INTRINSIC AND EXTRINSIC REGULATION OF POTENCY IN THE INTESTINAL STEM CELL NICHE

Adam David Gracz

A thesis submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Cell Biology and Physiology in the School of Medicine.

Chapel Hill
2013

Approved by:

Scott T. Magness, Ph.D.

P. Kay Lund, Ph.D.

John Rawls, Ph.D.

Michael Goy, Ph.D.

Robert Sandler, MD, MPH

ABSTRACT

ADAM DAVID GRACZ: Intrinsic and extrinsic regulation of potency in the intestinal stem cell niche
(Under the direction of: Scott T. Magness, Ph.D)

The intestinal epithelium is one of the most proliferative tissues in the adult body, undergoing near total renewal every 5-7 days. This remarkable turnover is driven by a small population of intestinal stem cells (ISCs), which maintain the physiological function and epithelial barrier integrity of the small intestine, and initiate repair following damage. While the anatomical location of ISCs has been appreciated for decades, the complex genetics and cellular behavior of these cells is still the subject of intense research and debate. Emerging research shows that patterns of ISC proliferation and differentiation are governed by highly complex interactions between the extrinsic signaling of the ISC niche and intrinsic genetic programs that regulate ISC behavior on the cell-autonomous level. In this dissertation, we aim to address the regulation of ISC potency at the intrinsic and extrinsic levels. We describe technologic approaches for the isolation and *in vitro* culture of two distinct human ISC populations, as well as advanced, high-throughput culture conditions for the study of ISC-niche interactions. To address ISC potency from the perspective of intrinsic genetic programming, we examine the role of Sry-box containing 4 (*Sox4*), which we demonstrate plays a role in ISC differentiation and proliferation, possibly through epigenetic mechanisms. Together, these studies provide valuable tools for examining the effects of extrinsic signaling on ISC potency *in vitro*, as well as describe a novel mechanistic regulator of ISC differentiation.

To my parents for their support and encouragement; to my friends and colleagues for their patience and assistance.

ACKNOWLEDGMENTS

The work presented in this dissertation would not have been possible without the help, guidance, and support of a number of individuals. I would like to first thank my mentors, Dr. Scott Magness and P Kay Lund for believing in me and facilitating my pursuit of an advanced education in biomedical science. I would also like to acknowledge the Gastrointestinal Stem Cell group in the Department of Medicine at UNC, including Drs. Susan Henning, Chris Dekaney, Laurianne vanLandeghem, and Victoria Newton. The cell culture microarray data presented here would have been possible without the help and guidance of Drs. Nancy Allbritton, Christopher Sims, and Yuli Wang in the UNC Department of Chemistry. Additionally, I would like to acknowledge my fellow graduate students in the Magness lab and Allbritton lab: Kyle Roche, Ian Williamson, Bailey Zwarycz, James McCann, Asad Ahmad, and Pete Attayek, for their contributions to the data presented here, as well as Danny Trotier for his significant contribution to phenotyping of *Sox4* knockout mice. Finally, acknowledgement is due to the exceptional undergraduates who contribute to research in the Magness lab, including Xiao Fu Liu, Keith Murphy, Casey Collins, and Liam Gaynor.

Due to the academic rigor and time commitment of a graduate education, I would be remiss to not acknowledge the support of my family and friends, who have been encouraging and patient throughout the completion of this dissertation.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

LIST OF ABBREVIATIONS

5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
APC	allophycocyanin
β-NF	beta naphthoflavone
BAC	bacterial artifical chromosome
bHLH	basic helix-loop-helix
BrdU	bromodeoxyuridine
CAG	cytomegalovirus/beta-actin/beta-globin
CBC	crypt base columnar
CD	cluster of differentiation
CDK	cyclin-dependent kinase
cDNA	complementary deoxyribonucleic acid
CNC	computer numerical control
CRC	colorectal cancer
DNA	deoxyribonucleic acid
DSL	Delta/Serrate/Lag2
DI	deionized
DPBS	Dulbecco's phosphate-buffered saline
DMSO	dimethyl sulfoxide
DTR	diphtheria toxin receptor
ECM	extracellular matrix
EDTA	ethylenediaminetetraacetic acid
EdU	5-ethynyl-2'-deoxyuridine

EE	enteroendocrine
EGF	epidermal growth factor
EGFP	enhanced green fluorescent protein
EpCAM	epithelial cell adhesion molecule
ESC	embryonic stem cell
FACS	fluorescent activated cell sorting
FBS	fetal bovine serum
FDR	false discovery rate
FSC	forward scatter
GBL	gamma-butyrolactone
GENSAT	gene expression nervous system atlas
GI	gastrointestinal
Gy	gray
HBSS	Hank's balanced salt solution
HDAC	histone deacetylase
HMG	high mobility group
hr	hour
IBD	inflammatory bowel disease
IgG	immunoglobulin G
iPSC	induced pluripotent stem cell
IRES	internal ribosomal entry site
ISC	intestinal stem cell
ISEMF	intestinal subepithelial myofibroblast
ISH	*in situ* hybridization

Lgr5	leucine-rich repeat-containing G-protein coupled receptor 5
lncRNA	long non-coding ribonucleic acid
LRC	label retaining cell
min	minute
mRNA	messenger ribonucleic acid
NDS	normal donkey serum
NGS	normal goat serum
NICD	Notch intracellular domain
OCT	optimal cutting temperature
PAA	poly(acrylic acid)
PBS	phosphate-buffered saline
PC	Paneth cell
PCA	principal component analysis
PCR	polymerase chain reaction
PDMS	polydimethylsiloxane
PEG	polyethylene glycol
PFA	paraformaldehyde
POU domain	Pit-1 Oct1/2 Unc-86 domain
PRC	Polycomb Repressor Complex
RNA	ribonucleic acid
RPM	revolutions per minute
RT-PCR	reverse transcriptase polymerase chain reaction
sec	second

SEM	standard error mean
SMA	smooth muscle actin
Sox	*Sry* Box-containing
Sry	Sex determining region Y
SSC	side scatter
TA	transit-amplifying cell
TDF	testis determining factor
TET	ten-eleven translocation
TGF-β	transforming growth factor β
wk	week
YFP	yellow fluorescent protein

CHAPTER 1

INTRODUCTION

Intestinal epithelial structure and function

The small intestine is a complex, multilayered tissue that is responsible for the processes of nutrient absorption and digestion critical for organismal survival. These layers consist of the intestinal mucosa, submucosa, muscularis, and serosa; each with a distinct cellular composition and function (Figure 1.1A&B). The intestinal epithelium, which is the central focus of this thesis, is a monolayer of columnar epithelium that lines the luminal surface of the small intestine and is responsible for a bulk of digestive processes. Additionally, since the intestinal lumen is constantly subjected to mechanical stress and ingested matter that may contain potential toxics and infectious agents, the epithelium provides an important barrier between the lumen and the body, and its integrity must be maintained in order preserve barrier function. The lamina propria lies immediately beneath the epithelium and consists mainly of fibroblasts, myofibroblasts, capillaries, and lymphocytes, which support the digestive and barrier function of the epithelium. Collectively, the epithelium and lamina propria form the intestinal mucosa. The outer layers of the small intestine, which include the submucosa, muscularis, and serosa, support the mucosa and contain the complex network of vasculature and nerves that regulate the movement of ingested matter down the gastrointestinal tract, and will not be further discussed here [1].

The intestinal epithelium is arranged along a crypt-villus axis (Figure 1.1C). The crypts, which exist as regularly dispersed invaginations in the epithelial monolayer, have long been appreciated as the proliferative compartment of the epithelium [2]. A multipotent stem cell population, located at the base of the crypts, maintains the epithelium and drives near-total tissue renewal approximately every 5-7 days, making it one of the most proliferative tissues in the body [2]. In addition to the intestinal stem cell (ISC) population, the crypts contain transit-amplifying progenitor cells (TA), which proliferate at a higher rate than ISCs, but exhibit a more restricted degree of potency and lack the ability to self-renew under physiologic conditions [3]. The villi, which serve as the main site of digestive function in the intestine, are made up mostly of absorptive enterocytes, with some secretory lineage cells (enteroendocrine and Goblet) dispersed throughout, but at much lower numbers [4-6]. The crypts are largely devoid of post-mitotic, differentiated cells, with the important exception of Paneth cells, which are hypothesized to be produced in the TA zone and migrate to the base of the crypt, where they intercalate between ISCs [7, 8]. Therefore, the hierarchy of proliferation in the intestinal epithelium flows from the ISCs, which divide relatively slowly, to the TAs, which undergo rapid, successive rounds of division and differentiation, to produce the post-mitotic cells of the villi. Because proliferation occurs continually in the intestinal epithelium, there is also a significant rate of cellular turnover. As they are produced in the crypts, cells migrate up the crypt-villus axis to the terminal end of the villus, where they are sloughed off into the lumen and undergo anoikis, a type of apoptosis that results from epithelial cells losing contact with their basement membrane (Figure 1.1C) [9]. The kinetics of intestinal epithelial proliferation and migration were established by studies that used tritiated thymidine to

"pulse label" actively dividing crypt cells and track their migration by observation at different time points following thymidine administration, and have since been confirmed using non-radioactive thymidine analogues such as BrdU and EdU [10-12].

