1*$^{+/-}$ (79) and *Apc*^{Min} *Kcnq1*$^{-/-}$ (23) mice and 0 out of 58 for *Apc*^{Min} *Kcnq1*$^{+/+}$ mice, a difference that is highly significant (p = 0.009, Fisher's Exact Test).

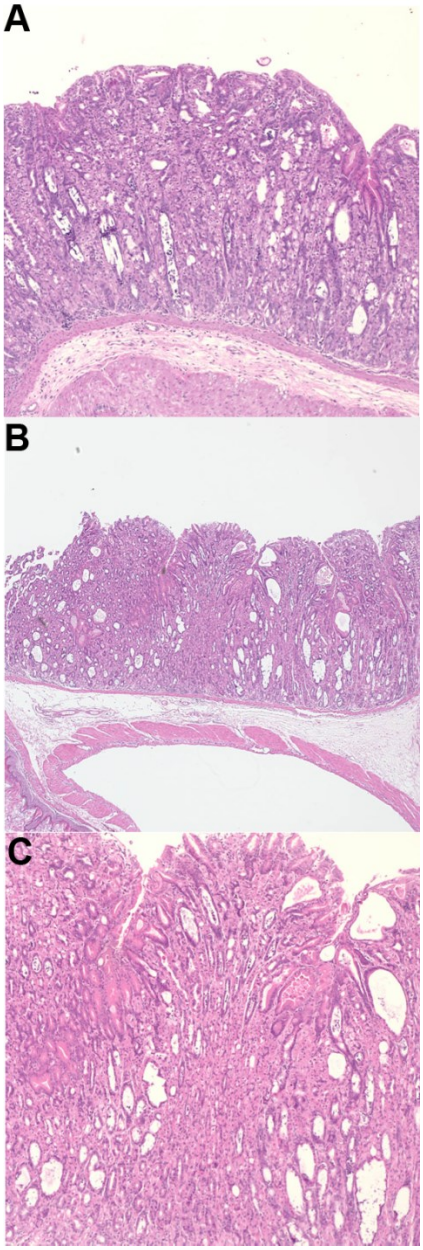


Figure 24. Stomach of *Apc*$^{+/+}$ *Kcnq1*$^{-/-}$ mice. Panels A & B, 10X; **Panel C**, 20X. Severe gastric hyperplasia that was characterized by mucosal hyperplasia, with parietal cells increased in size and crypt dilation with debris present. Hyperplastic adenomas developed in the pylorus (not shown). The incidence of pyloric tumors was 21 out of 23 *Apc*^{Min} *Kcnq1*$^{-/-}$ mice and 0 out of 58 *Apc*^{Min} *Kcnq1*$^{+/+}$ mice, a difference that is highly significant ($p < 2.2e\text{-}16$, Fisher's Exact Test).

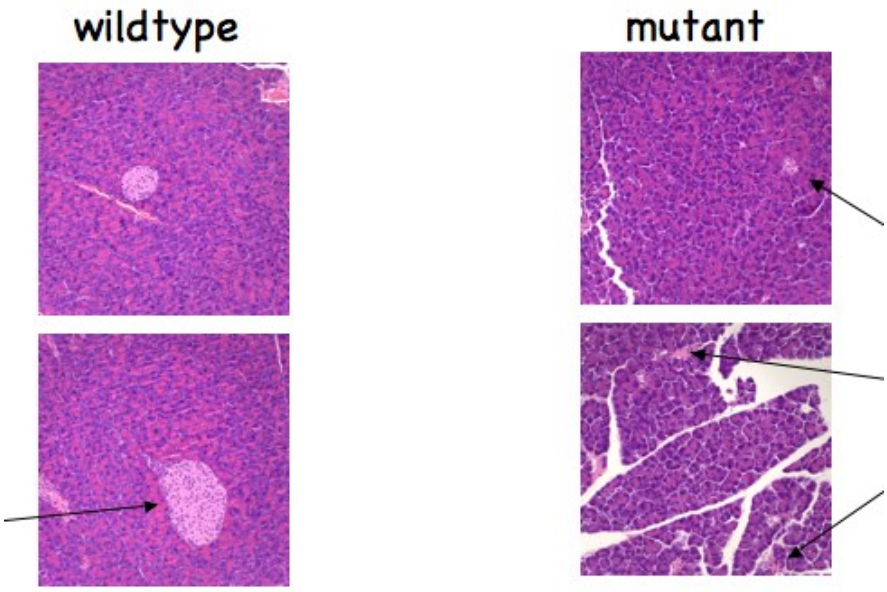


Figure 25. Pancreas. 20X. ***Kcnq1* Wildtype, left panels**; ***Apc^{Min} Kcnq1$^{-/-}$* right panels**. Zymogen granules were less eosinophilic and less prominent. Exocrine cells seemed to be smaller and islets appeared to be less prominent (arrows). The incidence of enlarged pancreases was 8 out of 23 *Apc^{Min} Kcnq1$^{-/-}$* mice and 0 out of 58 in *Apc^{Min} Kcnq1$^{+/+}$* mice, a difference that was high significant (p = 1.52e-05, Fisher's Exact Test). The maximum diameter of a total of 56 pancreatic islets was measured, 28 from *Apc^{Min} Kcnq1$^{-/-}$* mice and 28 from *Apc^{Min} Kcnq1$^{+/+}$* mice. The mean diameter of islets from *Apc^{Min} Kcnq1$^{-/-}$* mice was 100 microns and the mean diameter of islets from *Apc^{Min} Kcnq1$^{+/+}$* mice was 118 microns. While the difference in means was not significant, when the 56 islet diameters were ranked in size the top 7 islets were all from the *Kcnq1$^{+/+}$* group (246, 242, 212, 210, 199, 184, and 181 microns), while the largest islet from the *Kcnq1$^{-/-}$* group was 173 microns.

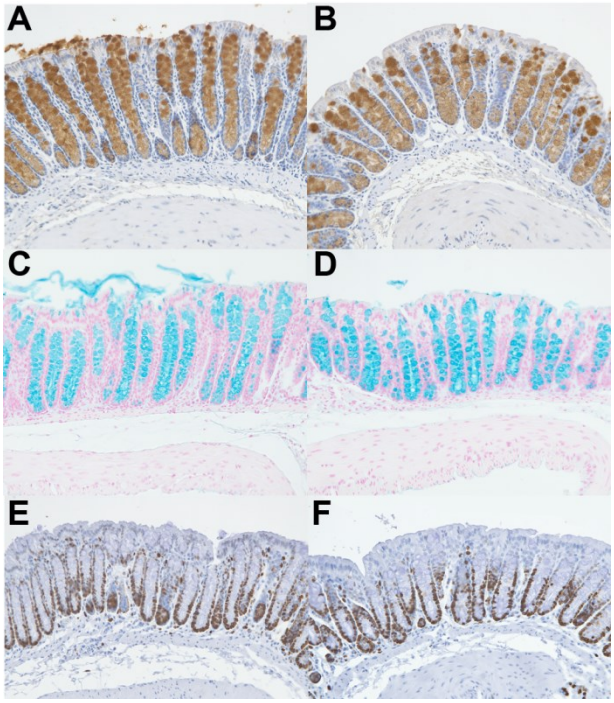
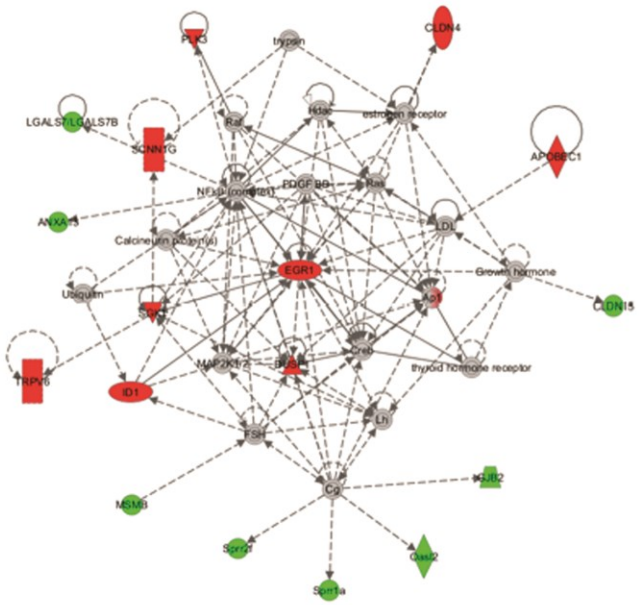


Figure 26. Immunohistochemistry for cellular markers in mouse colon. Age and gender matched *Kcnq1*$^{+/+}$ and *Kcnq1*$^{-/-}$ adult littermate mice were sacrificed and FFPE tissues processed as previously described (Fijneman *et al.*, 2012). Multiple sections were examined by immunostaining for the goblet cell markers Mucin2 and alcian blue (for detection of acidic mucopolysaccharides), and the proliferation marker Ki67. No significant differences were detected between genotypes. Representative colon section is shown above. Mucin2: Panel A (*Kcnq1*$^{+/+}$), Panel B (*Kcnq1*$^{-/-}$); Alcian Blue: Panel C (*Kcnq1*$^{+/+}$), Panel D (*Kcnq1*$^{-/-}$); Ki67: Panel E (*Kcnq1*$^{+/+}$), Panel F (*Kcnq1*$^{-/-}$).

Methods. 4 µm formalin-fixed, paraffin-embedded sections of small intestine were deparaffinized and rehydrated, followed by antigen retrieval using 10mM Citrate buffer pH 6.0 in a steamer. A rabbit polyclonal antibody was used for Mucin 2, Catalog number: SC15334, (Santa Cruz Biotechnology). A rabbit monoclonal antibody was used for Ki-67, Catalog number: CRM325, clone SP6, (Biocare Medical). Detection was achieved using a rabbit EnVision™+ Kit (catalog K4011, Dako) with DAB as chromogen. Alcian blue staining was performed using a standard protocol.

Figure 27. Colon Microarray - Top Functional Networks Identified by IPA

A. Cancer, Cellular Function and Maintenance, Molecular Transport



B. Immunology & Inflammatory Disease

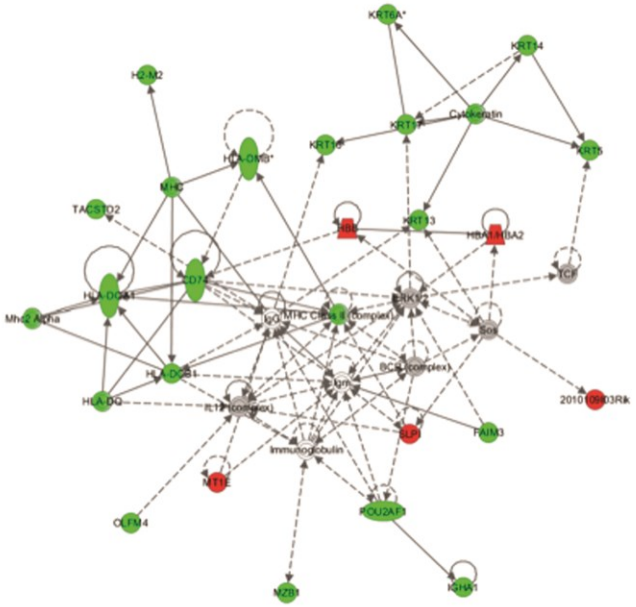


Figure 28. Proximal Small Intestine Microarray - Top Gene Networks Identified by IPA.

A. Molecular Transport, Cell Death and Survival, Cellular Assembly and Organization

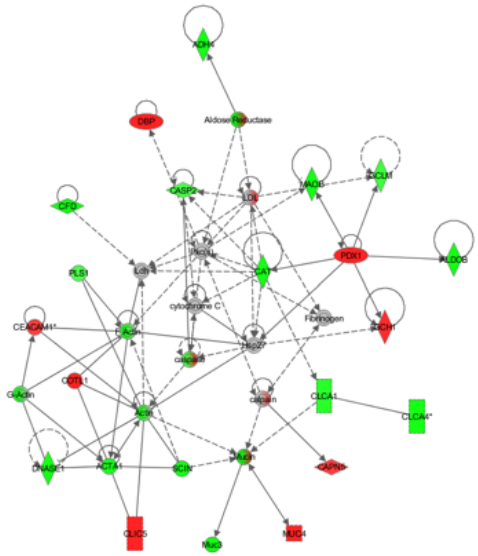
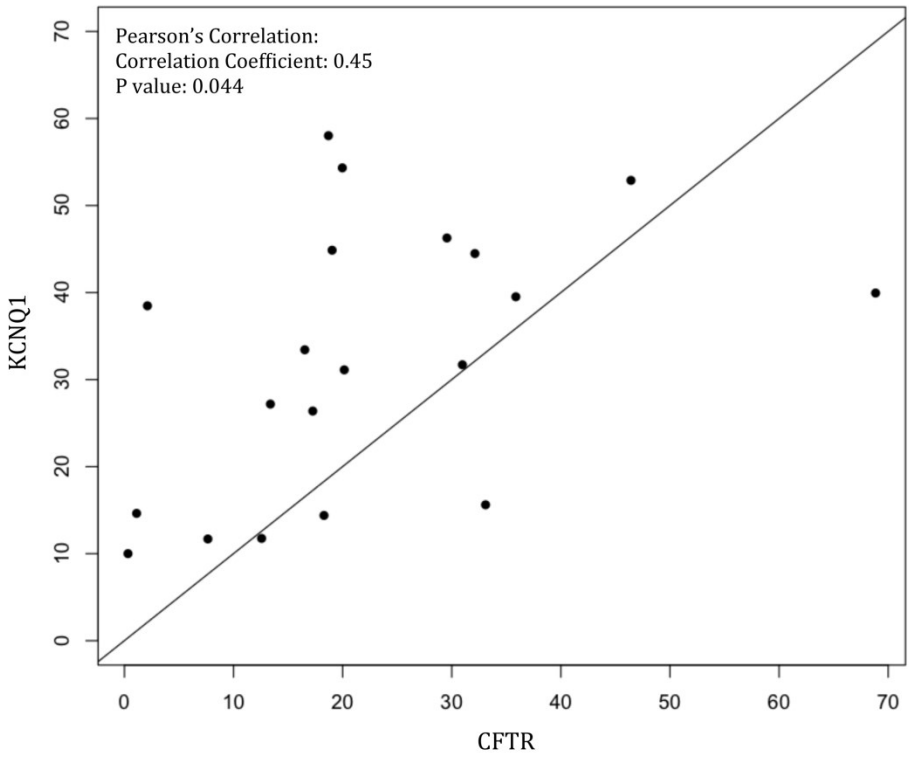
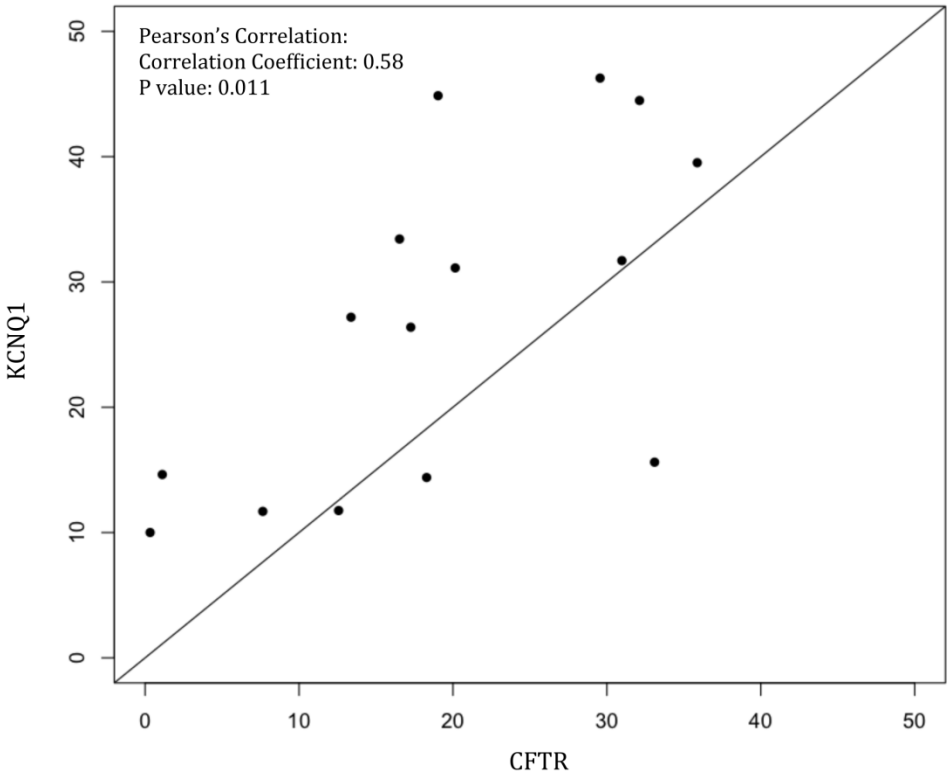


Figure 29. Concordance between expression of *KCNQ1* and *CFTR* in colorectal cancer liver metastases. Twenty stage IV CRC liver metastases were analyzed by RNA seq, followed by Pearson's Correlation analysis to determine concordance between the expression levels of *KCNQ1* and *CFTR*. Panel A. Using all twenty samples. Panel B. removal of two top and bottom outliers, N = 18.





RNA Seq analysis. Twenty fresh frozen stage IV CRC LM samples were obtained from the University of Minnesota BioNet Tissue Procurement facility. The samples were from 9 males and 11 females, average age 64 years. RNA was isolated using RNA LaterICE (Invitrogen) and Qiagen RNAEasy kits, and quantified using a Nanodrop-1000 Spectrophotometer (Thermo Scientific), and quantity and quality were confirmed using The extracted RNA-samples were sent to the University of Minnesota Genomics Center (UMGC) for sequencing. All raw reads that passed the CASAVA 1.8 P/F filter were directly deposited at the Minnesota Supercomputing Institute, where the data were analyzed. Sequence quality was assessed via fastqc (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Since the raw sequence reads were of sufficient quality, no data trimming or filtering was necessary. Hence, the raw paired-end reads were mapped to human genome (hg19 assembly) using Tophat 2.0.6 with the iGenomes reference (UCSC human hg19 annotation, downloaded from http://cufflinks.cbcb.umd.edu/igenomes.html March 9, 2012). Final mapping

percentages of the 20 samples range from 92% to 94%. The mapped samples were then analyzed via Cuffdiff (default parameters using the iGenomes UCSC hg19 GTF file previously mentioned) to quantify the expression level of each known gene in units of FPKM (Fragment mapped per kilobase of exons per million mapped reads). Additional statistical test of Pearson's correlation from R/Bioconductor was applied to selected gene pairs to check for expression concordance.

Table 3. List of top known genes 1.5 fold (A) up-regulated and (B) down-regulated in *Kcnq1*$^{-/-}$ mouse colons, represented as fold change vs control *Kcnq1*$^{+/+}$. The asterisk in the Fold Change column indicates a P value of 0.05 or less.

A) Top 23 Up-Regulated Known Genes:

Gene Symbol	Name	Fold Change	
Cfd	Complement factor D	3.0	
Hbb-b1	Hemoglobin, beta adult major chain	3.0	*
Hba-a1	Hemoglobin, beta adult major chain	2.6	
Egr1	Early growth response protein 1	2.4	
IGHV8S7_U23022_Ig_heavy_variable_8S7_163	Immunoglobulin heavy chain variable	2.3	

Areg	Amphiregulin	2.2	
IGHV8S6_U23021_Ig_heavy_variable_8S6_61	Immunoglobulin heavy chain variable	2.2	
Scnn1g	Amiloride-sensitive sodium channel subunit gamma	2.1	
Mt1	Metallothionein 1	2.0	*
Slc30a10	Solute carrier family 30, member 10	1.9	*
Trpv6	Transient receptor potential cation channel, subfamily V, member 6	1.9	*
Fxyd4	FXYD domain containing ion transport regulator 4	1.9	
Mrgprf	MAS-related GPR, member F	1.8	
Slpi	Antileukoproteinase	1.7	
Abcg5	ATP-binding cassette sub-family G member 5	1.7	
Dusp1	Dual specificity protein phosphatase 1	1.6	
Sgk1	Serine/threonine-protein kinase 1	1.6	

Apobec1	Apolipoprotein B mRNA editing enzyme, catalytic polypeptide 1	1.6	
Fos	FBJ murine osteosarcoma viral oncogene homolog	1.6	
Cyp2d9	Cytochrome P450, family 2, subfamily d, polypeptide 9	1.6	
Cldn4	Claudin 4	1.5	*
Id1	DNA-binding protein inhibitor ID-1	1.5	*
Plk3	Phosphatidylinositol (PI) 3-kinase	1.5	

B) Top 66 Down-regulated Known Genes:

Gene Symbol	Name	Fold Change
Retnlb	Resistin-like beta	-13.8
Ang4	Angiogenin, ribonuclease A family, member 4	-9.3

IGLC2_J00595_Ig_lambda_constant_2_14	Immunoglobulin lamda chain constant	-6.2
Krt14	Keratin 14	-4.1
Itlnb	Intelectin b	-3.6
Krt16	Keratin16	-3.1
Pnliprp2	Pancreatic lipase-related protein 2	-2.9
IGHV1S120_AF025443_Ig_heavy_variable_1S120_8	Immunoglobulin heavy chain variable	-2.8
Tgm4	Protein-glutamine gamma-glutamyltransferase 4	-2.8
IGKV1-88_AJ231206_Ig_kappa_variable_1-88_289	Immunoglobulin kappa chain variable	-2.7
Itln1	Intelectin-1	-2.7

IGHV1S35_M12376_Ig_heavy_variable_1S35_13	Immunoglobulin heavy chain variable	-2.5
IGHV10S3_AF064446_Ig_heavy_variable_10S3_9	Immunoglobulin heavy chain variable	-2.5
Clps	Colipase, Pancreatic	-2.5
Defb50	Defensin beta 50	-2.4
IGHV1S36_M13788_Ig_heavy_variable_1S36_40	Immunoglobulin heavy chain variable	-2.4
Krt6b	Keratin 6B	-2.3
IGKV2-137_AJ231263_Ig_kappa_variable_2-137_15	Immunoglobulin kappa chain variable	-2.3
Pla2g4c	Phospholipase A2, group IVC (cytosolic, calcium-independent)	-2.3
IGHV1S124_AF025449_Ig_heavy_variable_1S124_11	Immunoglobulin heavy chain variable	-2.3

IGKV9-120_V00804$J00566_Ig_kappa_variable_9-120_12	Immunoglobulin kappa chain variable	-2.2
IGKV8-31_AJ235957_Ig_kappa_variable_8-31_3	Immunoglobulin kappa chain variable	-2.2
IGHV1S28_X02460_Ig_heavy_variable_1S28_13	Immunoglobulin heavy chain variable	-2.1
IGHA_J00475$V00785_Ig_heavy_constant_alpha_135	Immunoglobulin heavy chain constant alpha	-2.1
Krt13	Keratin 13	-2.0
Lypd3	Ly6/PLAUR domain-containing protein 3	-2.0
Padi1	Peptidyl arginine deiminase, type I	-1.9
Ly6d	Lymphocyte antigen 6 complex, locus D	-1.9
IGHV5S18_AF290972_Ig_heavy_variable_5S18_125	Immunoglobulin heavy chain variable	-1.9

Krt17	Keratin 17	-1.9	
Igk-V5	Immunoglobulin kappa chain variable 5 (V5 family)	-1.9	
Krt5	Keratin 5	-1.8	
Tacstd2	Tumor-associated calcium signal transducer 2	-1.8	
Cd74	HLA class II histocompatibility antigen gamma chain	-1.7	*
Igh-6	Immunoglobulin heavy chain 6	-1.7	
IGHV1S133_AF304553_Ig_heavy_variable_1S133_89	Immunoglobulin heavy chain variable	-1.7	
Duoxa2	Dual oxidase maturation factor 2	-1.6	
Glycam1	Glycosylation-dependent cell adhesion molecule-1	-1.6	

Cldn15	Claudin 15	-1.6	
Igk-V38	Immunoglobulin kappa chain variable 38	-1.6	
Reg4	Regenerating islet-derived protein 4	-1.6	
H2-Ab1	Histocompatibility 2, class II antigen A, beta 1	-1.6	*
Faim3	Fas apoptotic inhibitory molecule 3	-1.6	*
H2-M2	Murine MHC class Ib	-1.6	*
Lgals7	Galectin-7	-1.6	
IGKV4-90_AJ231224_Ig_kappa_variable_4-90_22	Immunoglobulin kappa chain variable	-1.6	
Igk-C	Immunoglobulin kappa constant	-1.6	

Parp12	Poly [ADP-ribose] polymerase 12	-1.6
Indo	Indoleamine-pyrrole 2,3-dioxygenase	-1.6
B4galnt1	Beta-1,4 N-acetylgalactosaminyltransferase 1	-1.6
Cd52	Cluster of differentiation 52	-1.5
Krt6a	Keratin 6A	-1.5
Sprr2f	Small proline-rich protein 2F	-1.5
Pou2af1	POU domain class 2-associating factor 1	-1.5
H2-Aa	Histocompatibility 2, class II antigen A, alpha	-1.5
Anxa13	Annexin A13	-1.5

Igh-VJ558	Immunoglobulin heavy chain (J558 family)	-1.5
Serpinb11	Serpin peptidase inhibitor, clade B (ovalbumin), member 11	-1.5
Igh-4	Immunoglobulin heavy chain 4	-1.5
Gjb2	Gap junction beta-2	-1.5
Sprr1a	Small proline-rich repeat protein 1A	-1.5
Slc12a8	Solute carrier family 12 member 8	-1.5
IGHG1_J00453$V00793_Ig_heavy_constant_gamma_1_792	Immunoglobulin heavy chain constant gamma	-1.5
Igk-V33	Immunoglobulin kappa chain variable 33	-1.5
Olfm4	Olfactomedin 4	-1.5

H2-DMb1	Histocompatibility 2, class II, locus Mb1	-1.5	

Table 4. List of top known genes 1.5 fold (A) up-regulated and (B) down-regulated in *Kcnq1*$^{-/-}$ mouse small intestines, represented as fold change vs control *Kcnq1*$^{+/+}$. The asterisk in the Fold Change column indicates a P value of 0.05 or less.

A) Top 157 Up-Regulated Known Genes:

Gene Symbol	Name	Fold Change	
Ubd	Ubiquitin D	7.3	*
Dbp	D site of albumin promoter (albumin D-box) binding protein	5.8	
IGKV2-137_AJ231263_Ig_kappa_variable_2-137_15	Immunoglobulin kappa chain variable	4.4	
Pigz	Phosphatidylinositol glycan anchor biosynthesis, class Z	4.2	*
Akr1b8	Aldo-keto reductase family 1, member B8	4.1	*
Socs3	Suppressor of cytokine signaling 3	3.7	*

Slc2a6	Solute carrier family 2, facilitated glucose transporter member 6	3.3	*
Reg3g	Regenerating islet-derived protein 3 gamma	3.1	
Slc40a1	Solute carrier family 40 (iron-regulated transporter), member 1	2.9	
Cxcl16	Chemokine (C-X-C motif) ligand 6	2.9	*
Reg3b	Regenerating islet-derived protein 3 beta	2.9	*
Igl-V1	Immunoglobulin lambda variable 1-44	2.8	*
IGLC2_J00595_Ig_lambda_constant_2_14	Immunoglobulin constant chain lambda	2.8	
Slc11a2	Solute carrier family 11, member 2	2.8	
IGKV11-125_AJ231256_Ig_kappa_variable_11-125_15	Immunoglobulin kappa chain variable	2.8	*
IGHV10S1_AF064442_Ig_heavy_variable_10S1_100	Immunoglobulin heavy chain variable	2.7	*
Noxo1	NADPH oxidase organizer 1	2.5	*

Akr1b7	Aldo-keto reductase family 1, member B7	2.4	*
Barx2	Barx Homeobox 2	2.4	*
Igh-6	Immunoglobulin heavy constant mu	2.3	*
Gcnt3	Glucosaminyl (N-acetyl) transferase 3, mucin type	2.3	*
Cyba	Cytochrome b-245, alpha polypeptide	2.3	*
Cd74	HLA class II histocompatibility antigen gamma chain	2.2	
IGHV8S7_U23022_Ig_heavy_variable_8S7_163	Immunoglobulin heavy chain variable	2.2	
IGKV12-98_AJ235949_Ig_kappa_variable_12-98_12	Immunoglobulin kappa chain variable	2.2	
IGHV1S35_M12376_Ig_heavy_variable_1S35_13	Immunoglobulin heavy chain variable	2.2	
Nr1d1	Nuclear receptor subfamily 1, group D, member 1 (aka. Rev-ErbA alpha)	2.2	*
H2-DMb2	Histocompatibility 2, class II, locus Mb2	2.1	

Slc13a2	Solute carrier family 13 (sodium-dependent dicarboxylate transporter), member 2	2.1	
Il33	Interleukin 33	2.1	*
Cldn2	Claudin 2	2.0	*
Casp4	Caspase 4, apoptosis-related cysteine peptidase	2.0	*
IGKV9-128_AJ231245_Ig_kappa_variable_9-128_15	Immunoglobulin kappa chain variable	2.0	
Tns4	Tensin-4	2.0	*
IGHV1S120_AF025443_Ig_heavy_variable_1S120_8	Immunoglobulin heavy chain variable	2.0	
IGHV8S6_U23021_Ig_heavy_variable_8S6_61	Immunoglobulin heavy chain variable	2.0	
Defcr-rs7	Defensin-related cryptdin-related sequence 7	2.0	
B3galt5	Beta-1,3-galactosyltransferase 5	2.0	
Pglyrp1	Peptidoglycan recognition protein	2.0	

Ostb	Organic solute transporter beta	2.0	*
H2-Eb1	Histocompatibility 2, class II antigen E beta	2.0	
H2-Ab1	Histocompatibility 2, class II antigen A, beta 1	2.0	
Akp3	Alkaline phosphatase 3, intestine	1.9	
Tgm2	Transglutaminase 2	1.9	*
H2-DMa	Histocompatibility 2, class II, locus DMa	1.9	
Defcr4	Defensin-related cryptdin peptide-4	1.9	
Ctsh	Cathepsin H	1.9	*
Cotl1	Coactosin-like	1.9	*
Prss32	Protease, serine 32	1.9	*
Capg	Macrophage-capping (actin filament), gelsolin-like	1.9	

Npal2	NIPA-like domain containing 2	1.8	*
Ctrc	Chymotrypsin C (caldecrin)	1.8	
IGHV12S1_M22439_Ig_heavy_variable_12S1_339	Immunoglobulin heavy chain variable	1.8	
Smpdl3b	Sphingomyelin phosphodiesterase, acid-like 3B	1.8	*
Sult1c2	Sulfotransferase 1C2	1.8	
Pcsk9	Proprotein convertase subtilisin/kexin type 9	1.8	
Pou2af1	POU class 2 associating factor 1	1.8	
Noxa1	NADPH oxidase activator 1	1.8	*
Lyz2	Lysozyme 2	1.8	
Lcn2	Lipocalin-2	1.8	
Abcc3	ATP-binding cassette, sub-family C (CFTR/MRP), member 3	1.8	*

Cxcl9	Chemokine (C-X-C motif) ligand 9	1.8	*
Mmp9	Matrix metallopeptidase 9 (involved in the breakdown of extracellular matrix)	1.8	*
IGHV1S124_AF025449_Ig_heavy_variable_1S124_11	Immunoglobulin heavy chain variable	1.8	
Cirbp	Cold-inducible RNA-binding protein	1.8	*
Lgmn	Legumain	1.8	
Psg25	Pregnancy-specific glycoprotein 25	1.7	*
Lincr	Lung-inducible neuralized-related C3HC4 RING domain	1.7	*
Plk3	Polo-like kinase 3 (may play a role in regulation of cell cycle progression and tumorigenesis)	1.7	*
Rab31	Ras-related protein Rab-31	1.7	*
Aldh4a1	Aldehyde dehydrogenase 4 family, member A1	1.7	*
Fkbp11	FK506 binding protein 11	1.7	*

IGHV1S36_M13788_Ig_heavy_variable_1S36_40	Immunoglobulin heavy chain variable	1.7	
Tspan12	Tetraspanin-12	1.7	*
Rab32	Member RAS oncogene family	1.7	*
IGKV4-80_AJ231213_Ig_kappa_variable_4-80_91	Immunoglobulin kappa chain variable	1.7	
Apoe	Apolipoprotein E	1.7	
Slc37a1	Solute carrier family 37 (glycerol-3-phosphate transporter), member 1	1.7	
Tcfcp2l1	Transcription factor CP2-like 1	1.7	*
Tha1	Threonine aldolase 1	1.7	*
Ube1l	Ubiquitin-like modifier-activating enzyme 7	1.7	*
Prss27	Protease, serine 27	1.7	*
IGHV10S3_AF064446_Ig_heavy_variable_10S3_9	Immunoglobulin heavy chain variable	1.7	

Muc4	Mucin4	1.7	*
Clic5	Chloride intracellular channel protein 5	1.7	*
Areg	Amphiregulin	1.7	
Ltf	Lactotransferrin	1.7	
Cd52	Cluster of differentiation 52	1.7	
Ighg	Immunoglobulin heavy chain (gamma polypeptide)	1.6	
Nus1	Nuclear undecaprenyl pyrophosphate synthase 1 homolog	1.6	*
Tmem82	Transmembrane protein 82	1.6	*
Heatr5a	HEAT repeat containing 5A	1.6	
Mlkl	Mixed lineage kinase domain-like	1.6	
Iars2	Isoleucyl-tRNA synthetase 2, mitochondrial	1.6	*

Asns	Asparagine synthetase (glutamine-hydrolyzing)	1.6	
Car13	Carbonic anhydrase 13	1.6	*
Ly6d	Lymphocyte antigen 6 complex	1.6	*
Pdx1	Pancreatic and duodenal homeobox 1	1.6	
B4galt6	UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase, polypeptide 6	1.6	*
Slc30a10	Solute carrier family 30, member 10	1.6	
Axud1	AXIN1 up-regulated	1.6	*
Amica1	Adhesion molecule, interacts with CXADR antigen 1	1.6	*
Anpep	Alanine aminopeptidase	1.6	
Gdpd5	Glycerophosphodiester phosphodiesterase domain containing 5	1.6	*
Tm4sf5	Transmembrane 4 L6 family member 5	1.6	*

Tgfbr2	Transforming growth factor, beta receptor II	1.6	*
Ppt1	Palmitoyl-protein thioesterase 1 (a small glycoprotein involved in the catabolism of lipid-modified proteins during lysosomal degradation)	1.6	*
Abcc2	ATP-binding cassette, sub-family C (CFTR/MRP), member 2	1.6	
Mov10	Putative helicase MOV-10	1.6	*
Usp2	Ubiquitin carboxyl-terminal hydrolase 2	1.6	
Elovl1	ELOVL fatty acid elongase 1	1.6	*
Endod1	Endonuclease domain containing 1	1.6	*
Amdhd2	Amidohydrolase domain containing 2	1.6	*
Ccdc116	Coiled-coil domain containing 116	1.6	*
Gmds	GDP-mannose 4,6-dehydratase	1.6	
Fut4	Fucosyltransferase 4 (alpha (1,3) fucosyltransferase, myeloid-specific)	1.5	

Smpdl3a	Sphingomyelin phosphodiesterase, acid-like 3A	1.5	*
Gch1	GTP cyclohydrolase 1	1.5	*
Tnfrsf11a	Tumor necrosis factor receptor superfamily, member 11a, NFKB activator	1.5	*
Guk1	Guanylate kinase	1.5	*
Mep1b	Meprin A subunit beta	1.5	*
Ugcg	Ceramide glucosyltransferase	1.5	*
Defcr5	Defensin related cryptdin 5	1.5	
Upp1	Uridine phosphorylase 1	1.5	
Ptk6	Protein tyrosine kinase 6	1.5	
Ick	Serine/threonine-protein kinase	1.5	*
Acot10	Acyl-CoA thioesterase 10	1.5	*