In addition to established patterns of proliferation and migration, the differentiated cells of the intestinal epithelium have been well characterized by early studies into the cellular structure and function of the gastrointestinal tract [3-7]. Post-mitotic lineages in the epithelium are subdivided into two main categories, based on specialized function: absorptive and secretory. Absorptive enterocytes make up the majority of epithelial cells in the villus, and are identified by their columnar morphology and expression of enzymes associated with digestive processes, such as sucrase isolmaltase (*Sis*) and lactase (*Lct*) [5]. Secretory lineage cells are further subdivided into mucus-producing Goblet cells, hormone-producing enteroendocrine cells, and Paneth cells. Paneth cells are unique among differentiated intestinal epithelial cells in that they reside exclusively at the base of the crypts. It has long been hypothesized that Paneth cells differentiate in the TA zone and migrate downward to the crypt base, though this has not been tested experimentally [7]. The canonical function of Paneth cells is to secrete antimicrobial peptides into the intestinal lumen, to prevent extensive colonization by bacteria in the relatively sterile small intestine, as well as eliminate any ingested pathogens [13]. In addition to their antimicrobial function, there is intense and growing interest in the potential function of Paneth cells as "niche", or supporting, cells to the ISCs, which will be discussed in greater detail later in this thesis [8]. Emerging research has

characterized a fourth secretory cell lineage, tuft cells, which are identified by expression of COX1, COX2, and DCAMKL1, though their precise function remains unknown [14, 15].

Identification of intestinal stem and progenitor cells

Somatic stem cells are defined by their ability to meet the dual functional criteria of "stemness": self-renewal, or the ability to produce a daughter stem cell, and multipotency, or the ability to produce all post-mitotic lineages in a given tissue. Its high rate of physiological renewal, coupled with well-defined post-mitotic lineages, makes the intestinal epithelium an attractive tissue for the study of somatic stem cell maintenance and differentiation. While the dynamics of intestinal epithelial turnover have been appreciated for decades, direct studies on ISCs were long hampered by a lack of specific genetic biomarkers. Until the recent identification of ISC biomarkers, two predominant theories existed concerning the location and properties of ISCs. Cheng and Leblond hypothesized that the crypt-base columnar cell (CBC), located at the basal end of the crypt and intercalated between Paneth cells, represented the ISC population, based on observations made by light and electron microscopy and evidence of proliferation [3, 5]. An alternative hypothesis was generated based on early label-retention studies, which relied on tritiated thymidine-labeling followed by long "wash-out" periods to identify slowly dividing, or label retaining cells (LRCs, hypothesized to be stem cells), that were localized mainly to the "+4" position, relative to the base of the crypt [16, 17]. Unification of the two theories on ISC position has resulted in the general acceptance of two potential

ISC populations: an "active", rapidly-cycling ISC population of CBCs, and a "quiescent" LRC or reserve ISC population located at the +4 position [18]. However, proving, disproving, or reconciling these theories remained a significant hurdle for the ISC field for decades, as technological limitations prevented the functional testing of stemness.

In a seminal study, the G-protein coupled receptor *Lgr5* (*Gpr49*) became the first ISC biomarker to be validated by *in vivo* lineage-tracing [19]. Using a targeted knock-in of an enhanced green fluorescent protein (EGFP) reporter gene linked to tamoxifen-inducible Cre recombinase by an insertional ribosomal entry site (IRES), Barker, *et al*. demonstrated that *Lgr5* is expressed at high levels in CBCs, with low, but present expression in the TA region immediately above the Paneth cells [19]. Importantly, lineage tracing using the $Lgr5^{EGFP\text{-}CreERT2}$ allele demonstrated that derivatives of these cells were capable of forming long-lived clonal units that contained all post-mitotic lineages of the intestinal epithelium and persisted for 60 days or longer. These landmark studies established *Lgr5* as a specific marker of CBC ISCs through the gold standard of *in vivo* lineage tracing, and resulted in the rapid identification and characterization of other CBC ISC biomarkers, including *Olfm4*, *Ascl2*, and *Sox9* [20-23]. Further studies demonstrated that isolated $Lgr5^{High}$ cells were capable of forming structures containing all post-mitotic lineages of the small intestine, termed enteroids, *in vitro* [24, 25]. In addition to the obvious biological implications of these studies, the development of the $Lgr5^{EGFP\text{-}CreERT2}$ allele was an important technological contribution to the field, as it provided the first methodology for the isolation and downstream analysis of actively cycling ISCs. However, one important drawback of this mouse model is that the reporter gene is

expressed in a mosaic manner, with only some crypts exhibiting EGFP expression, despite the fact that *Lgr5* is expressed in all intestinal crypts [19].

The characterization of *Lgr5* as an ISC marker was quickly followed by a "boom" in ISC biomarker discovery. In addition to the growing number of CBC ISC markers, a number of studies were published characterizing putative markers of +4 ISCs, the first of which was *Bmi1*. Like *Lgr5*, lineage-tracing experiments in *Bmi1*^{CreERT2} mice indicated that *Bmi1*+ cells are capable of forming long-lived clonal units that contain all post-mitotic intestinal epithelial lineages *in vivo* [26]. However, examination of *Bmi1* expression by Cre induction in a Cre reporter mouse revealed that cells expressing *Bmi1* are located primarily at the +4 position, which provided the first direct evidence for a secondary, +4 ISC population in addition to the *Lgr5*+ CBC ISC [26, 27]. Subsequent studies identified other genes associated with the +4 position that also functioned as ISC biomarkers *in vivo*, including *mTert* and *Hopx* [28, 29]. The ErbB inhibitor *Lrig1* was shown to be an ISC marker *in vivo*, but appears to be expressed at both the CBC and +4 positions, with conflicting reports highlighting the need for further study of this specific marker [30, 31]. Interestingly, while all of the +4 ISC markers were shown to mark populations with some degree of stemness under normal physiological conditions, a majority of these studies also demonstrated increased lineage tracing from these markers following irradiation damage, resulting in a proposed model in which CBC ISCs represent an active stem cell population, and +4 ISCs are a reserve or quiescent population that activate to initiate repair following damage to the CBCs [27-29]. This model is currently under intense scrutiny,

due to a recent report that *Lgr5*^{High} ISCs are radioresistant and capable of initiating epithelial repair following irradiation damage [32].

The extent to which CBC and +4 ISC populations are distinct from one another remains controversial. A recent study that analyzed transcriptomic and proteomic data from intestinal epithelial cell populations expressing variable levels of *Lgr5* reported that all of the previously described +4 CBC markers were most significantly upregulated in the *Lgr5*^{High} population, which was initially characterized as a CBC-specific marker [19, 33]. Additionally, the study demonstrated that *Bmi1* is not exclusively expressed in +4 cells, but is also present in CBCs, as well as the TA zone of the crypts [33]. However, these data are subject to important caveats. The principle assumption of the Munoz, *et al* study is that *Lgr5*^{High} expression is restricted to CBCs, and it is upon this assumption that the authors conclude that previously characterized +4 markers are most strongly expressed in CBC ISCs. It is possible that, depending on their functional status, some +4 ISCs may express upregulated levels of *Lgr5* that are consistent with those found in CBCs. Examination of *Lgr5*^{EGFP} expression reveals that some crypts do express high levels of EGFP in the supra-Paneth cell position [19, 34]. This leaves open the possibility that heterogeneity within the *Lgr5*^{High} population could bias gene expression data aimed at determining the existence of two distinct ISC populations.

Our group previously demonstrated that low levels of *Sox9* expression are associated with actively cycling CBCs [21, 35]. A *Sox9*^{EGFP} BAC transgenic mouse was used to isolate cellular populations based on levels of EGFP expression, which faithfully

recapitulated expression of the endogenous mRNA and protein [35]. Based on gene expression studies, we demonstrated that the *Sox9*^{Sublow} population correlated with markers of TA cells; the *Sox9*^{Low} population correlated with markers of actively cycling CBCs; and the *Sox9*^{High} population correlated with markers of enteroendocrine cells, as well as previously described +4 ISC markers, including *Bmi1*, *Hopx*, and *mTert* (unpublished data and [21, 35]). Furthermore, we demonstrated that enteroid-forming capacity *in vitro* was restricted to the *Sox9*^{Low} population, consistent with the initial reports for *Lgr5*^{High} cells [21]. Compellingly, our gene expression data support a model in which CBC ISCs are distinct from +4 ISCs. That is, when isolating crypt populations based on *Sox9*^{EGFP} expression, as opposed to *Lgr5*^{EGFP} expression, the population expressing the highest levels of CBC ISC markers (*Ascl2*, *Olfm4*, and *Lgr5*) is genotypically distinct from the population expressing the highest levels of +4 ISC markers (*Bmi1*, *mTert*, and *Hopx*). Subsequent independent studies supported this observation, and demonstrated separation of CBC and +4 ISC marker genes between *Sox9*^{Low} and *Sox9*^{High} populations by RNA microarray, which also demonstrated differential clustering of the two populations by principal component analysis (PCA) [36]. Compellingly, these studies also demonstrated that *Sox9*^{High} cells were capable of forming enteroids *in vitro* following irradiation damage and partial gastrointestinal regeneration, consistent with *Sox9*^{High} cells representing a reserve or facultative ISC population, the role classically attributed to the +4 ISC [36]. Together, the emerging data on ISC biomarker and reported similarities and differences between CBC and +4 ISCs support a much more complex model for stemness and potency in the intestinal epithelium than originally postulated.