Hmox1	Heme oxygenase (decycling) 1	1.5	
Tbc1d15	TBC1 domain family, member 15	1.5	*
Mocs1	Molybdenum cofactor biosynthesis protein 1	1.5	*
Kctd5	Potassium channel tetramerisation domain containing 5	1.5	*
Nfkbia	Nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha	1.5	
Stk39	STE20/SPS1-related proline-alanine-rich protein kinase	1.5	*
Osbpl3	Oxysterol-binding protein-related protein 3	1.5	
Ppapdc1b	Phosphatidic acid phosphatase type 2 domain containing 1B	1.5	
Rrp15	Ribosomal RNA processing 15	1.5	*
Psat1	Phosphoserine aminotransferase	1.5	
Paox	Polyamine oxidase (exo-N4-amino)	1.5	

Npal1	NIPA-like domain containing 1	1.5	
Eif2ak2	Eukaryotic translation initiation factor 2-alpha kinase 2	1.5	*
Slc30a2	Solute carrier family 30 (zinc transporter), member 2	1.5	*
C1qc	Complement component 1, q subcomponent, C chain	1.5	
Pitpna	Phosphatidylinositol transfer protein alpha	1.5	*
Eif3s4	Eukaryotic translation initiation factor 3, subunit 4 delta	1.5	*
Ceacam1	Carcinoembryonic antigen-related cell adhesion molecule 1	1.5	
Capn5	Calpain 5	1.5	
Sec61a1	Sec61 alpha 1 subunit	1.5	
Pik3ap1	Phosphoinositide 3-kinase adapter protein 1	1.5	*
Pqlc3	PQ loop repeat containing 3	1.5	

S100a11	S100 calcium binding protein A11	1.5	
Serpina3g	Serine (or cysteine) peptidase inhibitor, clade A, member 3G	1.5	
Plscr1	Phospholipid scramblase 1	1.5	
Me2	Malic enzyme 2, NAD(+)-dependent, mitochondrial	1.5	
Myadm	Myeloid-associated differentiation marker	1.5	
Ms4a12	Membrane-spanning 4-domains, subfamily A, member 12	1.5	
Pusl1	Pseudouridylate synthase-like 1	1.5	*
Als2	Amyotrophic lateral sclerosis 2	1.5	*

B) Top 188 Down-Regulated Known Genes:

Gene Symbol	Name	Fold Change	
G6pc	Glucose-6-phosphatase	-7.5	*
Slc5a4b	Dolute carrier family 5 (neutral amino acid transporters, system A), member 4b	-5.3	*
Nt5e	5'-nucleotidase (Cd73)	-5.1	*
Muc3	Mucin 3	-4.5	*
Ces3	Carboxylesterase 3	-4.2	*
Itpr1	Inositol 1,4,5-trisphosphate receptor type 1	-3.8	*
Aqp8	Aquaporin 8	-3.7	
Slc28a1	Solute carrier family 28 (sodium-coupled nucleoside transporter), member 1	-3.2	*
Sepp1	Selenoprotein P	-3.1	*
Ang4	Angiogenin 4	-2.9	*

Slc17a4	Solute carrier family 17 (sodium phosphate), member 4	-2.8	*
Pmp22	Peripheral myelin protein 22	-2.8	*
Cd55	Complement decay-accelerating factor	-2.6	*
Cfd	Complement factor D	-2.5	
Cat	Chloramphenicol acetyltransferase	-2.5	
Dak	Bifunctional ATP-dependent dihydroxyacetone kinase/FAD-AMP lyase	-2.4	*
Pnliprp2	Pancreatic lipase-related protein 2	-2.4	*
Clca6	Chloride channel calcium activated 6	-2.3	*
Igsf9	Immunoglobulin superfamily, member 9	-2.3	*
Gprc5a	Retinoic acid-induced protein 3	-2.3	*
Rsad2	GTP-binding protein Rhes	-2.3	*

Arsg	Arylsulfatase G	-2.3	
Apol9b	Apolipoprotein L 9b	-2.3	*
Retnlb	Resistin-like beta	-2.3	
Dnase1	Deoxyribonuclease I	-2.2	
Rdh7	Retinol dehydrogenase 7	-2.2	
Osr2	Odd-skipped related 2	-2.2	
Erdr1	Erythroid differentiation regulator 1	-2.2	*
Adipor2	Adiponectin receptor 2	-2.2	*
Aldob	Aldolase B	-2.2	*
Gstk1	Glutathione S-transferase kappa 1	-2.2	*
Nudt4	Nudix (nucleoside diphosphate linked moiety X)-type motif 4	-2.1	*

Mmp15	Matrix metalloproteinase 15	-2.1	
Nmu	Neuromedin U	-2.1	*
Reg3a	Regenerating islet-derived 3 alpha	-2.1	
Iyd	Iodotyrosine deiodinase	-2.1	
Pmm1	Phosphomannomutase 1	-2.1	*
Susd2	Sushi domain containing 2	-2.1	
Trim38	Tripartite motif containing 38	-2.1	*
Bok	Bcl-2 related ovarian killer	-2.0	*
Slc16a10	Solute carrier family 16, member 10 (aromatic amino acid transporter)	-2.0	*
Aqp7	Aquaporin-7	-2.0	*
Fbp1	Fructose-1,6-bisphosphatase 1	-2.0	

Alox5ap	Arachidonate 5-lipoxygenase-activating protein	-1.9	*
Gsta3	Glutathione S-transferase alpha 3	-1.9	
Maob	Monoamine oxidase B	-1.9	*
Tsc22d1	TSC22 domain family protein 1	-1.9	
Cyp2d22	Cytochrome P450, family 2, subfamily d, polypeptide 22	-1.9	*
St5	Suppression of tumorigenicity 5	-1.9	
Edn1	Endothelin 1	-1.9	
Cisd1	CDGSH iron sulfur domain 1	-1.9	*
Gdpd2	Glycerophosphodiester phosphodiesterase domain containing 2	-1.9	*
Apol7a	Apolipoprotein L 7a	-1.9	*
Slc27a2	Solute carrier family 27 (fatty acid transporter), member 2	-1.9	

Cyp4f16	Cytochrome P450, family 4, subfamily f, polypeptide 16	-1.9	*
Adh4	Alcohol dehydrogenase 4	-1.8	
Rorc	RAR-related orphan receptor C	-1.8	*
Gpr155	G protein-coupled receptor 155	-1.8	
Socs2	Suppressor of cytokine signaling 2	-1.8	*
Lama3	Laminin, alpha 3	-1.8	*
Ephx1	Epoxide hydrolase 1	-1.8	
Tmem43	Transmembrane protein 43	-1.8	*
Bmp1	Bone morphogenetic protein 1	-1.8	*
Rai3	Retinoic acid-induced protein 3	-1.8	*
Lrp12	Low density lipoprotein receptor-related protein 12	-1.8	*

Aldh1a1	Aldehyde dehydrogenase 1 family, member A1	-1.8	
Car4	Carbonic anhydrase 4	-1.8	*
Scin	Scinderin	-1.8	*
Bdh2	3-hydroxybutyrate dehydrogenase, type 2	-1.8	*
Khk	Ketohexokinase (fructokinase)	-1.8	*
Paqr9	Progestin and adipoQ receptor family member IX	-1.8	
Hbegf	Heparin-binding EGF-like growth factor	-1.8	*
Cish	Cytokine-inducible SH2-containing protein	-1.8	
Cyp4v3	Cytochrome P450, family 4, subfamily v, polypeptide 3	-1.8	
Ccbl1	Cysteine conjugate-beta lyase 1	-1.8	*
Mib2	Mindbomb E3 ubiquitin protein ligase 2	-1.8	*

Pbld	Phenazine biosynthesis-like protein domain	-1.8	*
Hist1h2bc	Histone cluster 1, H2bc	-1.8	*
Slc28a2	Solute carrier family 28 (sodium-coupled nucleoside transporter), member 2	-1.8	
Bmp8a	Bone morphogenetic protein 8a	-1.7	*
Pcp4l1	Purkinje cell protein 4 like 1	-1.7	*
Zfand2a	Zinc finger, AN1-type domain 2A	-1.7	*
Ppme1	Protein phosphatase methylesterase 1	-1.7	*
Sh2d6	SH2 domain containing 6	-1.7	*
Dpep1	Dipeptidase 1	-1.7	*
Mlxipl	MLX interacting protein-like	-1.7	*
Adora2b	Adenosine A2b receptor	-1.7	

Asb13	Ankyrin repeat and SOCS box protein 13	-1.7	*
Cldn4	Claudin 4	-1.7	*
Hist1h2be	Histone cluster 1, H2be	-1.7	*
Hpgd	Hydroxyprostaglandin dehydrogenase 15-(NAD)	-1.7	
Pank3	Pantothenate kinase 3	-1.7	
Acta1	Actin, alpha 1, skeletal muscle	-1.7	
Eno1	Enolase 1	-1.7	
Slc6a3	Solute carrier family 6 (neurotransmitter transporter, dopamine), member 3	-1.7	
Hba-a1	Hemoglobin alpha, adult chain 1	-1.7	
Hist1h2bh	Histone cluster 1, H2bh	-1.7	*
Ncor1	Nuclear receptor co-repressor 1	-1.7	*

Naaladl1	N-acetylated alpha-linked acidic dipeptidase-like 1	-1.7	
Slc5a4a	Solute carrier family 5, member 4a	-1.7	
Avil	Advillin	-1.7	*
Aldh18a1	Aldehyde dehydrogenase 18 family, member A1	-1.7	
Hist1h2bj	Histone H2B type 1-J	-1.7	*
Amy2-2	Amylase 2-2, pancreatic	-1.7	
Fbln1	Fibulin-1	-1.7	*
Naip5	NLR family, apoptosis inhibitory protein 5	-1.7	*
Hist1h2bk	Histone H2B type 1-K	-1.7	*
Tinag	Tubulointerstitial nephritis antigen	-1.7	*
L2hgdh	L-2-hydroxyglutarate dehydrogenase	-1.7	*

Wbscr14	WBS critical region 14	-1.7	*
Scd1	Stearoyl-CoA desaturase 1	-1.7	
Gpr151	G protein-coupled receptor 151	-1.7	*
Sprr2a	Small proline-rich protein 2A	-1.6	
Fhl2	Four and a half LIM domains protein 2	-1.6	*
Btg1	B-cell translocation gene 1, anti-proliferative	-1.6	*
Rmnd1	Required for meiotic nuclear division 1	-1.6	*
Bmp3	Bone morphogenetic protein 3	-1.6	*
Hist1h2bm	Histone H2B type 1-M	-1.6	*
Clca3	Chloride channel, calcium activated, family member 3	-1.6	*
Glipr2	GLI pathogenesis-related 2	-1.6	

Npc1l1	Niemann-Pick C1-Like 1	-1.6	
Cnnm2	Cyclin M2	-1.6	*
C2cd3	C2 calcium-dependent domain containing 3	-1.6	*
Dusp6	Dual specificity phosphatase 6	-1.6	
Hist1h2bf	Histone cluster 1, H2bf	-1.6	
Cyp4f13	Cytochrome P450, family 4, subfamily f, polypeptide 13	-1.6	*
Lrmp	Lymphoid-restricted membrane protein	-1.6	*
Gpi1	Glucose phosphate isomerase 1	-1.6	*
Acaa1b	Acetyl-Coenzyme A acyltransferase 1B	-1.6	
Ugt2b36	UDP glucuronosyltransferase 2 family, polypeptide B36	-1.6	
Tnfrsf21	Tumor necrosis factor receptor superfamily member 21	-1.6	*

Tagln2	Transgelin-2	-1.6	*
Serpinh1	Serpin peptidase inhibitor, clade H, member 1	-1.6	
Tmie	Transmembrane inner ear	-1.6	*
Upb1	Beta-ureidopropionase	-1.6	*
Cyp2b23	Cytochrome P450, family 2, subfamily b, polypeptide 23	-1.6	
Mfsd7b	Major facilitator superfamily domain containing 7B	-1.6	
Tspan14	Tetraspanin 14	-1.5	*
Trak1	Trafficking protein, kinesin binding 1	-1.5	*
Bcl6	B-cell CLL/lymphoma 6	-1.5	
Tmem86a	Transmembrane protein 86A	-1.5	
Gstm2	Glutathione S-transferase mu 2	-1.5	

Ivns1abp	Influenza virus NS1A-binding protein	-1.5	*
Oas1d	2'-5' oligoadenylate synthetase 1D	-1.5	
Acot1	Acyl-CoA thioesterase 1	-1.5	
Gclm	Glutamate-cysteine ligase, modifier subunit	-1.5	
Slc5a9	Solute carrier family 5 (sodium/glucose cotransporter), member 9	-1.5	*
Ethe1	Ethylmalonic encephalopathy 1	-1.5	*
Klf6	Krueppel-like factor 6	-1.5	*
Slc31a1	Solute carrier family 31 (copper transporters), member 1	-1.5	
Spon2	Spondin 2, extracellular matrix protein	-1.5	
Ppp1r3c	Protein phosphatase 1, regulatory subunit 3C	-1.5	*
Plekhf1	Pleckstrin homology domain containing, family F, member 1	-1.5	

Hist1h2bn	Histone H2B type 1-N	-1.5	*
Gsdmc3	Gasdermin C3	-1.5	
Cyp3a11	Cytochrome P450, family 3, subfamily a, polypeptide 11	-1.5	
Bhlhb2	Basic helix-loop-helix family, member e40	-1.5	
Ephx2	Epoxide hydrolase 2, cytoplasmic	-1.5	
Sbk	Serine/threonine protein kinase	-1.5	
Tshz1	Teashirt zinc finger homeobox 1	-1.5	*
Tmem184a	Transmembrane protein 184A	-1.5	*
Mall	Mal, T-cell differentiation protein-like	-1.5	
Ypel3	Yippee like three	-1.5	*
Ctgf	Connective tissue growth factor	-1.5	*

Casp2	Caspase 2, apoptosis-related cysteine peptidase	-1.5	*
Gde1	Glycerophosphodiester phosphodiesterase 1	-1.5	*
Snurf	SNRPN upstream reading frame	-1.5	*
Acot4	Acyl-CoA thioesterase 4	-1.5	*
Cbr3	Carbonyl reductase [NADPH] 3	-1.5	
Slc46a3	Solute carrier family 46, member 3	-1.5	*
Zkscan14	Zinc finger with KRAB and SCAN domains 14	-1.5	
Oma1	OMA1 zinc metallopeptidase homolog	-1.5	*
Btbd12	BTB/POZ domain-containing protein 12	-1.5	
Ltc4s	Leukotriene C4 synthase	-1.5	
Mif4gd	MIF4G containing domain	-1.5	*

Retsat	Retinol saturase (all-trans-retinol 13,14-reductase)	-1.5	
Hmgcs2	Hydroxymethylglutaryl CoA synthase 2	-1.5	
Enpp7	Ectonucleotide pyrophosphatase/phosphodiesterase 7	-1.5	
Hist1h1c	Histone cluster 1, H1c	-1.5	*
Angptl4	Angiopoietin-related protein 4	-1.5	
Adck5	AarF domain containing kinase 5	-1.5	
Rps4y2	Ribosomal protein S4	-1.5	*
Myo5b	Myosin VB	-1.5	*
Chordc1	Cysteine and histidine-rich domain containing 1	-1.5	
Mycbpap	MYCBP associated protein	-1.5	
Hdac5	Histone deacetylase 5	-1.5	*

Plekha1	Pleckstrin homology domain-containing family A member 1	-1.5	*
Eml4	Echinoderm microtubule-associated protein-like 4	-1.5	*

Table 5. IPA Analysis of Colon Microarray.

Top Biological Functions

		# Molecules	Molecules	Range of p-Value
Diseases and Disorders	Dermatological Diseases and Conditions	23	*AREG/AREGB,CD52,CD74,CFD,CYP2E1,DUSP1,EGR1,FOS,GJB2,HBB,HBEGF,HLA-DQA1,HLA-DQB1,ID1,KRT13,KRT14,KRT16,KRT17,KRT5,KRT6A,SGK1,SLPI,TRPV6*	1.51E-10 - 1.55E-02

Immunological Disease	17	*Ang2 (includes others),CD52,CD74,CFD, DUSP1,EGR1,FAIM3,HBB,HBEGF,HLA-DQA1,HLA-DQB1,ID1,KRT16, KRT17,KRT5,KRT6A, MT1E*	1.51E-10 - 1.44E-02
Inflammatory Disease	18	*AREG/AREGB,CD52,CD74,CFD,DUSP1,EGR1,HBB,HBEGF,HLA-DMB,HLA-DQA1,HLA-DQB1,MT1E,KRT14,KRT16,KRT17,KRT5,KRT6A,TRPV6*	1.51E-10 - 7.76E-03
Cancer	39	*APOBEC1,AREG/AREGB,B4GALNT1, CD52,CD74,CFD, CLDN4,CYP2D6,CYP2E1,DUSP1,EGR1,FAIM3,FOS,GJB2,HBA1/HBA2,HBB,HBEGF,HOXD10,ID1,ITLN1,KRT13,KRT14,KRT16,KRT17,KRT5,KRT6A,LGALS7/LGALS7B,MRGPRF,MSMB, MT1E,OLFM4,PLK3,REG4,SGK1,SLC12A8,SLPI,TACSTD2,TRPV6*	6.93E-08 - 1.89E-02

	Developmental Disorder	15	*AREG/AREGB,DUOXA2, ID1,GJB2,HBA1/HBA2,HBB,HBEGF, HOXD10,KRT14,KRT16,KRT17,KRT5, KRT6A,SLC30A10, TACSTD2*	2.24E-07 - 1.89E-02
Molecular and Cellular Functions	Cell Morphology	22	*AREG/AREGB,B4GALNT1,CLDN15,CYP2E1,EGR1,FOS, HBA1/HBA2,HBB, HBEGF,HLA-DQA1,HLA-DQB1,HOXD10,ID1,KRT14,KRT16,KRT17,KRT6A,MT1E,PLK3,POU2AF1,SCNN1G,TRPV6*	6.23E-06 - 1.93E-02
	Cellular Movement	23	*Ang2 (includes others),AREG/AREGB,CD74,CLDN4, DUSP1,EGR1,FOS, GJB2,HBEGF,HOXD10,ID1,KRT16,KRT6A,LY6D,LYPD3, OLFM4,POU2AF1, Retnlb,SCNN1G,SGK1,SLPI,TACSTD2, TRPV6*	9.07E-06 - 1.56E-02

	Cell Death and Survival	27	*2010109I03Rik,Ang 2 (includes others),AREG/ARE GB,B4GALNT1,CD 74,CLDN4,CYP2E1 ,DUSP1,EGR1,FAI M3,FOS,GJB2,HBA 1/HBA2,HBB,HBE GF,ID1,KRT14,KR T6A,LGALS7/LGAL S7B,MT1E,MZB1,P LK3,POU2AF1,SG K1,SLC30A10,SLPI ,Sprr1a*	6.58E-05 - 1.93E-02
	Drug Metabolism	3	*B4GALNT1,CYP2D 6,CYP2E1*	1.48E-04 - 1.55E-02
	Molecular Transport	16	*ABCG5,APOBEC1, B4GALNT1,CFD,C LPS,CYP2E1,DUS P1,EGR1,FOS,FXY D4,HBA1/HBA2,H BB,MT1E,PLA2G4 C,SCNN1G,SGK1*	2.21E-04 - 1.16E-02
Physio-logical System Develop-ment and Function	Organ Morphology	20	*AREG/AREGB,B4G ALNT1,CLDN15,C LPS,DUOXA2,EGR 1,FOS,HBA1/HBA2 ,HBB,HBEGF,HLA -DQA1,HLA-DQB1,HOXD10,ID 1,KRT14,KRT16,K RT17,KRT5,KRT6A ,MT1E*	1.25E-06 - 1.93E-02
	Organismal Functions	8	*B4GALNT1,EGR1, HBEGF,KRT6A,MT 1E,SCNN1G,SLPI,T ACSTD2*	4.93E-06 - 7.76E-03

	Tissue Morphology	26	*ABCG5,AREG/AREGB,B4GALNT1,CD74,CLDN15,DUSP1,EGR1,FAIM3,FOS,HBA1/HBA2,HBB,HBEGF,HLA-DQA1,HLA-DQB1,HOXD10,ID1,KRT14,KRT16,KRT5,KRT6A,MT1E,POU2AF1,SCNN1G,SLPI,SGK1,TACSTD2*	4.93E-06 - 1.90E-02
	Digestive System Development and Function	11	*AREG/AREGB,ABCG5,CLDN15,CYP2E1,DUSP1,EGR1,FOS,KRT16,KRT5,KRT6A,MT1E*	8.89E-05 - 1.16E-02
	Organismal Development	23	*AREG/AREGB,CD74,CLPS,DUSP1,DUOXA2,EGR1,FOS,FXYD4,HBA1/HBA2,HBB,HBEGF,HLA-DQB1,HOXD10,ID1,KRT13,KRT14,KRT16,KRT17,KRT5,KRT6A,MT1E,POU2AF1,TACSTD2*	8.89E-05 - 1.93E-02

Top Canonical Pathways

		# Molecules	Molecules	Range of p-Value
1	B-Cell Development	3	*H2-DMb1,H2-Aa,H2-Ab1*	9.59E-02 - 4.51E-02

2	Cytotoxic T Lymphocyte-mediated Apoptosis of Target Cells	4	*H2-M2,H2-DMb1,H2-Aa,H2-Ab1*	4.83E-02 - 4.51E-02
3	Allograft Rejection Signaling	4	*H2-Ab1,H2-Aa,H2-DMb1,H2-M2*	4.83E-02 - 4.51E-02
4	OX40 Signaling Pathway	4	*H2-M2,H2-DMb1,H2-Aa,H2-Ab1*	4.83E-02 - 4.51E-02
5	Antigen Presentation Pathway	3	*Cd74,H2-DMb1,H2-Aa*	3.40E-02 - 9.78E-02

Top Molecules

	Log Ratio Up-Regulated		**Exp. Value**
1	*CFD*	(+1.57)	
2	*HBB*	(+1.56)	
3	*HBA1/HBA2*	(+1.37)	
4	*EGR1*	(+1.29)	
5	*AREG/AREGB*	(+1.15)	
6	*SCNN1G*	(+1.05)	

7	*MT1*	(+1.03)
8	*SLC30A10*	(+0.93)
9	*TRPV6*	(+0.93)
10	*FXYD4*	(+0.91)

Log Ratio Down-Regulated

1	*Retnlb*	(-3.79)
2	*Ang2 (includes others)*	(-3.26)
3	*KRT14*	(-2.02)
4	*KRT16*	(-1.63)
5	*PNLIPRP2*	(-1.55)
6	*TGM4*	(-1.48)
7	*ITLN1*	(-1.43)
8	*CLPS*	(-1.30)
9	*KRT6A*	(-1.22)
10	*PLA2G4C*	(-1.19)

Top Upstream Reg-ulators

	Upstream Regulator	**p-value of overlap**	**Target molecules in dataset**	**Mechanistic Network**

1	beta-estradiol	1.17E-10	*Ang2* (includes others),*AREG/AREGB,CD74,CLDN4,DUSP1,EGR1,FOS,GJB2,HBA1/HBA2,HBB,HBEGF,HLA-DQA1,ID1,KRT13,KRT16,KRT17,KRT5,KRT6A,MSMB,PADI1,SGK1,SLPI,Sprr1a,Sprr2f,TACSTD2*	*Creb,EGFR,ERBB2,ERK1/2,ESR1,FOS,NFkB (complex),NR3C1,RNA polymerase II,SMAD3,SP1,SP3,STAT3,TGFB1,TP53,TP63,beta-estradiol,estrogen receptor*
2	Diethylstilbestrol	6.73E-09	*AREG/AREGB,EGR1,FOS,GJB2,HBEGF,KRT13,MSMB,SLPI,Sprr2f,TGM4*	*Creb,ELK1,ERBB2,ERK,ESR1,FOS,NFKB1,SMAD3,STAT1,TGFB1I1,TP53,diethylstilbestrol,estrogen receptor,testosterone*
3	FGF19	1.97E-08	*CYP2D6,EGR1,FOS,HLA-DQB1,ID1,SGK1,SLPI*	*Creb,ERK1/2,FGF19,NR3C1,SMAD3,SP3,TP53*
4	progesterone	2.24E-08	*AREG/AREGB,CFD,CLDN4,CYP2E1,FOS,GJB2,HBEGF,HLA-DQB1,ID1,KRT5,SCNN1G,SGK1,TACSTD2*	*CEBPB,Creb,D-glucose,ERK1/2,ESR1,NR3C1,PDX1,SMAD3,SP1,SP3,STAT3,TGFB1,TP53,progesterone*
5	17-alpha-ethinylestradiol	6.21E-08	*ABCG5,CD74,CYP2D6,FOS,FXYD4,KRT14,PLK3,SGK1*	

Table 6. IPA Analysis of Proximal Small Intestine Microarray.

Top Biological Functions

		# Molecules	Range of p-Value	Molecules
Diseases and Disorders	Cancer	111	4.42E-07 - 1.42E-02	*ABCC2,ABCC3,ACOT9,ADORA2B, AKR1B10,ALDH1A1,ALDOB,ANGPTL4,ANPEP,APOE,AREG/AREGB,ATP2B1,BCL6,BMP3,BTG1,CAPG,CAPN5,CASP2,CASP4,CAT,CD52,CD74,CEACAM1,CFD,CIRBP,CLDN2,CLDN4,CLIC5,CNNM2,COTL1,CTGF,CTSH,CXCL16,CXCL9,CYBA, CYP2B6,CYP3A5,DPEP1,DUSP6,EDN1,EIF2AK2,ENO1,EPHX1,EPHX2,FBLN1,FHL2,FUT4,G6PC,GCH1, GPI,GPRC5A,GSTK1,GSTM1,GUK1,HBA1/HBA2,HBEGF,HDAC5,HIST1H1C,HIST1H2BD,HIST2H2BE (includes others),HMOX1,HPGD,IGSF9,IL33, ITPR1,KLF6,LAMA3,LCN2,LTF,Lyz1/Lyz2,MMP9,MUC4,NFKBIA,NMU,NT5E,OSR2,PGLYRP1,PITPNA,PLK3,PMP22,PPAPDC1B,PPP1R3C,PSAT1,PTK6,RAB31,RDH16,RMND1, RORC,SCD,SLC17A4,SLC27A2,SLC40A1,SLC46A3,SLC6A3,SOCS2,SOCS3,SPON2,ST5,SULT1C2,SUSD2, TGFBR2,TGM2,TNFRSF11A,TNS4, TRAK1,TSHZ1,TSPAN12,UGCG,UGT2B7,UPP1,USP2*
	Infectious Disease	67	9.77E-07 - 1.07E-02	*ABCC2,ALOX5AP,AMDHD2,ANPEP,APOE,AREG/AREGB,BCL6,BMP1,CAPG,CASP2,CASP4,CAT,CD74, CEACAM1,CFD,CHORDC1,CISH, CTGF,CXCL9,CXCL16,DAK,EIF2AK2,ETHE1,FUT4,GPI,GSTM1,HIST1H1C,HIST1H2BD,HIST2H2BE (includes others),HLA-DMA,HLA-DQB1,HMOX1,HPGD,IL33,LCN2,LTF,LY6D,Lyz1/Lyz2,MMP9,MOV10, MYO5B,NAIP,NFKBIA,OSBPL3,PA*

			NK3,PGLYRP1,PLK3,PMM1,POU2AF1,PRSS27,RAB31,RAB32,Reg3g,RORC,RSAD2,SEPP1,SLC31A1,SLC6A3,SOCS2,SOCS3,SPON2,STK39,TAGLN2,TGFBR2,TGM2,TRIM38,UBD
Neurological Disease	57	3.94E-06 - 1.04E-02	*ADORA2B,ALOX5AP,APOE,ARSG,ATP2B1,BCL6,CA13,CA4,CASP2,CD52,Cd55/Daf2,CD74,CIRBP,CTGF,CXCL16,DBP,EDN1,EIF2AK2,ELOVL1,EPHX2,ETHE1,FLVCR1,G6PC,GCH1,GPI,HDAC5,HLA-DMA,HLA-DMB,HLA-DQB1,HLA-DRB1,HMGCS2,HMOX1,ITPR1,IVNS1ABP,LCN2,LTF,MAOB,MMP9,NAIP,NR1D1,PITPNA,PMP22,PPT1,PSAT1,RSAD2,SCD,SEPP1,SLC17A4,SLC30A10,SLC6A3,TGFBR2,TGM2,TSHZ1,UGCG,USP2*
Psychological Disorders	23	3.94E-06 - 8.72E-03	*ADORA2B,APOE,CA13,CA4,CASP2,CASP4,CD74,CFD,CTGF,CXCL16,EIF2AK2,HLA-DQB1,HLA-DRB1,HMOX1,LCN2,MAOB,NAIP,SCD,SEPP1,SLC6A3,TGFBR2,TGM2,TSHZ1*
Skeletal and Muscular Disorders	64	8.75E-06 - 1.04E-02	*ACTA1,ADORA2B,ALOXA5AP,APOE,AREG/AREGB,ATP2B1,BCL6,BMP1,CAT,CASP2,CD52,Cd55/Daf2,CD74,CEACAM1,CIRBP,CTGF,CXCL16,CXCL9,CYBA,CYP2B6,CYP4F3,DBP,DNASE1,EDN1,ELOVL1,ENO1,EPHX2,FHL2,G6PC,GCH1,GLIPR2,GPI,HBEGF,HDAC5,HLA-DMA,HLA-DMB,HLA-DQB1,HLA-DRB1,HMGCS2,HMOX1,IL33,ITPR1,IVNS1ABP,LCN2,LTF,MAOB,MMP9,NFKBIA,NMU,NR1D1,PGLYRP1,PMP22,PSAT1,RSAD2,SCD,SLC17A4,SLC6A3,SOCS3,TGFBR2,TGM2,TNFRSF11A,TNFRSF21,UGCG,USP2*

Molecular and Cellular Functions	Cell Death and Survival	110	1.88E-10 - 1.43E-02	*2010109I03Rik,ABCC3,ADORA2B,ALDH1A1,Ang2 (includes others),ANGPTL4,ANPEP,APOE,AQP7,AREG/AREGB,ASNS,ATP2B1,BCL6,BMP1,BMP8B,BOK,BTG1,CA4,CASP2,CASP4,CAT,CCBL1,Cd55/Daf2,CD74,CEACAM1,CISH,CLDN4,CTGF,CXCL9,CYBA,CYP2B6,DNASE1,DUSP6,EDN1,EIF2AK2,ENO1,EPHX1,EPHX2,ETHE1,FBLN1,FHL2,FUT4,G6PC,GCLM,GPI,GSTM1,HBA1/HBA2,HBEGF,HDAC5,HIST1H1C,HLA-DMA,HLA-DRB1,HMOX1,IL33,ITPR1,IVNS1ABP,KHK,KLF6,LAMA3,LCN2,LGMN,LTF,Lyz1/Lyz2,MAOB,MIF4GD,MLKL,MMP9,MUC4,NAIP,NFKBIA,NR1D1,NT5E,PCSK9,PDX1,PGLYRP1,PIK3AP1,PITPNA,PLEKHF1,PLK3,PMP22,POU2AF1,PPT1,PTK6,RAB32,REG3A,RORC,SCD,SCIN,SEPP1,SIRT3,SLC11A2,SLC28A1,SLC30A10,SLC31A1,SLC40A1,SLC6A3,SOCS3,TAGLN2,TBC1D15,TGFBR2,TGM2,TINAG,TNFRSF11A,TNFRSF21,TNS4,TSC22D1,UBD,UGCG,USP2*
	Cellular Movement	61	1.49E-06 - 1.42E-02	*Aldh1a7,ALOX5AP,AMICA1,Ang2 (includes others),ANGPTL4,ANPEP,APOE,AREG/AREGB,BARX2,CAPG,CAT,Cd55/Daf2,CD74,CEACAM1,CLDN4,CTGF,CXCL16,CXCL9,DUSP6,EDN1,FHL2,FUT4,G6PC,GPI,HBEGF,HMOX1,IL33,ITPR1,LAMA3,LCN2,LGMN,LTC4S,LTF,LY6D,Lyz1/Lyz2,MEP1B,MMP9,MUC4,MYO5B,NFKBIA,NMU,NR1D1,NT5E,NUS1,PMP22,POU2AF1,PTK6,REG3A,Retnlb,RORC,SOCS2,SOCS3,SPON2,TAGLN2,TGFBR2,TGM2,TM4SF5,TNFRSF11A,TNFRSF21,TNS4,USP2*

Lipid Metabolism	71	1.67E-06 - 1.49E-02	*ABCC2,ABCC3,Acot1,ACOT4,ACOT9,AHD4,ADIPOR2,ADORA2B,Akr1b7,AKRB10,ALDH1A1,Aldh1a7,ALOX5AP,ALPI,ANGPTL4,ANPEP,APOE,AQP7,AQP8,B4GALT6,BDH2,CAT,CD74,CEACAM1,CFD,COTL1,CYP2B6,CYP3A5,Cyp4F16/Gm9705,CYP4F3,DBP,EDN1,ELOVL1,EPHX1,EPHX2,G6PC,GSTK1,HMOX1,HPGD,IL33,LCN2,LTC4S,LTF,MLXIPL,MMP9,NCOR1,NKFBIA,NMU,NUS1,NPC1L1,OMA1,PCSK9,PITPNA,PLEKHA1,PNLIPRP2,PPP1R3C,RETSAT,RDH16,SCD,SEPP1,SIRT3,SLC27A2,SLC51B,SOCS3,STK39,TGFBR2,TGM2,UGCG,UGT2B7*
Small Molecule Biochemistry	82	1.67E-06 - 1.49E-02	*ABCC2,ABCC3,ADH4,ADIPOR2,ADORA2B,Akr1b7,AKR1B10,ALDH1A1,ALOX5AP,ALPI,ANGPTL4,APOE,AQP7,AQP8,CAT,CEACAM1,CFD,Cyp4f16/Gm9705,CYP4F3,EDN1,EPHX2,GMDS,GUK1,G6PC,GSTK1,HMOX1,HPGD,IL33,LCN2,LTC4S,MLXIPL,MMP9,NMU,NUS1,NPC1L1,NT5A,OMA1,PCSK9,PITPNA,PPP1R3C,RDH16,SCD,SIRT3,SLC28A1,SLC28A2,STK39,TGFBR2,TGM2,UGCG,Acot1,ACOT4,ACOT9,B4GALT6,CD74,COTL1,DBP,LTF,MLXIPL,NFKBIA,NPC1L1,PITPNA,SLC27A2,SLC51B,ADH4,Aldh1a7,CYP2B6,CYP3A5,RRETSAT,NCOR1,NUS1,GPI,ANPEP,ELOVL1,SOCS3,HPGD,LTC4S,MAOB,SLC6A3,SULT1C2,DUSP6,USP2,NOXA1*