Stem cell plasticity and interconversion between CBCs and +4 ISCs

Due to the complexities of reconciling myriad ISC biomarker studies, each focused on individual biomarkers, there has been intense interest in generating a unifying theory for how CBC and +4 ISCs contribute to maintenance of the intestinal epithelium, and if they function independently or in a cooperative manner. There is reasonable evidence that interconversion between CBCs and +4 ISCs occurs on a semi-regular basis, as *Hopx*+ +4 cells are capable of giving rise to $Lgr5^{High}$ CBCs both *in vivo* and *in vitro* [29]. To further support the concept of ISC interconversion and distinct CBC and +4 populations, Tian, *et al* demonstrated that *Lgr5*-negative, *Bmi1*+ cells were capable of maintaining intestinal epithelial homeostasis following complete ablation of *Lgr5*+ cells [37]. The group developed a novel $Lgr5^{EGFP}$ allele that also expressed the diphtheria toxic receptor (DTR). Importantly, this model exhibited constitutive expression of the knock-in gene across all crypts, unlike the mosaicism of EGFP expression seen in the original $Lgr5^{EGFP}$ model developed by Barker, *et al* [19, 37]. The authors noted rare lineage tracing from *Bmi1*+ cells under physiological conditions, consistent with previous reports [26, 37]. When mice were injected with diphtheria toxin, all *Lgr5*+ cells underwent apoptosis, and the number of lineage tracing events from *Bmi1*+ cells increased substantially [37]. Surprisingly, there were no obvious defects in proliferation and differentiation following full ablation of *Lgr5*+ cells, though studies were limited to 5 days, due to organ failure in other tissues dependent on *Lgr5*+ populations [37]. This study was the first definitive evidence that intestinal epithelial homeostasis is not dependent on the CBC ISC, and

supported the existence of multiple populations with stemness potential in the intestinal crypt.

As an alternative explanation to the existence of two distinct ISC populations, emerging evidence suggests that the +4 population does not represent *bona fide* ISCs, but rather one or more populations of TA progenitor cells capable of dedifferentiating and adopting stem cell fate following damage. Early studies theorized that TA cells may be able to function as ISCs, and recent experiments have begun to support these hypotheses [3, 38]. In a recent study, the *Notch* ligand *Dll1* was shown to mark a population of TA cells, concentrated around the +5 position [39]. Under physiological conditions, lineage tracing from *Dll1*+ cells produced small, short-lived clones that contained secretory lineage cells, demonstrating that *Dll1* marks TA cells committed to the secretory differentiation [39]. However, when *Dll1* lineage tracing was conducted following exposure to irradiation damage, the cells were able to function as ISCs, producing long-lived clones that contained both secretory and absorptive lineage cells [39]. Labeling of *Dll1*+ cells immediately before irradiation produced these long-lived clones, but labeling 5 or more days prior to irradiation failed to produce any significant lineage tracing events [39]. This final observation revealed another important characteristic of this facultative *Dll1* population, in that it is short-lived, and the window of time between cell "birth" and differentiation in which *Dll1*+ cells can dedifferentiate is limited to less than 5 days. While informative, this study did not address the possibility of a long-lived LRC that was capable of acting as an ISC.

A central problem in studying the functional output of the putative quiescent LRC, or +4 ISC, lies in the fact that label-retaining methods rely on the incorporation of thymidine analogues into DNA during S-phase. Unfortunately, all of these techniques require detection methods that must be carried out on fixed cells or tissue, precluding the ability to isolate LRCs for downstream culture and gene expression analysis. Furthermore, despite rapid and substantial ISC biomarker discovery, the overlapping expression and conflicting reports of ISC biomarker specificity resulted in no specific genetic signature that could be used to isolate LRCs. In a very recent and landmark study, Buczacki, *et al* employed an elegant strategy to genetically label LRCs through the use of a *Cyp1a1-H2B*^{YFP} allele [34]. In this model, expression of *YFP* is induced in all cells of the intestinal epithelium by injection of β-naphthoflavone (βNF), followed by a "washout" period, during which the YFP label is lost in short-lived, proliferating cells through multiple rounds of divisions and migration. The end result is retention of YFP expression exclusively in long-lived cells, effectively recapitulating label retaining studies that relied on thymidine analogues. Buczacki, *et al* found that the majority of LRCs were found at the +3 position following two weeks of post-induction washout [34]. Interestingly, gene expression analysis of these cells revealed that while they expressed secretory lineage markers, they also shared significant overlap in highly expressed genes with *Lgr5*^{High} cells, and *in vitro* culture demonstrated that they were capable of forming enteroids [34]. Additionally, the authors demonstrated that LRCs contributed to the production of Paneth and enteroendocrine cells *in vivo* [34]. In order to meet the "gold standard" of *in vivo* lineage tracing to demonstrate stemness, the authors developed a novel mouse model, in which one half of a dimerizable Cre recombinase was induced by βNF in the same

manner as in the *Cyp1a1-H2B*^{YFP} model. This Cre was then present, but inactive, in LRCs following a washout period. Administration of a dimerizing agent, AP20187, caused the retained half of the Cre recombinase to combine with the remaining half, which was constitutively expressed, in order to induce recombination of a reporter allele [34]. Using this model, which allowed the authors to lineage trace from functionally identified LRCs in the absence of genetic biomarkers, it was shown that LRCs were capable of acting as ISCs following irradiation damage, but not under normal conditions [34]. Together with evidence from *Dll1* studies, these results suggest that there are at least two populations of TA progenitors that have secondary potential as facultative ISCs: one that is short-lived (*Dll1*+), and one that is long-lived (LRCs). However, the comprehensive genetic signatures of these populations and the degree to which they overlap with populations identified by single canonical ISC biomarkers (*Lgr5*, *Ascl2*, *Olfm4*, *Bmi1*, *Hopx*, *mTert*, *Lrig*, *Sox9*) remains unknown. Further research is needed into the functional role of ISC biomarkers in potency and stemness, and may shed light into the cellular dynamics of ISC biology. These emerging data in the ISC field increasingly support a model in which multiple populations of genotypically and/or phenotypically distinct cells collectively share the burden of stemness. Though not in direct conflict with earlier theories concerning ISC populations, recent studies suggest that ISC populations are more diverse and complex than originally thought.

Fate decisions and clonal expansion of intestinal stem cells

In addition to efforts focused on the identification and characterization of ISC biomarkers, there is also a strong interest in understanding the proliferative dynamics and clonal behavior of ISCs. As stemness is characterized by a cell's ability to self-renew and differentiate in a multipotent manner, the clonal behavior of a single ISC is the functional outcome of these processes as they pertain to the production and maintanence of a given tissue area. The survival of an ISC clone is entirely dependent on the ability of the parental ISC to self-renew and produce daughter ISCs that can continue to proliferate, differentiate and maintain the epithelial monolayer. Since inadequate or excessive self-renewal could theoretically lead to a compromised epithelium or tumorigenesis, respectively, the cellular dynamics underlying ISC self-renewal and differentiation are critical in understanding intestinal epithelial homeostasis and disease.

The classical model for self-renewal is invariant asymmetry, in which ISCs consistently divide to give rise to one progenitor cell, which will undergo one or more cycles of division and terminal differentiation, and one stem cell, which will serve the purpose of replacing the parent cell and maintaining homeostasis of the stem cell population (Figure 1.2A). This model was proposed as the mechanism for self-renewal in ISC populations, largely based on theoretical models of tissue renewal and maintanence of the stem cell pool [9]. Potten argued against the term "self-renewal", in that it could imply expansion of stem cell numbers, and instead employed the term "self-maintanence"

to describe the production of new stem cells without an overall net change in stem cell numbers [9]. Additionally, Potten supported his model of invariant asymmetry by arguing that expansion of ISC numbers by symmetric division, while a feasible possibility in the context of tissue regeneration following damage, would be impractical under physiological conditions, as it would lead to aberrant tissue growth and, perhaps, tumorigenesis [9]. It was acknowledged that an alternative model could exist, in which some ISCs divide symmetrically to produce two daughter ISCs (Figure 1.2B). This expansion could exist in homeostatic conditions under which the symmetric division of other ISCs toward differentiated lineages compensated for it, though Potten argued that this stochastic process would be subject to random and aberrant tissue loss and/or expansion since random shift in ISC population dynamics could be difficult to balance (Figure 1.2C) [9]. Since these theories about ISC population dynamics and self-renewal were formed long before established ISC biomarkers, they remained impossible to test for some time.

Taking advantage of the validated *Lgr5* biomarker, recent studies have tested Potten's hypothesis regarding invariant asymmetry and, surprisingly, demonstrated that the kinetics associated with clonal expansion driven by ISCs support a model of population asymmetry, also referred to as neutral drift [40, 41]. In population asymmetry, stem cells divide stochastically, resulting in asymmetric division or symmetric division, producing either two stem cells or two differentiated progeny (Figure 1.2A-C) [42]. To establish this experimentally, Snippert, *et al* generated a novel Cre reporter transgene, called *R26-Confetti*, based on the $Thy1^{Brainbow}$ allele. In both $R26^{Confetti}$ and $Thy1^{Brainbow}$

mice, expression of Cre recombinase results in the removal of a floxed stop sequence preceding a number of different fluorescent protein reporter cassettes [41, 43]. Opposing flox sequences flank these cassettes, such that Cre activity results in the random expression of a single fluorescent reporter. In the *R26*^{Confetti} model, cells that undergo recombination can be labeled by GFP, RFP, YFP, or BFP [41]. Importantly, as with lineage tracing, all progeny of these cells will express the same color as their parent cell. To test the kinetics of ISC self-renewal, *Lgr5*$^{EGFP\text{-}CreERT2}$ mice were crossed to the *R26*^{Confetti} allele, lineage tracing was initiated, and expansion of clonal units (independently identified by color differences) was observed at different time points following induction [41]. Under the assumptions of the invariant asymmetry hypothesis, each ISC should expand clonally at approximately the same rate, such that each clone always contains a single ISC and persists over a long period of time to produce labeled, differentiated progeny. In this model, the number of clones observed at early time points post-induction would be equivalent to that observed at late time points, and all clones would expand at approximately the same rate. Surprisingly, Snippert, *et al* observed a decreasing number of clones over time, with increasing diversification of clone size [41]. Through the application of mathematical modeling, the authors showed that the dynamics of ISC clone expansion followed neutral drift, an observation that was independently confirmed by another group that conducted similar experiments using a different genetic model for clone expansion [40, 41]. These data are in direct conflict with Potten's model of invariant asymmetry, and instead demonstrate that stochasticity in self-renewal is the predominant mode of ISC maintenance [9]. Although some stem cells are lost to differentiation under this model, the net outcome of all stem cell divisions is the same as it would be under

invariant asymmetry and is sufficient to maintain homeostasis in the stem cell pool. In the absence of the present experimental validation of neutral drift ISC dynamics, logic would follow Potten's original postulation that stochasticity in ISC self-renewal seems impractical [9]. However, despite seemingly random expansion and loss of ISC clones, there appear to no pathological consequences to intestinal epithelial homeostasis. Therefore, rather than behaving in a manner consistent with cell autonomous asymmetry, ISCs exhibit the property of population asymmetry, in that the net outcome of all stochastic symmetric and asymmetric divisions is such that the overall number of ISCs in the intestinal epithelium remains at equilibrium. In further support of this model, a very recent study has demonstrated that only a minority of ISCs identified by the expression of ISC markers actually exhibits functional stemness *in vivo* [44]. These data provide experimental validation of the self-limiting nature of population asymmetry in the intestinal epithelium.

While these recent studies reveal that ISCs undergo clonal expansion patterns that are consistent with population asymmetry and neutral drift, it remains unproven whether the processes leading to this behavior are cell-autonomous or entirely influenced by the stem cell niche, the complex network of support cells that may provide fate-deterministic signals to ISCs. Though niches have been well described in other stem cell populations, especially in the hematopoietic system, little is known about the ISC niche [45, 46]. Recent studies on the interaction between Paneth cells and ISCs have suggested that the former may provide critical mitogenic and morphogenic signals to the latter and greatly enhance

ISC survival and growth *in vitro* [8]. This observation provides proof of principle that extrinsic factors can, at very least, significantly modulate ISC behavior.

Regulation of proliferation and differentiation in the intestinal stem cell niche

Though much of our current knowledge surrounding molecular signaling networks in the intestinal epithelium precedes the advent of ISC biomarkers, there is a considerable amount of data concerning the various pathways that regulate proliferation and differentiation in the intestinal epithelium. The most substantially studied pathways pertaining to intestinal epithelial proliferation and differentiation are Wnt, Notch, and Bmp [47]. The molecular functions of these pathways are reviewed at great length elsewhere, and will be summarized here [48-50].

The Wnt signaling pathway is one of the most comprehensively studied pathways, and provides many of the critical proliferative signals to stem and progenitor cell populations throughout a wide range of cell populations and tissues, both during development and in adult organisms. Canonical Wnt signaling occurs when WNT ligands interact with FZD or LRP receptors. This ligand-receptor interaction stabilizes the cytoplasmic protein β-catenin, which is normally bound to a destruction complex consisting of APC, GSKβ, AXIN, DVL, and CK1, which targets it for ubiquitinization and degradation by the proteasome. Receptor activation sequesters the destruction complex through direct interaction with LRP, and allows the accumulation of unbound β-catenin in the cytoplasm. This free β-catenin then translocates to the nucleus, where it

derepresses the transcription factor TCF, and recruits myriad co-factors to activate context-dependent targets, generally associated with cellular growth and proliferation (reviewed by [48]). In addition to canonical Wnt signaling, non-canonical Wnt signaling can be activated by WNT ligands, but induces downstream effects that are unique and independent from stabilization of β-catenin [51]. Because proliferation is an important steady-state process in the intestinal epithelium, the Wnt pathway is especially active throughout adult life, and has been the subject of many studies aimed at understanding ISC biology. Additionally, mutations in the Wnt pathway are common across a wide range of cancers. Inactivating mutations in *Apc* are one of the most common mutations found in colorectal cancer (CRC), being detected in 85% of sporadic CRCs, and resulting in the aberrant accumulation of β-catenin due to defective destruction complexes [52-55].

Numerous studies have demonstrated that WNT is essential for the maintanence of ISC populations, and genomic analyses of the effects of disruption in Wnt signaling were the basis for early ISC biomarker discovery, as many ISC biomarkers are downstream targets of WNT/β-catenin. One of the most striking results from these early studies was the observation that complete loss of Wnt signaling, through genetic ablation of *Tcf712* (which encodes the TCF4 protein), resulted in a lack of proliferation *in vivo* [56]. Strikingly, the absence of Wnt signaling was sufficient to completely ablate ISCs, as *Tcf712*$^{-/-}$ mice died within 24 hr of birth [56]. Histological examination of embryonic intestinal tissue showed a complete lack of proliferation in intervillus regions, which are analogous to crypts in the adult intestine, indicating a loss of the ISC population [56]. Interestingly, the authors note that transition from endoderm to intestinal epithelium

occurs normally in both wild type and *Tcf712*$^{-/-}$ mice at E14.5, demonstrating that Wnt signaling is required for the establishment of ISCs and TAs, but dispensable for the formation of a differentiated epithelium [56]. Later studies examined the role of *Tcf1*, *Tcf3*, and *Tcf4* in adult murine intestine using conditional, floxed alleles, and demonstrated that loss of Wnt signaling through *Tcf* results in the loss of *Lgr5*/*Olfm4*+ ISCs in adult intestine [57]. In addition to the establishment of intestinal crypts, it was shown that Wnt signaling is required for CRC cell proliferation through the expression of dominant-negative TCF1 and TCF4 in CRC cell lines [58]. Inducible expression of the dominant-negative proteins was used to show that loss of functional TCF results in a depletion of canonical Wnt signaling, a halt in proliferation, and G1 cell cycle arrest in cell lines [58]. The authors showed that loss of proliferation was due to decreased TCF4 signaling causing a decrease in C-MYC and subsequent depression of $p21^{CIP1/WAF1}$, a cyclin-dependent kinase (CDK) inhibitor [58]. Importantly, these studies identified a number of direct genetic targets of Wnt signaling, which were in turn associated with ISC identification and function.

Extending these findings toward the identification of broader genetic programs underlying the formation of ISCs, the Clevers lab conducted gene expression microarray on human colonic adenomas, which are known to have increased levels of Wnt signaling, and normal human colon, as well as several CRC cell lines expressing dominant-negative forms of TCF4 [59]. By comparing these analyses and localizing anatomical expression of significantly regulated genes to the crypt base, the authors generated a list of 17 genes that were regulated by Wnt signaling and represented a putative ISC signature [59].

Included in this list were *Lgr5*, *Ascl2*, and the subject of a portion of this thesis, *Sox4* [59]. Further studies by the Clevers group demonstrated that, in addition to its role in ISC establishment and proliferation, Wnt signaling exerts an influence on cell position in the crypts. Dominant-negative TCF experiments by van der Wetering, *et al*. demonstrated that Wnt signaling positively regulated the expression of *EphB* receptors, while negatively regulating the expression of their ligands, *ephrinB* [58, 60]. In normal tissues, Eph-ephrin interactions result in repulsive forces that effectively "sort" cells into discrete compartments. While *EphB*s and *ephrinB*s were expressed as an opposing gradient in the intervillus regions of normal embryonic intestines and adult intestinal crypts, they were absent in *Tcf712*$^{-/-}$ embryonic intestines [60]. Genetic or pharmacological disruption of *EphB* expression in the intestine resulted in the displacement of Paneth cells, as well as a decrease in crypt proliferation, and less clearly separated proliferative and differentiated zones between the crypts and villi [60, 61].

Recent studies have demonstrated that multiple sources in the intestinal epithelium produce WNT ligands, including the Paneth cells and mesenchymal cells of the lamina propria [8, 62]. Paneth cells have been shown to produce sufficient niche signals, including WNT3, to support the growth of isolated ISCs *in vitro*, but appear to be dispensable for ISC survival *in vivo* [63, 64]. A series of *in vitro* co-culture experiments demonstrated that Wnt signaling can also be supplied to ISCs by subepithelial intestinal myofibroblasts (ISEMFs), providing a partial explanation for the persistence of ISCs in the absence of Paneth cells [62]. Interestingly, Wnt signaling also appears to drive differentiation of Paneth cells, an unexpected result since Wnt/β-catenin is most often

associated with proliferative cell populations [62, 65]. Several Paneth cell genes are dependent on *Tcf4* signaling, and proper Paneth cell differentiation and localization to the crypt base is dependent on expression of the WNT receptor *Fzd5* [65]. Together, these data demonstrate that WNT provides a significant mitogenic and morphogenic signal in the intestinal epithelium that is critical to ISC survival and proliferation. Additionally, Wnt signaling works in a highly context-dependent manner, directing ISC self-renewal and proliferation, while simultaneously controlling Paneth cell differentiation. Wnt signaling appears to arise from cells that constitute the ISC niche, including ISEMFs and Paneth cells, suggesting that dynamic crosstalk between the intestinal epithelium and its underlying mesenchyme is critical for controlling ISC self-renewal and differentiation.