	Molecular Transport	78	1.81E-06 - 1.49E-02	*ABCC2,ABCC3,ADIPOR2,ADORA2B,Akr1b7,ALDH1A1,ALOX5AP,ALPI,ANGPTL4,APOE,AQP7,AQP8,ATP2B1,BCL6,CA4,CAT,CD74,CEACAM1,CLCA1,CLCA4,CLDN2,CLDN4,CLIC5,CYP2B6,EDN1,EPHX1,EPHX2,FLVCR1,GCH1,G6PC,GSTK1,HBA1/HBA2,HLA-DMA,HMOX1,HPGD,IL33,ITPR1,LCN2,LTC4S,LTF,MLXIPL,MMP9,MYO5B,NCOR1,NFKBIA,NMU,NPC1L1,NUDT4,OMA1,PCSK9,PDX1,PITPNA,PNLIPRP2,PPME1,PPP1R3C,PPT1,RDH16,SCD,SEPP1,SIRT3,SLC11A2,SLC13A2,SLC16A10,SLC17A4,SLC27A2,SLC28A1,SLC28A2,SLC30A2,SLC31A1,SLC40A1,SLC51B,SLC6A3,STK39,TGFBR2,TGM2,TRAK1,UGCG*
Physiological System Development and Function	Tumor Morphology	28	8.52E-07 - 1.42E-02	*ABCC2,ANPEP,AREG/AREGB,APOE,BOK,CASP2,CAT,CD74,CEACAM1,CLDN4,CTGF,EDN1,EIF2AK2,ENO1,FHL2,HBEGF,HMOX1,ITPR1,LAMA3,MMP9,MUC4,NFKBIA,RORC,SOCS3,TGFBR2,TNFRSF21,TSC22D1,UGCG*
	Hematological System Development and Function	55	1.49E-06 - 1.49E-02	*ANPEP,APOE,BCL6,CD74,CISH,CLCA1,CXCL16,EIF2AK2,FLVCR1,FUT4,HBA1/HBA2,HBEGF,HLA-DMA,HLA-DQB1,HLA-DRB1,HMOX1,IL33,LAMA3,LCN2,LGMN,LTC4S,MMP9,NCOR1,NFKBIA,PIK3AP1,POU2AF1,Reg3g,RORC,SLC11A2,SOCS3,SPON2,TGFBR2,TGM2,TNFRSF11A,TNFRSF21,Cd55/Daf2,CEACAM1,CTSH,CXCL9,DNASE1,EDN1,HDAC5,HLA-DMB,LTF,NT5E,REG3A,SOCS2,ALOX5AP,AMICA1,CAPG,CTGF,FHL2,G6PC,NMU*

	Immune Cell Traffickin g	41	1.49E-06 - 1.49E-02	*ALOX5AP,AMICA1,APOE,CAPG,C d55/Daf2,CD74,CEACAM1,CISH,C TGF,CTSH,CXCL16,CXCL9,DNASE 1,EDN1,FHL2,FUT4,G6PC,HBEGF ,HDAC5,HLA-DMA,HLA-DMB,HLA-DQB1,HMOX1,IL33,LCN2,LTC4S,L TF,Lyz1/Lyz2,MEP1B,MMP9,NFKB IA,NMU,NT5E,POU2AF1,REG3A,R ORC,SOCS3,SPON2,TGFBR2,TGM 2,TNFRSF21*
	Digestive System Developm ent and Function	26	1.56E-05 - 1.04E-02	*ABCC2,Akr1b7,APOE,AREG/AREG B,ARSG,BMP1,CEACAM1,EDN1,E NPP7,G6PC,KLF6,LAMA3,LGMN, MLXIPL,NCOR1,NPC1L1,PDX1,PI TPNA,PPT1,PTK6,RORC,SCD,SOC S3,TGFBR2,TNFRSF11A,TSHZ1*
	Hepatic System Developm ent and Function	12	1.56E-05 - 1.04E-02	*ABCC2,Akr1b7,APOE,EDN1,G6PC, KLF6,LGMN,MLXIPL,PITPNA,SCD ,SOCS3,TGFBR2*

Top Canonical Pathways

		# Molecules	Range of p-Value	Molecules
1	LPS/IL-1 Mediated Inhibition of RXR Function	17	2.19E-01 - 2.56E-01	*ABCC2,ABCC3,ALDH18A1,ALDH1 A1,ALDH4A1,APOE,CAT,CYP2B6, CYP3A5,GSTA3,GSTK1,GSTM1,HM GCS2,IL33,MAOB,SLC27A2,SULT1 C2*
2	PXR/RXR Activation	10	3.83E-01 - 2.56E-01	*ABCC2,ABCC3,ALDH1A1,Aldh1a7, CYP2B6,CYP3A5,G6PC,GSTM1,H MGCS2,SCD*

3	Xenobiotic Metabolism Signaling	16	1.24E-01 - 2.56E-01	*ABCC2,ABCC3,ALDH18A1,ALDH1A1,ALDH4A1,CAT,CYP2B6,CYP3A5,GSTA3,GSTK1,GSTM1,HDAC5,HMOX1,MAOB,SULT1C2,UGT2B7*
4	Stearate Biosynthesis I (Animals)	6	6.89E-03 - 9.26E-02	*ACOT4,ACOT9,BDH2,ELOVL1,PPT1,SLC27A2*
5	T Helper Cell Differentiation	6	2.88E-02 - 2.14E-01	*H2-Dma,H2-DMb2,H2-Ab1,H2-Eb1,Rorc,Tgfbr2*

Top Molecules

	Log Ratio Up-Regulated	Exp. Value
1	UBD	(+2.86)
2	DBP	(+2.53)
3	AKR1B10	(+2.04)
4	SOCS3	(+1.88)
5	SLC2A6	(+1.72)
6	Reg3g	(+1.64)
7	SLC40A1	(+1.56)
8	CXCL16	(+1.53)
9	REG3A	(+1.52)
10	SLC11A2	(+1.48)

		Log Ratio Down-Regulated
1	G6PC	(-2.91)
2	Slc5a4b	(-2.41)
3	NT5E	(-2.34)
4	Muc3	(-2.19)
5	2010109I03Rik	(-2.11)
6	ITPR1	(-1.92)
7	AQP8	(-1.89)
8	SLC28A1	(-1.68)
9	SEPP1	(-1.62)
10	Ang2 (includes others)*	(-1.52)

Top Upstream Regulators

	Upstream Regulator	Predicted Activation State	Activation z-score	p-value of overlap	Target molecules in dataset	Mechanistic Network
1	CFTR	Inhibited (14 of 17 genes have expression direction consistent with inhibitio	-2.578	1.34E-14	*Acaa1b,ASNS,CA4,CAPG,CYP2B6,CYP3A5,CYP4V2,EPHX2,HPGD,ITPR1,NFKBIA,PGLYRP1,Reg3g,Retnlb,SCD,SLC27A2,SOCS3*	

		n of CFTR)				
2	POR		-1.265	1.68E-14	*ABCC3,Acot1,ACTA1,ADH4,ALDH1A1,Aldh1a7,AQP8,ASNS,CD74,CYP2B6,ELOVL1,ETHE1,GSTA3,GSTM1,HLA-DMA,HLA-DQB1,HLA-DRB1,HMOX1,NAIP,RETSAT,SCD,SUSD2,UBD*	
3	TO-901317		-0.589	6.59E-14	*Acaa1b,Acot1,ADIPOR2,Akr1b7,APOE,CASP4,CYP3A5,ELOVL1,ENO1,EPHX1,FBP1,G6PC,GPI,HMGCS2,LTC4S,MMP9,NAIP,NFKBIA,NR1D1,PCSK9,PGLYRP1,SCD,SLC27A2,TGM2*	*JUN,NFKB1,NFkB (complex),NR1H3,PPARA,PPARG,PPARGC1A,RELA,RXRA,SREBF2,STAT1,TO-901317,cholesterol*

4	IFNG	1.946	1.21E-12	*Acot1,ADORA2B,ALOX5AP,ANGPTL4,BTG1,C1QC,CASP2,CASP4,CAT,CCBL1,CD74,CEACAM1,CIRBP,CISH,CLIC5,CTGF,CTSH,CXCL16,CXCL9,CYBA,DBP,EDN1,EIF2AK2,FBLN1,FBP1,FHL2,GCH1,GLIPR2,HBEGF,HLA-DMA,HLA-DMB,HLA-DQB1,HLA-DRB1,HMOX1,KLF6,LCN2,MMP9,MUC4,NFKBIA,NR1D1,PDX1,RORC,RSAD2,SEPP1,SERPINH1,SLC11A2,SLC28A1,SLC28A2,SLC40A1,SOCS2,SOCS3,TGFBR2,TNFRSF11A,UBD*	*CEBPA,CSF2,FOS,IFNG,IL1B,JUN,NFkB (complex),NR3C1,PPARG,RELA,STAT,STAT1,STAT3,STAT5A,STAT5B,TNF*

5	PPARA	-0.711	1.77E-12	*Acaa1b,Acot1,ACTA1,Akr1b7,ALDOB,ANGPTL4,AQP7,AQP8,ASNS,CAT,CYBA,CYP2B6,EDN1,ETHE1,G6PC,GSTK1,HIST1H1C,HMGCS2,HMOX1,KHK,LY6D,MMP9,NFKBIA,NPC1L1,NR1D1,PAQR9,RETSAT,SCD,SLC27A2,SOCS2*	*ADIPOQ,JUN,NFKB1,NFkB (complex),NR1H3,PPARA,PPARD,PPARG,PPARGC1A,RXRA,SP1,STAT3,bezafibrate,fenofibrate,pirinixic acid*

Table 7. Core genes enriched in *Kcnq1* KO.

Gene	Position in *Kcnq1* KO ranked gene list	Expression ratio, *Kcnq1* KO/ *Kcnq1*	Gene	Position in *Kcnq1* KO ranked gene list	Expression ratio, *Kcnq1* KO/ *Kcnq1*	Gene	Position in *Kcnq1* KO ranked gene list	Expression ratio, *Kcnq1* KO/ *Kcnq1*
upregulated in *Cftr*			**downregulated in *Muc2***					
Socs3	6	3.70	*Cyp3a11*	29162	0.66			
Prss32	28	1.86	*Mod1*	29334	0.82			
Tspan12	30	1.70	*Gsta2*	29363	0.71			
Pglyrp1	102	1.97	*Gstm2*	30064	0.65			
Reg3g	113	3.11	*Cyp2b10*	30069	0.68			
Blnk	131	1.35	*Gstm4*	30120	0.77			

Ptk6	153	1.52	*Reg3a*	30408	0.48			
S100a11	155	1.46	*Gstm6*	30496	0.75			
Tmsb10	196	1.40	*Gclm*	30521	0.65			
Asns	213	1.62	*Gstt1*	30548	0.75			
Ceacam1	225	1.47	*Ephx2*	30557	0.66			
Il7	267	1.12	*Cyp4v3*	30575	0.57			
Slc2a1	424	1.34	*Cat*	30666	0.41			
Pycr2	685	1.25	*Gstk1*	30687	0.46			
Hddc2	941	1.09						
Cxcl10	1134	1.06						
Myl7	1176	1.14						
Idi1	1183	1.12						
Ssbp4	1250	1.17						
downregulated *in Cftr*			*Bst1*	28894	0.85	*Osbpl1a*	30189	0.92
Ugt2b34	26831	0.93	*Cyp3a11*	29162	0.66	*Slc22a4*	30243	0.77
Slc7a8	26840	0.92	*Guca2b*	29196	0.77	*Crat*	30345	0.85
Slc6a20b	26933	0.90	*Lamb3*	29368	0.85	*Hpgd*	30381	0.59
Mtap7d1	26962	0.98	*Scd1*	29425	0.60	*Ace2*	30541	0.71
Zfp28	27098	0.97	*Calb2*	29585	0.87	*Tsc22d1*	30544	0.53
Adh1	27155	0.89	*Angptl4*	29678	0.68	*Gstt1*	30548	0.75
Sgk1	27221	0.71	*Aqp1*	29687	0.70	*Ephx2*	30557	0.66
Xpot	27520	0.98	*Pdk4*	29737	0.71	*Cyp4v3*	30575	0.57
Slc46a1	27675	0.86	*Hsd17b4*	29752	0.79	*Sec14l2*	30617	0.71
Reln	27747	0.96	*Pdzk1*	29754	0.79	*Slc27a2*	30628	0.54
Peg3	28183	0.96	*Arg2*	29768	0.70	*Pmp22*	30648	0.36
Cyb5r3	28216	0.94	*Abcc6*	29809	0.82	*Fbln1*	30658	0.60
Trp53bp1	28382	0.95	*Samhd1*	29845	0.80	*Cat*	30666	0.41
Cox6b2	28426	0.95	*Slc5a11*	29923	0.83	*Fbp1*	30701	0.51
Pgrmc2	28436	0.92	*Fbp2*	29987	0.82	*Aldh1a1*	30720	0.56
Slc5a6	28519	0.81	*Cyp2b10*	30069	0.68	*Rdh7*	30791	0.45
Ace	28551	0.80	*Acaa1b*	30099	0.63	*Sema7a*	30796	0.70
Tfam	28662	0.89	*Rfx1*	30102	0.87	*Bmp1*	30821	0.55
Cmbl	28673	0.88	*Retsat*	30132	0.67	*Itpr1*	30850	0.26
Fam134b	28786	0.82	*Adora2b*	30142	0.58			
Faah	28857	0.90	*Pepd*	30175	0.80			

Table 8. Expression of *Areg*

	Whole colon – Fold change	Method
***Kcnq1* KO**	↑ 2.2	Illumina WG6 mouse bead array
***Muc2* KO**	↑ 11.7	Agilent 44K mouse microarray
***Cftr* KO**	↑ 1.67	qRT-PCR

Kcnq1 KO: Table 3

Muc2 KO: Fijneman & Cormier, unpublished data, employing *Muc2* KO mouse (Velcich *et al.* 2002)

Cftr KO: Than, unpublished data, employing conditional KO of *Cftr* (Hodges *et al.* 2008).

Table 9. Patient characteristics

	Total Population	
Clinicopathological variable	**n / mean / median**	**% / st.dev. / range**
General characteristics		
All	507	
Gender		
male (%)	320	63.1
female (%)	187	36.9
Age at liver resection		

mean (st.dev.)	63.3	0.5
median (range)	64.3	27.6-83.9
OS after liver resection in months		
mean (st.dev.)	37.2	1.7
median (range)	27	3.0-212.0
Maximal CRCLM diameter in cm		
mean (st.dev.)	4.1	0.1
median (range)	3.5	0.2-22.0
Number of CRCLMs		
mean (st.dev.)	2	0.1
median (range)	2	1-12
Systemic treatment of liver metastases		
preoperative (%)	60	11.8
perioperative (%)	12	2.4
postoperative (%)	97	19.1
none (%)	321	63.3
unknown (%)	17	3.4

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank Dr Karl Pfeifer at the NIH for generously providing *Kcnq1* mutant mice and reviewing the manuscript. Research was supported by a grant to RC and DL (NCI R01 CA134759-01A1) and to JG and RF (Center for Translational Molecular Medicine, DeCoDe project grant 03O-101) and to TKS (NCI R00 4R00CA151672-02).

Personal Contribution. B.L.N.T contributed to the design of the research. B.L.N.T contributed to all mouse husbandry and tumor analysis. B.L.N.T isolated/processed RNAs for Illumina microarrays and analyzed data. B.L.N.T designed primers, performed qRT-PCR for gene validation, analyzed data and statistical tests. B.L.N.T carried out IPA on microarray datasets, conducted organoid experiments and performed statistical analysis.

CHAPTER THREE

The cystic fibrosis transmembrane conductance regulator (*CFTR*) is a tumor suppressor in the gastrointestinal tract

Than, B.L.N.[1,2], Luczak, T.[1], O'Sullivan, M.G.[3], Rod, A.[1], Zhang, Y[4], Starr, T.K.,[5] Largaespada, D.A.[6], Scott, P.M.[1], and Cormier, R.T.[1]

[1]Department of Biomedical Sciences, University of Minnesota Medical School, Duluth, MN, USA; [2]Toxicology Graduate Program, University of Minnesota, Duluth, MN, USA; [3]College of Veterinary Medicine, University of Minnesota, St Paul, MN, USA; [4]University of Minnesota Supercomputing Institute, Minneapolis, MN, USA; [5]Department of Obstetrics, Gynecology and Women's Health, Masonic Cancer Center, University of Minnesota Medical School; Minneapolis, MN, USA [6]Department of Genetics, Cell Biology and Development, Center for Genome Engineering, Masonic Cancer Center, University of Minnesota, Minneapolis, MN, USA.

August 2013. Manuscript in preparation.

Abstract. *Sleeping Beauty* DNA transposon-based forward genetic screens in mice identified the cystic fibrosis transmembrane conductance regulator gene, *Cftr*, as a gastrointestinal (GI) candidate cancer gene. CFTR promotes chloride and bicarbonate secretion, playing an essential role in ion and acid-base homeostasis in the GI tract. To investigate the effects of CFTR dysregulation on GI cancer, we generated Apc^{Min} mice that carried an intestinal-specific knockout of *Cftr*. Our results indicate that *Cftr* is a tumor suppressor gene in the intestinal tract as both heterozygous and homozygous *Cftr KO* mice developed significantly more tumors in the colon and the entire small intestine. In $Apc^{+/+}$ mice aged to ~ 1 year, *Cftr* deficiency alone caused the development of intestinal tumors as well as pancreatic abnormalities. Colon organoid formation was significantly increased in organoids created from *Cftr* mutant mice compared to wildtype

controls, suggesting a potential role of *Cftr* in regulating the intestinal stem cell compartment. To identify the changes in gene expression from the loss of *Cftr*, we performed microarray studies in the colon and the small intestine. Microarray data identified dysregulated genes that belong to groups of immune response, ion channel, intestinal stem cell and other growth signaling regulations. These associated clusters of genes were confirmed by pathway analysis using Ingenuity Pathway Analysis (IPA). Defining the tumor suppressive mechanism of CFTR can lead to its use as a potential therapeutic target for GI cancer.

INTRODUCTION

CFTR encodes for a member of the ATP-Binding Cassette (ABC) protein family and plays an essential role in anion regulation and tissue homeostasis of various epithelia. CFTR proteins are found primarily at the apical surfaces of epithelial cells in the airways, pancreas, stomach, intestine, kidney and testis (Hodges *et al.* 2008) where their major role is to regulate chloride secretion, and in the sweat glands where they modulate the reabsorption of chloride from sweat. In the GI tract, *CFTR* expression is weak in the stomach but strong along the intestine. Within the small intestine, *CFTR* expression is the strongest in the duodenum (including high expression in the mucus secreting Brunner's glands (Collaco *et al.* 2013)) with a decreasing expression gradient towards the ileum (Strong *et al.* 1994). In the large intestine, *CFTR* expression is moderate. Along the crypt-villi axis, *CFTR* expression is highest in the crypts of the small intestine and near the crypt bases of the large intestine. More recent studies have reported CFTR's expression on the apical membrane of intestinal crypt cells (Jakab *et al.* 2011, Linley *et al.* 2013).

More than 1900 *CFTR* mutations have been identified in human diseases (Ferec and Cutting 2012). The predominant mutation is a deletion of phenylalanine at position 508 (*CFTR F508Δ*) that causes misfolding, retention in the ER and early degradation of CFTR proteins, ultimately disrupting their function at the cell membrane. However, it has been noted that a subpopulation of *F508Δ-CFTR* patients still display residual chloride

secretion. Other mutations of *CFTR* have also been found to impair protein folding, gating and conductance. *CFTR* mutations cause Cystic Fibrosis (CF), an autosomal recessive genetic disorder with a basic defect caused by malfunction of chloride conductance. Due to CFTR's crucial role in fluid homeostasis of diverse tissues, clinical features of CF ranges from meconium ileus to pancreatic exocrine insufficiency, liver disease, chronic lung infection, GI complications and shortened lifespan (Hostos *et al.* 2011). Major phenotypes seen in CF patients occur in the lungs and airways with characteristics including impaired mucus clearance, infection and inflammation. In a normal condition, anion secretions provides the airways with a thin airway surface liquid layer adjacent to the apical membrane and a high-viscosity gel layer where secreted mucins trap debris and microorganisms. In CF, defective or absent CFTR leads to dysfunctional anion secretion which affects the formation and function of these protective layers, subsequently leading to the accumulation of mucus in the airways. Disruption in the protective mucus layer in the lung can potentially cause colonization of bacteria such as *Pseudomonas aeruginosa* and *Staphylococcus* (Tiddens *et al.* 2013, Cohen-Cymberknoh *et al.* 2013). This leads to an exaggerated airway inflammation while the protective immune system fails to eradicate invading pathogens. Once advanced, chronic bacterial infection can progressively destroy the lung and ultimately ends with respiratory failure.

In the GI tract, CFTR plays a multifunctional role in the process of ion and acid-base transport crucial for GI physiology. Considered as a major ion transporter in the intestinal crypt epithelium, CFTR regulates the transepithelial export of Cl^- and HCO_3^-. The conductive properties of CFTR exhibit a Cl^-:HCO_3^- ratio of approximately 4:1. The major role of HCO_3^- export by CFTR in maintaining the alkaline intracellular pH levels in the duodenal epithelium is illustrated by the abnormally low pH level in CF patients due to the absence of CFTR and defective bicarbonate secretion from the apical membranes of intestinal epithelial cells. (Wilschanski M *et al.* 2007). In addition, because of its role in chloride secretion, CFTR also indirectly regulates water homeostasis which is important for the protective function of the mucus layer. Overall, CFTR channel

dysregulation leads to a wide array of changes in the GI tract. *First*, CFTR plays an important role in maintaining proper pH levels. This is supported by the gradient of CFTR localization and expression in the small intestine as CFTR levels and secretion of bicarbonate are highest in the proximal small intestine – the region that receives high acid influx from the stomach. The small intestine also receives a high volume of bicarbonate-rich fluid from the pancreas. Disruption of CFTR function can therefore lead to gastric acid not being properly neutralized, causing ineffective digestive and absorptive function of the small intestine. *Second*, besides its absorptive function, the intestine also secrets a large volume of fluid to flush out unwanted materials and maintain a relatively sterile protective environment for the intestinal crypts. Thus, small intestinal bacterial overgrowth (SIBO) is found in one half of CF patients. *Third*, in the proximal large intestine, CFTR disruption depletes water hydration and impairs intestinal expulsion of toxic waste. Therefore, a high percentage of infants with CF suffer from meconium ileus (MI) – a severe intestinal blockade which can lead to rupture and sepsis if untreated. (The adult form of MI is known as distal intestinal obstructive syndrome (DIOS)). *Fourth*, it has been suggested that CFTR is necessary for the proper function of the intestinal mucus layer. Several lines of previously reported evidence support this observation. It was found that CF mice with a conditional knockout of *Cftr* displayed intestinal mucus accumulation (Hodges *et al.* 2008). Bicarbonate and functional CFTR were concluded to be crucial for proper secretion of intestinal mucin, hence the link to abnormal mucus phenotype in CF (Gustafsson *et al.* 2012).

With its highly important role in the normal physiology of various organs such as the GI tract, disruption of CFTR can potentially contribute to a wide array of cancers including esophageal, gastric, hepatobiliary, and GI cancers relative to the U.S. population (Gelfond and Borowitz 2013). Specifically, a 10-year study in CF patients revealed an increased risk of digestive tract cancer, with large numbers of cancers detected in the small intestine, colon and biliary tract (Maisonneuve *et al.* 2003). Defective expression of CFTR in GI tissues may contribute to increased inflammation and increased cell turnover. Other high-risk factors such as defective mucus function and

excessive acid exposure can predispose patients to GI cancer. In fact, CF patients have been recorded to experience malabsorption or intestinal stasis, and in some cases chronic inflammation in the small intestine that may lead to preinvasive changes including tumor development.

Cftr was identified as a top candidate GI cancer gene in multiple *Sleeping-Beauty* (SB) DNA-tranposon based forward mutagenesis screens in mouse intestine. In an *Apc* wild-type screen, *Cftr* mutations were found in 8 tumors (6%) (Starr *et al.* 2009) where SB transposon insertions were found in on both forward and reverse strand orientation, consistent with a well-accepted model that the SB transposon acts to disrupt function of the gene. An SB screen using mice on the Apc^{Min} background found *Cftr* mutations in 102 tumors (23%), ranking it #70 of 641 CIS genes (March *et al.* 2011). A screen in a $p53^{R27OH}$ mutant background identified *Cftr* mutations in 12 tumors (20%) (TKS, manuscript in preparation).

To investigate the role of Cftr in GI cancer, we measured GI cancer phenotypes in Apc^{Min} mice carrying an intestinal-specific knockout of *Cftr*. We also examined a smaller number of $Apc^{+/+}$ *Cftr* KO mice that were aged to ~ 1 year. We performed microarrays using RNA isolated from the colon and proximal small intestine and identified gene sets and potential signaling pathways associated with the loss of *Cftr*. Our results indicate that *Cftr* is a tumor suppressor in the GI tract.

RESULTS

Loss of *Cftr* enhances tumor multiplicity in Apc^{Min} mice.

To investigate the role of *Cftr* in mouse GI cancer, we crossed $Cftr^{fl10}$ mice with Villin-Cre mice where Cre recombinase expression is driven by the Villin promoter, creating an intestinal specific knockout of *Cftr* ($Cftr^{fl/fl}$ Cre^{+}). These mice were then introgressed into the Apc^{Min} model of GI cancer. Crosses between female $Cftr^{fl/fl}$ mice and male Apc^{Min} $Cftr^{fl/+}$ Cre^{+} mice generated a variety of Apc^{Min} test mice that were either

wildtype, heterozygous or homozygous mutant for *Cftr*. Apc^{Min} $Cftr^{fl/fl}$ Cre^{+} are also referred to as Apc^{Min} $Cftr^{-/-}$.

We found that Apc^{Min} *Cftr* mutant mice developed significantly more tumors than Apc^{Min} *Cftr* wildtype mice. This phenotype was observed in both *Cftr* heterozygous and *Cftr* homozygous mutants, with the tumor phenotype strongest in *Cftr* homozygous mutants. The increase in tumor multiplicity occurred in both the small intestine (***Figure 30***) and the colon (***Figure 31***). Apc^{Min} $Cftr^{-/-}$ mice became moribund at an earlier age (on average 30 days ealier) than Apc^{Min} $Cftr^{wt}$ mice due to their high tumor burden. Our data also indicated a strong bias against survival of male *Cftr* homozygous KO mice versus females. This was true for both $Apc^{+/+}$ and Apc^{Min} mice, regardless of Gavilyte C treatment. Overall, female KO mice demonstrated more than twofold survival advantage for reasons that are currently being investigated.

Apc^{Min} $Cftr^{-/-}$ mice in our study manifested a range of previously reported GI phenotypes including meconium ileus, intestinal blockade and shortened lifespan. The rectum of some Apc^{Min} $Cftr^{-/-}$ mice demonstrated adenomatous rectal hyperplasia. No substantial abnormalities were observed in other GI regions such as the stomach and pancreas of Apc^{Min} *Cftr* mutant mice.

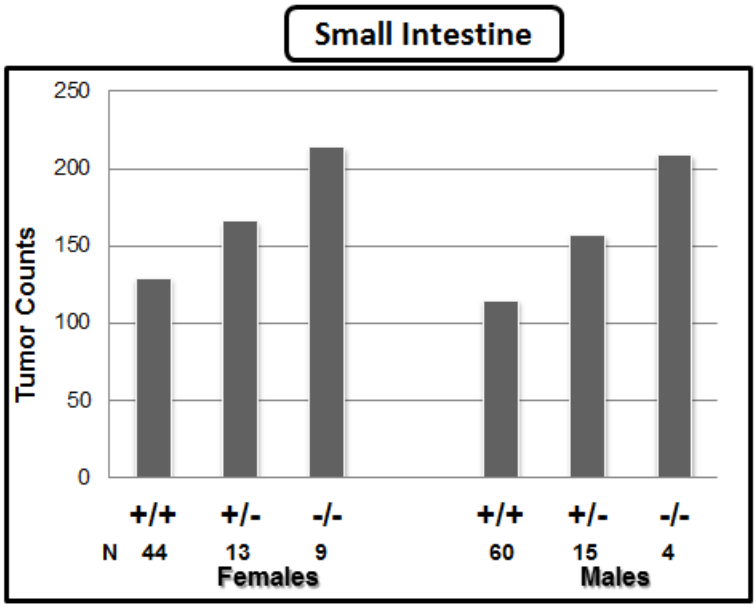


Figure 30. Loss of *Cftr* enhances tumor multiplicity in the entire small intestine of *Apc*^{Min} mice. *Apc*^{Min} *Cftr*$^{+/+}$, *Apc*^{Min} *Cftr*$^{+/-}$ and *Apc*^{Min} *Cftr*$^{-/-}$ mice were sacrificed at 120 days of age and tumors counted and measured.

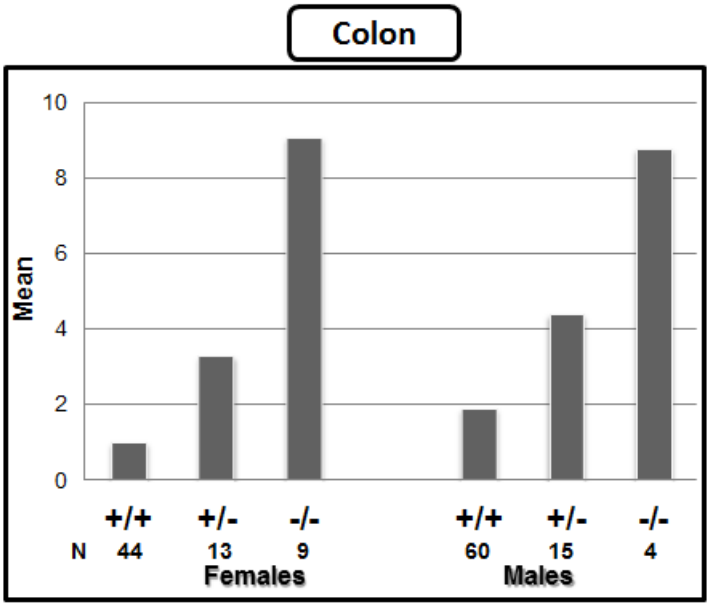


Figure 31. Loss of Cftr enhances tumor multiplicity in the colon of *Apc^{Min}* mice.
Apc^{Min} Cftr$^{+/+}$, *Apc^{Min} Cftr$^{+/-}$* and *Apc^{Min} Cftr$^{-/-}$* mice were sacrificed at 120 days of age and tumors counted and measured.

Loss of *Cftr* causes intestinal tumorigenesis in *Apc$^{+/+}$* mice.

Upon aging a cohort of *Apc^{wt} Cftr$^{-/-}$* mice to 12 months, we found that loss of *Cftr* alone could induce tumor formation in the intestine of the *Apc^{wt}* mouse (***Figure 32***). ~50% of *Apc$^{+/+}$ Cftr$^{-/-}$* mice developed up to 3 tumors, all in the small intestine. Thus far, all tumors observed have been adenomas. Sections of these tissues also show evidence of adenoma merging with Brunner's gland (***Figure 32c***). Considering the important role of Brunner's gland in acid neutralization and intestinal nutrient absorption via its bicarbonate secretion and the previously mentioned role of Cftr in gastrointestinal bicarbonate regulation, this suggests a potential link between Brunner's gland's abnormality and intestinal tumorigenesis.

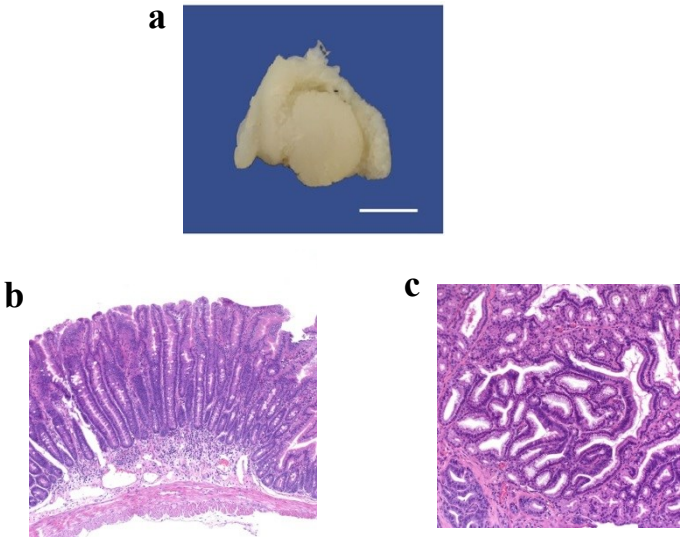


Figure 32. Histopathology of an *Apc$^{+/+}$ Cftr$^{-/-}$* tumor. (a) Duodenual adenoma from an *Apc$^{+/+}$ Cftr$^{-/-}$* mouse. Bar = 2.5 mm. Sections from this lesion were taken from various regions and all showed

evidence of dysplatic tumor. Two of these adenoma sections are depicted in (b–c). **(b)** Flat adenoma. 10X. **(c)** Adenoma merges imperceptibly with Brunner's glands (possibly originates from the BGs). 20X.

Loss of Cftr causes other GI tract abnormalities.

Loss of *Cftr* in Apc^{Min} mice causes rectal prolapse (***Figure 33a***) with signs of extensive rectal hyperplasia (***Figure 33b***). Interestingly, we also observed malignant transformation into rectal adenocarcinoma (***Figure 33c***).

Notably, while $Cftr^{-/-}$ mice have been reported to be pancreatic sufficient (Snouwaert *et al.* 1992, Bijvelds *et al*. 2005), in at least one case a severely enlarged pancreas (***Figure 34a***) was observed in a 12 month old $Apc^{+/+}$ $Cftr^{-/-}$ mouse. Histopathology indicated that the pancreas was engulfed by a large round cell tumor mass, consistent with lymphoma, arising from a pancreatic lymph node (***Figure 34b***).

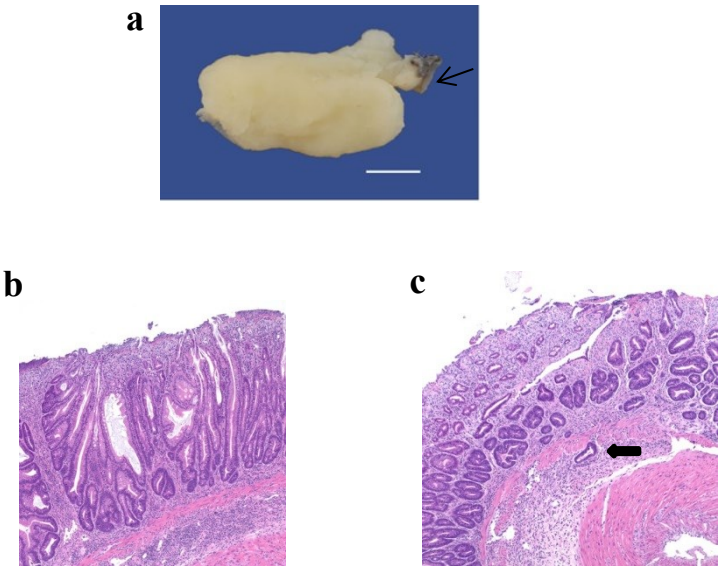


Figure 33. Histopathology of a rectum from an Apc^{Min} $Cftr^{-/-}$ mouse. **(a)** Rectum from an Apc^{Min} $Cftr^{-/-}$ mouse. Bar = 2.5 mm. Arrow points to an area of normal rectum. Enlarged area to the left indicates abnormal transformation; sections from this lesion were taken from various regions and they

showed evidence of hyperplasia and adenocarcinoma. **(b)** Areas of adenomatous hyperplasia. 10X. **(c)** Focal invasion with block arrow indicating malignant transformation (adenocarcinoma). 20X.

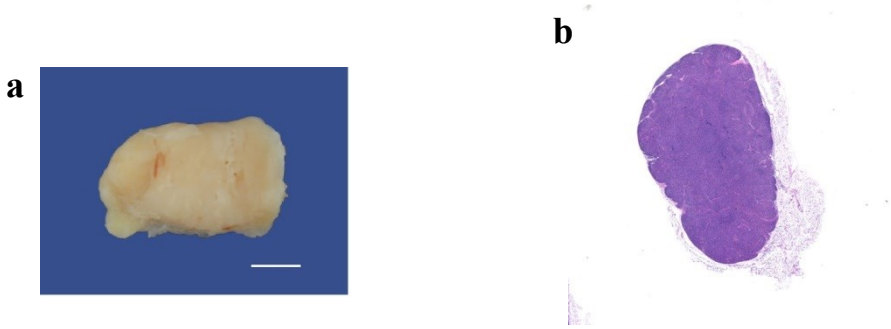


Figure 34. Histopathology of pancreas from an *Apc*$^{+/+}$ *Cftr*$^{-/-}$ mouse. Pancreatic leasion started from a lymph node that just engulf the whole pancreas. No evidence of systemic disease (such as similar lesion in the thymus, spleen or elsewhere). **(a)** Pancreas from an *Apc*$^{+/+}$ *Cftr*$^{-/-}$ mouse. Bar = 2.5 mm. **(b)** A pancreatic lymph node infiltrated by round cell tumor. 10X.