While much less is known about the role of Bmp signaling in the intestinal epithelium, evidence suggests that this pathway provides pro-differentiation cues that serve as a counterbalance to Wnt-induced proliferation. BMP proteins belong to the TGF-β superfamily, and act by binding to type II and type I serine/threonine kinase receptors, which form heteromeric complexes and activate downstream transcriptional effectors belonging to the Smad family [66]. Mutations in the Bmp pathway, both at the receptor (BMPR1A) and effector (SMAD4) level, are associated with juvenile polyposis syndrome, a heritable disorder that predisposes individuals to gastrointestinal cancer [67, 68]. In the intestine, *Bmp4* is strongly expressed by the mesenchymal cells immediately underlying the epithelium, suggesting that these stromal cells play an important role in the induction of epithelial differentiation [69]. As expected, *Bmp4* is most strongly expressed in the villus mesenchyme, with only a few ISEMFs expressing the transcript in

the crypt mesenchyme [69, 70]. One of the most compelling studies demonstrating the role of Bmp signaling in the intestinal epithelium employed the use of transgenic mice which over-expressed *Noggin* in epithelial cells [69]. TGF-β inhibitory proteins, such *Noggin*, can negatively regulate Bmp signaling. Upon overexpression of *Noggin* in the epithelium, *Bmp4* signals from the mesenchyme were effectively inactivated, as measured by a reduction in SMAD phosphorylation, a downstream effect of *Bmp4* signaling [69]. The striking physiological consequence of this inhibition was the *de novo* formation of crypts along the length of the intestinal villi, which the authors likened to the hyperplastic growth observed in patients with juvenile polyposis syndrome [69]. Further studies examined the role of *Bmpr1a* in the intestinal epithelium through genetic ablation. Surprisingly, while these studies produced defects that included increased epithelial proliferation and decreased differentiation that eventually led to a polyposis phenotype, they did not phenocopy the severe *de novo* crypt phenotype observed by Haramis, *et al* [70, 71]. Together, these studies suggest that the severe phenotypes associated with clinical polyposis disorders are the result of multiple mutations in the Bmp signaling pathway, which plays a major role in regulating intestinal epithelial proliferation through the induction of differentiation [67-71]. Genetic evidence indicates that the opposing regulatory elements of Wnt and Bmp signaling may act through extensive crosstalk, as a number of loci containing regulatory response elements for both Wnt (*Tcf*/*Lef*) and Bmp (*Smad*) have been identified [72].

In addition to the relatively long-range cell-cell signaling provided by Wnt and Bmp pathways in the intestinal epithelium, Notch provides localized signals that regulate

cell fate and differentiation through direct cell-cell contact. Notch receptors and their ligands, which belong to the Delta/Serrate/Lag-2 (DSL) family, are both expressed on the cell surface and, unlike Wnt and Bmp signaling, require direct contact between neighboring cells to induce regulatory responses [50]. Interaction of Notch receptors and DSL ligands induces the release of the Notch intracellular domain (NICD) through cleavage by ADAM family metalloproteases, which allows NICD to localize to the nucleus and interact with cofactors to induce transcriptional responses [50]. In addition to *trans* signaling, Notch receptors and ligands expressed on the same cell exhibit *cis* inhibition through unknown mechanisms [50]. *Cis* inhibition is thought to preserve the integrity of the *trans* Notch signal, by ensuring that both neighboring cells do not activate the downstream pathway. In the intestinal epithelium, Notch signaling is mainly restricted to the crypts, which express *Notch1* and *Notch2* receptors, as well as *Dll1*, *Dll4*, and *Jag1* ligands [73]. *Notch1/2* receptors are associated with ISCs, and appear to play important roles in ISC maintenance [74]. Ablation of Notch signaling in the intestine, through genetic deletion of the NICD transcriptional cofactor, or pharmacologic inhibition of γ-secretase (which acts upstream of NICD) results in a complete loss of ISCs, and conversion of a majority of crypt cells into differentiated Goblet cells [75]. However, there appears to be some redundancy in Notch receptors and ligands, as dual deletion of *Notch1* and *Notch2* is required to convert proliferating cells to secretory lineages, and dual deletion of *Dll1* and *Dll4* is required to achieve the same effect through modulation of ligand expression [76, 77]. These results point to a role for Notch signaling in ISC maintanence as well as secretory differentiation. Indeed, active Notch signaling results in the expression of the target gene *Hes1*, the loss of which results in increased

secretory lineage differentiation and decreased production of absorptive enterocytes [78]. Loss of *Hes1* can be partially compensated for by *Hes3* and *Hes5*, and loss of all three intestinal *Hes* genes results in decreased proliferation, but does not fully ablate ISCs [79]. Taken in context of the conversion of all proliferating cells to secretory fates following broad ablation of Notch signaling, these data suggest that other Notch targets may sustain ISC cell fate, while *Hes1* is more important for absorptive lineage commitment. *Olfm4*, which is a CBC ISC marker, is also a target of Notch signaling and may be a candidate Notch-regulated effector of ISC fate [80]. In addition to participating in absorptive fate specification, expression of *Hes1* also results in the repression of transcription factor *Atoh1*, which drives cells toward enteroendocrine, goblet, or Paneth cell secretory fates in the absence of Notch signaling [78, 81]. Loss of *Atoh1* results in defective secretory lineage differentiation, and represents a reciprocal phenotype compared to loss of *Hes1*, strengthening the evidence for a balanced interaction between the two pathways in normal intestinal physiology [81]. Therefore, the repression of Notch through *cis* inhibition or pharmacologic intervention results in early specification toward a secretory cell fate. Through this mechanism, Notch may provide contact-dependent regulation of ISC numbers in the intestinal crypts, by preventing two neighboring cells from adopting ISC fates. Interestingly, Paneth cells, which are intercalated between CBC ISCs, express DSL ligands, suggesting that they may contribute to the establishment of an ISC niche by providing *trans*-activation of Notch in ISCs [8]. Additionally, *cis* inhibition by receptor/ligand interaction may help maintain or establish Paneth cell fates.

Studying intestinal stem cells *in vitro*

The inability to culture primary intestinal tissue, including ISCs, was a significant hurdle in the field of intestinal biology for decades. Due to the lack of an *in vitro* model system for studying ISCs, much of the early data regarding molecular signaling in the intestine relied on expensive and time-consuming *in vivo* mouse models, or CRC cell lines, which often fail to accurately recapitulate physiologically relevant properties. However, the increased understanding of signaling networks in the intestinal epithelium led to the development of a novel, three-dimensional culture system for intestinal crypts and primary-isolated ISCs [24]. These culture conditions rely on the extensive use of small molecules and growth factors targeting the Wnt, Bmp, and Notch pathways. While originally established for murine small intestinal epithelium, these conditions have now been adapted and expanded to support the growth of murine colon and single colonic stem cells, human small intestinal, colonic crypts, and colon stem cells, as well as murine stomach and adenomatous tissue [24, 82, 83]. Single isolated ISCs are capable of producing complex, multicellular organoids without the need for co-culture with ISEMFs [24]. These long-lived structures, termed enteroids, produce crypt-like proliferative buds and contain all of the post-mitotic lineages found in the intestinal epithelium [24, 25]. Significant modifications to the original growth factor cocktail used to support ISC growth has resulted in increased survival rates that potentiate expanded *in vitro* studies [84].

The ISC culture system is an invaluable technology that has been applied to elegant *in vitro* studies clarifying and supporting observations made *in vivo*. Our group

recently used the system to demonstrate that specific levels of the transcription factor *Sox9* are associated with ISCs, TAs, and enteroendocrine cells [21]. While *in vivo* lineage tracing remains the gold standard for establishing the stemness of cellular populations, the *in vitro* assay adds an additional level of precision for genes like *Sox9*, which exhibit broad but variable expression levels, in that lineage tracing is unable to distinguish between cell types based on gene expression level. Since the *in vitro* conditions result in the production of multipotent enteroids that can be maintained through multiple rounds of serial passaging, they provide an alternative platform that can be used to establish the stemness criteria of multipotency and self-renewal in isolated cells [21, 24]. Additionally, the culture system has been used to demonstrate that Paneth cells are capable of supporting the development of ISCs into enteroids *in vitro*, suggesting that Paneth cells constitute a niche cell for ISCs [8]. These experiments were in turn used as a basis to assess the effect of calorie restriction on ISC-niche interactions, which demonstrated that decreased caloric intake correlated with an increased ability of Paneth cells to promote enteroid formation from isolated ISCs [85]. However, while the conventional system has facilitated a handful of functional experiments examining ISC biology, significant problems prevent it from being applied to elegant molecular and genetic assays [8, 36, 80]. For one, simple quantification of ISC development into enteroids is significantly hindered by the three-dimensional nature of the system and the tendency of cells to move through the extracellular matrix and fuse to form single enteroids from multiple cells [8]. Additionally, Paneth cell co-culture experiments provide proof of principle that non-ISC cell types can considerably influence ISC behavior *in vitro*, raising significant concerns about cell-cell signaling in ISC culture assays and highlighting the need advanced technology aimed at

studying ISCs *in vitro* [8]. Despite these limitations, the development of an *in vitro* culture system for ISCs has provided the field with a critical tool for examining and modulating the effect of extrinsic signaling on ISC stemness and differentiation.

Sry-box (SOX) transcription factors and cellular potency

In addition to understanding how ISCs respond to extrinsic signaling cues, there is a strong interest in determining what intrinsic regulatory pathways affect stemness and how these cues prime ISCs for renewal or differentiation, as well as how they might become dysregulated in the context of cancer. The *Sry*-box (Sox) family of transcription factors is implicated in a broad range of developmental processes and, as such, is an attractive target for the study of intrinsic regulation of stemness in ISCs. Sox genes code for strongly conserved HMG-box containing proteins, that were originally identified in mammals through the discovery that Y-linked *Sry* is responsible for sex-determination and testicular initiation in embryos [86, 87]. All other Sox factors are identified by homology to the SRY HMG-box, and further subdivided into specific classes by homology in protein structure and function [88, 89]. Initially, 30 Sox factors were characterized, but redundancies between different studies and differences in nomenclature resulted in a final identification of 20 unique Sox factors [88, 90]. All Sox factors share a HMG binding domain motif, and although compensation between Sox factors appears to be common, especially between members of the same subfamilies, it appears that interaction with binding partners helps direct individual Sox factors to specific transcriptional regulatory sites [88, 89, 91]. Beyond classical roles in trans-activation and repression, Sox factors possess the capacity to direct dramatic conformational changes in chromatin structure through

binding the minor groove of the DNA helix and bending chromatin [92, 93]. These conformational effects on chromatin structure have been shown to be essential for sex-determination by *Sry* and are thought to function through joining enhancer and promoter regions that are normally separated by large distances across the genome [91, 94]. Because Sox factors share a common binding motif, functional compensation has been observed between several pairs of Sox factors. Both *Sox5* and *Sox6* are independently dispensable for chondrogenesis, but genetic deletion of both results in defective cartilage formation [95]. Similarly, *Sox4* and *Sox11*, which both belong to the SoxC subfamily, exhibit overlapping expression in the developing nervous system and can compensate for one another in spinal cord development [96].

Sox factors are generally studied in the context of broad transcriptional regulatory programs, are known to regulate diverse cellular processes such as proliferation, differentiation, and apoptosis, and are found in stem cell populations throughout the body, both during development and in adult tissues [97]. *Sox2* is one of the most studied Sox factors, as well as one of the master regulatory transcription factors involved in maintaining embryonic stem cell (ESC) identity [98]. Additionally, *Sox2* is one of the four transcription factors, along with *Oct4*, *c-Myc*, and *Klf4*, that together are capable of reprogramming post-mitotic cells to an induced-pluripotent stem cell (iPSC) state [99, 100]. POU domain proteins, such as OCT4, are well-characterized binding partners for Sox factors, and the synergistic power of *Sox2-Oct4* in programming cellular potency has been demonstrated by the ability of these two factors to generate iPSCs in the absence of *Klf4* and *c-Myc*, through the combined used of a histone deacetylase (HDAC) inhibitor [101-103]. As development proceeds, *Oct4* may partner with different Sox factors to promote

specification toward tissue-specific genes. For example, during the generation of cardiac progenitors from iPSCs, *Oct4* switches from driving *Sox2* expression to driving *Sox17* expression, which results in the promotion of an endodermal cell fate [104]. Further studies into the interaction between *Sox2* and *Oct4* have yielded significant insight into specificity and redundancy between Sox factors. Under normal conditions, SOX2 and SOX17 both act as OCT4 cofactors, but SOX2-OCT4 targets remain distinct from SOX17-OCT4 targets [105]. Substitution of a single amino acid in the OCT4-interaction domains of SOX2 and SOX17 reverses this specificity and results in a novel SOX17 protein that is capable of inducing pluripotency, and a novel SOX2 protein that is capable of activating genes associated with endodermal specification [105]. Expanded studies revealed that similar results could be achieved by modifying SOX7, but not other Sox factors, though further modification of C-terminal transactivation domains could confer some reprogramming function to SOX4 and SOX18 [106]. Together, these studies indicate that Sox factors have strong potential for directing genetic programs associated with stemness and potency, and despite having similar binding motifs, possess context-dependent capabilities based on association with partner proteins.

Sox factors are expressed throughout the gastrointestinal tract, and play important roles in the developmental and stem cell biology of the esophagus, stomach, small intestine, colon, pancreas, and liver [107]. *Sox9* is the most comprehensively studied Sox factor in the small intestine, and has been shown to have important roles in both ISC proliferation and differentiation, consistent with the broad functionality observed for Sox factors in ESCs and iPSCs. *Sox9* is expressed throughout the intestinal crypts, encompassing ISCs, TAs, Paneth cells, and crypt-based enteroendocrine cells [20, 21, 35, 108,

[109]. Loss of Wnt signaling results in a total loss of *Sox9* expression in the crypts, indicating that *Sox9* is a Wnt target gene [108]. Interestingly, *Sox9* is expressed at different levels throughout the crypt, with $Sox9^{Sublow}$ expression marking TAs, $Sox9^{Low}$ expression marking ISCs, and $Sox9^{High}$ expression marking enteroendocrine cells [21, 35]. Some $Sox9^{High}$ cells are present in the villi and express the enteroendocrine marker *Chga*, as well as the tuft cell marker *Dclk1*, though the heterogeneity of the $Sox9^{High}$ cells in the villus has not been fully examined [14, 15, 35]. Interestingly, $Sox9^{High}$ cells in the crypts are commonly observed in the +4 position and express markers associated with reserve ISCs [35, 36]. Though normally not observed to be proliferative, the $Sox9^{High}$ cells of the crypt enter the cell cycle and are capable of forming enteroids *in vitro* following irradiation damage [36]. Together, these data suggest that $Sox9^{High}$ cells are at least partially consistent with the LRC secretory precursors described by Buczacki, *et al* [34]. Conditional genetic ablation of *Sox9* specific to the intestinal epithelium results in crypt hyperplasia and reduced production of Paneth cells, demonstrating that *Sox9* has a role in tempering proliferation in the ISC zone [108, 110]. Overexpression of *Sox9* in cell lines results in reduced proliferation and a endocrine-like phenotype, consistent with observations made *in vivo* [35, 110]. Despite the phenotypic changes in *Sox9* knockout mice, the animals are healthy and viable, with no overt deleterious effects, suggesting that compensation by other Sox factors may ameliorate genetic dysfunction in the absence of *Sox9* [110]. A number of other Sox factors, including *Sox4*, *Sox7*, *Sox17*, and *Sox18* are expressed in the adult intestinal epithelium, but their functions remain largely undetermined [111-114].

Sox4 as a regulator of development and differentiation

Sox4 has been studied mainly in the contexts of hematopoietic and neural development. It was originally established as a lymphocyte-associated transcription factor expressed in T and pre-B lymphocytes as well as the thymus, and was shown to have potent transactivating capabilities that led to the conclusion that *Sox4* might be important for the establishment of cell fate in lymphocytes [115, 116]. Later, *Sox4* was established as a critical factor in cardiac development, as mice carrying constitutive deletion of *Sox4* die at E14.5 due to significant defects in the cardiac outflow tract, which resemble common arterial trunk defects in humans [117]. Consistent with its early association with lymphocytes, hematopoietic cells isolated from *Sox4* knockout embryos and transplanted into wild-type recipients produced fewer numbers of pre- and pro-B cells, but normal number of T lymphocytes [117]. Despite the normal numbers of T lymphocytes in recipient mice following transplantation, further studies on thymic explant cultures from *Sox4*$^{-/-}$ embryos demonstrated decreased efficiency in T lymphocyte differentiation as well [118]. Studies in adult mice demonstrated that induced ablation of *Sox4* results in apoptosis and loss of pro-B cells, indicating that *Sox4* plays an important role in fate specification in the hematopoietic system even after development [119]. These phenotypic observations generated an interest in the mechanistic effects of *Sox4* in B and T cell differentiation. *Sox4* is known to be a target of TGF-β signaling in hematopoietic and neural tissues [120, 121]. Detailed studies into the role of *Sox4* in T cell development showed that *Sox4* is activated downstream of TGF-β to suppress differentiation into T helper type II cells, both by binding of SOX4 protein directly to GATA3, and by inhibitory binding of SOX4

to genes required for T helper type II cells [121]. Through these mechanisms, *Sox4* is able to effectively block *Gata3*-dependent T helper type II differentiation, both through classical transcriptional repression and by direct inhibition of the GATA3 transcription factor. In pre- and pro-B cell lineages, *Sox4* was shown to bind and enhance the activity of the B cell-specific *λ5-VpreB1* locus, providing mechanistic evidence for how *Sox4* may specify B cell fate in hematopoietic precursors [122]. Interestingly, this locus was bound by *Sox2* in ES cells, where it carried histone marks associated with both active and repressed gene transcription [122]. These combinatorial histone marks, known as bivalent domains, are associated with inactive, but primed transcription and are thought to represent regions of the genome that are prepared for rapid initiation of transcription following specific developmental cues [123]. Loss of *Sox2* was associated with a loss of the active histone mark in ES cells, and SOX4 colocalized with the same active mark in pre-B cells [122]. Together these data implicate a *Sox2-Sox4* cascade in the epigenetic regulation of B cell differentiation, suggesting that Sox factors may affect large-scale transcriptional changes associated with development and differentiation through epigenetic mechanisms.