Loss of *Cftr* alters gene expression.

Upon performing microarray expression studies on the epithelial fraction of the colon and the proximal small intestine of *Apc*$^{+/+}$ *Cftr*$^{+/+}$ and *Apc*$^{+/+}$ *Cftr*$^{-/-}$ mice, we were able to identify more than 100 genes that showed a greater than 1.5 fold change in expression in the small intestine (***Table 10***) and 58 in the colon (***Table 11***). Expression of 20 of these genes was confirmed by RT-PCR (still in process). Among top gene clusters identified are immune cell homeostasis (*C3, Hbb-a1, Hbb-b1, Ly6a, Ly6g6c, Reg3g, Reg3b, Cfd, Igk-V34*), ion channels (*Clca6, Fxyd4, Slc4a4, Slc40a1, Slc5a8, Slca4b*), mucins (*Muc3, Muc13*), Inflammation (*Slpi, S100A11, Defa20, Defa3, Defa26, Defa6, Defa1, Defb37, Retnlb, B3galt5, Soc3, Rnf186, Sprr2a1, Dmbt1, Nfkbiz, Nt5e, Chst4, Acaa1b, Ifi27l2a, Gcnt3*), growth regulation and cell signaling (*Dlk1, Soc3, Areg, Fos,*

Dusp6, Gprc5a, Errfi1, Elf3, Ndrg1, Mxd1), migration and invasion (*Anxa4, S100a11, Phlda1, Tmsb10*), stress response and detoxification (*Duoxa2, Cyp2d9, Cd177, Gsta3, Cyp4b1, Gpx2, Duox2, Cyp2c55, Aldh1a1, Cyp2e1, Gstm1, Cyp3a25, Adh4, Ephx1, Acaa1b, Ugt2b36, Gsta2, Agmat, Cyp2c65, Cyp2b23, Bbox1, Cyp2b10, Gsta1, Dbp, Acot1, Maob, Gstm2, Adh1, Cyp4v3, Gstm4, Dio1*), water homeostasis (*Aqp1, Aqp4*), and stem cell regulation (*Phlda1, Aldh1a1, Aqp1*). Many of these genes overlap with genes disrupted by the loss of *Kcnq1,* as recently reported by our group (Than *et al.* 2013), including groups of mucin genes (*Muc3, Muc13*), stress response genes (*Cyp2c55, Gstm2*) and stem cell target genes (*Aldh1a1, Clca4*).

Table 10. Top upregulated and downregulated genes identified by microarray in the small intestine.

Upregulated Genes	Downregulated Genes
Reg3g	*Cfd*
Reg3a	*Scd*
Amy2a	*Defa1*
Retnlb	*Igk-V8*
Anxa4	*Cyp2C9*
B3galt5	*Lpl*
Muc3a	*Slc5a4b*
Cela3a	*Adipoq*
Defa3	*Reg3g*
Cldn4	*Aldh1a1*

Table 11. Top upregulated and downregulated genes identified by microarray in the colon.

Upregulated Genes	Downregulated Genes
Defa6	*Sycn*
Defa3	*Gsdmc*
Slpi	*Clps*
Reg3g	*Fxyd4*
Clca4	*Ces1d*
Defa-rs1	*Rn18s*
Lyz1/Lyz2	*Cyp4b1*
C3	*Actb*
Reg3a	*Prlr*
Clca6	*Pbld*

Loss of *Cftr* leads to alteration in potential signaling pathways.

To confirm and further identify potential signaling pathways disrupted by the loss of *Cftr*, we performed IPA analysis on the microarray datasets of the small intestine and the colon. Several major networks were identified. In the small intestine, these include genes involved in drug metabolism-small molecule biochemistry-lipid metabolism (***Figure 35a***) and molecular transport-GI disease-inflammatory disease (***Figure 35b***). In the colon, major networks cover genes involved in molecular transport-lipid metabolism-small molecule biochemistry (***Figure 36a***) and cellular growth/proliferation-renal and urological system development and function (***Figure 36b***). Among top biological functional groups, altered genes that play important roles in tissue morphology and immune trafficking are *Muc3, Fos, Areg, Socs3, Ly6a, Slpi*. These gene clusters point to alterations in biological processes that involve the interactions of CFTR-expressing cells with other cell types such as mucin-secreting goblet cells and immune cells.

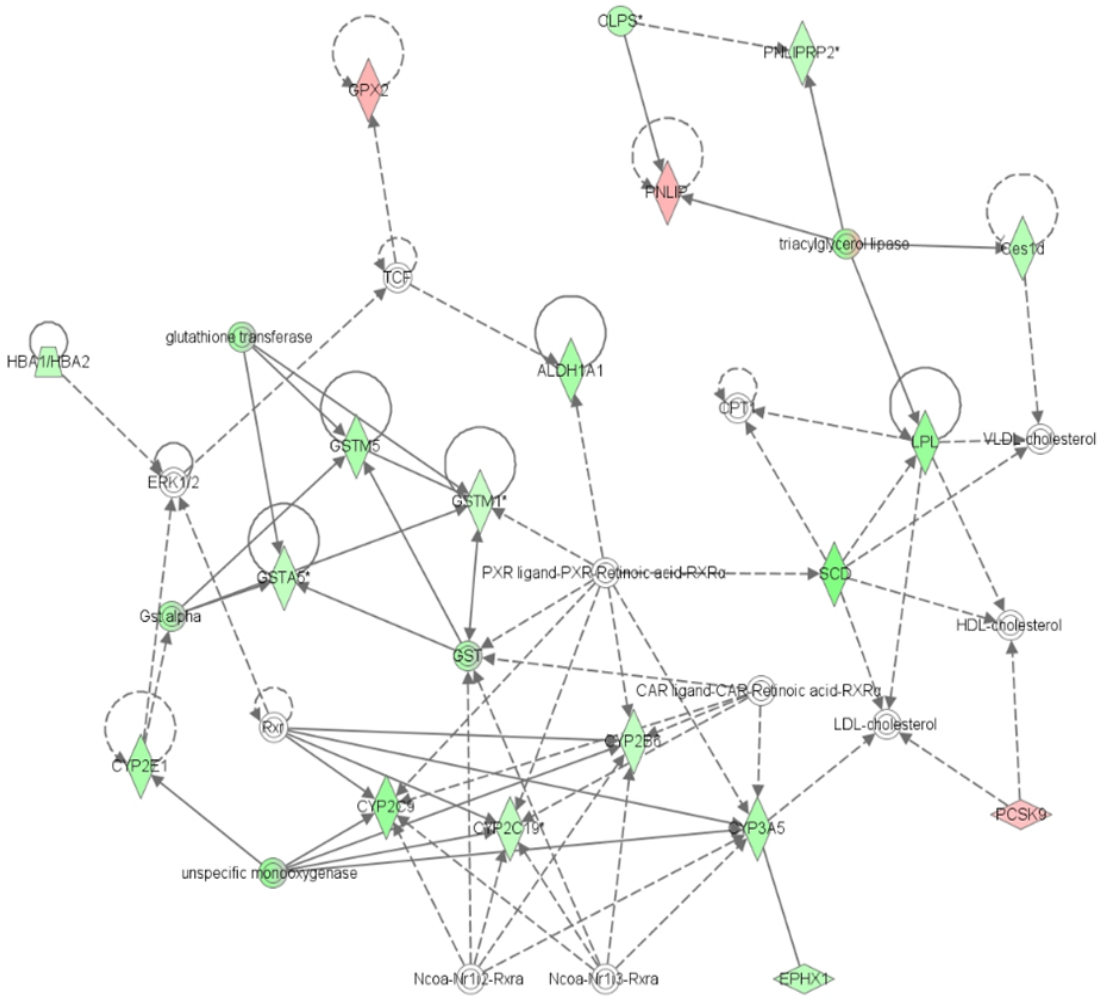


Figure 35a. Associated network functions in the small intestine identified by IPA: Drug Metabolism, Small Molecule Biochemistry and Lipid Metabolism.

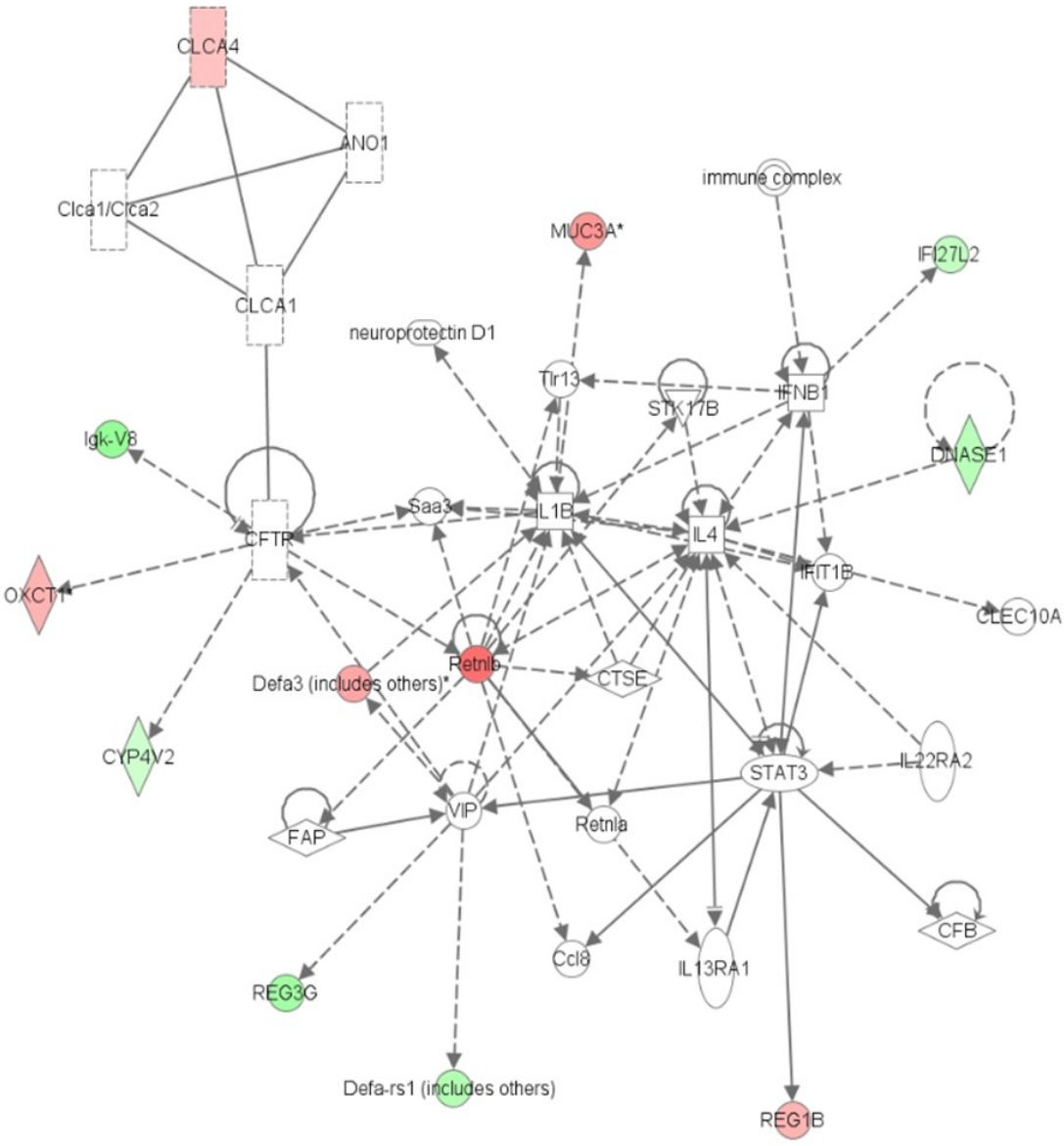


Figure 35b. Associated network functions in the small intestine identified by IPA: Molecular Transport, Gastrointestinal Disease and Inflammatory Disease.

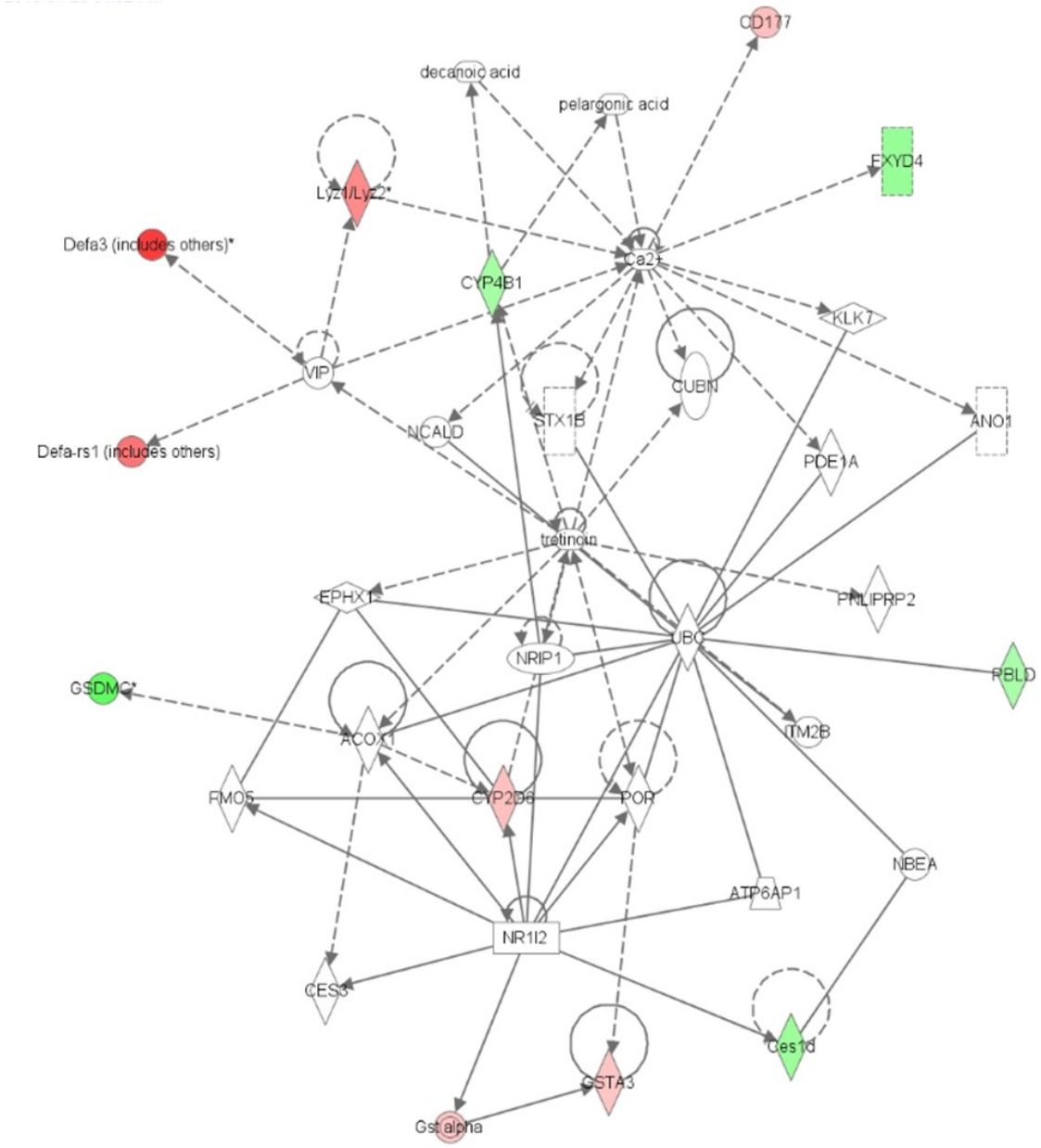


Figure 36a. Associated network functions in the colon identified by IPA: Molecular Transport, Lipid Metabolism and Small Molecule Biochemistry.

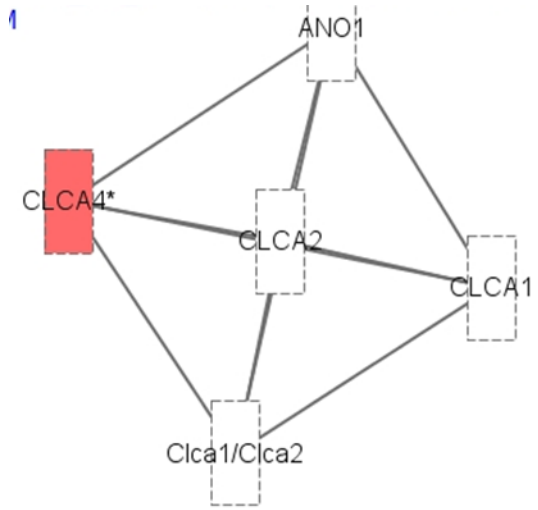


Figure 36b. Associated network functions in the colon identified by IPA: Cellular Growth and Proliferation, Renal and Urological System Development and Function.

Loss of *Cftr* promotes colon organoid development.

To investigate a role for *Cftr* in the intestinal stem cell compartment we created intestinal organoids from the colons of a group of $Apc^{+/+}$ $Cftr^{+/+}$ and $Cftr^{-/-}$ age and gender matched littermate mice. At days 2 through 5 after plating, organoid growth was monitored and organoid number counted on day 5. Loss of *Cftr* increased the number of colon organoids by greater than 2 fold, suggesting a potential regulatory effect of *Cftr* on the intestinal stem cell compartment (Rothenberg *et al.* 2012) (***Figure 37***).

(Statistical analysis for organoid experiments is in progress).

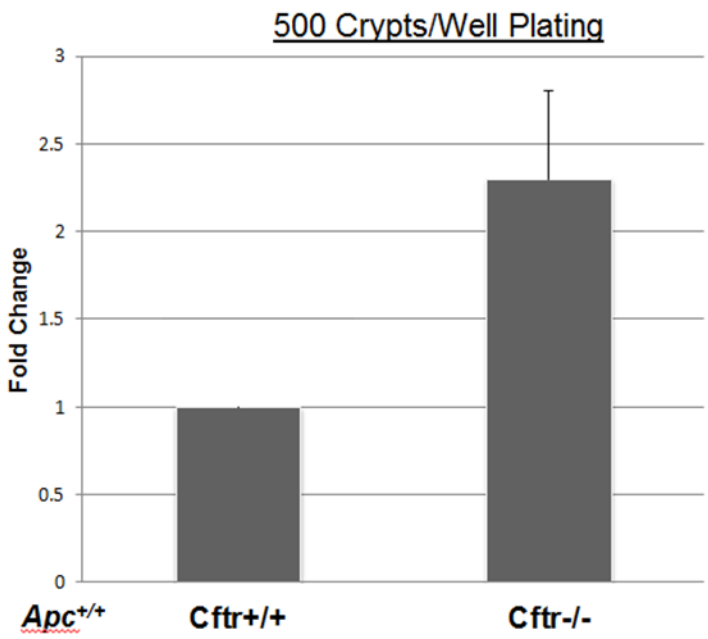


Figure 37. Colon Organoids. Colon organoids were created from 500 crypt bottoms and plated in triplicate in 24-well plates. Crypt bottoms were isolated on separate days from 5 of each, age, gender and littermate matched *Apc*$^{+/+}$ *Cftr*$^{-/-}$ and *Apc*$^{+/+}$ *Cftr*$^{+/+}$ mice (one matched pair per day). Organoids were examined daily and counted at day 5 post-plating.

DISCUSSION

Identified by *Sleeping Beauty* transposon-based forward genetic screens in mice as a candidate GI cancer gene, *Cftr* may play an important role in GI cancer that remains to be further explored. *CFTR* encodes for a transmembrane channel responsible for chloride secretion and fluid homeostasis in various epithelia. Mutation of *CFTR* causes disruption in Cl^- secretion and in turn results in CF. Though the median survival age has improved steadily over the past half century, CF is still a life-threatening disease that affects ~30,000 children and adults in the US (David and Ferkol 2013). In the lung and airways where major CF phenotypes are observed, dysregulated CFTR causes impaired mucus and bacterial clearance, and subsequently airway obstruction.

In the GI tract, CFTR plays an important role in various biological processes, thus its disruption can lead to GI failure. In the large intestine, particularly in the colon, chloride secretion and fluid secretion is crucial for maintain hydrated mucosa and electrolyte homeostasis. Normally, chloride transport out of the apical membrane of the mucosa creates an osmotic gradient that draws water movement into the lumen. CF patients however secret less fluid, thus potentially causing intestinal blockade by improperly hydrated stools. In addition to intestinal blockade, CF patients also experience intestinal bacterial overgrowth due to thick mucus secretion and abnormal acid-base balance. In the small intestine, specifically in the duodenum, CFTR is suggested to be an important pH regulator besides its direct role in chloride secretion. Duodenal pH is controlled by bicarbonate-rich pancreatic secretions and intestinal bicarbonate secretion that neutralize acid from the stomach. An unbalanced acid-base environment coupled with impaired mucus protection can lead to abnormal bacterial overgrowth and intestinal inflammation, symptoms that once become severe can cause whole system failure and predisposition to GI cancer.

Our study indicates *CFTR* is a tumor suppressor in the GI tract. This was observed by the increased tumor phenotype in the colon and small intestine of Apc^{Min} $Cftr^{-/-}$ mice and appearance of intestinal tumors in aged Apc^{wt} $Cftr^{-/-}$ mice. Consistent with the findings in our study, *CFTR* has been reported to be frequently hypermethylated in both cancer cell lines (e.g. promyelocytic leukemia) and primary lung tumor samples (Denamur and Chehab 1995, Son JW *et al.* 2011). *CFTR* was also found to be a tumor suppressor in prostate cancer, in part through its ability to alter miR-193 targeting the urokinase-type plasminogen activator (Xie *et al.* 2013). The mechanism by which *CFTR* exerts its tumor suppressive role in GI cancer remains to be elucidated. Our microarray studies point to several potential signaling pathways including immune mediation, inflammation, stress responses and stem cell regulation. On the basis of our studies and others, a proposed model is presented as below, whereby KCNQ1 and CFTR act to mediate chloride secretion, regulate water homeostasis and contributes to the protective function of the intestinal mucus layer.

Proposed model: KCNQ1/KCNE3 channels which are expressed in intestinal crypt mediate basolateral K^+ recycling required for Cl^- secretion by CFTR, an important regulator of water homeostasis.

Chloride secretion is a complex mechanism that involves several ion transporters, including K^+ (KCNQ1) and Cl^- (CFTR) channels which have both been suggested to express in the colonocytes (Rapetti-Mauss *et al.* 2013, Linley *et al.* 2013). The process is tightly regulated by signaling peptides in order to prevent dehydration of colonic mucosa or diarrhea. First, chloride is transported into the cell basolaterally via a Na^+-K^+-$2Cl^-$ (NKCCl) transporter. K^+ recycling, via KCNQ1, across the basolateral membrane is the electrochemical driving force for Cl^- to be secreted across the apical membrane through CFTR. In addition, these channels are responsible for preventing depolarization of membrane potential, which is required for the maintenance of chloride secretion. Chloride secretion in turn presents CFTR and KCNQ1 as important regulators of water homeostasis. Aquaporins (Aqp) are water channels that play very important roles in normal intestinal function. It was reported that Aqp5 co-localized with Cftr on the apical domains of acinar cells in rat Brunner's glands and are suggested to play an important role in water transport by the gland (Matsuzaki T *et al.* 2004, Collaco e*t al.* 2013).

When this process is disrupted due to mutation in CFTR, intestinal mucus protection is compromised, subsequently leading to alterations in downstream pathways that contribute to GI oncogenesis. Evidence on the link between CFTR disruption and GI cancer has been reported but there is yet a national consensus due to the small number of digestive tract tumors developed during the 10-year study period in the North American CF population (Maisonneuve *et al.* 2013). Regardless, our functional study of Cftr suggests a potential connection between mutation in CFTR and GI tumorigenesis. Because CFTR function is highly relevant to cellular development and human physiological processes, understanding CFTR's underlying mechanism can lead to its use in the clinic as a prognostic predictor and potential therapeutic target.

Materials and Methods

Mice. C57BL/6J mice and C57BL/6J-*Apc*^{Min} mice were obtained from the Jackson Laboratory (Bar Harbor, ME). Conditional C57BL/6-*Cftr* knockout mice were obtained from Dr. Mitchell Drumm (Case Western Reserve University). Details of mouse husbandry are as previously described (Hodges *et al.* 2008, Fijneman *et al.* 2012, Than *et al.* 2013). The genotype of the *Apc and Cftr* loci were determined by PCR assays as previously described (Hodges *et al.* 2008, Fijneman *et al.* 2012).

***Cftr* mutant mouse model.** We obtained a conditional knockout mouse model of *Cftr* from Prof. Mitchell Drumm at Case Western Reserve University. These mice carry a conditional null allele of *Cftr*, in which loxP sites were inserted around exon 10 of *Cftr* (*Cftr*^{fl10}) (Hodges *et al.* 2008). Disruption of exon 10 in mice is known to create severe CF phenotypes similar to those with a *F508Δ* mutation in human patients. While the lung pathology is dominantly observed in human CF patients, lung phenotypes are less common in CF mouse models probably due to species differences. However, *Cftr*-mutant mice display multiple other phenotypes similar to those in CF humans such as poor growth, delayed puberty and intestinal obstruction. On the C57BL/6 background, intestinal obstruction phenotype is very severe as over 95% of *Cftr* homozygous knockouts do not survive pass weaning (De Lisle and Borowitz 2013). To bypass this early lethality, for more than a decade investigators have added Colyte, a commercially available osmotic laxative to the drinking water of *Cftr* KO mice (Clarke *et al.* 1996). In our study we used an identical solution called Gavilyte C per the manufacturer's instructions (Gavis Pharmaceuticals).

Tumor analysis. Scoring of GI tumor tissues was performed as previously described (Fijneman *et al.* 2012). Two-sided P values for tumor counts were determined by use of the Wilcoxon Rank Sum Test comparing gender and age-matched classes produced in the same genetic crosses.

Histopathology. Histopathological analysis of tumors and adjoining normal tissue was performed on FFPE tissues by an A.C.V.P.-certified veterinary pathologist (Dr. M. Gerald O'Sullivan) from the University of Minnesota Masonic Cancer Comparative Pathology Shared Resources facility using protocols as previously described (Fijneman *et al.* 2012).

Organoid Culture: Gender, age, and littermate-matched C57Bl/6J *Cftr*$^{fl/fl}$ and *Cftr*$^{fl/fl}$ *Cre*$^{+}$ mice were sacrificed between 8 and 12 weeks of age. Colons were removed, cut open and washed in cold PBS. Colon organoids were then cultured using the protocol of Sato *et al.* 2011, following the plating of 500 crypt bottoms per well in triplicate per sample.

Epithelial RNA processing for Illumina Bead Arrays. Mouse intestinal and colon tissues were removed, opened longitudinally, and rinsed in PBS. For the small intestine, villi were removed from the distal quarter of the proximal quarter, tissue was then flash frozen in liquid nitrogen. Tissue was transferred from liquid nitrogen to pre-cooled RNALaterIce (Ambion) at -80^{o}C for 30min, and then was transferred to -20^{o}C for at least 48hrs. Immediately before RNA isolation, tissue in either RNALater or RNALaterIce was placed in 2-4 mL of RLT/14.3M β-mercaptoethanol (BME) buffer. Tissue was then homogenized using an IKA Ultra-Turrax T25 digital homogenizer (Fisher Scientific). RNA isolation was then conducted using an RNeasy Mini Kit with an additional DNAse Digestion step (Qiagen). For the colon, epithelial RNA processing and isolation were performed as previously described (Fijneman *et al.* 2012). RNA sample concentrations were measured using a Nanodrop-1000 Spectrophotometer (Thermo Scientific). RNA was stored at -80^{o}C.

Illumina Bead Microarray and Data Analysis. RNA labeling, microarray hybridization and scanning were performed at the University of Minnesota – BioMedical Genomics Center using Illumina MouseWG-6 v2.0 Expression BeadChips (Illumina, San Diego, CA, USA), according to the manufacturer's instructions. Differential expression of genes

was quantified by the moderated t-statistic. Data were analyzed using GeneData Expressionist Software (GeneData Inc, San Francisco, CA); genes that showed statistically significant differences in expression between groups were determined using the two-group *t*-test and ANOVA. Gene expression data has been submitted to the Gene Expression Omnibus (GEO), Accession number pending.

qRT-PCR. Quantitative RT-PCR was performed as previously described (Fijneman *et al.* 2012). Significant differences in expression between groups were determined using the two-group *t*-test.

List of primers used: All primers were designed using Primer-BLAST (National Center for Biotechnology Information) and obtained from Integrated DNA Technologies.

Mouse Primers

RT-PCR is in process

Pathway Analysis

Ingenuity Pathway Analysis (IPA). Genes showing a 1.5 fold change in expression in whole colon (N = 58) and proximal small intestine (N = 119) microarrays were analyzed using IPA.

Conflict of interest. The authors state that there are no conflicts of interest to declare.

Acknowledgements. We wish to thank Dr. Drumm and Dr. Hodges at the Case Western Reserve University for generously providing the conditional knockout *Cftr* mice and reviewing the manuscript. Research was supported by a grant to RC and DL (NCI R01 CA134759-01A1).

Personal Contribution. B.L.N.T contributed to the design of the research. B.L.N.T contributed to all mouse husbandry and tumor analysis. B.L.N.T isolated/processed RNAs for Illumina microarrays and analyzed data. B.L.N.T designed primers, performed

qRT-PCR for gene validation, analyzed data and statistical tests. B.L.N.T carried out IPA on microarray datasets, conducted organoid experiments and performed statistical analysis.

CHAPTER FOUR

Discussion

Challenges in GI cancer

Despite decades of research effort, epithelial cancers remain as a major source of mortality in the US population. For example, CRC remains the second leading cause of cancer-related deaths in the US. Thus, there is an urgent need for the identification and development of new molecular targets for CRC therapeutic interventions. In this dissertation, two candidate ion-channel genes, *Kcnq1* and *Cftr*, identified in *SB* mutagenesis screens in mice, were identified as potent tumor suppressors in GI cancer and thus are potential therapeutic targets for CRC.

The GI tract plays an indispensable role in digesting and absorbing nutrients while eliminating indigestible and toxic materials. GI cancer is heterogeneous in that it displays unique histopathology, different levels of tumor aggressiveness and diverse clinical outcomes. This is due to the effects of both specific genetic alterations and environmental damage to the intestinal tract.

At the genetic level, GI cancer and specifically CRC arise through the accumulation of multiple mutations in genes essential for normal cellular development and a broad range of normal cellular processes including cell proliferation, metabolism and apoptosis. In the most common model of the intestinal adenoma to carcinoma sequence, normal tissue is transformed into dysplastic tissue after the loss in function of *APC* or activation of WNT signaling. The evolution from a dysplastic tissue to an invasive carcinoma occurs with additional activating mutations in driver oncogenes (ie. *RAS*) and inactivating mutations in driver tumor suppressor genes (ie. *P53*). Importantly, each CRC also acquires and requires genetic and epigenetic alterations in a larger set of low frequency genes that collectively represent major cellular processes whose dysregulation is necessary for cancer development. Notably, these low frequency

alterations are also likely to represent a significant source of tumor heterogeneity but also can be diagnostic and prognostic biomarkers and potential therapeutic targets.

Environmental Influences on GI cancer

In addition to genetic factors, environmental elements including diet, chemical agents as well as xenobiotics may target important driver genes (such as ion-channel genes), and, in turn, influence the susceptibility to various cancers including GI cancer. For example, in respect to diet, it is reported that the incidence of CRC is higher in developed countries than developing countries. The geographic differences in CRC incidence appear to be attributed to the differences in exposures to a Western high fat diet. Chemical agents, for example estrogen (E2) and E2 mimics represent environmental factors that can influence CRC development. It was reported that E2 confers a protective effect in GI cancer (Giroux, Betnatchez and Carrier 2011), with estrogen receptor (ER)-β specifically as the principal mediator of E2 effect in the colon. Interestingly, recent studies have reported a potential link between E2 and the two ion-channels KCNQ1 and CFTR. E2 was found to regulate the internalization of colonic KCNQ1 by endocytosis (Rapetti-Mauss *et al.* 2013) while CFTR was shown to regulate ovarian E2 biosynthesis via amplification of *Follicle-stimulating-hormone* stimulated signal (Chen *et al.* 2012). Since certain chemical compounds (such as soy supplementation, synthetic estrogens found in drinking water from sources such as manufacturing (e.g. plastic additive Bisphenol-A), synthetic fertilizers, and hormones given to livestock) can mimic the actions of E2, this may potentially affect targeted signaling including that of KCNQ1 and CFTR upon the alteration in the activity of ER. Xenobiotics are well-known as potential oxidative stress inductors that lead to the production of reactive oxygen species and in turn cause alterations in important biological processes such as cell apoptosis and proliferation. Therefore, identification and characterization of ion-channel genes in this dissertation can help eludicate how the effects of certain potential exogenous toxicants on

ion channel signaling may contribute to human cancers including GI and colorectal cancers.

KCNQ1 and CFTR

The GI tract contains a variety of ion channels including two, KCNQ1 and CFTR, that regulate the secretion of K^+ and Cl^-, respectively. CFTR plays a major role in the outflow of chloride across the apical membrane of the intestinal crypt cells, and KCNQ1 is suggested to regulate the outflow of potassium across the basolateral membrane of the same cell type. As discussed in Chapter 1, inherited mutations in the *KCNQ1* gene encoding the basolateral membrane potassium channel are associated with cardiovascular disease syndromes such as LQTS and JLNS. The second gene, *CFTR,* encodes an apical membrane chloride channel, mutations of which cause cystic fibrosis (CF). (The largest number of patients with CF exhibit the *F508Δ* mutation in *CFTR*, which was closely modeled by $Cftr^{fl10}$ mice in our studies described in Chapter 3.) Both *Kcnq1* and *Cftr* were identified as candidate intestinal cancer genes in our *SB* forward genetic screens in mice. Collectively, our data described in Chapter 2 and 3 indicate that *KCNQ1* and *CFTR* act as tumor suppressors in GI cancer. We demonstrated the tumor suppressive function of *Kcnq1* and *Cftr* in Apc^{Min} mice, where both hetero- and homozygous inactivation of these ion-channel genes resulted in significantly more intestinal adenomas. In the case of Apc^{Min} *Kcnq1* mutant mice, we also reported progression to adenocarcinoma. For *Cftr*, we found that *Cftr* deficiency alone was sufficient for the development of tumors in the small intestine of mice aged to ~ 1 year. We also observed other GI phenotypes including rectal adenomatous hyperplasia, pancreatic abnormalities, gastric hyperplasia (caused by *Kcnq1* deficiency) and pancreatic abnormalities (caused by *Cftr* deficiency). Strikingly, our human study showed that in CRC patients with liver metastasis, low KCNQ1 expression correlated with an almost 2 year reduction in their overall survival, compared to those with high KCNQ1 expression. Notably, prior to the work described in Chapter 2 and 3,

very little was known about the roles of KCNQ1 (especially) and CFTR in GI cancer. Thus our findings are novel and of potential diagnostic, prognostic and therapeutic valule.

Limitations in Our Studies

First, due to the germline mutation, our *Kcnq1* mouse model exhibited a variety of phenotypes in other organs that potentially led to a confounding effect on the GI tumor phenotype. Similar to human patients with *KCNQ1* mutations, our *Kcnq1* homozygous knockout mice demonstrated hyperkinetic behavior due to an imbalance in the inner ear. This resulted in constant circling, causing the mice to be leaner and smaller than their littermates. A consequence of this hyperkinetic activity was no increase (or a slight reduction) in tumors in the distal small intestine of Apc^{Min} $Kcnq1^{-/-}$ mice. This phenocopies Apc^{Min} mice who exercise on a running wheel which causes a significant reduction (~30%) in tumors in the distal small intestine and colon (Ju *et al.* 2008, Colbert *et al.* 2006). Therefore, we had countervailing influences on tumorigenesis in the distal small intestine and colon: loss of *Kcnq1* promoted tumor development while their hyperkinetic behavior and physical activity acted in the opposite direction. (Interestingly, as we emphasized in Chapter 2, the proximal small intestine did not appear to be affected by the hyperkinetic activity as Apc^{Min} $Kcnq1^{-/-}$ mice developed up to a 5-fold increase in tumors along with the development of adenocarcinomas.) Therefore, we would predict the use of an intestinal specific KO allele of *Kcnq1* in future studies should produce a stronger tumor multiplicity phenotype in the GI tract. *Second*, defining the underlying mechanisms of KCNQ1 and CFTR action in GI cancer has posed some challenges, especially with using the typical approach of performing functional assays in human cell culture. In fact, we did explore disruption of *KCNQ1* and *CFTR* in human cell lines using siRNA knockdown transfection. Unfortunately, the drawback to modeling the effects of ion channel genes such as *KCNQ1* and *CFTR* in cell culture is that many of their targets are not found or poorly detected. For example, identified in our microarrays is a group of immune, Paneth and goblet cell genes. We found that many commonly employed human

CRC cells, such as SW480, co-express very low to undetectable levels of those target genes. Fortunately, T84, a lung CRC metastasis cell line expresses decent levels of *KCNQ1* and *CFTR*, and has been used to study electrolyte transport, especially CFTR-mediated transport. However, upon performing transient siRNA experiments, we found that knockdown of ~50-60% of both *KCNQ1* and *CFTR* still did not cause any significant changes in gene expression of any of the major genes identified in our *Kcnq1* and *Cftr* mouse microarrays. Given species and cell culture differences, these observations may be understandable. This also suggests that dosage is important for *KCNQ1* and *CFTR*'s effects at the molecular and cellular level since residual expression of the genes after siRNA-mediated knockdown could rescue its normal regulation of target genes in T84 cells. Future studies with stable knockdown employing shRNAs and lentiviral vectors can potentially eliminate this limitation. Alternatively, we have used a well-known chemical inhibitor of CFTR, CFTR(inh)-172*,* to inhibit the function of the gene in T84 cells and have started to detect some changes in genes identified in the *Cftr* mutant mice. Experimental phenotypic assays, such as cell proliferation and cell anoikis, to investigate this effect of CFTR disruption in human cell culture are currently in progress. The same line of experiments in cell culture will also be conducted for KCNQ1.

Implications for Human Diseases

It is important to discuss what could be applied from our preclinical and clinical studies of CFTR and KCNQ1 to the potential implications for human diseases. *First*, our human KCNQ1 data clearly indicated the significant role of KCNQ1 expression in maintaining almost two years of overall survival for CRC patients with liver metastasis. Thus, expression of KCNQ1 could become a factor used to decide eligibility for hepatic resection in CRC patients. Currently, only ~ 25% of CRC patients with liver metastases are judged to be eligible for hepatic resection, with the main criteria for eligibility being the number and size of metastatic lesions. Yet, even with these stringent criteria, many patients do not benefit from surgery and die rapidly, a finding confirmed in our data in

Chapter 2. Therefore, expression levels of KCNQ1 could be added to the oncology decision making algorithm for hepatic resection. *Second*, it would be of great interest to investigate whether this same protective mechanism of KCNQ1 expression occurs during the early stages of tumor development. Therefore, targeting KCNQ1 would lead to a wide range of therapeutic treatments that might delay CRC progression, possibly small molecules activating drugs (Xiong *et al*. 2008, Wulff *et al.* 2009) that could be cancer-specific when still at its primitive stage, or, for advanced CRC, prolong survival. *Third*, with our conditional knockout of *Cftr*, we found that heterozygous knockout of *Cftr* could induce tumor multiplicity in the intestinal tract of Apc^{Min} mice. This suggests that human patients with a heterogygous deficiency of *CFTR* could a have higher risk of developing intestinal tumors in their lifetime. (Indeed, CF is the most common autosomal recessive disease in Causasians with a carrier frequency of 4%. Thus, as many as 10 million people in the US are heterozygous for CF mutations.) Therefore, it would be of interest to examine CRC incidence in CF family registries to determine if CF carriers are at a higher risk of developing CRC, a finding that could lead to enhanced surveillance. *Fourth*, KCNQ1 and CFTR function may potentially be epigenetically regulated in CRC. Epigenetic silencing of tumor suppressor genes is well established in CRC, for example, the silencing of MLH2 in MSI CRCs. According to the current TCGA data (Appendix, Section 3), mutation of both *KCNQ1* and *CFTR* occurs at a rate of ~5% in all human intestinal cancers. This suggests that there may be another mechanism of regulation, perhaps via epigenetic silencing. Interestingly, as mentioned in Chapter 1, *KCNQ1* is imprinted during the developmental stages (paternally imprinted *KCNQ1* and maternally imprinted *KCNQ1ot1* lncRNA) and become bi-allelically activated in adults. Thus, a potential link between the loss of *KCNQ1* and adult CRC might be through its sudden gain of imprinting, which causes silencing of *KCNQ1* at the epigenetic level. Similar speculation can be made for *CFTR* as there has been some evidence that *CFTR* may be subjected to methylation-dependent epigenetic regulation. It has been found that the promoter of *CFTR* is GC rich and subjected to methylation in various cell lines (Denamur and Chehab 1995). A study using human non-small lung cancer tissues also reported the link of *CFTR* promoter hypermethylation and associated clinical features (Son JW *et al.*

2011). Thus, there is possibly a nonmutational mechanism for inactivation of *CFTR*, leading to CF and potentially GI cancer. *Fifth*, our functional studies of *KCNQ1* and *CFTR* have demonstrated the power of *Sleeping Beauty* (SB) mutagenesis as a system for identifying candidate cancer genes in human GI tract. Sequencing of human CRC has revealed a relatively small number of genes mutant at a high frequency, e.g. the *APC* mutations in chromosomal instable CRC and the mismatch repair *MLH1/MSH2* mutations in microsatellite instable CRC but these analyses have also demonstrated the existence of a much larger constellation of genes dysregulated at a much lower frequency. Our SB system of somatic intestinal mutagenesis thus offers innovative and distinct advantages over other CRC mouse models in that it examined sets of low frequent SB-induced mutant candidate genes in hundreds of individual tumors and on different host genetic backgrounds (i.e. Apc^{wt}, Apc^{Min}, $p53^{R27OH}$). This has helped us develop a much fuller list of potential candidate driver genes and will assist us in eludicating more candidate gene cooperative networks that underlie CRC. SB mutagenesis screens therefore offer a powerful tool to identify important cancer genes in various types of cancer.

Potential Mechanisms of Action of KCNQ1 and CFTR

Loss of *Kcnq1* and *Cftr* in our studies led to changes in expression of genes important for various cellular processes and signaling pathways essential in both normal and cancer intestinal epithelial cells, as detected in our mouse microarrays. Analysis of microarray datasets from the colon and small intestine of *Kcnq1*, *Cftr* and *Muc2* KO mice detected a significant overlap of altered genes. For *Kcnq1* and *Cftr*, the link was further supported by IPA, GSEA analysis and RNA Seq data from 20 human CRC metastatic samples. Based on our collective data and our comparison of genes regulated in KO versus WT mice in the current dissertation, we propose the following mechanisms to explain how loss of *KCNQ1* and *CFTR* may contribute to the development of CRC.

It is possible that the tumor suppressive role of KCNQ1 and CFTR in GI cancer lies in their regulation of important biological processes and pathways including inflammation, stress responses and stem cell regulation (including possibly epithelial-mesenchymal interactions).

Inflammation: Bacterial species such as the *Bacteroides*, *Prevotella* and *Ruminococcus* play an important role in the GI tract, generally acting as commensal organisms that are involved in immunity, digestion and protection against diseases including cancer. In the intestinal tract, especially the colonic crypt, a layer of mucus lines the entire epithelial surface, that helps to maintain a symbiotic relationship with commensal flora. However, escape of bacteria across the intestinal mucosal barrier can elicit an inflammatory response. In a normal host, epithelial cells in the epithelium and immune cells in the stroma work together to downregulate this process and block the entry of flora into deeper layers of the intestine. In a susceptible host, genetics factors, a leaky epithelial barrier and environmental triggers can converge to cause inflammation. One of the major characteristics of chronic inflammation is an increase in fibrosis in the submucosal and muscularis layers of the intestinal wall, disturbing its normal function. Patients with chronic inflammation such as intestinal bowel disease (IBD) are predisposed to GI and colorectal cancers.

A hypothetical model of how this process develops starts with mutations in ion channel genes, *KCNQ1* and *CFTR*, which leads to dysregulated potassium and chloride secretion. This subsequently can lead to water transport imbalance and impaired mucus protection, which results in luminal factors, such as bacteria, toxic substances and apoptotic cells and factors, gaining access to the mucosa. Once advanced, chronic inflammation can predispose patients to CRC. Therefore with an understanding of the roles of KCNQ1 and CFTR as the potential genetic basis for this type of disease, we will be able to identify, target, and successfully treat specific patient populations with abnormalities pertaining to inflammation.

Stress response: A second potential mechanism of KCNQ1 and CFTR in CRC oncogenesis may involve their participation in the stress response. Since oxidative stress can be produced by intestinal microbiota both directly and indirectly, it is reportedly associated with both genotoxicity and carcinogenesis. In both of our KCNQ1 and CFTR studies, we found a large number of overlapping genes associated with detoxification and stress response, especially in the *Muc2* knockout mice, such as *Gstk1, Aldh1a1* and *Cyp2c55*. Surprisingly, cellular stress also affects intestinal stem cell maintenance and differentiation. A recent study showed that endoplasmic reticulum (ER) stress and activity of the unfolded protein response (UPR) were induced in the transit amplifying compartment, which causes the loss of stemness and thus induction of differentiation (Heijmans *et al.* 2013). Inhibition of ER stress, however, resulted in stem cell accumulation. Our preliminary data on the delayed differentiation in a *Kcnq1 KO* small intestinal organoid experiment support a link between the stress response and stem cell differentiation (Appendix, Section 2). Since oxidative stress factors can be released during an inflammatory response, inflammation and stress response are also tightly co-regulated. An example is the role of the UPR transcription factor *Xbp1* in the maintenance of secretory cell lineages and its association with the risk of developing IBD (Kaser *et al.* 2008). Thus, impaired immune-mediated oxidative stress, potentially resulting from the loss of *KCNQ1* and *CFTR*, may alter systemic genotoxicity and ultimately lead to the development of CRC.

Stem cell regulation/EMT: Another mechanistic model of KCNQ1 and CFTR in CRC oncogenesis may be their regulation of the intestinal stem cell compartment and their influence on the epithelial-mesenchymal environment. This idea is partially based on the previous findings that both *Kcnq1* and *Cftr* are major genes located in the intestinal stem cell zone (Munoz *et al.* 2012, Rothenberg *et al.* 2012). As discussed in Chapter 1, the intestinal stem cell compartment in the epithelium interacts with and is influenced by the mesenchymal (stromal) compartment. Intestinal epithelial-stromal interactions present an important oncologic-molecular paradigm in which other biological processes and pathways such as inflammation can be involved. The dysregulated

interplay between epithelial and mesenchymal compartments can potentially promote tumor progression and may result in the formation of intestinal adenocarcinomas (as presented in Chapter 2 on *Kcnq1* mouse tumor phenotypes). We found that deficiency for both *Kcnq1* and *Cftr* significantly influenced the survival, growth and differentiation of intestinal organoids, results consistent with a role for these genes in regulation of the stem cell compartment. For example, we found that deficiency for *Kcnq1* caused a strong increase in the number of colon organoids while significantly delaying the differentiation of organoids in the small intestine.

In summary, the mechanism by which KCNQ1 and CFTR contributes to CRC oncogenesis may be via individual pathways as proposed above, or potentially the interplay of all three, or others.

In conclusion, the studies in this dissertation provide the rationale to pursue further investigations into the exact mechanisms of KCNQ1 and CFTR in CRC tumorigenesis. Although medical technologies and treatments for CRC have greatly advanced in the last decade, the most efficient approach remains early detection. When cancer becomes advanced, few effective treatment options are available. Even in the case of available therapies, while some patients respond well to a certain therapy, others do not. As an example, currently CRC patients with activating mutations in exon 2 of *KRAS* do not seem to benefit from receiving cetuximab, a monoclonal antibody targeting EGFR (Meric-Bernstam and Mills 2012). Thus, more precise, individualized treatment strategies are needed. The potential use of biomarkers can be seen in predicting resistance to treatment or lack of treatment benefits. Our work on KCNQ1 clearly exemplifies the efforts of identifying potential biomarkers and therapeutic targets urgently needed as 80% of CRC patients die from liver metastases (a common site of recurrence in CRC). Our human CRC study found that maintenance of KCNQ1 expression was associated with a significant 23 month increase in overall survival in CRC patients who had undergone hepatic resection for liver metastasis. Therefore, KCNQ1 expression could act as a biomarker to guide clinical decisions and targeting KCNQ1 could offer a potential therapeutic strategy for late stage CRC patients.

Based on our KCNQ1 and CFTR studies, we believe that the outcomes of further preclinical investigations could potentially guide the personalized therapeutic trial designs that target specific, subtype-selective sensitivity in CRC patients who in turn would receive optimal clinical benefit. In addition, the characterization of the two CIS candidate genes on the Apc^{Min} mouse background can offer unique insights into the development of *Apc*-deficient colorectal tumors in the human population. Our functional studies of KCNQ1 and CFTR, therefore, can lead to direct analysis in both human CRC and other human cancers. In line with the development of new genetic manipulations such as TALEN and CRISPR, we may soon be able to develop new effective correctors and activators for KCNQ1 and CFTR to achieve greater targeted therapies.

BIBLIOGRAPHY

Abbott GW, Sesti F, Splawski I, Buck ME, Lehmann MH *et al*. MiRP1 forms IKr potassium channels with HERG and is associated with cardiac arrhythmia. *Cell*. 1999 Apr 16;97(2):175-87.

Adam Humphries and Nicholas A. Wright. Colonic crypt organization and tumorigenesis. Nat Rev Cancer 2008; Volume 8: 415-424

American Cancer Society . "Colorectal cancer key statistics". http://www.cancer.org/cancer/colonandrectumcancer/detailedguide/colorectal-cancer-key-statistics

Ammar I, Izsvák Z, Ivics Z. The Sleeping Beauty transposon toolbox. *Methods Mol Biol*. 2012;859:229-40.

Andersen DH. Cystic fibrosis of the pancreas and its relation to celiac disease: a clinical and pathological study. *Am J Dis Child* 1938; 56:344-399.

Bajwa PJ, Lee JW, Straus DS, Lytle C. Activation of PPARgamma by rosiglitazone attenuates intestinal Cl- secretion. *Am J Physiol Gastrointest Liver Physiol*. 2009 Jul;297(1):G82-9.

Bardelli A, Parsons DW, Silliman N, Ptak J, Szabo S *et al*. Mutational analysis of the tyrosine kinome in colorectal cancers. *Science*. 2003 May 9;300(5621):949.

Bardou M, Barkun AN, Martel M. Obesity and colorectal cancer. *Gut*. 2013 Jun; 62(6):993-47.

Bartlett JR, Friedman KJ, Ling SC, Pace RG, Bell SC *et al*. Genetic modifiers of liver disease in cystic fibrosis. *JAMA*. 2009 Sep 9;302(10):1076-83.

Belt EJT, Fijneman RJA, van den Berg EG, *et al.* Loss of lamin A/C expression in stage II and III colon cancer is associated with disease recurrence. *Eur J Cancer* 2011; 47:1837-1845.

Bijvelds MJ, Bronsveld I, Havinga R, Sinaasappel M, de Jonge HR, Verkade HJ. Fat absorption in cystic fibrosis mice is impeded by defective lipolysis and post-lipolytic events. *Am J Physiol Gastrointest Liver Physiol* 2005;288:G646–53.

Boivin GP, *et al*. (2003) Pathology of mouse models of intestinal cancer: Consensus report and recommendations. *Gastroenterology* 124:762–777.

Burger-van Paassen N, Loonen LM, Witte-Bouma J, Korteland-van Male AM, de Bruijn AC, van der Sluis M *et al*. Mucin Muc2 deficiency and weaning influences the expression of the innate defense genes Reg3β, Reg3γ and angiogenin-4. *PLoS One* 2012; 7:e38798.

Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature*. 2012 Jul 18;487(7407):330-7.

Casimiro MC, Knollmann BC, Ebert SN, Vary JC Jr, Greene AE, Franz MR *et al*. Targeted disruption of the Kcnq1 gene produces a mouse model of Jervell and Lange-Nielsen Syndrome. *Proc Natl Acad Sci U S A* 2001; 98:2526-31.

Casimiro MC, Knollmann BC, Yamoah EN, Nie L, Vary JC Jr, Sirenko SG *et al*. Targeted point mutagenesis of mouse Kcnq1: phenotypic analysis of mice with point mutations that cause Romano-Ward syndrome in humans. *Genomics* 2004; 84:555-64.

CFTR protein structure. Oxford University GeneMedicine. July 25, 2012. http://www.genemedresearch.ox.ac.uk/cysticfibrosis/protein.html

Chen H, Guo JH, Lu YC, Ding GL, Yu MK *et al*. Impaired CFTR-dependent amplification of FSH-stimulated estrogen production in cystic fibrosis and PCOS. *J Clin Endocrinol Metab*. 2012 Mar;97(3):923-32.

Cho NL, Javid SH, Carothers AM, Redston M, Bertagnolli MM (2007) Estrogen receptors alpha and beta are inhibitory modifiers of Apc-dependent tumorigenesis in the proximal colon of Min/+ mice. *Cancer Res* 67:2366–2372.

Claude Ferec and Garry Cutting. Assessing the Disease-Liability of Mutations in CFTR. *Cold Spring Harb Perspect Med*. 2012.

Cleveland AG, *et al*. (2009) Disruption of estrogen receptor signaling enhances intestinal neoplasia in Apc(Min/+) mice. *Carcinogenesis* 30:1581–1590.

Cohen-Cymberknoh M, Kerem E, Ferkol T, Elizur A. Airway inflammation in cystic fibrosis: molecular mechanisms and clinical implications. *Thorax*. 2013 May 23.

Colbert LH, Mai V, Tooze JA, Perkins SN, Berrigan D *et al*. Negative energy balance induced by voluntary wheel running inhibits polyp development in APCMin mice. *Carcinogenesis*. 2006 Oct; 27(10):2103-7.

Collaco AM, Jakab RL, Hoekstra N, Mitchell K, Brooks A, Ameen NA. Regulated traffic of anion transporters in mammalian Brunner's glands: a role for water and fluid transport. *Am J Physiol Gastrointest Liver Physiol.* 2013 Jun 6.

Collier LS, Carlson CM, Ravimohan S, Dupuy AJ, Largaespada DA (2005) Cancer gene discovery using Sleeping Beauty transposon-based somatic mutagenesis in the mouse. *Nature* 436:272–276.

Copeland NG, Jenkins NA. Harnessing transposons for cancer gene discovery. Nat Rev Cancer. 2010 Oct;10(10):696-706.

Cormier RT, *et al*. (1997) Secretory phospholipase Pla2g2a confers resistance to intestinal tumorigenesis. *Nat Genet* 17:88–91.

Crosnier C, Stamataki D, Lewis J. Organizing cell renewal in the intestine: stem cells, signals and combinatorial control. *Nat Rev Genet*. 2006 May;7(5):349-59.

Cystic Fibrosis Foundation. Cystic Fibrosis Foundation Patient Registry, 2011 annual data report to the center directors. Bethesda, MD. Cystic Fibrosis Foundation, 2011.

David H. Raulet and Nadia Guerra. Oncogenic stress sensed by the immune system: role of natural killer cell receptors. *Nat Rev Immunology* 9, 569-580. August 2009

de Hostos EL, Choy RK, Nguyen T. Developing novel antisecretory drugs to treat infectious diarrhea. *Future Med Chem*. 2011 Aug;3(10):1317-25.

De Lisle RC, Borowitz D. *Cold Spring Harb Perspect Med*. 2013 Jun 20.

De Lisle, R. C., Mueller, R. and Boyd, M. Impaired mucosal barrier function in the small intestine of the cystic fibrosis mouse. *J Pediatr Gastroenterol Nutr*. 2011; 53, 371-9.

DeBerardinis RJ, Lum JJ, Hatzivassiliou G, Thompson CB. The biology of cancer: metabolic reprogramming fuels cell growth and proliferation. *Cell Metab*. 2008 Jan; 7(1):11-20.

Dedek K, Waldegger S. Colocalization of KCNQ1/KCNE channel subunits in the mouse gastrointestinal tract. *Pflugers Arch*. 2001 Sep;442(6):896-902.

Dekkers JF, Wiegerinck CL, de Jonge HR, Bronsveld I, Janssens HM, de Winter-de Groot KM *et al.* A functional CFTR assay using primary cystic fibrosis intestinal organoids. *Nat Med*. 2013 Jun 2.

Demolombe S, Franco D, de Boer P, Kuperschmidt S, Roden D, Pereon Y *et al*. Differential expression of KvLQT1 and its regulator IsK in mouse epithelia. *Am J Physiol Cell Physiol* 2001; 280:C359-72.

Denamur E, Chehab FF. Methylation status of CpG sites in the mouse and human CFTR promoters. *DNA Cell Biol*. 1995 Sep;14(9):811-5.

Derks S, *et al*. (2008) Integrated analysis of chromosomal, microsatellite and epigenetic instability in colorectal cancer identifies specific associations between promoter

methylation of pivotal tumour suppressor and DNA repair genes and specific chromosomal alterations. *Carcinogenesis* 29:434–439.

Dietrich WF, *et al*. (1993) Genetic identification of Mom-1, a major modifier locus affecting Min-induced intestinal neoplasia in the mouse. *Cell* 75:631–639.

Duerkop BA, Vaishnava S, Hooper LV. Immune responses to the microbiota at the intestinal mucosal surface. *Immunity*. 2009 Sep 18; 31(3):368-76.

Dupuy AJ, *et al*. (2009) A modified sleeping beauty transposon system that can be used to model a wide variety of human cancers in mice. *Cancer Res* 69:8150–8156.

Dupuy AJ, Akagi K, Largaespada DA, Copeland NG, Jenkins NA.
Mammalian mutagenesis using a highly mobile somatic Sleeping Beauty transposon system. *Nature*. 2005 Jul 14;436(7048):221-6.

Dziarski R, Gupta D. Review: Mammalian peptidoglycan recognition proteins (PGRPs) in innate immunity. *Innate Immun*. 2010 Jun; 16(3):168-74.

el Marjou F, *et al*. (2004) Tissue-specific and inducible Cre-mediated recombination in the gut epithelium. *Genesis* 39:186–193.

Elso CM, Lu X, Culiat CT, Rutledge JC, Cacheiro NL, Generoso WM *et al*. Heightened susceptibility to chronic gastritis, hyperplasia and metaplasia in Kcnq1 mutant mice. *Hum Mol Genet*. 2004; 13:2813-21.

Figure 2.13. Schematic overview of alimentary tract and abdominal viscera. “Essential Clinical Anatomy, 4th Edition” by Keith L. Moore, Anne M. R. Agur and Arthur F. Dalley, 2011. Copyright 2011 by the Lippincott Williams & Wilkins publishing group.

Fijneman RJ, Anderson RA, Richards E, Liu J, Tijssen M *et al.* Runx1 is a tumor suppressor gene in the mouse gastrointestinal tract. *Cancer Sci*. 2012 Mar;103(3):593-9.

Fodde R (2002) The APC gene in colorectal cancer. *Eur J Cancer* 38:867–871.

Fodde R, Smits R (2001) Disease model: Familial adenomatous polyposis. *Trends Mol Med* 7:369–373.

Fong Y, Fortner J, Sun RL, *et al.* Clinical score for predicting recurrence after hepatic resection for metastatic colorectal cancer: analysis of 1001 consecutive cases. *Ann Surg* 1999; 230:309-318.

Forbes S, *et al*. (2006) COSMIC 2005. *Br J Cancer* 94:318–322.

Frizzell RA, Hanrahan JW. Physiology of epithelial chloride and fluid secretion. *Cold Spring Harb Perspect Med*. 2012 Jun;2(6):a009563.

Gelfond D, Borowitz D. Gastrointestinal complications of cystic fibrosis. *Clin Gastroenterol Hepatol*. 2013 Apr;11(4):333-42.

Gholami K, Muniandy S, Salleh N. Progesterone downregulates oestrogen-induced expression of CFTR and SLC26A6 proteins and mRNA in rats' uteri. *J Biomed Biotechnol*. 2012; 2012:596084.

Giroux V, Bernatchez G, Carrier JC. Chemopreventive effect of ERβ-Selective agonist on intestinal tumorigenesis in Apc(Min/+) mice. *Mol Carcinog*. 2011 May;50(5):359-69.

Grady WM, Markowitz SD (2002) Genetic and epigenetic alterations in colon cancer. *Annu Rev Genomics Hum Genet* 3:101–128.

Grahammer F, Herling AW, Lang HJ, Schmitt-Graff A, Wittekindt OH, Nitschke R *et al*. The cardiac K+ channel KCNQ1 is essential for gastric acid secretion. *Gastroenterology* 2001; 120:1363-71.

Grahammer F, Warth R, Barhanin J, Bleich M, Hug MJ. The small conductance K+ channel, KCNQ1: expression, function, and subunit composition in murine trachea. *J Biol Chem*. 2001 Nov 9;276(45):42268-75.

Grunnet M, Olesen SP, Klaerke DA, Jespersen T. hKCNE4 inhibits the hKCNQ1 potassium current without affecting the activation kinetics. *Biochem Biophys Res Commun*. 2005 Mar 25;328(4):1146-53.

Gustafsson JK, Ermund A, Ambort D, Johansson ME, Nilsson HE, Thorell K *et al.* Bicarbonate and functional CFTR channel are required for proper mucin secretion and lin k cystic fibrosis with its mucus phenotype. *J Exp Med*. 2012 Jul 2;209(7):1263-72

Habermann JK, *et al*. (2007) Stage-specific alterations of the genome, transcriptome, and proteome during colorectal carcinogenesis. *Genes Chromosomes Cancer* 46: 10–26.

Haigis KM, Caya JG, Reichelderfer M, Dove WF (2002) Intestinal adenomas can develop with a stable karyotype and stable microsatellites. Proc Natl Acad Sci USA 99 Futreal PA, *et al*. (2004) A census of human cancer genes. *Nat Rev Cancer* 4:177–183.

Haigis KM, Dove WF (2003) A Robertsonian translocation suppresses a somatic recombination pathway to loss of heterozygosity. *Nat Genet* 33:33–39.

Hall EH, Crowe SE. Environmental and lifestyle influences on disorders of the large and small intestine: implications for treatment. *Dig Dis*. 2011;29(2):249-54.

Han SW, Kim HP, Shin JY, Jeong EG, Lee WC, Lee KH *et al*.
Targeted sequencing of cancer-related genes in colorectal cancer using next-generation sequencing. *PLoS One*. 2013 May 21;8(5):e64271.

Hans Clevers. The intestinal crypt, a prototype stem cell compartment. *Cell*. Volume 154, Issue 2. 18 July 2013, Pages 274–284

Hardwick JC, Kodach LL, Offerhaus GJ, van den Brink GR. Bone morphogenetic protein signalling in colorectal cancer. *Nat Rev Cancer*. 2008 Oct;8(10):806-12.

Hashimoto T, Perlot T, Rehman A, Trichereau J, Ishiguro H, Paolino M *et al*. ACE2 links amino acid malnutrition to microbial ecology and intestinal inflammation. *Nature*. 2012 Jul 25; 487(7408):477-81.

Heath JK. Transcriptional networks
and signaling pathways that govern vertebrate intestinal development. *Curr Top Dev Biol*. 2010;90:159-92.

Heijmans J, van Lidth de Jeude JF, Koo BK, Rosekrans SL, Wielenga MC *et al.* ER stress causes rapid loss of intestinal epithelial stemness through activation of the unfolded protein response. ER stress causes rapid loss of intestinal epithelial stemness through activation of the unfolded protein response. *Cell Rep*. 2013 Apr 25;3(4):1128-39.

Hodges CA, Cotton CU, Palmert MR, Drumm ML. Generation of a
conditional null allele for Cftr in mice. *Genesis*. 2008 Oct;46(10):546-52.

Hogan AM, Collins D, Baird AW, Winter DC. Estrogen and gastrointestinal malignancy. *Mol Cell Endocrinol*. 2009 Aug 13;307(1-2):19-24.

IPA Ingenuity Systems. Available at http://www.ingenuity.com. Accessed February 2, 2013.

Ivics Z, Kaufman CD, Zayed H, Miskey C, Walisko O, Izsvák Z.
The Sleeping Beauty transposable element: evolution, regulation and genetic applications. *Curr Issues Mol Biol*. 2004 Jan;6(1):43-55.

Jakab RL, Collaco AM, Ameen NA. Physiological relevance of cell-specific distribution patterns of CFTR, NKCC1, NBCe1, and NHE3 along the crypt-villus axis in the intestine. *Am J Physiol Gastrointest Liver Physiol*. 2011 Jan;300(1):G82-98.

Jan Paul Medema and Louis Vermeulen. Microenvironmental regulation of stem cells in intestinal homeostasis and cancer. *Nat Rev* 2011; Vol 474: 318-326.

Je Y, Jeon JY, Giovannucci EL, Meyerhardt JA. Association between physical activity and mortality in colorectal cancer: A meta-analysis of prospective cohort studies. *Int J Cancer*. 2013 Apr 12.

Jespersen T, Grunnet M, Olesen SP. The KCNQ1 potassium channel: from gene to physiological function. *Physiology (Bethesda)*. 2005 Dec;20:408-16.

Ju J, Nolan B, Cheh M, Bose M, Lin Y *et al*. Voluntary exercise inhibits intestinal tumorigenesis in Apc(Min/+) mice and azoxymethane/dextran sulfate sodium-treated mice. *BMC Cancer*. 2008 Nov 2 ;8:316.

Kalluri R and Weinberg RA. The basics of epithelial-mesenchymal transition.

Kaltenbach S, Capri Y, Rossignol S, Denjoy I, Soudée S *et al*. Beckwith-Wiedemann syndrome and long QT syndrome due to familial-balanced translocation t(11;17)(p15.5;q21.3) involving the KCNQ1 gene. *Clin Genet*. 2013 Jul;84(1):78-81.

Kang S, Okuno T, Takegahara N, Takamatsu H, Nojima S, Kimura T *et al*. Intestinal epithelial cell-derived semaphorin 7A negatively regulates development of colitis via αvβ1 integrin. *J Immunol*. 2012 Feb 1;188(3):1108-16.

Kaser, A., Lee, A.H., Franke, A., Glickman, J.N., Zeissig, S., Tilg, H., Nieuwenhuis, E.E., Higgins, D.E., Schreiber, S., Glimcher, L.H., and Blumberg, R.S. (2008). XBP1 links ER stress to intestinal inflammation and confers genetic risk for human inflammatory bowel disease. *Cell*. 134, 743–756.

Kemp Z, *et al*.; ColoRectal tumour Gene Identification (CoRGI) Study Consortium (2006) Evidence for a colorectal cancer susceptibility locus on chromosome 3q21-q24 from a high-density SNP genome-wide linkage scan. *Hum Mol Genet* 15:2903–2910.

Keng VW, *et al*. (2009) A conditional transposon-based insertional mutagenesis screen for genes associated with mouse hepatocellular carcinoma. *Nat Biotechnol* 27: 264–274.

Kent WJ, *et al*. (2002) The human genome browser at UCSC. *Genome Res* 12:996–1006.

Knobloch M, Braun SM, Zurkirchen L, von Schoultz C, Zamboni N, Araúzo-Bravo MJ *et al*. Metabolic control of adult neural stem cell activity by Fasn-dependent lipogenesis. *Nature*. 2013 Jan 10; 493(7431):226-30.

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

Kosinski C, Li VS, Chan AS, *et al*. Gene expression patterns of human colon tops and basal crypts and bmp antagonists as intestinal stem cell niche factors. *Proc Natl Acad Sci U S A*. 2007;104:15418–15423.

Kunzelmann K, Mall M. Electrolyte transport in the mammalian colon: mechanisms and implications for disease. *Physiol Rev*. 2002 Jan;82(1):245-89.

Kurashina K, *et al*. (2008) Chromosome copy number analysis in screening for prognosis-related genomic regions in colorectal carcinoma. *Cancer Sci* 99:1835–1840.

Lang F and Shumilina. Regulation of ion channels by the serum- and glucocorticoid-inducible kinase SGK1. *FASEB J*. 2013 Jan;27(1):3-12.

Lassmann S, *et al*. (2007) Array CGH identifies distinct DNA copy number profiles of oncogenes and tumor suppressor genes in chromosomal- and microsatellite-unstable sporadic colorectal carcinomas. *J Mol Med* 85:293–304.

Lee MP, Ravenel JD, Hu RJ, Lustig LR, Tomaselli G, Berger RD *et al*. Targeted disruption of the Kvlqt1 gene causes deafness and gastric hyperplasia in mice. *J Clin Invest* 2000; 106:1447-55.

Lee JT, Bartolomei MS. X-inactivation, imprinting, and long noncoding RNAs in health and disease. *Cell.* 2013 Mar 14;152(6):1308-23.

Levy DB, *et al*. (1994) Inactivation of both APC alleles in human and mouse tumors. *Cancer Res* 54:5953–5958.

Linley J, Loganathan A, Kopanati S, Sandle GI, Hunter M.
Evidence that two distinct crypt cell types secrete chloride and potassium in human colon. *Gut*. 2013 Jun 5.

Lock H, Valverde MA (2000) Contribution of the IsK (MinK) potassium channel subunit to regulatory volume decrease in murine tracheal epithelial cells. *J Biol Chem* 275:34849–34852

Lubbe WJ, Zuzga DS, Zhou Z, Fu W, Pelta-Heller J *et al*. Guanylyl cyclase C prevents colon cancer metastasis by regulating tumor epithelial cell matrix metalloproteinase-9. *Cancer Res*. 2009 Apr 15;69(8):3529-36.

Luongo C, Moser AR, Gledhill S, Dove WF (1994) Loss of Apc+ in intestinal adenomas from Min mice. *Cancer Res* 54:5947–5952.

Maisonneuve P, Marshall BC, Knapp EA, Lowenfels AB. Cancer risk in cystic fibrosis: a 20-year nationwide study from the United States. *J Natl Cancer Inst*. 2013 Jan 16;105(2):122-9.

Maisonneuve, P., FitzSimmons, S. C., Neglia, J. P., Campbell, P. W., 3rd and Lowenfels, A. B. Cancer risk in nontransplanted and transplanted cystic fibrosis patients: a 10-year study. *J Natl Cancer Inst*. 2003; 95, 381-7.

Mall M, Wissner A, Schreiber R, Kuehr J, Seydewitz HH, Brandis M, Greger R, Kunzelmann K (2000) Role of K(V)LQT1 in cyclic adenosine monophosphate-mediated Cl(–) secretion in human airway epithelia. *Am J Respir Cell Mol Biol* 23:283–289

Mancini-Dinardo D, Steele SJ, Levorse JM, Ingram RS, Tilghman SM. Elongation of the Kcnq1ot1 transcript is required for genomic imprinting of neighboring genes. *Genes Dev*. 2006 May 15;20(10):1268-82.

Mann KM, Ward JM, Yew CC, Kovochich A, Dawson DW, Black MA *et al*. Sleeping Beauty mutagenesis reveals cooperating mutations and pathways in pancreatic adenocarcinoma. *Proc Natl Acad Sci U S A* 2012; 109:5934-41.

March HN, Rust AG, Wright NA, ten Hoeve J, de Ridder J, Eldridge M *et al*. Insertional mutagenesis identifies multiple networks of cooperating genes driving intestinal tumorigenesis. *Nat Genet*. 2011; 43:1202-9.

Matsuzaki T, Tajika Y, Ablimit A, Aoki T, Hagiwara H, and Takata K. Aquaporins in the digestive system. *Med Electron Microsc*. 37: 71-80, 2004.