In addition to its roles in hematopoietic differentiation, multiple studies have established *Sox4* as a regulator of differentiation in the nervous system. In the neurons of the spinal cord, *Sox4* exhibits overlapping expression with *Sox11* and *Sox12*, which together constitute the SoxC subgroup of Sox factors [88, 96, 124]. Together, *Sox4* and *Sox11* activate a number of basic helix-loop-helix (bHLH) transcription factors that are associated with neural lineages and required for fate specification into terminally differentiated neurons [124]. Since *Sox4* and *Sox11* activate neural bHLHs, they are

considered very early differentiation factors in neural development, a conclusion that is supported by the observation that a complex known to inhibit neural differentiation directly blocks expression of *Sox4* and *Sox11* in neural progenitors [124]. Independent studies found that concomitant loss of *Sox4* and *Sox11* resulted in widespread apoptosis in the early neural tube, in addition to defects in organogenesis in cardiac, skeletal, and some abdominal organs [125]. These studies were further supported by the observation that deletion of both *Sox4* and *Sox11* results in apoptosis in the developing spinal cord, but that individual deletion of each gene has no observable phenotypic effect [96]. Expression of SOX4 and SOX11 was found to be restricted to neural progenitors, supporting a role for the factors as regulators of differentiation [96]. Interestingly, loss of *Sox4* and *Sox11* in the sympathetic nervous system does not result in differentiation defects, but rather in significantly decreased progenitor proliferation [126]. Perhaps more surprisingly, studies found a decrease in genes associated with myelin differentiation, severe hypomyelination, and ataxia in mice carrying transgenic overexpression of *Sox4* in oligodendrocyte precursors [127, 128]. Together, studies on *Sox4* in the nervous system assign disparate roles for the gene in different cell and tissue types, with some lineages appearing dependent on expression of *Sox4* for differentiation and others appearing to require downregulation of *Sox4* before differentiation can take place. However, common themes are found in the restricted expression of *Sox4* in progenitor populations, functional compensation for loss of *Sox4* by other SoxC factors, and the general involvement of *Sox4* in the process of differentiation.

In the gastrointestinal tract, *Sox4* is most appreciated as an important regulator of endocrine differentiation in the pancreas [129, 130]. Pancreatic explant cultures grown from *Sox4*$^{-/-}$ embryos exhibit decreased expression of insulin and glucagon, as well as key bHLH transcription factors required for endocrine differentiation, including *Ngn3* and *NeuroD1* [130]. In the intestinal epithelium, *Sox4* was identified as a Wnt target gene in early studies examining the role of Wnt signaling in epithelial development and crypt formation [56, 59]. Like *Sox9*, *Sox4* is expressed in the base of the intestinal crypt, though the number of *Sox4* transcripts per cell appears to be slightly lower than that of *Sox9*, based on single molecule fluorescent *in situ* hybridization (ISH) [131]. Studies in CRC cell lines demonstrate that *Sox4* positively regulates canonical Wnt signaling by increasing the activity of a TOP-flash reporter construct [114]. *In vitro* studies also indicate that SOX4 may bind directly to β-catenin and TCF/LEF proteins, suggesting that it may stabilize the former and act as a transcriptional cofactor with the latter [114]. However, the role of *Sox4* in intestinal stemness and differentiation *in vivo* remains unknown. Though *in vitro* studies in cell lines can produce meaningful data aimed at understanding the molecular biology of signaling pathways, they often fail to recapitulate physiological conditions, as immortalized CRC cell lines carry multiple mutations that may not be found in normal tissue or even tumors *in vivo*. Additionally, CRC cell lines do not accurately model the compartmentalized proliferation and differentiation of the intestinal epithelium. Due to its complex and conserved role in differentiation during embryonic development and across multiple different tissue types, understanding the role of *Sox4* in the intestinal epithelium may provide valuable mechanistic insight into the regulation of stemness and differentiation in ISCs.

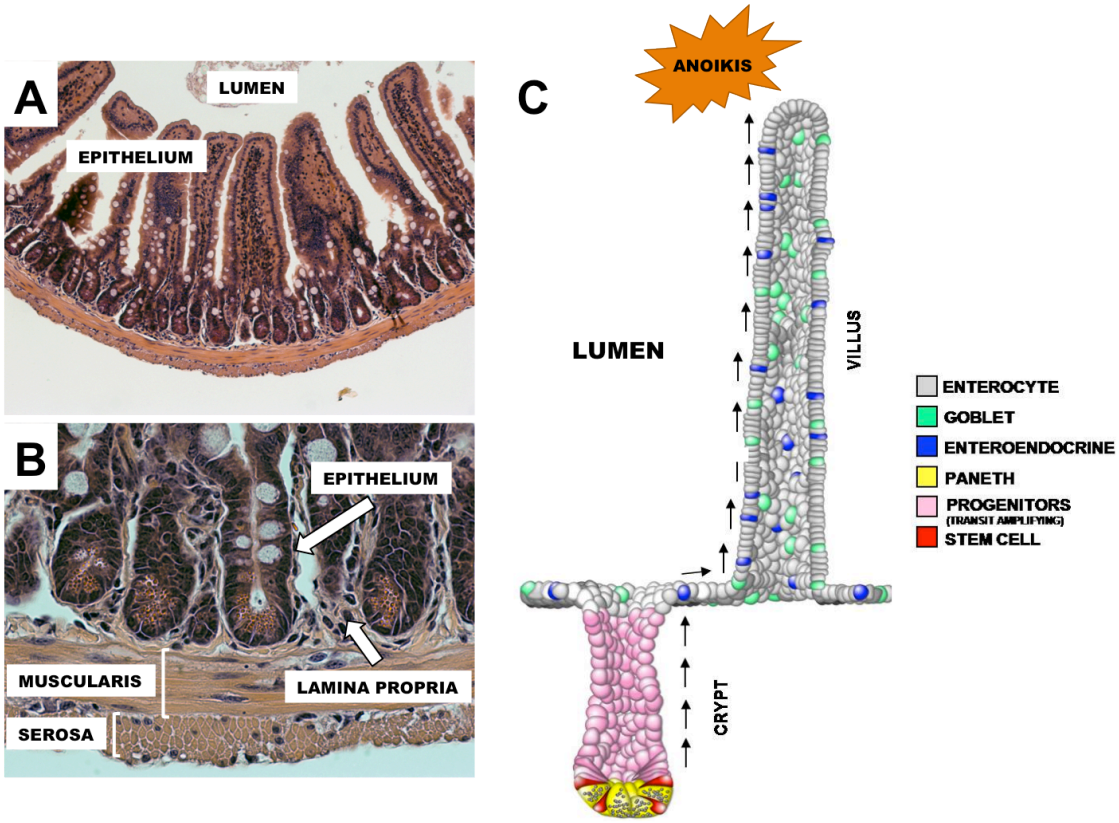


***Figure 1.1*: Anatomy of the small intestine.** The intestinal lumen is lined by a monolayer of columnar epithelial cells, called the intestinal epithelium, which is responsible for critical nutrient processing and barrier function (A). The intestine is a complex, multilayered tissue, with the lamina propria, muscularis, and serosa acting together to support the intestinal epithelium and aid in digestive and absorptive processes (B). The intestinal epithelium forms a regular micro anatomical landscape consisting of crypts and villi (C). New epithelial cells are produced by ISCs in the base of the intestinal crypts at a rapid rate under normal conditions. These cells then undergo multiple round of differentiation and lineage specification as they migrate toward the villus tip, where they are sloughed off after ~5-7 days and undergo programmed cell death, called anoikis (C).