Mattison J, van der Weyden L, Hubbard T, Adams DJ. Cancer gene discovery in mouse and man. *Biochim Biophys Acta*. 2009 Dec;1796(2):140-61.

McAlpine CA, Barak Y, Matise I, Cormier RT (2006). Intestinal-specific PPARgamma deficiency enhances tumorigenesis in ApcMin/+ mice. *Int J Cancer* 119:2339–2346.

Meric-Bernstam F, Mills GB. Overcoming implementation challenges of personalized cancer therapy. *Nat Rev Clin Oncol*. 2012 Sep;9(9):542-8.

Mootha VK, Lindgren CM, Eriksson KF *et al.* PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* 2003; 34: 267-73.

Moser AR, Pitot HC, Dove WF (1990) A dominant mutation that predisposes to multiple intestinal neoplasia in the mouse. *Science* 247:322–324.

Mueller PR, Wold B (1989) In vivo footprinting of a muscle specific enhancer by ligation mediated PCR. *Science* 246:780–786.

Munoz J, Stange DE, Schepers AG, van de Wetering M, Koo BK, Itzkovitz S *et al*. The Lgr5 intestinal stem cell signature: robust expression of proposed quiescent '+4' cell markers. *EMBO J* 2012; 31:3079-91.

Nakajo K, Nishino A, Okamura Y, Kubo Y. KCNQ1 subdomains involved in KCNE modulation revealed by an invertebrate KCNQ1 orthologue. *J Gen Physiol*. 2011 Nov;138(5):521-35.

Nakao K, *et al*. (2004) High-resolution analysis of DNA copy number alterations in colorectal cancer by array-based comparative genomic hybridization. *Carcinogenesis* 25:1345–1357.

Natalwala A, Spychal R, Tselepis C. Epithelial-mesenchymal transition mediated tumourigenesis in the gastrointestinal tract. *World J Gastroenterol*. 2008 Jun 28;14(24):3792-7.

Neal AM, Taylor HC, Millar ID, Kibble JD, White SJ, Robson L. Renal defects in KCNE1 knockout mice are mimicked by chromanol 293B in vivo: identification of a KCNE1-regulated K+ conductance in the proximal tubule. *J Physiol*. 2011 Jul 15;589(Pt 14):3595-609.

Neklason DW, *et al.* (2008) Common familial colorectal cancer linked to chromosome 7q31: A genome-wide analysis. *Cancer Res* 68:8993–8997.

Netchine I, Rossignol S, Azzi S, Le Bouc Y. Epigenetic anomalies in childhood growth disorders. *Nestle Nutr Inst Workshop Ser*. 2013;71:65-73.

Neyroud N, Tesson F, Denjoy I, Leibovici M, Donger C, Barhanin J, Faure S, Gary F, Coumel P, Petit C, Schwartz K, and Guicheney P. A novel mutation in the potassium channel gene KVLQT1 causes the Jervell and Lange-Nielsen cardioauditory syndrome. *Nat Genet* 15: 186–189, 1997.

Nikou GC, Toubanakis C, Moulakakis KG, Pavlatos S, Kosmidis C, Mallas E *et al*. Carcinoid tumors of the duodenum and the ampulla of Vater: current diagnostic and therapeutic approach in a series of 8 patients. Case series. *Int J Surg* 2011; 9:248-53.

Nishisho I, *et al.* (1991) Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. *Science* 253:665–669.

Norkina O, Kaur S, Ziemer D, De Lisle RC. Inflammation of the cystic fibrosis mouse small intestine. *Am J Physiol Gastrointest Liver Physiol*. 2004 Jun;286(6):G1032-41.

Odorizzi G (2006). The multiple personalities of Alix. *J Cell Sci* 119:3025–3032.

O'Mahony F., Alzamora R., Chung H. L., Thomas W., Harvey B. J. (2009). Genomic priming of the antisecretory response to estrogen in rat distal colon throughout the estrous cycle. *Mol. Endocrinol*. 23, 1885–1899.

Ostedgaard LS, Meyerholz DK, Vermeer DW, Karp PH, Schneider L *et al*. Cystic fibrosis transmembrane conductance regulator with a shortened R domain rescues the intestinal phenotype of CFTR-/- mice. *Proc Natl Acad Sci U S A*. 2011 Feb 15;108(7):2921-6.

Ousingsawat J, Spitzner M, Puntheeranurak S, Terracciano L, Tornillo L *et al*. Expression of voltage-gated potassium channels in human and mouse colonic carcinoma. *Clin Cancer Res*. 2007 Feb 1;13(3):824-31.

Palmer BR, Frampton CM, Skelton L, Yandle TG, Doughty RN *et al*. KCNE5 polymorphism rs697829 is associated with QT interval and survival in acute coronary syndromes patients. *J Cardiovasc Electrophysiol*. 2012 Mar;23(3):319-24.

Pedersen SF, Stock C. Ion channels and transporters in cancer: pathophysiology, regulation, and clinical potential. *Cancer Res* 2013 Mar 15;73(6):1658-61.

Perez-Mancera PA, Rust AG, van der Weyden L, Kristiansen G, Li A, Sarver AL *et al*. The deubiquitinase USP9X suppresses pancreatic ductal adenocarcinoma. *Nature* 2012; 486:266-70.

Peroz D, Rodriguez N, Choveau F, Baro, I, Merot J, Loussouarn G. Kv7.1 (KCNQ1) properties and channelopathies. *J Physiol* 2008; 586:1785-9.

Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29:2003–2007

Pott J, Hornef M. Innate immune signalling at the intestinal epithelium in homeostasis and disease. *EMBO Rep*. 2012 Aug;13(8):684-98.

Powell DW, Adegboyega PA, Di Mari JF, Mifflin RC. Epithelial cells and their neighbors I. Role of intestinal myofibroblasts in development, repair, and cancer. *Am J Physiol Gastrointest Liver Physiol* 2005;289:G2–G7.

Powell SM, *et al*. (1992) APC mutations occur early during colorectal tumorigenesis. *Nature* 359:235–237.

Preston P, Wartosch L, Günzel D, Fromm M, Kongsuphol P, Ousingsawat J, Kunzelmann K, Barhanin J, Warth R, Jentsch TJ. Disruption of the K+ channel beta-subunit KCNE3 reveals an important role in intestinal and tracheal Cl- transport. *J Biol Chem*. 2010; 285, 7165-75.

Preston P, Wartosch L, Günzel D, Fromm M, Kongsuphol P *et al.* Disruption of the K+ channel beta-subunit KCNE3 reveals an important role in intestinal and tracheal Cl- transport. *J Biol Chem*. 2010 Mar 5;285(10):7165-75.

Rabinovich RA, MacNee W. Chronic obstructive pulmonary disease and its comorbidities. *Br J Hosp Med (Lond)*. 2011 Mar;72(3):137-45.

Rapetti-Mauss R, O'Mahony F, Sepulveda FV, Urbach V, Harvey BJ. Oestrogen promotes KCNQ1 potassium channel endocytosis and postendocytic trafficking in colonic epithelium. *J Physiol.* 2013 Jun 1;591(Pt 11):2813-31.

Restrepo-Angulo I, De Vizcaya-Ruiz A, Camacho J. Ion channels in toxicology. *J Appl Toxicol*. 2010 Aug;30(6):497-512.

Reya T, Clevers H. Wnt signalling in stem cells and cancer. *Nature*. 2005 Apr 14;434(7035):843-50.

Rice KS, Dickson G, Lane M, Crawford J, Chung SK, Rees MI *et al*. Elevated serum gastrin levels in Jervell and Lange-Nielsen syndrome: a marker of severe KCNQ1 dysfunction? *Heart Rhythm* 2011; 8:551-4.

Rice P, Longden I, Bleasby A (2000) EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet* 16:276–277.

Ried T, *et al*. (1996) Comparative genomic hybridization reveals a specific pattern of chromosomal gains and losses during the genesis of colorectal tumors. *Genes Chromosomes Cancer* 15:234–245.

Riordan JR, Rommens JM, Kerem B, Alon N, Rozmahel R *et al*. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science*. 1989 Sep 8;245(4922):1066-73.

Rivas A, Francis HW. Inner ear abnormalities in a Kcnq1 (Kvlqt1) knockout mouse: a model of Jervell and Lange-Nielsen syndrome. *Otol Neurotol*. 2005 May;26(3):415-24.

Roepke TK, Kontogeorgis A, Ovanez C, Xu X, Young JB *et al*. Targeted deletion of kcne2 impairs ventricular repolarization via disruption of I(K,slow1) and I(to,f). *FASEB J*. 2008 Oct;22(10):3648-60.

Rohrmann S, Hermann S, Linseisen J. Heterocyclic aromatic amine intake increases colorectal adenoma risk: findings from a prospective European cohort study. *Am J Clin Nutr.* 2009 May;89(5):1418-24.

Rosalyn Carson-DeWitt. Gastrointestinal Cancers. *NYU Division of Hematology and Medical Oncology*. http://medicine.med.nyu.edu/medonc/node/491

Rothenberg ME, Nusse Y, Kalisky T, Lee JJ, Dalerba P, Scheeren F *et al.* Identification of a cKit(+) colonic crypt base secretory cell that supports Lgr5(+) stem cells in mice. *Gastroenterology*. 2012 May;142(5):1195-1205

S. Ullrich, J. Su, F. Ranta, O.H. Wittekindt, F. Ris, M. Rösler, U. Gerlach, D. Heitzmann, R. Warth, F. Lang, Effects of IKs channel inhibitors in insulinsecreting INS-1 cells. *Pflugers Arch*. 451 (2005) 428–436.

Saif MW and Chu E. Biology of colorectal cancer. *Cancer J*. 2010 May-Jun;16(3):196-201.

Samuels Y, Velculescu VE. Oncogenic mutations of PIK3CA in human cancers. *Cell Cycle*. 2004 Oct;3(10):1221-4.

Sato T, Stange DE, Ferrante M, Vries RG, Van Es JH, Van den Brink S *et al*. Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. *Gastroenterology* Nov; 141:1762-72.

Schee K, Lorenz S, Worren MM, Günther CC, Holden M, Hovig E *et al*. Deep Sequencing the MicroRNA Transcriptome in Colorectal Cancer. *PLoS One*. 2013 Jun 18;8(6):e66165.

Schmidt MH, *et al*. (2004) Alix/AIP1 antagonizes epidermal growth factor receptor downregulation by the Cbl-SETA/CIN85 complex. *Mol Cell Biol* 24:8981–8993.

Schroeder BC, Waldegger S, Fehr S, Bleich M, Warth R, Greger R, Jentsch TJ. A constitutively open potassium channel formed by KCNQ1 and KCNE3. *Nature*. 2000 403:196–199

Sebastian A, Rishishwar L, Wang J, Bernard KF, Conley AB *et al*. Origin and evolution of the cystic fibrosis transmembrane regulator protein R domain. *Gene*. 2013 Jul 10;523(2):137-46.

Seber GAF (2002). The Estimation of Animal Abundance and Related Parameters (Blackburn, Caldwell, NJ) Ed 2.

Shalapour S, Deiser K, Sercan O, Tuckermann J, Minnich K, Willimsky G *et al*. Commensal microflora and interferon-gamma promote steady-state interleukin-7 production in vivo. *Eur J Immunol*. 2010 Sep; 40(9):2391-400.

Shao J, Sheng GG, Mifflin RC, *et al*. Roles of myofibroblasts in prostaglandin E2-stimulated intestinal epithelial proliferation and angiogenesis. *Cancer Res* 2006; 66:846–855.

Shi Y, Massagué J. Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *Cell*. 2003 Jun 13;113(6):685-700.

Shih IM, *et al*. (2001) Evidence that genetic instability occurs at an early stage of colorectal tumorigenesis. *Cancer Res* 61:818–822.

Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. *CA Cancer J Clin*. 2013 Jan;63(1):11-30.

Simon R, Mirlacher M, Sauter G: Tissue microarrays. *Biotechniques* 2004; 36:98-105.

Snouwaert JN, Brigman KK, Latour AM, Malouf NN, Boucher RC, Smithies O, Koller BH. An animal model for cystic fibrosis made by gene targeting. *Science* 257: 1083–1088, 1992.

Son JW, Kim YJ, Cho HM, Lee SY, Lee SM *et al.* Promoter hypermethylation of the CFTR gene and clinical/pathological features associated with non-small cell lung cancer. *Respirology*. 2011 Nov;16(8):1203-9.

Stabenau A, *et al*. (2004) The Ensembl core software libraries. *Genome Res* 14: 929–933.

Starr TK, Allaei R, Silverstein KA, Staggs RA, Sarver AL, Bergemann TL *et al*. A transposon-based genetic screen in mice identifies genes altered in colorectal cancer. *Science* 2009; 323:1747-50.

Starr TK, Scott PM, Marsh BM, Zhao L, Than BL, O'Sullivan MG *et al*. A Sleeping Beauty transposon-mediated screen identifies murine susceptibility genes for adenomatous polyposis coli (Apc)-dependent intestinal tumorigenesis. *Proc Natl Acad Sci U S A* 2011; 108:5765-70.

Stichting FMWV Rotterdam: Code Goed Gebruik van lichaamsmateriaal, 2011, ISBN 978-90-817510-0-1

Strong TV, Boehm K, Collins FS. Localization of cystic fibrosis transmembrane conductance regulator mRNA in the human gastrointestinal tract by in situ hybridization. *J Clin Invest*. 1994 Jan;93(1):347-54.

Su LK *et al.* (1992). Multiple intestinal neoplasia caused by a mutation in the murine homolog of the APC gene. *Science* 256:668–670.

Subramanian A, Tamayo P, Mootha VK *et al*. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005; 102: 15545-50.

Takagi T, Nishio H, Yagi T, Kuwahara M, Tsubone H, Tanigawa N *et al*. Phenotypic analysis of vertigo 2 Jackson mice with a Kcnq1 potassium channel mutation. *Exp Anim* 2007; 56:295-300.

Takayama T, Miyanishi K, Hayashi T, Sato Y, Niitsu Y (2006). Colorectal cancer: Genetics of development and metastasis. *J Gastroenterol* 41:185–192.

Teng S, Ma L, Zhen Y, Lin C, Bähring R *et al*. Novel gene hKCNE4 slows the activation of the KCNQ1 channel. *Biochem Biophys Res Commun*. 2003 Apr 11;303(3):808-13.

Ueo T, Imayoshi I, Kobayashi T, Ohtsuka T, Seno H *et al*. The role of Hes genes in intestinal development, homeostasis and tumor formation. *Development*. 2012 Mar;139(6):1071-82.

Unoki H, Takahashi A, Kawaguchi T, Hara K, Horikoshi M, Andersen G *et al*. SNPs in KCNQ1 are associated with susceptibility to type 2 diabetes in East Asian and European populations. *Nat Genet* 2008; 40:1098-102.

Vaishnava S, Yamamoto M, Severson KM, Ruhn KA, Yu X, Koren O *et al*. The antibacterial lectin RegIIIgamma promotes the spatial segregation of microbiota and host in the intestine. *Science*. 2011 Oct 14; 334(6053):255-8.

Vallon V, Grahammer F, Richter K, Bleich M, Lang F, Barhanin J, Volkl H, and Warth R. Role of KCNE1-dependent K+ fluxes in mouse proximal tubule. *J Am Soc Nephrol* 12: 2003–2011, 2001.

Vallon V, Grahammer F, Volkl H, Sandu CD, Richter K, Rexhepaj R *et al*. KCNQ1-dependent transport in renal and gastrointestinal epithelia. *Proc Natl Acad Sci U S A* 2005; 102:17864-9.

Van Laer L, Carlsson PI, Ottschytsch N, Bondeson ML *et al*. The contribution of genes involved in potassium-recycling in the inner ear to noise-induced hearing loss. *Hum Mutat*. 2006 Aug;27(8):786-95.

Venkatasubramanian J, Ao M, Rao MC. Ion transport in the small intestine. *Curr Opin Gastroenterol*. 2010 Mar;26(2):123-8.

Vito P, Lacanà E, D'Adamio L. (1996) Interfering with apoptosis: Ca(2+)-binding protein ALG-2 and Alzheimer's disease gene ALG-3. *Science* 271:521–525.

Vogelstein B, *et al*. (1989) Allelotype of colorectal carcinomas. *Science* 244:207–211.

Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA Jr, Kinzler KW. Cancer genome landscapes. *Science*. 2013 Mar 29;339(6127):1546-58.

Wang H, Miao K, Zhao J, Liu L, Cui G *et al*.
Common Variants in KCNQ1 Confer Increased Risk of Type 2 Diabetes and Contribute t o the Diabetic Epidemicin East Asians: A Replication and Meta-Analysis. *Ann Hum Genet*. 2013 Jun 21.

Wang Q, Curran ME, Splawski I, Burn TC, Millholland JM *et al*. Positional cloning of a novel potassium channel gene: KVLQT1 mutations cause cardiac arrhythmias. *Nat Genet*. 1996 Jan;12(1):17-23.

Warth R, Garcia Alzamora M, Kim JK, Zdebik A, Nitschke R, Bleich M *et al*. The role of KCNQ1/KCNE1 K(+) channels in intestine and pancreas: lessons from the KCNE1 knockout mouse. *Pflugers Arch* 2002; 443:822-8.

Watson AJ, Collins PD. Colon cancer: a civilization disorder. *Dig Dis*. 2011;29(2):222-8.

Wilschanski M, and Durie PR. Patterns of GI disease in adulthood associated with mutations in the CFTR gene. *Gut* 56: 1153-1163, 2007.

Winbo A, Sandstrom O, Palmqvist R, Rydberg A. Iron-deficiency anaemia, gastric hyperplasia, and elevated gastrin levels due to potassium channel dysfunction in the Jervell and Lange-Nielsen Syndrome. *Cardiol Young* 2012; 18:1-10.

Winkler GS, Mulder KW, Bardwell VJ, Kalkhoven E, Timmers HT (2006) Human Ccr4-Not complex is a ligand-dependent repressor of nuclear receptor-mediated transcription. *EMBO J* 25:3089–3099.

Wonderlin WF, Strobl JS. Potassium channels, proliferation and G1 progression. *J Membr Biol*. 1996 Nov;154(2):91-107.

Wood LD, *et al.* (2007) The genomic landscapes of human breast and colorectal cancers. *Science* 318:1108–1113.

Wu X, Li Y, Crise B, Burgess SM. (2003) Transcription start regions in the human genome are favored targets for MLV integration. *Science* 300:1749–1751.

Wulff, H, Castle, NA, Pardo, LA. Voltage-gated potassium channels as therapeutic targets. *Nat Rev Drug Discovery*. 2009; 8: 982-1001.

Xie C, Jiang XH, Zhang JT, Sun TT, Dong JD *et al.* CFTR suppresses tumor progression through miR-193b targeting urokinase plasminogen activator (uPA) inprostate cancer. *Oncogene*. 2013 May 2;32(18):2282-91

Xiong Q, Gao Z, Wang W, Li M. Activation of Kv7 (KCNQ) voltage-gated potassium channels by synthetic compounds. *Trends Pharmacol Sci*. 2008 Feb;29(2):99-107.

Yang, K, Popova, NV, Yang, W, Lozonschi, I, Tadesse, S, Kent, S *et al*. Interaction of Muc2 and Apc on Wnt signaling and in intestinal tumorigenesis: potential role of chronic inflammation. *Cancer Research* 2008; 68: 7313-7322.

Yasuda K, Miyake K, Horikawa Y, Hara K, Osawa H, Furuta H *et al*. Variants in KCNQ1 are associated with susceptibility to type 2 diabetes mellitus. *Nat Genet* 2008; 40:1092-7.

Yun KE, Chang Y, Jung HS, Kim CW, Kwon MJ *et al*. Impact of body mass index on the risk of colorectal adenoma in a metabolically healthy population. *Cancer Res*. 2013 Jul 1;73(13):4020-7.

Zambrowicz BP, *et al*. (1997) Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. *Proc Natl Acad Sci USA* 94:3789–3794.

Zhou L, Dey CR, Wert SE, DuVall MD, Frizzell RA *et al*.
Correction of lethal intestinal defect in a mouse model of cystic fibrosis by human CFTR. *Science*. 1994 Dec 9;266(5191):1705-8.

Zlobec I, Steele R, Terracciano L, *et al.* Selecting immunohistochemical cut-off scores for novel biomarkers of progression and survival in colorectal cancer. *J Clin Pathol* 2007; 60:1112-1116.

APPENDIX 1

A *Sleeping Beauty* transposon-mediated screen identifies murine susceptibility genes for adenomatous polyposis coli (*Apc*)-dependent intestinal tumorigenesis

Timothy K. Starr[a], Patricia M. Scott[b], Benjamin M. Marsh[b], Lei Zhao[b], Bich L. N. Than[b], M. Gerard O'Sullivan[a,c], Aaron L. Sarver[d], Adam J. Dupuy[e], David A. Largaespada[a], and Robert T. Cormier[b]

[a]Department of Genetics, Cell Biology and Development, Center for Genome Engineering, Masonic Cancer Center, University of Minnesota, Minneapolis, MN 55455; [b]Department of Biochemistry and Molecular Biology, University of Minnesota Medical School, Duluth, MN 55812; [c]Department of Veterinary Population Medicine, College of Veterinary Medicine, University of Minnesota, St. Paul, MN 55108; [d]Department of Biostatistics and Informatics, Masonic Cancer Center, University of Minnesota, Minneapolis, MN 55455; and [e]Department of Anatomy and Cell Biology, University of Iowa, Iowa City, IA 52242

Starr TK, Scott PM, Marsh BM, Zhao L, Than BL, O'Sullivan MG et al. A Sleeping Beauty transposon-mediated screen identifies murine susceptibility genes for adenomatous polyposis coli (Apc)-dependent intestinal tumorigenesis. *Proc Natl Acad Sci USA* 2011; 108: 5765–5770.

Abstract. It is proposed that a progressive series of mutations and epigenetic events leads to human colorectal cancer (CRC) and metastasis. Furthermore, data from resequencing of the coding regions of human CRC suggests that a relatively large number of mutations occur in individual human CRC, most at low frequency. The functional role of these low-frequency mutations in CRC, and specifically how they may cooperate with high-frequency mutations, is not well understood. One of the most common rate-limiting mutations in human CRC occurs in the adenomatouspolyposis coli (*APC*) gene. To identify mutations that cooperate with mutant *APC*, we performed a forward genetic screen in mice carrying a mutant allele of *Apc* (*Apc*^{Min}) using *Sleeping Beauty (SB)* transposon-mediated mutagenesis. *Apc*^{Min} *SB*-mutagenized mice developed three times as

many polyps as mice with the Apc^{Min} allele alone. Analysis of transposon common insertion sites (CIS) identified the Apc locus as a major target of *SB*-induced mutagenesis, suggesting that *SB* insertions provide an efficient route to biallelic Apc inactivation. We also identified an additional 32 CIS genes/loci that may represent modifiers of the Apc^{Min} phenotype. Five CIS genes tested for their role in proliferation caused a significant change in cell viability when message levels were reduced in human CRC cells. These findings demonstrate the utility of using transposon mutagenesis to identify low-frequency and cooperating cancer genes; this approach will aid in the development of combinatorial therapies targeting this deadly disease.

Human colorectal cancers (CRC) generally can be divided into two classes based on whether they display chromosomal instability (CIN) or microsatellite instability (MSI). The majority of CRC (~80–90%) have a CIN phenotype; the remaining cases are characterized by MSI (Grady and Markowitz 2002). CRC displaying CIN frequently harbor allelic losses or mutations in adenomatous polyposis coli (*APC*), v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (*KRAS*), SMAD family member 4 (*SMAD4*), and tumor protein p53 (*TP53*), whereas MSI-type CRC usually have a mutation in one of six DNA mismatch repair genes (Takayama *et al.* 2007). In both CIN and MSI CRC complete functional loss of a gatekeeper tumor suppressor gene typically is the rate-limiting event in intestinal cell transformation. For CIN CRC, *APC* plays the key gate-keeping role, and its loss underlies the great majority of CIN CRC and >80% of all CRC. Although both classes of CRC are characterized by high-frequency mutations, such as those in *APC*, it is evident that many more low-frequency mutations are required for CRC development, and the majority of these low-frequency mutations are unknown (Wood *et al.* 2007).

To identify these low-frequency mutations, we performed a forward genetic screen in mice using the *Sleeping Beauty (SB*) DNA transposon as a mutagen in intestinal epithelial cells. To focus on mutations that contribute to the CIN phenotype, we

conducted the screen in mice carrying the *Apc*^{Min} allele. *Apc*^{Min} mice harbor a T→A nonsense mutation in the *Apc* gene (Moser *et al.* 1990, Su *et al.* 1992) that results in a truncated protein product that is unable to bind β-catenin and promote its degradation, thus leading to abnormal levels of β-catenin protein and up-regulation of β-catenin target genes such as cyclin D1 (*Ccnd1*) and myelocytomatosis oncogene (*C-Myc*). The *Min* mutation corresponds to a mutational hotspot in the human *APC* ortholog, and these mutations similarly result in dysregulation of the Wnt/β-catenin signaling pathway. There is strong evidence that β-catenin dysregulation is a common transformative event in tumorigenesis in the *Apc*^{Min} mouse and in both the inherited form of *APC*-deficient CRC (familial adenomatous polyposis, FAP) and in sporadic CRC (Fodde R 2002). Thus, the *Apc*^{Min} mouse is an informative genetic model for *APC*-deficient intestinal cancer. *Apc*^{Min} mice on the C57BL/6J background strain rarely survive beyond 120 d and can develop >100 tumors throughout the small and large intestine, with the phenotype dependent on diet, mouse strain, and other environmental factors (Dietrich *et al.* 1993, Cormier *et al.* 1997).

As in human CRC patients, loss of heterozygosity (LOH) leading to inactivation of both alleles of *Apc* is necessary for tumorigenesis to commence in *Apc*^{Min} mice (Levy *et al.* 1994, Luongo *et al.* 1994). However, in contrast to LOH events in many human CRC, LOH in *Apc*^{Min} tumors occurs predominantly by homologous somatic recombination (Haigis and Dove 2003). In this study we screened for mutations that cooperate with the *Apc*^{Min} mutation by randomly mutating genes through selective activation of SB transposition in intestinal cells of *Apc*^{Min} mice. The results of our screen support the importance of the loss of the second allele of *Apc*, because the great majority of tumors analyzed contained a transposon insertion in *Apc*, in particular in tumors in which there was maintenance of heterozygosity (MOH) for the *Min* allele. In addition to *Apc*, we identified 32 other genes and loci that probably facilitate the development of intestinal cancer in an *Apc*^{Min} model. The function of these additional mutations could be to remove the requirement for *Apc* LOH, or they may function in some other manner. The majority of these genes have not been associated with CRC previously. To confirm that

these genes play a causal role in tumor development, we used siRNA to knock down message levels of nine of the candidate genes in human colon cancer cell lines and demonstrated that five of these genes affected the growth rate of these cells.

Results

Design of a Forward Genetic Screen for CRC Genes. In a previous study we demonstrated that *SB* transposon-mediated mutagenesis in the intestinal tract of C57BL/6J $Apc^{+/+}$ mice resulted in polyp formation (Starr *et al.* 2009). By mapping transposon insertions in DNA extracted from these tumors, we were able to identify 77 genetic loci which probably harbored genes that, when mutated, contributed to tumor development. Because *APC* loss is rate limiting in the development of most human CRC (Powell *et al.* 1992), we reasoned that *SB* mutagenesis in a mouse already harboring a mutation in *Apc* might generate more tumors with a shorter latency and reveal mutations that cooperate with *Apc* during tumor development.

To identify these genes, we performed a forward genetic screen using *SB* transposon-mediated mutagenesis in Apc^{Min} mice. The screen consisted of a cohort of Apc^{Min} SB transgenic test mice along with three groups of Apc^{Min} control mice. The Apc^{Min} SB test mice harbored three transgenes required for targeting *SB* mutagenesis to the gastrointestinal tract (***Figure 39***). The first transgene was a concatamer of oncogenic transposons (T2/Onc) that were resident on chromosome 1 (Collier *et al.* 2005). To enhance the mutagenic potential of the transposon, T2/Onc contains a strong viral promoter, splice acceptors in both orientations, and a bidirectional polyA signal. The second transgene was a conditionally expressed knockin SB11 transposase allele downstream of the Rosa26 promoter (Rosa26-LsL-SB11) (Zambrowics *et al.* 1997, Dupuy *et al.* 2009). Because of the presence of a floxed stop cassette, the transposase allele is not expressed unless Cre recombinase protein is present. The third transgene was Cre recombinase driven by the gastrointestinal tract-specific Villin promoter (Vil-Cre) (el Marjou *et al.* 2004). We have shown previously that these three transgenes effectively

limit SB mutagenesis to the intestinal tract (Starr *et al.* 2009). All mice were heterozygous for the *Apc*^{Min} allele, and all the transgenes were fully congenic on the C57BL/6J genetic background. The first control group contained Rosa26-LsL-SB11 and either Vil-Cre or T2/Onc; the second control group contained T2/Onc and/or Vil-Cre but not Rosa26-LsL-SB11; and the third control group harbored only the *Apc*^{Min} allele. Mice were killed when moribund or at 120 d.

Intestinal Tumorigenesis Is Enhanced Significantly in *Apc*^{Min} SB Test Mice. *Apc*^{Min} mice that harbored all three transgenes (Rosa26- LsL-SB11, T2/Onc, and Vil-Cre) developed an average of 360 polyps (test mice, ***Table 12***). In contrast, mice carrying the *Apc*^{Min} allele alone developed an average of 112 polyps (control group 3, ***Table 12***), a result that is consistent with the phenotype of *Apc*^{Min} mice in our colony (McAlpine *et al.* 2006). Surprisingly, we also observed an enhanced rate of polyp development in control group 1 that carried the Rosa-26-LsL-SB11 allele but not the complete combination of alleles required for transposition. This control group developed an average of 182 polyps (control group 1, ***Table 12***), a result that was unexpected based on previous screens. It is possible that the increased polyp number in these animals is caused by one or more modifiers linked to the Rosa26-LsL-SB11 transgene, because strain-specific modifiers are known to exist (Fodde and Smits 2001). Control animals carrying the *Apc*^{Min} allele, T2/Onc, and/or Vil-Cre, but not Rosa26-LsL-SB11 (control group 2, ***Table 12***) developed the same number of polyps as the control mice carrying the *Apc*^{Min} allele alone. Although the Rosa26-LsL-SB11 allele alone contributes to polyp formation, the effect of active SB transposition was much greater, resulting in twice as many polyps in the test mice. In addition, the tumor burden was so extensive that *Apc*^{Min} SB test mice became moribund earlier than any of the three control groups (***Table 12***). Indeed, in a subset of *Apc*^{Min} SB test mice the tumor load was very severe, with some animals developing as many as 700 tumors.

Group*	Number per group	Average no. polyps per mouse†			Date of death‡
		Large intestine	Small intestine	Total	
Test	38	14.7	345	360	85
Control 1	57	4.2	178	182	110
Control 2	9	2.4	113	115	120+
Control 3	100	2	110	112	120+

*Groups: Test = Apc^{Min} × Rosa26-LsL-SB11 × T2/Onc × Vil-Cre; Control 1 = littermates harboring either Apc^{Min} × Rosa26-LsL-SB11 or Apc^{Min} × Rosa26-LsL-SB11 × T2/Onc or Apc^{Min} × Rosa26-LsL-SB11 × Vil-Cre; Control 2 = littermates harboring either Apc^{Min} × T2/Onc or Apc^{Min} × Vil-Cre or Apc^{Min} × T2/Onc × Vil-Cre; Control 3 = contemporaneous mice harboring Apc^{Min} only.
†Average number of polyps per mouse by large intestine, small intestine, and total.
‡Control groups 2 and 3 were killed at 120 d whether they were moribund or not.

Table 12. Polyp number and age of death for transgenic mice.

Although polyp number was greatly increased by SB mutagenesis, there was no evidence of local or systemic metastasis in experimental or control mice. We performed histopathologic analysis of tumors collected from 10 animals. These analyses identified numerous microadenomas and adenomas in the small intestine and a much smaller number of these lesions in the large intestines. No adenocarcinomas were identified, perhaps because of the short lifespan of *Apc^{Min}* SB test mice. Immunohistochemistry for β-catenin was performed on 24 adenomas from seven animals. There was increased expression of β-catenin in all tumors compared with the adjacent normal mucosa epithelium (***Figure 38***).

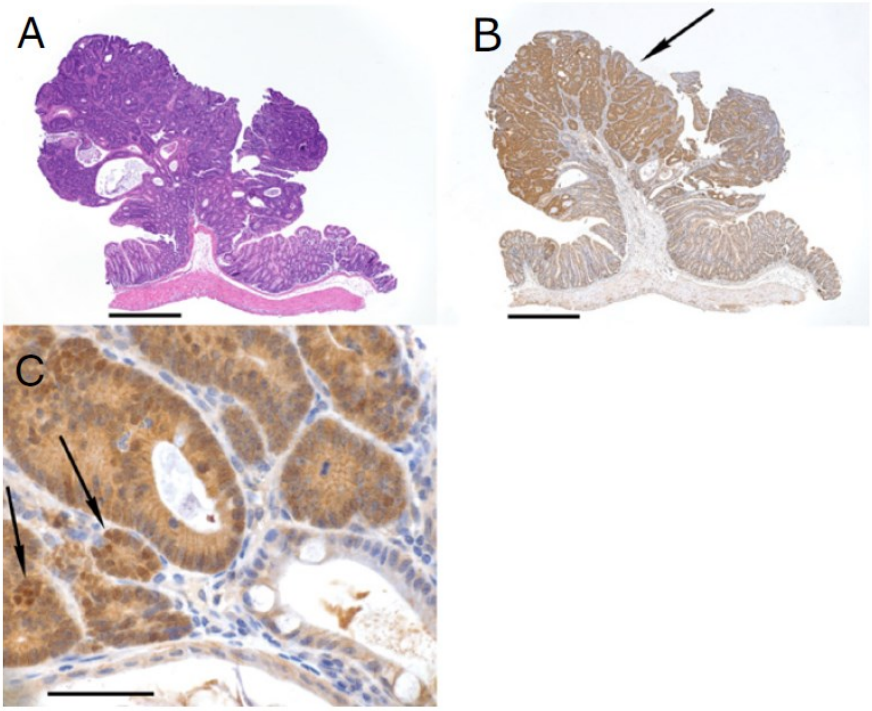


Figure 38. A pedunculated adenoma stained with H&E (A) or immunostained for β-catenin (B and C). **(B)** There is increased staining for β-catenin (arrow) in the adenoma. **(C)** Higher-power magnification of a different section showing increased cytoplasmic and nuclear (arrows) staining for β-catenin in tumor cells compared with adjacent normal tissue seen in lower right and bottom of picture. (Scale bars: A and B, 500 μm; C, 50 μm.)

Analysis of Common Insertion Sites Identifies 30 Candidate Cancer Genes. To identify genes that contribute to tumor initiation and development in Apc^{Min} mice, we analyzed transposon insertions in 96 polyps, representing all regions of the intestines, from 12 mice to find common insertion sites (CIS). A CIS is defined by analyzing transposon insertions in many tumors and identifying genomic loci that contain transposon insertions at a higher rate than would be expected by chance (SI Materials and Methods). The presence of a CIS indicates that a transposon-mediated mutation in that locus probably has contributed to tumor development. By analyzing the genes within the CIS, one can identify candidate cancer genes. To map transposon insertions, we isolated

DNA from the 96 tumors, digested the DNA with restriction enzymes, and performed ligation-mediated PCR (LM-PCR) to amplify transposon-genomic fragments specifically (Mueller and Wold 1989). Barcodes and fusion sequences were attached to the LM-PCR primers to enable pooling of the amplicons, which then were sequenced using the Roche GS FLX pyrosequencing machine. Six separate sequencing runs produced 347,993 sequence reads, 93% of which contained a barcode, the transposon sequence, and sufficient genomic sequence (>16 bp) for BLAST analysis. We were able to map more than half of these sequences (53%) unambiguously to the mouse genome. Of the 173,101 mapped sequences, 100,171 (67%) were redundant, leaving 72,930 nonredundant mapped insertions. Roughly half of the nonredundant insertions mapped to the same chromosome as the donor transposon concatamer (Chr 1), as expected because of the phenomenon of local hopping seen in other SB screens (Starr *et al.* 2009, Collier *et al.* 2005, Keng *et al.* 2009). To eliminate statistical bias in the dataset, these sequences were eliminated along with a smaller number of insertions that probably represent PCR artifacts (SI Materials and Methods). The remaining 30,088 insertions (***Table 18***) were analyzed to determine CIS. We used Monte Carlo simulations to find insertion rates in a given genomic window size that would not be expected to occur by chance (Starr *et al.* 2009). For example, based on a random assignment of 30,088 insertions to the mouse genome, one would not expect to find five or more insertions within a 12-kb window. Using these Monte Carlodefined parameters, we identified 37 CIS. Two of these CIS were removed from further analysis because all the tumors contributing to these two CIS originated from a single mouse, indicating the tumors may be clonally related. Two more CIS were removed because they also were identified in a control dataset of tail-snip DNA from mice harboring unselected SB insertions and may represent hotspots for SB insertions (Starr *et al.* 2009) Because this control dataset was generated from tail snips, it is possible that other hotspots exist in other types of cells. After removal of these possible artifacts, 33 CIS remained (***Table 13***).

We assigned a candidate gene to each CIS if the majority of the insertions were in or near a single gene (***Table 13***). Four of the 33 CIS did not have an annotated gene

within 40 kb and were not assigned a candidate gene. Another CIS contained two overlapping genes, SET domain-containing 5 (*Setd5*) and lipoma HMGIC fusion partner-like 4 (*Lhfpl4*), and all insertions in this CIS were in both genes. Notably, this CIS is located adjacent to the Rosa26 locus where the conditional SB11 knockin is located, and eight of nine insertions in this CIS are oriented with the internal promoter in the direction that would cause overexpression of the transgene. Rather than tagging an endogenous cancer gene, this CIS could represent selective pressure for increased mutagenesis via overexpression of SB11 transposase. Whether the transposon insertion caused a gain- or loss-of-function mutation sometimes can be predicted by analyzing the location and orientation of the insertions in all the tumors that comprise a single CIS. If all the tumors in a single CIS have transposon insertions in the same intron, and all the transposons are oriented in the direction of transcription, we predict the insertion causes a gain-of-function mutation. If the distribution of transposon insertions in all the tumors of a CIS is apparently random, and there is no bias in orientation, we predict a loss-of-function effect. ***Table 13*** lists the predictions for the CIS. In total we identified 30 genes and four genomic loci with no annotated genes that probably contribute to intestinal tract cancer when mutated.

Candidate gene	Chromosome	Start address*	End address*	Number of insertions†	Tumors with insertions‡	Predicted effect on gene§
4930422G04Rik	3	127241424	127431287	23	12	NP
AC115907.7	3	127697338	127730035	8	8	NP
AC131780.5	9	3001410	3030207	7	7	NP
Adamts6	13	105093822	105104050	5	3	NP
Ap1ar	3	127454370	127633377	17	10	NP
Apc	18	34324389	34514767	185	72	Loss
Atf2	2	73708689	73773260	9	9	NP
Atl2	17	80239637	80306988	10	7	Loss
Cnot1	8	98254541	98300340	8	8	Loss
Csnk1a1	18	61726208	61750047	7	6	Loss
Elac1	18	73903365	73913215	5	3	Gain
Emcn	3	136918543	137092521	17	9	NP
Esco1	18	10578932	10773953	15	11	Loss
Fnbp1l	3	122217801	122319064	11	11	Loss
Itgam	7	135180978	135240752	11	10	NP
Myo5b	18	74613029	74796289	18	17	Gain
No Gene 16	16	29019897	29031672	5	3	NP
No Gene 18	18	26169184	26282443	12	7	NP
No Gene 4	4	131170124	131238909	12	8	NP
No Gene Y	Y	2781406	2897989	16	15	NP
Nsd1	13	55352993	55372133	6	6	Gain
Pdcd6ip	9	113560446	113723693	14	14	Loss
Pde4dip	3	97593369	97718572	14	11	NP
Pigl	11	62203495	62369763	16	16	NP
Setd5 or Lhfpl4	6	113039133	113100176	9	9	NP
Sfi1	11	3004743	3179859	21	19	Loss
Snx24	18	53440600	53638364	19	12	Loss
Srfbp1	18	52654288	52675270	6	5	NP
Stag1	9	100583003	100698850	12	12	Loss
Tmem132b	5	126077980	126153976	11	4	Loss
Wac	18	7855248	8046126	17	15	Loss
Zfp397	18	24110107	24157877	8	7	NP
Zfp609	9	65561432	65607064	8	8	Loss

*Genomic address based on National Center for Biotechnology Information Mouse genome Build 37.

†Number of nonredundant *SB* transposon insertions within the locus.

‡Number of independent tumors with an insertion within the locus.

§Predicted effect is based on an analysis of the location and orientation of SB transposon insertions in all tumors in a single CIS (see text for discussion). Gain, gain of function; Loss, loss of function; NP, no prediction.

Table 13. List of 33 CIS.

Transposon Insertions Implicated in LOH of the Wild-Type Allele of *Apc*. The most commonly mutated gene in this study was *Apc* (in 72 of 96 tumors), reflecting the strong selective pressure for loss of the wild-type allele in Apc^{Min} mice. Previous studies have demonstrated that loss of the Apc^{+} allele is an early event that occurs in almost every adenoma in Apc^{Min} mice (Levy *et al.* 1994, Luongo *et al.* 1994). In addition, inactivation of the wild-type *Apc* allele is caused predominantly by homologous somatic recombination events, leading to the replacement of the *Apc* wild-type allele with a second Apc^{Min} allele. We reasoned that in our transgenic model LOH could be accomplished by an inactivating transposon insertion, as opposed to duplication of the *Min* allele. To test this hypothesis, we performed PCR on DNA from tumors to amplify the region surrounding the *Min* mutation (T2860A). By sequencing the PCR amplicon, LOH can be ascertained in Apc^{Min} mice by measuring the ratio of the T:A trace peak heights at the location of the *Min* mutation. In heterozygous tissue the T:A ratio is between 0.8 and 1.2, which is considered MOH, but in tissue that has lost the wildtype allele the ratio drops below 0.5 (***Figure 40***). Ratios between 0.5– 0.8 and >1.2 are considered uninformative, most likely caused by contamination from nontumor tissue. Of the 96 tumors tested, 47 gave informative results (***Table 14***). Of these 47 tumors, 32 had an identified transposon insertion in the *Apc* locus, and 15 did not. The majority (73%) of tumors lacking a transposon insertion in *Apc* had T:A ratios < 0.5, indicating LOH probably caused by loss of the entire allele. In support of our hypothesis, 53% of the tumors that had a transposon insertion in the *Apc* locus had T:A ratios between 0.8 and 1.2, indicating maintenance of the wild-type *Apc* locus at the site of the *Min* mutation. This result suggests that in these tumors the wild-type *Apc* allele is inactivated by the transposon more frequently than by duplication of the *Min* allele.

Set of CIS Identified in Apc^{Min} Mice Differs Significantly from Those Found in *Apc* Wild-Type Mice. We compared the list of genes identified in this study with the 77 genes identified in the screen we performed on an *Apc* wild-type background (Starr *et al.* 2009). Surprisingly, only four genes were identified in both studies: *Apc*, nuclear receptor binding SET domain protein 1 (*Nsd1*), Sfi1 homolog, spindle assembly associated (*Sfi1*),

and WW domain containing adaptor with coiled-coil (*Wac*). There are several reasons that could explain why the overlap between the two studies was low. First, the total number of genes that could contribute to tumor formation may be large enough that the size of these two studies is not sufficient to saturate the candidate genes. Second, because we use a statistical method to identify cancer genes, it is likely that transposon insertions contributed to carcinogenesis in some of the tumors, but the insertions did not occur at a rate high enough to qualify as a CIS. In support of this hypothesis, 70% of the loci identified as CIS in this study (23 of 33) also had one or more insertions in the same locus in the previous study. Third, the overlap may be small because selection pressure for specific genetic mutations in cells that already have an *Apc* mutation is biased toward a different set of cancer genes than in cells with a different initial mutation. Fourth, because of technical limitations, our method of amplifying and sequencing transposon insertions does not identify all transposon insertions, so a portion of driver mutations will not be identified. For example, analysis of replicate sequencing runs indicates that 20–40% of the PCR amplicons in a given library are not sequenced in a given GS FLX sequencing run (SI Materials and Methods and ***Table 15***).

Relevance to Human Disease. To determine the relevance of these findings to human cancers, we analyzed the regions of human orthology to the CIS loci and the orthologous human genes. Of the 30 candidate mouse genes associated with a CIS, 28 had human orthologs. We queried the literature for mutations and recurrent copy number changes in these genes in human CRC. Three of the genes, *APC, NSD1*, and phosphodiesterase 4D interacting protein (*PDE4DIP*), are considered bona fide cancer genes based on the cancer gene census maintained by the Wellcome Trust Sanger Institute (Futreal *et al.* 2004). Eight genes, *APC*, activating transcription factor-2 (*ATF2*), atlastin GTPase 2 (*ATL2*), casein kinase 1, alpha 1 (*CSNK1A1*), integrin, alpha M (*ITGAM*), programmed cell death 6-interacting protein (*PDCD6IP*), and *WAC*, have documented mutations in human cancers cataloged in the COSMIC database (Forbes *et al.* 2006).

We found strong concordance between the CIS mouse loci and orthologous regions in the human genome showing recurrent chromosomal losses and gains in human

CRC (Vogelstein *et al.* 1989, Reid *et al.* 1996, Nakao *et al.* 2004, Habermann *et al.* 2007, Lassmann *et al.* 2007, Shih *et al.* 2001, Derks *et al.* 2008, Kurashina *et al.* 2008). Of the 33 identified CIS, 31 can be mapped to an orthologous human locus. Of these 31 candidate cancer loci, 24 are found in regions that commonly are lost or gained in human CRC (***Table 16***), including several of the CIS that are orthologous to human chromosomal arms 18q, 17p, 5q, and 4q. Interestingly, one CIS that has no annotated genes nearby (CIS No gene 16) is in an orthologous region (3q21–24) that is associated with CRC based on genome-wide linkage analyses (Kemp *et al.* 2006, Neklason *et al.* 2008). To determine the significance of this overlap, we performed the analysis using randomly generated CIS lists and a single dataset of regions that are lost recurrently in human CRC (Nakao *et al.* 2004). Roughly 250 genomic regions in this dataset were lost in > 5% of the human samples tested, and 22 of the 31 CIS were located within these regions. In 10,000 simulations using equivalent-sized randomly generated CIS lists, we find an overlap of this magnitude < 0.3% of the time. These results suggest that our SB screen may be capable of pinpointing the affected genes in these regions.

Candidate Genes Regulate Proliferation of Human CRC Cell Lines. We tested nine genes [CCR4-NOT transcription complex, subunit 1 (*CNOT1*), *PDE4DIP*, *PDCD6IP*, *ATF2, SFI1*, *formin-binding protein 1-like (FNBP1L*), *myosin VB (MYO5B), sorting nexin 24 (SNX24)*, and *stromal antigen 1 (STAG1)*] for their effect on proliferation of the human CRC cell line SW480 by knocking down message levels using siRNA. We used the SW480 line because it has an *APC* gene-truncation mutation similar to the Apc^{Min} mutation (Nishisho *et al.* 1991). Cells were transfected two times at 48-h intervals with siRNA targeting the human genes. Knockdown efficiency was at least 50% for all nine genes as measured by quantitative real-time PCR. Cell proliferation was measured using a tetrazolium-based colorimetric assay on days two and six after the second transfection. Depletion of five (*CNOT1, PDE4DIP, PDCD6IP, ATF2*, and *SFI1*) of the nine genes tested resulted in a significant decrease in cell viability compared with a control siRNA of at least 33% at day six after transfection (***Table 17***).

Discussion

Using a transposon-based forward genetic screen in mice, we identified 33 genomic loci that probably cooperate with a germline mutation in the *Apc* gene to cause intestinal tumorigenesis. The most frequently mutated locus was the *Apc* locus, a result that supports the hypothesis that there is strong selective pressure to lose the wild-type copy during tumor formation. SB insertions in *Apc* were found in 72 of 96 tumors (75%), suggesting that 75% of the tumors in Apc^{Min} SB test mice undergo LOH at the *Apc* locus via SB insertional mutagenesis. This hypothesis is supported by sequencing of the region spanning the 1-bp *Min* mutation, which indicated that the majority of tumors containing SB insertions at the *Apc* locus maintained heterozygosity at the location of the T:A *Min* mutation. In contrast, the majority of tumors lacking a transposon insertion in *Apc* showed loss of the wild-type sequence at the *Min* mutation site. These results suggest that the majority of tumors underwent biallelic loss of *Apc* activity through transposon insertion or somatic recombination. However, it also is possible that in some cases activity was lost through other mechanisms of *Apc* inactivation or through transposon insertion substituting for loss of the wildtype allele.

Although LOH at the *Apc* locus is the rate-limiting event in tumor initiation in the Apc^{Min} mouse and in familial and sporadic *APC*-deficient CRC, loss of *APC* probably is insufficient for the survival and growth of transformed cells into adenomas and, eventually, adenocarcinomas. SB-mutagenized animals showed increased polyp number but no evidence of adenocarcinoma or metastasis. Thus, it is likely that the CIS candidate genes identified in this SB screen contribute in a diverse fashion to initiation, establishment, and survival of adenomas. Moreover, depending on the complexity of the mechanism or pathway, the CIS candidates discovered in our screen, like the relatively large number of genes reported to be mutant in individual human CRC (Wood *et al.* 2007), might be expected to occur at a low frequency if mutations at any one of multiple genes in a complex pathway can contribute equally to tumorigenesis.

Aside from *Apc*, only a few of the remaining 28 known genes that we identified as CIS in our screen have been implicated directly in CRC development, although > 90% are located in genomic regions that are lost or gained in human CRC. To eliminate false positives, we removed CIS that also were identified in a control dataset of unselected transposon insertions mapped in tail snips. However, it is possible that other tissuespecific hotspots could result in false positives. Nevertheless, the known functions of several of the CIS candidate genes make them plausible candidates for drivers of human CRC. For example, *CNOT1* is a member of the Ccr4-Not complex, which is implicated in mRNA decay and transcriptional repression. In human cells, CNOT1 has been reported to be a repressor of nuclear receptor-mediated transcription (Winkler *et al.* 2006). One target of CNOT1 repression appears to be estrogen receptor alpha (ERα), via interactions between CNOT1 and the ligand-binding domain of ERα. Inhibition of CNOT1 caused an increase in the expression of ERα target genes in breast cancer cells, and *ERα* has been shown to be a tumor suppressor gene in the intestinal tract (Cho *et al.* 2007); in particular, knockout of *ERα* in Apc^{Min} mice caused a significant increase in intestinal tumorigenesis (Cleveland *et al.* 2009).

Another gene identified in this study, *Pdcd6ip* (also known as ALG-2 interacting protein X, Alix) is involved in membrane trafficking and apoptosis (Odorizzi G 2006). *Pdcd6ip* produces a protein that binds to the protein product of *Pdcd6*, a proapoptotic gene involved in T-cell receptor–, Fas-, and glucocorticoid-induced cell death (Vito *et al.* 1996). *Pdcd6ip* also can block down-regulation of the EGF receptor (EGFR), thereby having a positive effect on growth factor signaling (Schmidt *et al.* 2004). These contradictory roles could explain why loss of *Pdcd6ip* in the mouse tumors promoted growth (via the loss of the proapoptotic function) but the loss of *Pdcd6ip* in SW480 cells caused decreased proliferation (via increased down-regulation of EGFR). Further functional studies are required to elucidate the role of this adaptor protein.

In summary, our approach identified 30 genes that probably modify tumorigenesis in the Apc^{Min} model of human CRC. Further functional analysis of these CIS candidate

genes may provide insights into the etiology and treatment of human CRC, especially those cancers arising downstream of *APC* deficiency.

Materials and Methods (Detailed protocols are given in SI Materials and Methods.)

Mice. Mice containing Apc^{Min}, Rosa26-LsL-SB11, Villin-Cre, and T2/Onc were reared using Institutional Animal Care and Use Committee-approved protocols. All mice were on an isogenic C57BL/6J background. Mice were monitored daily and killed and necropsied when moribund or after 120 d.

Histopathology and Immunohistochemistry. Formalin-fixed tissues were embedded in paraffin, and standard techniques were used to stain tissue sections with H&E. Standard immunohistochemistry techniques were used to detect β-catenin.

Linker-Mediated PCR. Linkers [described previously (Wu *et al.* 2003)] were ligated to NlaIII- (right-side) or BfaI- (left side) digested genomic DNA using T4 DNA ligase. A secondary digest (XhoI, right side; BamHI, left side) was performed to destroy concatamer-generated products. Primary and secondary PCR was performed using primers specific for linker and SB transposon sequences along with Fusion and barcode sequences. PCR amplicons were sequenced using the GS FLX (Roche).

Sequence Analysis. Sequences were analyzed for the presence of the barcode, inverted repeat/direct repeat (IR/DR) sequences required for transposition, and linker sequences. Genomic sequence was blasted against the mouse genome using BLASTN at 95% stringency and requiring a single match. Of the 324,898 sequences analyzed, 53% could be uniquely mapped to the mouse genome. Sequences were removed if they were redundant, on the donor concatamer resident chromosome (Chr 1), in the En2 gene (because the En2 sequence is present in the transposon), and when a single TA dinucleotide contained multiple insertions from several tumors from multiple mice (because of the possibility of a PCR artifact). The remaining 30,088 nonredundant

sequences were used to identify CIS. A CIS was defined by Monte Carlo simulations using a random dataset of 30,088 insertions.

***Apc* LOH Analysis**. To measure LOH for the Apc^{Min} mutation, DNA was isolated from individual polyps, and PCR was performed using primers that flank the mutation (sense primer: CGGAGTAAGCAGAGACACAA; antisense primer: GGGAGGTATGAATGGCTGAT). The PCR product was purified using Qiagen 96 MinElute vacuum purification plates per the manufacturer's protocol and was sequenced using the sense primer as the sequencing primer. Trace peak heights at the location of the mutation were measured for each tumor, and the ratio of the T peak to the A peak was calculated.

Comparisons with Human Data. Eight publicly available studies measuring DNA copy number in CRC compared with normal tissue were analyzed. Mutations in human tumors were examined using the Catalog of Somatic Mutations in Cancer database (Forbes *et al.* 2006), and cancer gene status was based on the Census of Human Cancer Genes maintained by the Wellcome Trust Sanger Institute (Futreal *et al.* 2004).

Knockdown of CIS Candidate Genes using siRNA in SW480 Cells. SW480 cells were obtained from ATCC (catalog no. CCL-228) and were cultured under recommended conditions. Transient siRNA transfection was used to deplete expression of *Min* CIS genes.

Cell Viability Assay. Viability of siRNA-treated SW480 cells was determined using a 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay (Cell Viability Kit 1; Roche Applied Sciences).

Supporting Information

SI Materials and Methods

Mice. *Apc*^{Min} mice (C57BL/6J-*Apc*^{Min} /J), which harbor a T→A nonsense mutation in the *Apc* gene that results in a truncated protein product, were obtained from a breeding colony at the University of Minnesota Medical School Animal Services facility. Rosa26-LsL-SB11 mice (backcrossed to C57BL/6J) were a generous gift from Adam Dupuy (University of Iowa, Iowa City, IA). Villin-Cre mice [B6.D2-Tg(Vil-Cre)20Syr], strain 01XE7, were purchased from the National Cancer Institute Mouse Repository. T2/Onc mice (mixture of C57BL/6J and FVB) were described previously (Collier *et al.* 2005). Mice were necropsied, and both normal and tumor tissues were collected by snap-freezing in liquid nitrogen or overnight fixation in 10% buffered formalin followed by 70% ethanol. PCR primer sequences were as follows: T2/Onc forward: CGCTTCTCGCTTCTGTTCGC, T2/Onc reverse: CCACCCCCAGCATTCTAGTT; Villin-Cre forward: CAAGCCTGGCTCGACGGCC, Villin-Cre reverse: CGCGAACATCTTCAGGTTCT; Rosa26-lox-stop-lox-SB11 and Rosa26-SB11 knock-in3- primer: wild-type forward: CTGTTTTGGAGGCAGGAA, wildtype reverse: CCCCAGATGACTACCTATCCTCCC, knock-in reverse: CTAAAAGGCCTATCACAAAC. *Apc*^{Min} mice were genotyped as described previously (Dietrich *et al.* 1993).

Histopathology and Immunohistochemistry. Histopathological analysis of tumors and adjoining normal tissue was performed on tissues that were fixed in 10% neutral buffered formalin, routinely processed into paraffin, sectioned at a thickness of 4μm, and stained with H&E. Multiple H&E sections were obtained from tumors from the colon and duodenum. All tissues were analyzed by an American College of Veterinary Pathologists-certified veterinary pathologist (M.G.O.) from the University of Minnesota Masonic Cancer Comparative Pathology Shared Resources facility and using the standardized nomenclature of the 2003 Consensus Report and Recommendations for pathology of mouse models of intestinal cancer (Boivin *et al.* 2003). Immunohistochemistry was

performed on 4-μm formalin-fixed, paraffin-embedded sections of small intestine which were deparaffinized and rehydrated, followed by antigen retrieval using 10 mM citrate buffer, pH 6.0, in a steamer. Staining for β-catenin was performed on a Dako Autostainer using a goat anti-human β-catenin polyclonal antibody (catalog no. sc-1496; Santa Cruz) as primary antibody (after blocking endogenous peroxidase and application of a protein block), with detection by a biotinylated donkey anti-goat antibody (Jackson ImmunoResearch Laboratories) and streptavidin-linked horseradish peroxidase (Dako) using diaminobenzidine (Dako) as the chromogen. Mayer's hematoxylin (Dako) was used as the counterstain. Small intestinal adenomas from $Apc^{Min/+}$ mice were used as a positive control tissue, and for negative control slides the primary antibody was substituted with Super Sensitive Goat Negative Serum (Biogenex).

Linker-Mediated PCR. Linkers used to sequence insertions were described previously (Wu *et al*. 2003). Genomic DNA was digested with NlaIII (for sequencing from the right side of T2/Onc) or BfaI (for sequencing from the left side of T2/Onc) and ligated to the linker using T4 DNA ligase. A secondary digestion was performed to destroy concatamer-generated products (XhoI for right-side cloning and BamHI for left-side cloning). These two enzymes do not cut the transposon distal to the NlaIII or BfaI sites but do cut the plasmid backbone present in the transgene concatamer, effectively destroying concatamer linker amplicons. Primary PCR was performed using primers that flank the inverted repeat/di direct repeat (IR/DR) sequences and the linker. Primer sequences are listed below. Primary PCR products were diluted 1:75 and used in a secondary PCR with nested primers. Secondary PCR was performed using FusA+BC+Left2°Primer or FusA+BC+ Right2°Primer and FusB+linker2°Primer primers (see below). FusA and FusB are sequences required for pyrosequencing using the 454 GSFlex machine (BC, barcode). PCR products were quantified using QuantIT picogreen assay (Invitrogen) and diluted to a concentration of 200,000 molecules/μL. All tumor samples were combined and diluted by the number of samples added, for a final concentration of 200,000 molecules/μL. Samples were sequenced using the Roche Genome Sequencer FLX using 454 pyrosequencing technology (Roche Applied Science)

by the University of Minnesota Biomedical Genomics Center. Sequences (all sequences are listed in the 5′ → 3′ direction) used were

Left linker+: GTAATACGACTCACTATAGGGCTCCGCTTAAGGGAC

Left linker−: TAGTCCCTTAAGCGGAG

Right linker+: GTAATACGACTCACTATAGGGCTCCGCTTAAGGGACCATG

Right linker−: GTCCCTTAAGCGGAGCC

Linker 1°Primer: GTAATACGACTCACTATAGGGC

Linker 2°Primer: AGGGCTCCGCTTAAGGGAC

(Note, linker primers work for both the Left linker and the Right linker)

Left 1°Primer: CTGGAATTTTCCAAGCTGTTTAAAGGCACAGTCAAC

Left 2°Primer: GGACATCTACTTTGTGCATGACACAAGTC

Right 1°Primer: GCTTGTGGAAGGCTACTCGAAATGTTTGACCC

Right 2°Primer: CCACTGGGAATGTGATGAAAGAAATAAAAGC

FusA+BC+Left2°Primer:
GCCTCCCTCGCGCCATCAGAATGCCGCATTTAAGTGTATGTAAACTTC

Fusion A: GCCTCCCTCGCGCCATCAG

Example barcode (each tumor is different): AATGCCGCAT.

FusA+BC+Right2°Primer:
GCCTCCCTCGCGCCATCAGAATGCCGCATTAAGGTGTATGTAAACTTC

FusB+linker2°Primer: GCCTTGCCAGCCCGCTCAGAGGGCTCCGCTTAAGGGAC

Fusion B: GCCTTGCCAGCCCGCTCAG

A full list of barcodes is available upon request.

Processing of Sequence Files. PCR amplicons were generated so that the 10-bp library-identifying barcode always appeared in the beginning of the sequence in the sense orientation. Sequence quality was outstanding even up to the first base. Using a custom Perl script, we scanned positions 1–12 of all reads for the presence of the library barcode, allowing 0 or 1 mismatch. We typically found perfect matches to a single barcode at positions 1–11, with matches at 2–12 occurring rarely. We did not find barcode sequences (0–1 mismatch) anywhere else in the read sequences. We successfully assigned 98% of all sequence reads to a library barcode. All barcodes differed by at least 2 bp, and no sequence read matched two or more barcodes in a given region. This process was carried out separately for each of the six 454 pyrosequencing regions used in these analyses. Following barcode identification, the data for each of the six runs were merged to allow uniform handling of the entire dataset. Note that we did not attempt to assemble reads into contigs for several reasons: (i) The read quality was outstanding, matching the consensus with >99.9% accuracy when assembly was performed; (ii) the contigs tended not to tile at all, because the reads all were primed at the same location and are of similar length, conferring little advantage to using contigs; and (iii) assembly might introduce chimeric artifacts, particularly when two relatively closely spaced insertion sequences appeared on opposite strands.

To identify and remove IR/DR and linker sequences from each read, we applied EMBOSS Vectorstrip (Rice *et al.* 2000) with custom-designed modifications for pipeline application and assessed the best mismatch parameters to use. We sequentially attempted to match both construct elements (IR/DR and linker) in sense and antisense orientations with four successively less stringent parameter sets:

i) 10% mismatch allowed, long-construct elements (17–32 bp)

ii) 10% mismatch allowed, short-construct elements (<26 bp)

iii) 15% mismatch allowed, short-construct elements (<26 bp)

iv) 20% mismatch allowed, short-construct elements (<26 bp)

Note that the short-construct elements affect only the IR/ DR sequence, because the linker element already is shorter than 26 bp. As we reduced the stringency, more construct elements could be detected, but the risk of finding spurious chance matches increased. To guard against the introduction of spurious matches, we used generic scripts to assign a label to each recognized construct that was encountered from the 5′end to the 3′ end. We assigned the following labels: Ideal: a construct with both IR/DR and linker elements in the same orientation

No-linker: a construct with an IR/DR but missing a linker

No-IR/DR: a construct missing the IR/DR

Bad: a construct with no recognizable elements or multiple IR/DRs

Unknown: anything else

For example, an ideal construct might look like 11-[+IR/DR]- 42-[+linker]-4, where, in this case, the 42-bp insertion is in the sense orientation (as indicated by the + signs for IR/DR and linker). The numbers between elements in this representation indicate the number of base pairs between elements. Insertions < 16 bp were considered empty or unmappable. Ideal and nearly ideal construct configuration counts were monitored as the stringency parameters were relaxed. We expected the number of ideal configurations detected to increase and the unknown count to remain steady until we hit a parameter set that was too lax. We started with the strictest set of parameters (stringency level 0) and collected all ideal configurations. Then we took all sequences with nonideal configurations through stringency levels 1–3 to see if we could move them into the ideal category. In our final merged summary sequence table, the best status (ideal > no-linker > no-IR/DR > bad) was reported for each sequence along with the data for the most stringent run level (0–3) where it achieved this label.

Mapping Insertion Sequences. To map sequences to the mouse genome [National Center for Biotechnology Information (NCBI) Build 37], we used BLASTN (DeCypher's TeraBLASTN, Active Motif, http://timelogic.com), requiring query sequences to align within 1 bp of the start of right-IR/DR sequences or within 1 bp of the end of left-IR/DR sequences (i.e., within 1 bp of the transposon insertion site with both types of reads). Additionally, the query was required to match with at least 95% identity. Lower thresholds of 90% and 85% identity were tested but failed to yield sufficiently higher percentages of newly mappable insertions to warrant lowering the matching stringency. Because we were most interested in the IR/DR position, we were careful to ensure that the query matched within 1 bp of the IR/DR insertion site, but we did not require the 3′ end of the query to match, in case cloning artifacts had altered that end of the sequence. If secondary genome hits were found that were at least 95% as long as the first match, their count was recorded, and the insertion location was considered ambiguous. However, if all secondary hits appeared within 5,000 bp of the primary hit on the same chromosome, we considered the insertion to be uniquely mappable to that locus. Of the 324,898 sequences analyzed by BLASTN, 173,101 (53%) could be uniquely mapped to the mouse genome. We removed redundant sequences that arose from the same tumor and mapped to the same TA dinucleotide insertion site in the genome. Of the 173,101 mapped sequences, 100,171 (67%) were redundant, leaving 72,930 nonredundant mapped insertions. Three more filtering steps were performed:

i) To avoid the bias of "local hopping," all nonredundant insertions mapping to the chromosome containing the transposon donor concatamer (Chr 1) were removed.

ii) The T2/Onc transposon contains sequence from an intron and splice acceptor of the murine En2 gene. Any insertions mapping to this sequence were removed because they might represent a transposition event back into the concatamer and not an insertion into the genomic En2 locus.

iii) Three-primer PCR was used to validate a sampling of nonredundant mapped insertions. We were successful in validating mapped insertions except in one

circumstance. When a single TA dinucleotide contained multiple insertions from several tumors from multiple mice, we could not validate the insertions with three-primer PCR in some cases. We hypothesize that some of these TA dinucleotides are neasequences that cause PCR artifacts and do not represent true transposon insertions. We chose to adopt a conservative approach to avoid these artifacts, so we eliminated all insertions in TA dinucleotides containing insertions from two or more tumors from two or more different mice. These three filtering steps removed 42,842 (59%) insertions, leaving a total of 30,088 nonredundant mapped insertions. This set was used to determine CIS.

Statistics Used to Identify CIS. To identify CIS, nonredundant insertions were assigned to clusters if the local density of insertions in a given window size exceeded that which would be expected by chance. Window sizes were determined by exact Monte Carlo simulation (see below). Based on a dataset of 30,088 insertions, the significance thresholds obtained are:

Five or more insertions within 12,000 bp

Six or more insertions within 22,000 bp

Seven or more insertions within 34,000 bp

Eight or more insertions within 50,000 bp,

Nine or more insertions within 65,000 bp

Ten or more insertions within 82,000 bp

Eleven or more insertions within 105,000 bp

Twelve or more insertions within 124,000 bp

Thirteen or more insertions within 150,000 bp

Fourteen or more insertions within 175,000 bp

Fifteen or more insertions within 200,000 bp

The assumption of standard Poisson statistics that potential insertion sites are randomly distributed throughout the genome is not strictly correct, because (i) TA dinucleotides are naturally clustered in genomes, and (ii) numerous unfinished regions in the mouse genome are "off-limits" because they are long tracts of Ns. For example, the initial telomeric region of every chromosome except Y is padded with 3 million consecutive Ns. Both these factors lead standard analytical approaches to underestimate the size and number of clusters that actually would be encountered by simply picking randomly chosen real TA sites. In other words, by ignoring the natural clustering of TA sites in the genome, the number of false-positive CIS that will be predicted is increased systematically. The magnitude of deviation gets larger as more and more insertion sites are scattered about the genome, as one would expect intuitively. Hence, we wrote a program to compute exactly the expected number of CIS of a given size in a specified window across all the chromosomes that one would encounter by chance via Monte Carlo simulation. The observed number of unambiguous mappable nonredundant insertions was used for each chromosome separately as input. For example if chromosomes 1 and 2 had 2,100 and 1,420 insertions, respectively, then we randomly distributed 2,100 insertions among the real TA dinucleotide sites on mouse chromosome 1 and another 1,420 among the TA sites of chromosome 2. Once the total count of insertions was distributed randomly among the real TA sites across the whole genome, a tally of the number of CIS of size $\geq 3, \geq 4, \ldots, \geq 15$ was recorded within windows of 10,000 bp, 20,000 bp, . . . 150,000 bp. This process was repeated 100 times, and the average counts over those 100 iterations were computed. Four independent simulations of 100 iterations each were performed, yielding SE bars between simulations of $< 1\%$, indicating sufficient convergence. The values obtained can be interpreted as expected values (E-values), because they indicate the expected number of CIS of a given number of insertions that would be observed within a given window size merely by chance. We chose a threshold with an E-value < 1. Thus, finding 11 CIS of ≥ 15 insertions within 200,000 bp, when not even a single CIS was expected, is highly significant. We compared the thresholds obtained by this method with the thresholds obtained using standard Poisson statistics with the assumption of random insertion in the genome. We found our method was

uniformly more stringent and yielded fewer false positives than the standard Poisson statistics. A modified Poisson model that takes into account the local density of TA sites in the genome yielded excellent agreement with the Monte Carlo calculations.

Annotation and Sequence Information Management.We created two primary annotation files: one outlining details of each unique insertion, and one describing each CIS. These files provide information on the chromosomal mapping position of each insertion or CIS, redundancy information on each insertion, and characteristics of the nearest EnsEMBL gene that flanks the insertion. EnsEMBL mappings were identified by a custom Perl script that uses the published Application Programmer Interface (Stabenau *et al.* 2004). To facilitate the management of all sequence information, an MySQL relational database was constructed to store (i) genotypic and phenotypic information on all mice and tumors from which the insertion sequences were derived; (ii) metainformation on the sequencing runs themselves; (iii) raw read sequences; (iv) construct element matching characteristics; (v) final processed insertion sequences; (vi) mapping information for each processed insert sequence to the mouse genome; (vii) clustering assignments of inserts into CIS; and (viii) annotation information on all mapped inserts and CIS. SQL queries were performed to facilitate the merging of distinct gastrointestinal tract tumor datasets and the annotation process.

Analysis of Replicate Sequencing Runs to Determine Percentage of Library Capture. To estimate the extent of undersampling of transposon insertions in our study, we analyzed the GS FLX sequencing replicates separately. Using a four-region plate, one sequencing run of the GS FLX machine can sequence four separate samples. In one of the sequencing runs we ran two aliquots of the left-side ligation-mediated (LM)-PCR pool in separate regions (1 and 4) and two aliquots of the right-side LMPCR in separate regions (2 and 3). We analyzed the sequence reads by custom Perl script to determine the extent of overlap (***Table 15***). For example, region 1 returned 68,371 total reads, of which 56,435 contained a perfect match to a barcode, the transposon-specific sequence, and a genomic TA dinucleotide along with at least 16 bases of genomic DNA. When duplicate sequences were combined, 20,654 unique transposon insertion reads remained. Region 4,

which was a replicate, had 18,863 unique reads. The overlap between these two regions was 10,448 reads, a little more than 50%. Because this process is similar to a mark-recapture experiment, one can use the Lincoln–Peterson method (Seber GAF 2002) to estimate the number of amplicons in the total population (see below). Based on the overlap between the replicates sampled in regions 1 and 4, we estimate that there were 37,289 unique amplicons in the original pool. As a rough approximation, our protocol sampled about 78% of the amplicons present in the original left-side LM-PCR pool. If we apply the same analysis to the right-side LM-PCR pool, where the overlap was lower (~35%), we estimate that we sampled only 58% of the amplicons in the original right-side LM-PCR pool. In either case, it is apparent that by increasing the number of sequencing runs, or perhaps by using a different sequencing platform, we could find more transposon insertions and perhaps more candidate cancer genes.

Lincoln–Peterson Method to Estimate Total Population in a Mark-and- Recapture Experiment. A mark-and-recapture experiment can be used to estimate population size in an ecological setting where the researcher marks all animals captured during a first visit. The researcher then returns and makes a second capture, noting how many of the animals in the second capture were marked in the first capture. To use the Lincoln–Peterson method (Seber GAF 2002) in our study, we assume that the duplicate pools of amplicons contain equal numbers of the same amplicons. This assumption could be invalid, because the concentration in the amplicon pool is very small (~200,000 molecules/μL). Randomly removing a small volume could result in the two volumes containing different amplicons. Another caveat to using the Lincoln–Peterson method is that our use of unique reads instead of actual reads is not directly analogous to a mark-and-recapture experiment. Nevertheless, if we assume the aliquots are similar, then the number of reads in the first region (M) is analogous to the number of animals caught and marked in the first capture of a mark-andrecapture experiment. The number of reads in the second region (C) is analogous to the number of animals caught in the second capture, and the number that overlaps (R) is analogous to the number of animals marked in the first capture that are caught in the second capture. To estimate the total population (N),

the Lincoln–Peterson method states that the proportion of marked individuals in the second capture to the number in the first capture (R/M) should equal the proportion of the number of animals in the second capture to the total population (C/N). Rearrangement of this equation gives n = (M × C)/R. Using the number of unique reads in regions 1 and 4 (***Table 15***) equates to (20,654 × 18,863)/10,448 = 37,289. The total number of unique reads we captured in both sequencing regions (20,654 + 18,863 − 10,448 = 29,069) represents 78% of the estimated total population of 37,289. Using the same analysis of the unique reads in region 2 and 3 (***Table 15***) equates to an estimated total populationof (25,753 × 25,804)/9,079 = 73,194. In this case, we have sampled only (25,753 + 25,804 − 9,079 = 42,478) 58% of the total population.

Apc **Loss of Heterozygosity Analysis**. To measure loss of heterozygosity (LOH) for the Apc^{Min} mutation, DNA was isolated from individual polyps, and PCR was performed using primers that flank the mutation (sense primer: CGGAGTAAGCAGAGACACAA; antisense primer: GGGAGGTATGAATGGCTGAT). The PCR product was purified using Qiagen 96 MinElute vacuum purification plates according to the manufacturer's protocol and was sequenced using the sense primer as the sequencing primer. Trace peak heights at the location of the mutation were measured for each tumor. Wild-type DNA will only have a single peak, representing thymidine, whereas Apc^{Min} nontumor DNA will have two peaks of equal height representing the wild-type thymidine base and the mutant adenine base. If significant LOH has occurred, the ratio of the height of the thymidine peak to the adenine peak will be < 0.5, whereas a ratio > 0.8 indicates maintenance of heterozygosity. Ratios between 0.5 and 0.8 and ratios > 1.2 are considered to be contaminated with nontumor tissue.

Chromosomal Copy Number Analysis. Eight studies measuring DNA copy number in CRC compared with normal tissue were analyzed (Vogelstein *et al.* 1989, Ried *et al.* 1996, Nakao *et al.* 2004, Habermann *et al.* 2007, Lassmann *et al.* 2007, Shih *et al.* 2001, Derks *et al.* 2008, Kurashina *et al.* 2008). Of the 33 CIS loci, 31 were mapped to the homologous human region using the Batch Coordinate Conversion (Lift- Over) utility from the University of California, Santa Cruz (UCSC) genome browser (Kent *et al.*

2002). The genome build appropriate to each study was analyzed. A CIS locus was determined to be lost or gained recurrently based on the individual study methodology. To determine the chances of this overlap occurring randomly, we ran 10,000 simulations using randomly generatedCIS listscompared with the recurrently lost regions in Nakao *et al*. (Nakao *et al.* 2004). In this study they hybridized genomic DNA from 135 human CRC along with normal lymphocytic DNA to a BAC array (HumArray1.14, University of California San Francisco; UCSF) and used Spot/Sproc analysis software (UCSF) to determine log2 ratios. They published an Excel file with the log2 ratios for each BAC clone for each human sample. Chromosomal losses were based on a log2 ratio lower than −0.225. The genomic coordinates of the BAC clones were downloaded from the UCSF website (http://cancer.ucsf.edu/_docs/cores/array/analysis/HA1.14_clonepos_May04.20060811.txt) and were mapped to the Human genome build HG19 using the Batch Coordinate Conversion utility from the UCSC genome browser (http://genome.ucsc.edu/cgi-bin/hgLiftOver). After BAC clones that did not have at least 65% of samples giving a reading and positions that did not map to the HG19 build were removed, log2 ratios for 2,075 BAC clones remained. We selected the subset of these clones that had log2 ratios lower than −0.225 in at least 5% of the samples. Genomic regions recurrently lost were defined based on the genome coordinates of consecutive BAC clones that had log2ratios lower than −0.225. For example, if three adjacent clones all showed a loss, we assigned the region based on the start coordinate of the first clone and the end coordinate of the last clone. If the neighboring clones on both sides of a BAC were not decreased, the region was defined as the coordinates of the single BAC. Next, we wrote a Perl script that created 10,000 random sets of 31 CIS using the UCSC Golden Path coordinates. Each set of 31 CIS contained 31 randomly generated genomic regions that were the same length as the actual Apc^{Min} dataset CIS. The Perl script then counted the overlap between the random CIS list and the Nakao recurrently lost regions, and the process was repeated 10,000 times. In 10,000 simulations, only three random CIS lists had an overlap of 22 regions or more.

Knockdown of CIS Genes Using siRNA in SW480 Cells. SW480 cells were purchased from American Type Culture Collection. Cells were maintained in DMEM supplemented with 10% FBS, 2 mmol/L glutamine, 100 U/mL penicillin, and 100 μg/mL streptomycin and were incubated in a humidified atmosphere of 95% air and 5% CO_2 at 37 °C. Cells were switched to antibiotic-free medium before siRNA transfection. Cells at 70% confluence were transfected twice, 48 h apart, with siRNA oligonucleotides (siRNA oligos) targeting the human ortholog of the CIS candidate gene (CIS siRNA) or with a nontargeting control (control siRNA), at a final concentration of 50 nmol/L, using Lipofectamine 2000 transfection reagent. Oligos targeting CIS genes were obtained from Qiagen (Hs_*ATF2*_3, Hs_*PDCD6IP*_5 Hs_*SFI1*_7, and Hs_ *PDE4DIP*_15) and from Dharmacon (*CNOT1* On-TARGETplus SMARTPool). The control siRNA oligo (OnTARGETplus Nontargeting siRNA#3) was obtained from Dharmacon. Total RNA was harvested fromCIS siRNA- and control siRNA-treated cells 2 or 3 d after the second siRNA transfection using the RNeasy kit (Qiagen). For each experimental sample, 1.5 μg RNA was converted to cDNA with random nonomer primers (Integrated DNA Technologies) and recombinant Omniscript Reverse Transcriptase using the Omniscript RT kit (Qiagen) according to manufacturer's instructions. CIS cDNA and 18S ribosomal cDNA (used as internal control) were amplified from total cDNA by PCR. PCR was carried out in 10 μL using 2 μL cDNA from a 10× dilution of the 20 μL RT reaction (equivalent to 15 ng reversetranscribed RNA) for CIS genes and a 100× dilution for 18S rRNA, 500 nM of each primer, and 5 μL LightCycler 480 SYBR Green 1 Master Mix (Roche Applied Science). PCR was performed on a LightCycler 480 System (Roche Diagnostics) in 96-well plates using the amplification protocol: one cycle of preincubation, 5 m at 95°C; 45 cycles of amplification each consisting of denaturation at 95°C for 5 s, annealing at 60°C for 5 s, and elongation at 72°C for 10 s; one cycle melting at 95°C for 5 s, 65°C for 1 m, heating to 97°C; one cycle cooling at 40°C for 30 s. Water was used as a template for negative control amplifications for each PCR run. All reactions were performed in duplicate. Standards were generated by reverse transcription of total RNA from untreated cells followed by PCR amplification to generate template DNA of the same sequence as the predicted CIS gene or 18S gene PCR product. Serial

dilutions of template DNA were amplified in parallel with experimental samples and used to generate a standard curve for each gene. Data were analyzed using Roche LightCycler 480 software, and crossing points (CPs) were calculated using the absolute quantification-second derivative maximum method. The standard curve was used to determine efficiency of PCR amplification (E) for each gene. Relative mRNA levels represent the expression of the CIS gene in CIS siRNA-treated cells relative to expression in control siRNAtreated cells. CIS gene expression was normalized to 18S rRNA levels, and relative mRNA levels were calculated as described by Pfaffl (Pfaffl MW 2001): relative mRNA levels = ([Etarget] ΔCPtarget(control-treated))/[Eref] ΔCPref(control-treated)) × 100, where E is real-time PCR efficiency and CP is defined as the point at which fluorescence rises appreciably above background.

Primer sequences are as follows: activating transcription factor-2 (*ATF2*), 5′-TGACCGAAAGGATCATGAACTA-3′ and 5′-GCAGTCCTTTCTCAAGTTTCCA-3′; CCR4-NOT transcription complex, subunit 1 (*CNOT1*), 5′ CTTTCAACCCCCAATCAGACC-3′ and 5′-AGGTTTCATCTTACTCTGCTGGA-3′; programmed cell death 6-interacting protein (*PDCD6IP*), 5′ AGGTGTTCCCTGTCTTGGCTGC-3′ and 5′ TTCATCATAGCGAGATGCCACTGTTT-3′; phosphodiesterase 4D-interacting protein (*PDE4DIP*), 5′-GAGAACTCCAGGACAAGAAACAGCAT-3′ and 5′-GGATTCCTCCTGCAGAAGCTGG-3′; Sfi1 homolog, spindle assembly associated (*SFI1*), 5′-AGCAGCAGGAGATGAGGAACAAG-3′ and 5′-CGAACAACCACGTAGATCAACCAG-3′.

Cell Viability Assay. Viability of siRNA-treated SW480 cells was determined using a 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay (Cell Viability Kit 1; Roche Applied Sciences) in which absorbance at 595 nm is proportional to viable cell number. One day after the second siRNA transfection, cells were replated into 96-well plates at 1,250 cells per well (*ATF2, PDCD6IP*, and *PDE4IP*) or 5,000 cells per well (*CNOT1*) in triplicate in regular growth medium. Cell viability was determined on days 2–6 after the second transfection. Day 6 A595 values for each treatment were normalized

to day 2 values for the same treatment to correct for differences in plating. Relative cell viability represents the ratio of normalized absorbance at 595 nm of cells treated with CIS siRNA to absorbance of cells treated with control siRNA on day 6. Results shown for mRNA and cell viability are the mean ± SD of at least two experiments.

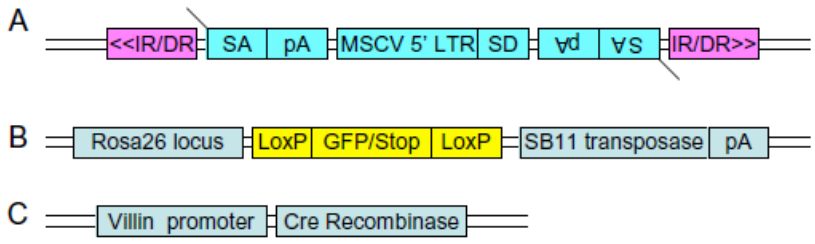


Figure 39. Three alleles used to target SB mutagenesis to the intestinal tract. (A) T2/Onc transposon. IR/DR, inverted repeat/direct repeat sequences required for transposition; MSCV 5′ LTR, murine stem cell virus 5′ long terminal repeat; SA pA, splice acceptor with a polyA signal. **(B)** Conditional SB allele. Rosa26 locus, endogenous Rosa26 locus [Gt(Rosa)26Sor]; LoxP-GFP/Stop-LoxP, GFP cDNA flanked by LoxP sites; SB11 transposase pA, SB11 transposase cDNA with polyA signal. **(C)** Cre recombinase cDNA driven by the Villin promoter.

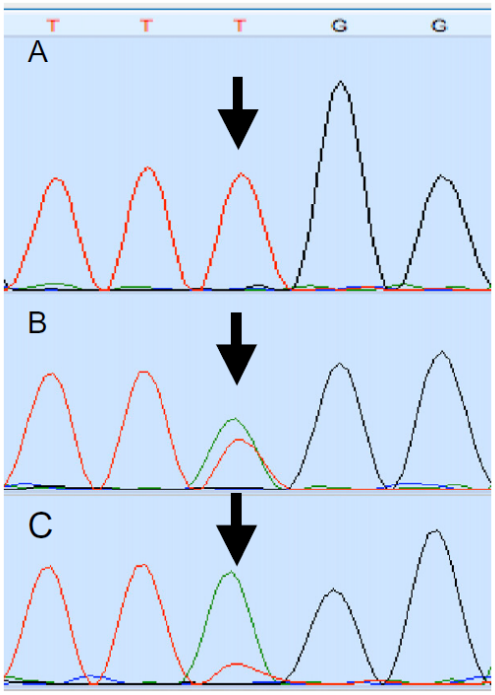


Figure 40. PCR technique for detecting *Apc* LOH. Trace peak heights of the T→A *Min* mutation in **(A)** wild-type mice, **(B)** heterozygous Apc^{Min} mice, and **(C)** an adenoma with LOH. Arrows indicate *Min* mutation. Red peaks = T; green peaks = A.

Table 14. LOH and MOH in *Apc*^{Min} tumors based on the ratio of T:A trace peaks.

Transposon insertion in *Apc*	Number of informative tumors*	% tumors with T:A ratio <0.5 (LOH)	% tumors with T:A ratio between 0.8–1.2 (MOH)
Yes[†]	32	47	53
No	15	73	27

MOH, maintenance of heterozygosity.
*To qualify as an informative tumor, the peak height ratio T:A must be <0.5 or between 0.8–1.2.
[†]Yes indicates the tumors had a mapped *SB* transposon insertion in the *Apc* gene.

Table 15. Sequence read overlap between duplicate regions of a single GS FLX sequencing run.

	Region 1 (left)[†]	Region 4 (left)[†]	Region 2 (right)[†]	Region 3 (right)[†]
Average read length (bases)	95	94	94	94
Total no. reads	68,371	59,228	54,405	113,982
No. reads with a perfect barcode[‡]	66,607	57,631	53,168	111,016
No. reads with IR/DR+TA match[§]	56,435	49,006	49,531	102,534
No. reads with unique 16 bases after TA[¶]	20,654	18,863	25,753	25,804
No. unique reads that overlap	10,448	10,448	9,079	9,079
Percent unique reads that overlap	51	55	35	35

*A four-region plate was used in the sequencing run.
[†]Left-side LM-PCR duplicate aliquots were run in regions 1 and 4; right-side duplicates were in regions 2 and 3.
[‡]The first 10 bases of the read perfectly matched one of the 96 barcodes.
[§]The bases immediately following the barcode perfectly matched the nested primer, the remainder of the IR/DR, and a TA dinucleotide.
[¶]The 16 bases immediately following the TA dinucleotide were unique compared with all other sequence reads.

Table 16. Human orthologous regions to the mouse CIS with recurrent chromosomal copy number changes based on published data.

Mouse address*	Mouse gene	Mouse Entrez ID	Human address†	Human gene	Human Entrez ID	Human band	In region of loss/gain‡
chr9:3001410–3030207	AC131780.5	114673?	Deleted in humans	No homology	N/A	N/A	N/A
chrY:2781406–2897989	No Gene Y	N/A	Deleted in humans	No homology	N/A	N/A	N/A
chr3:127241424–127431287	4930422G04Rik	71643	chr4:113299994–113576517	C4orf21	55345	4q25	Yes
chr3:127697338–127730035	AC115907.7	100043382	chr4:112926160–112964108	No Gene 3	N/A	4q25	Yes
chr13:105093822–105104050	Adamts6	108154	chr5:64747713–64758225	ADAMTS6	11174	5q12.3	Yes
chr3:127454370–127633377	Ap1ar	211556	chr4:113034449–113260728	AP1AR	55435	4q25	Yes
chr18:34324389–34514767	Apc	11789	chr5:111979623–112236985	APC	324	5q22.2	Yes
chr17:80239637–80306988	Atl2	56298	chr2:38501789–38613026	ATL2	64225	2p22.2–22.1	Yes
chr18:61726208–61750047	Csnk1a1	93687	chr5:148873515–148912809	CSNK1A1	1452	5q32	Yes
chr18:73903365–73913215	Elac1	114615	chr18:48495579–48509486	ELAC1	55520	18q21.2	Yes
chr3:136918543–137092521	Emcn	59308	chr4:101323972–101541428	EMCN	51705	4q24	Yes
chr18:10578932–10773953	Esco1	77805	chr18:19121974–19393499	ESCO1	114799	18q11.2	Yes
chr3:122217801–122319064	Fnbp1l	214459	chr1:93916974–94045997	FNBP1L	54874	1p22.1	Yes
chr7:135180978–135240752	Itgam	16409	chr16:31267174–31289751	ITGAM	3684	16p11.2	Yes
chr18:74613029–74796289	Myo5b	17919	chr18:47508944–47707289	MYO5B	4645	18q21.1	Yes
chr18:26169184–26282443	No Gene 18	N/A	chr18:35353815–35494000	No Gene 18	N/A	18q12.2	Yes
chr4:131170124–131238909	No Gene-4	N/A	chr1:29768033–29870348	No Gene 4	N/A	1p35.3	Yes
chr13:55352993–55372133	Nsd1	18193	chr5:176657455–176671302	NSD1	64324	5q35.3	Yes
chr3:97593369–97718572	Pde4dip	83679	chr1:144945932–145112522	PDE4DIP	9659	1q21.1	Yes
chr11:62203495–62369763	Pigl	327942	chr17:16040912–16299937	PIGL	9487	17p11.2	Yes
chr6:113039133–113100176	Setd5/Lhfpl4	72895/269788	chr3:9450610–9516832	SETD5 LHFPL4	55209 375323	3p25.3	Yes
chr11:3004743–3179859	Sfi1	78887	chr22:31790517–32022116	SFI1	9814	22q12.2	Yes
chr18:53440600–53638364	Snx24	69226	chr5:122221887–122440150	SNX24	28966	5q23.2	Yes
chr18:52654288–52675270	Srfbp1	67222	chr5:121370968–121397902	SRFBP1	153443	5q23.1	Yes
chr18:24110107–24157877	Zfp397	69256	chr18:32819631–32871403	ZNF397	84307	18q12.2	Yes
chr9:65561432–65607064	Zfp609	214812	chr15:64859323–64946983	ZNF609	23060	15q22.31	Yes
chr2:73708689–73773260	Atf2	11909	chr2:175997880–176076969	ATF2	1386	2q31.1	No
chr8:98254541–98300340	Cnot1	234594	chr16:58568279–58622838	CNOT1	23019	16q21	No
chr16:29019897–29031672	No Gene 16	N/A	chr3:192750830–192764409	No Gene 16	N/A	3q29	No
chr9:113560446–113723693	Pdcd6ip	18571	chr3:33684948–33911522	PDCD6IP	10015	3p22.3	No
chr9:100583003–100698850	Stag1	20842	chr3:136220568–136384331	STAG1	10274	3q22.3	No
chr5:126077980–126153976	Tmem132b	208151	chr12:125776600–125970181	TMEM132B	114795	12q24.31–24.32	No
chr18:7855248–8046126	Wac	225131	chr10:28820138–28913566	WAC	51322	10p12.1	No

*Based on mouse July 2007 Assembly (mm9) UCSC genome browser.

†Based on human GRCh37 Assembly (hg19) UCSC genome browser.

‡Region was identified in one of the following studies: Nakao *et al.* 2004; Derks *et al.* 2008; Lassmann *et al.* 2007; Habermann *et al.* 2007; Ried *et al.* 1996; Vogelstein *et al.* 1989; Shih *et al.* 2001.

Table 17. Knockdown of *Apc*^{Min} CIS candidate genes affects viability of human colon cancer cells.

Gene	Relative mRNA levels of cells depleted for CIS gene*	Relative cell viability of cells depleted for CIS gene†
ATF2	17.2 ± 5.6%	61.0 ± 1.5%
CNOT1	13.7 ± 7.4%	18.5 ± 16%
PDCD6IP	10.5 ± 2.6%	64.2 ± 28%
PDE4DIP	27.9 ± 10%	55.0 ± 19%
SFI1	40.4 ± 13.2%	60.0 ± 14%

*Relative mRNA levels calculated using the ΔΔCt method (1) and represent levels of CIS mRNA in CIS siRNA-treated cells relative to levels in control, nontargeting siRNA-treated cells. 18S RNA was used as a housekeeping control. Results shown are the mean ± SD of at least two experiments.

†Relative cell viability represents the ratio of absorbance at 595 nm of cells treated with the indicated CIS gene siRNA to absorbance of cells treated with nontargeting control siRNA. Results shown are the mean ± SD of at least two experiments.

Table 18. Mapped transposon insertions in 96 tumors. Please see the weblink: http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1018012108/-/DCSupplemental/sd01.xls

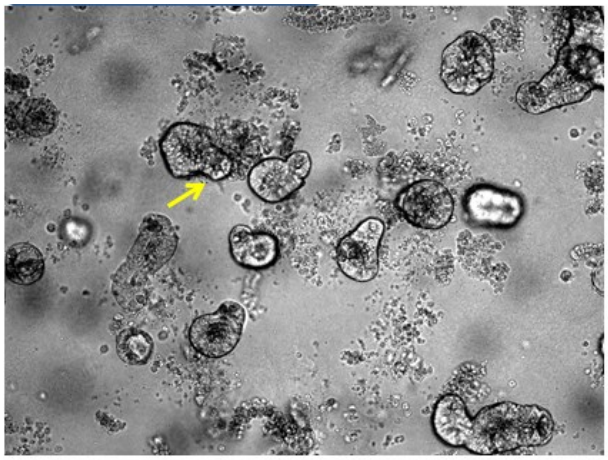
Personal Contribution. B.L.N.T performed MTT experiments analyzed data to determine the knockdown effect of the following genes on proliferation of human (SW480) CRC cells: *ATF2, MYO5B*, and *SNX24.*

APPENDIX 2

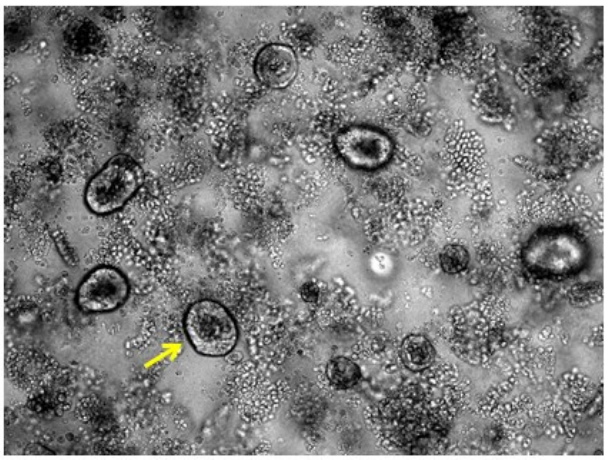
Loss of *Kcnq1* delayed organoid differentiation in the small intestine of Apc^{wt} mice

Unlike in the colon where there was an increase in organoid numbers resulting from the the loss of *Kcnq1*, in the small intestine we found no significant changes in the organoid numbers of the $Apc^{+/+}$ $Kcnq1^{-/-}$ group compared to that of the control $Apc^{+/+}$ $Kcnq1^{+/+}$ group. Interestingly, we observed a difference in morphology of organoids from the small intestine of $Apc^{+/+}$ $Kcnq1^{-/-}$ mice and those from the small intestine of the control $Apc^{+/+}$ $Kcnq1^{+/+}$ mice (***Figure 41***). On Day 2, while small intestinal organoids from the $Apc^{+/+}$ $Kcnq1^{+/+}$ control group had started to grow additional crypts, small intestinal organoids from the $Apc^{+/+}$ $Kcnq1^{-/-}$ group were still in their original sphere forms. On Day 4, organoids from the small intestine of $Apc^{+/+}$ $Kcnq1^{+/+}$ mice had reached full maturity with multiple crypts; on the other hand, organoids from the small intestine of $Apc^{+/+}$ $Kcnq1^{-/-}$ mice just now started to extend additional crypts. Then on Day 6, small intestinal organoids from the $Apc^{+/+}$ $Kcnq1^{+/+}$ control group underwent apoptosis, shredding multiple apoptotic cells and factors into the lumen, whereas small intestinal organoids from the $Apc^{+/+}$ $Kcnq1^{+/+}$ just reached full maturity with multiple crypts branching out. This indicates that loss of *Kcnq1* affects small intestinal differentiation, further suggesting a role of Kcnq1 in regulating the intestinal crypt.

A. Day 2

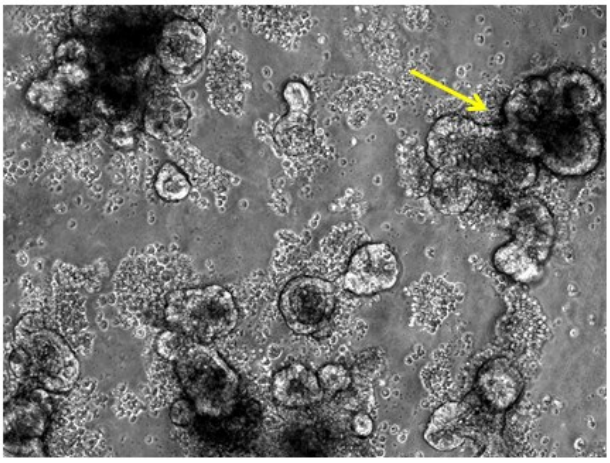


***Kcnq1*$^{+/+}$ (10x)**

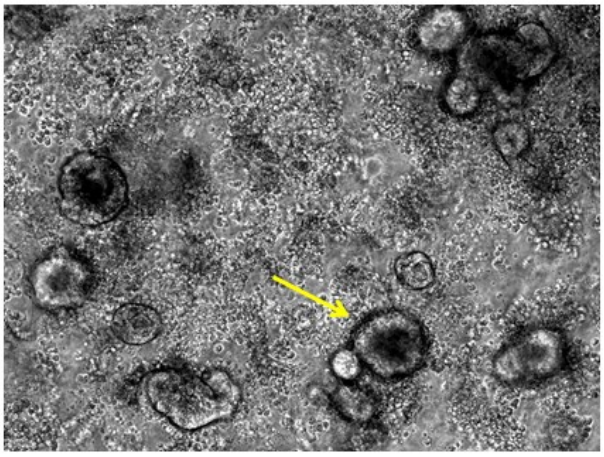


***Kcnq1*$^{-/-}$ (10x)**

B. Day 4



***Kcnq1*$^{+/+}$ (10x)**



***Kcnq1*$^{-/-}$ (10x)**

C. Day 6

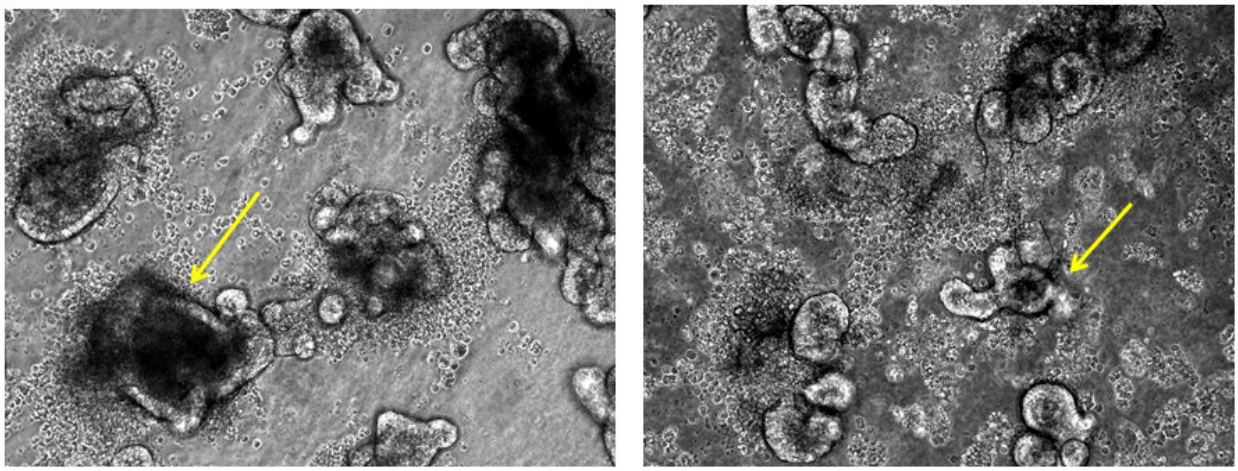


Figure 41. Loss of Kcnq1 affects differentiation in the small intestine. Small intestinal organoids were created from 500 crypt bottoms and plated in triplicate in 24 well plates. Crypt bottoms were isolated on separate days from 2 of each, age- and littermate-matched *Apc*$^{+/+}$ *Kcnq1*$^{+/+}$ and *Apc*$^{+/+}$ *Kcnq1*$^{-/-}$ male mice (one matched pair per day). Organoids were examined daily from Day 1 to Day 6 post plating. Pictures were taken on Day 2 (A), Day 4 (B) and Day 6 (C), at 10X magnification. Yellow arrow points to an example of an individual organoid.

Materials and Methods. Two pairs of age, and littermate-matched C57Bl/6J *Kcnq1*$^{+/+}$ and *Kcnq1*$^{-/-}$ male mice were sacrificed between 8 and 12 weeks of age. Small intestines were removed, cut open and washed in cold PBS. Villi were removed (by using a glass slide). Small intestine organoids were then cultured using the protocol of Sato *et al.* 2011, following the plating of 500 crypt bottoms per well in triplicate per sample. Organoids were monitored daily from Day 1 to Day 6 after plating. (Growth factors were added on Day 2 and 6, media were changed on Day 4). Pictures were taken on Day 2, 4 and 6 using Nikon Photometrics CoolSNAP ES microscope and processed by ImageJ program.

APPENDIX 3

The Cancer Genome Atlas (TCGA) Data on *KCNQ1* and *CFTR* Mutations in Human Cancers

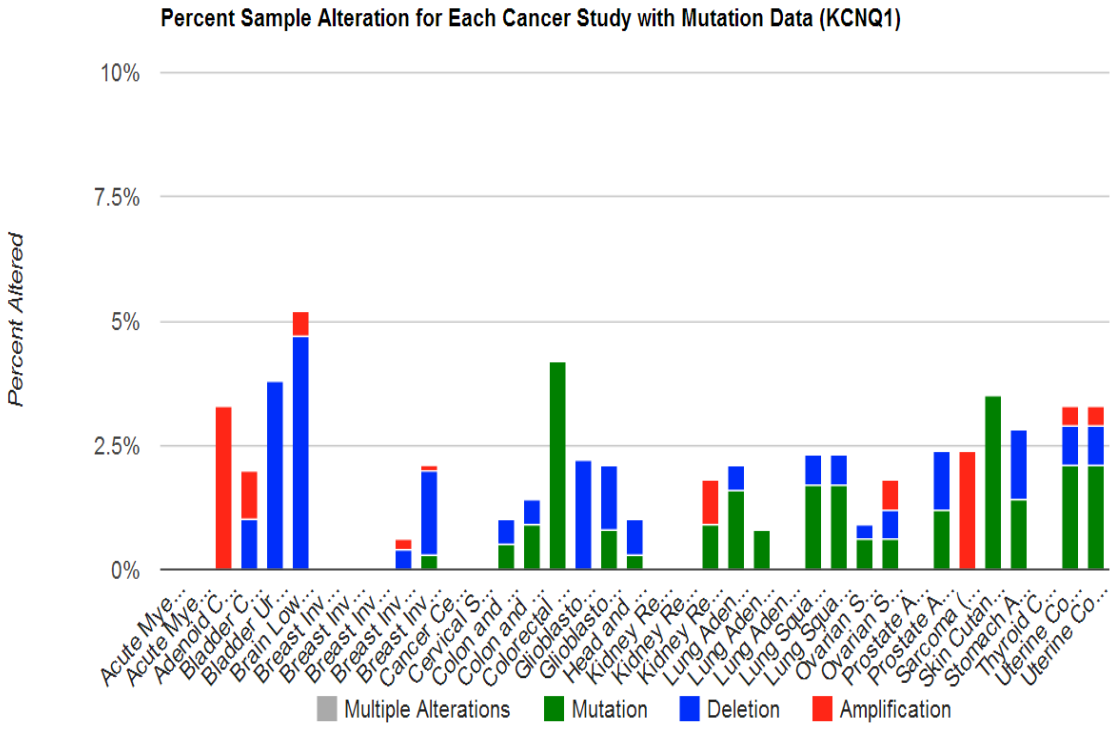


Figure 42. TCGA Report on *Kcnq1* mutations in human cancers.

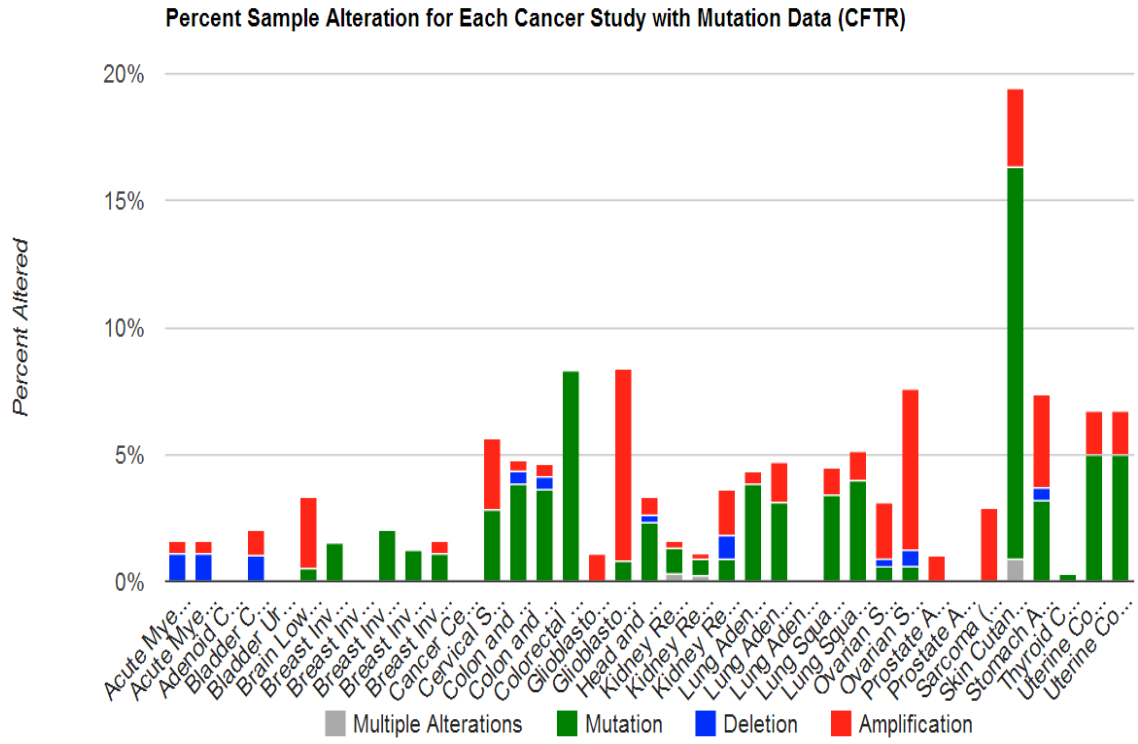


Figure 43. TCGA Report on *Cftr* mutations in human cancers.