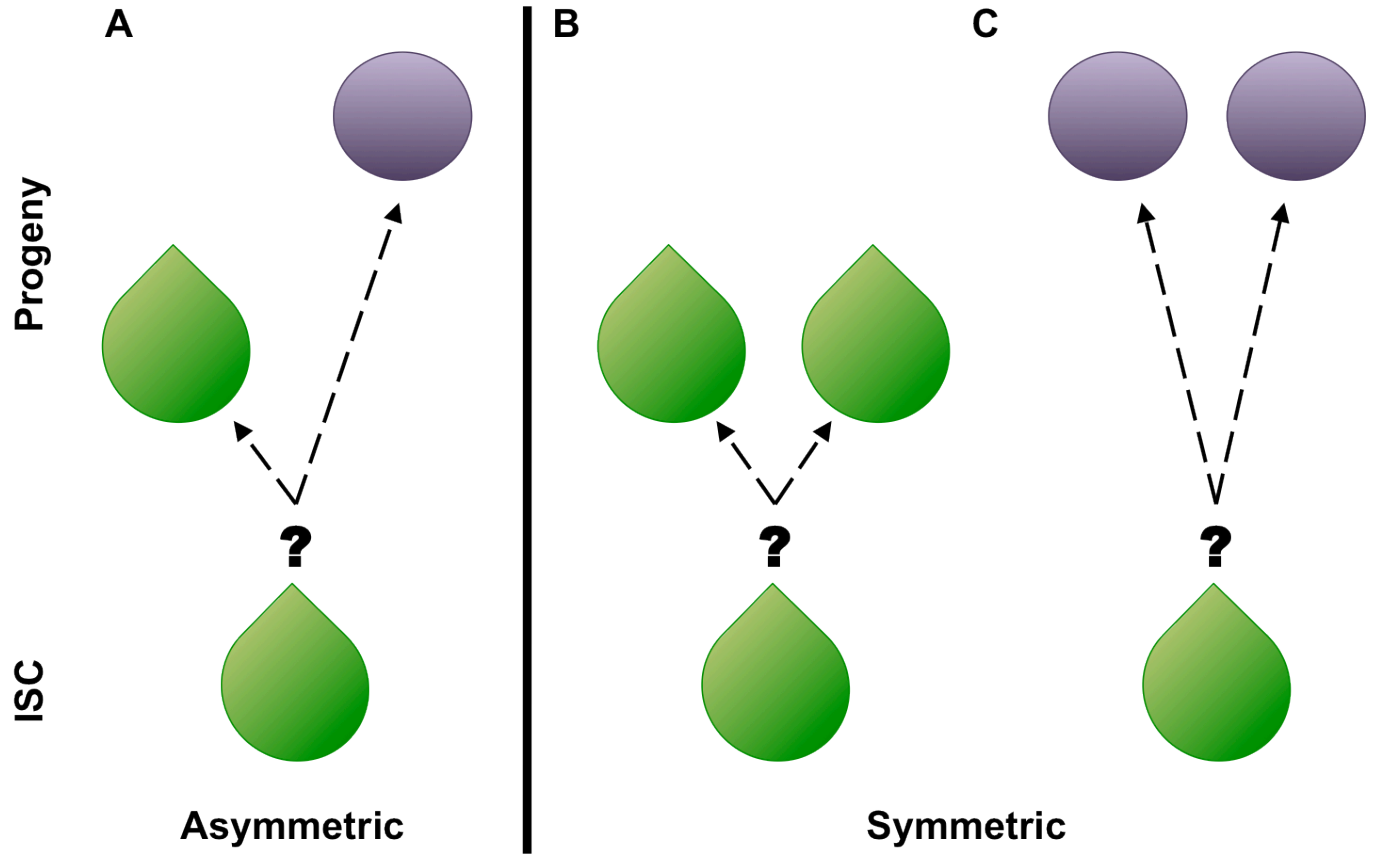


***Figure 1.2*: ISCs can adopt asymmetric or symmetric fate decisions.** Asymmetric division of ISCs results in self-renewal and differentiation, with the production of one daughter ISC and one TA progenitor (A). Symmetric division can result in either self-renewal of ISCs (B) or loss of the clonal unit by differentiation into TAs (C). The mechanisms underlying each fate decision remain unknown.

CHAPTER 2

CD24 AND CD44 MARK HUMAN INTESTINAL CELL POPULATIONS WITH CHARACTERISTICS OF ACTIVE AND FACULTATIVE STEM CELLS[1,2]

OVERVIEW

Recent seminal studies have rapidly advanced the understanding of intestinal epithelial stem cell (ISC) biology in murine models. However, the lack of techniques suitable for isolation and subsequent downstream analysis of ISCs from human tissue has hindered the application of these findings toward the development of novel diagnostics and therapies with direct clinical relevance. This study demonstrates that the cluster of differentiation genes *CD24* and *CD44* are differentially expressed across *LGR5* positive "active" stem cells as well as *HOPX* positive "facultative" stem cells. Fluorescence-activated cell sorting enables differential enrichment of *LGR5* cells (CD24-/CD44+) and *HOPX* (CD24+/CD44+) cells for gene expression analysis and culture. These findings provide the fundamental methodology and basic cell surface signature necessary for isolating and studying intestinal stem cell populations in human physiology and disease.

---

[1] previously published material: John Wiley & Sons, Inc., used with permission. The original citation is as follows: Gracz, *et al*. "CD24 and CD44 mark human intestinal cell populations with characteristics of active and facultative stem cells", *Stem Cells* 31:9, 2024-2030.

[2] full list of contributing authors: Gracz AD, Fuller MK, Wang F, Li L, Stelzner M, Dunn JCY, Martin MG, and Magness ST.

INTRODUCTION

Lgr5 was the first validated ISC biomarker shown to be expressed in actively cycling mouse crypt base columnar cells (CBCs) [19]. Subsequent studies demonstrated a secondary, "reserve" population of mouse ISCs marked by *Bmi1*, *Hopx*, *mTert*, and *Lrig* [27-30]. Emerging evidence indicates overlapping expression of *Lgr5* with these "reserve" ISC biomarkers; however, *Lgr5*-negative cell populations have also been shown to dedifferentiate in response to damage, suggesting the existence of one or more functionally competent 'facultative' ISC populations [33, 36, 39]. Despite these advances in ISC biomarker discovery, FACS isolation and functional characterization of putative '"active" and "reserve/facultative" ISC populations from human intestinal tissue has been limited by the lack of validated human ISC biomarkers and *in vitro* assays to functionally test stemness at the single cell level.

Investigators in other stem cell fields have utilized FACS-based approaches, which rely on multiple cell surface antigens, to isolate target stem cell populations of varying purity. Notably, biomarkers comprised of cluster-of-differentiation (CD) genes have long been used to identify hematopoietic stem cells and their progenitors [132]. We recently adopted a similar strategy to demonstrate that low levels of CD24 facilitate FACS of *Sox9*^{Low}*/Lgr5*+ murine ISCs capable of forming enteroids *in vitro* [21]. Similarly, CD44 is expressed in the stem cell zone of the murine small intestine and can be used to enrich for *Lgr5*+ CBCs [133]. In this study we explored whether CD24 and CD44 could be used to FACS-isolate human ISCs.

RESULTS AND DISCUSSION

CD24 and CD44 expression was assessed on human jejunum derived from patients who had undergone *roux-en-y* gastric-bypass surgery. Immunostaining demonstrates that CD44 is expressed on cells from the base of the crypt to the crypt-villus junction (Figure 2.1A, C, & E). By contrast, the villus epithelium does not show appreciable CD44 staining (Figure 2.1B & D). CD24 demonstrates similar expression to CD44 in the epithelium with the notable exception that staining is primarily distributed along the apical membrane (Figure 2.1F & H, arrows). A minority population of crypt-based cells expresses high levels of cytoplasmic CD24 (Figure 2.1G & I, arrows). While CD24/44 expression is highly restricted to the stem cell zone in the epithelium, there is broad expression in non-epithelial cells in the lamina propria, sub-mucosa, and muscle (Figure 2.1A-I). EpCAM (CD326) expression is unique to all crypt and villus epithelial cells, and was deemed useful for positive FACS selection to separate epithelial from non-epithelial CD24/44-expressing cells (Figure 2.1J-L).

Next, we analyzed CD24, CD44, CD45, and CD326 on dissociated single epithelial cells (Figure 2.2). Positive selection for epithelial cells (EpCAM+) and negative selection for lymphocytes (CD45-) was included in the FACS parameters for robust epithelial enrichment (Figure 2.3C). Flow cytometry data demonstrate four distinct epithelial cell populations based on CD24/44 expression status (Figure 2.3D). Each of the four populations (CD24-/44-; CD24+/44-; CD24-/44+; CD24+/44+) was collected for gene expression analysis. Semi-quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) validated the FACS by showing enriched gene expression for

CD24/44 in the appropriate populations (Figure 2.4A & B). The data demonstrate that CD24-/CD44+ populations are most enriched for the "active" cycling ISC markers, *LGR5* and *OLFM4* (Figure 2.5A & B); and the CD24+/CD44+ population is most enriched for the "reserve/facultative" ISC marker *HOPX* (Figure 2.5C). Interestingly, all CD24+ and CD44+ populations demonstrate significant de-enrichment for Paneth cell markers DefensinA6 (*DEFA6*) (Figure 2.6A) and Lysozyme (*LYZ*) (Figure 2.6B), and goblet cell marker, Mucin2 (*MUC2*) (Figure 2.6C), all of which associate with CD24 expression in mice [8, 39, 134]. However, the CD24+/CD44+ population is highly enriched for the enteroendocrine cell marker Chromogranin A (*CHGA*) which is consistent with observations made in mice (Figure 2.6D) [21].

To test if the CD24-/CD44+ and CD24+/CD44+ populations had functional properties of stemness, we subjected isolated single cells to culture conditions similar to those that have been successfully used to grow small intestine crypts and single colonic stem cells from human tissue [82, 83]. Both ISC populations formed appreciable cystic enterosphere structures by 48 hrs (Figure 2.7A), while the other populations (CD24-/CD44- and CD24+/CD44-) failed to do so. Enterospheres derived from CD24-/CD44+ cells continued to develop over the first week of culture, while CD24+/CD44+ derived enterospheres exhibited a limited increase in size (Figure 2.7A). Nevertheless, both populations persisted in culture at 7 days (Figure 2.7B & C). In an attempt to increase culture efficiency, we added GSK-inhibitor CHIR99021, which promotes self-renewal of embryonic stem cells was recently used to enhance *in vitro* growth and survival of human colon cancer stem cells and intestinal crypts [84, 135, 136]. Interestingly, initial GSK-

inhibition significantly improved 7-day enterosphere survival in the CD24-/CD44+ population, but had no appreciable effect on the CD24+/CD44+ population (Figure 2.7B & C). Initial GSK-inhibition greatly increased 14-day survival of the CD24-/CD44+ population, but was insufficient for the maintenance of CD24+/CD44+ derived enteroids, which did not survive to 14 days, regardless of GSK-inhibition (Figure 2.7C). Previous studies have shown that *Lgr5*-negative populations can de-differentiate and function as "facultative" stem cells when presented with the proper extrinsic cues [36, 39]. Additionally, co-culture of primary human intestinal crypts with myofibroblasts enhances culture efficiency [137]. In an attempt to "activate" the CD24+/CD44+ population, we co-cultured the cells with myofibroblasts isolated from human jejunal submucosa. Surprisingly, when grown in these conditions, both CD24-/CD44+ and CD24+/CD44+ cells produced long-lived enteroids, at rates of 0.07% and 0.76%, respectively (Figure 2.8).

To assess multipotency, we retrieved and processed enteroids for immunohistochemistry at 21 days. CD24-/CD44+ derived enteroids were epithelial in nature (Figure 2.9) and produced enteroendocrine (CHGA), goblet (MUC2), and Paneth cells (LYZ), as well as absorptive enterocytes (SIM) (Figure 2.10B). CD24+/CD44+ cells grown in co-culture with primary human myofibroblasts exhibited the same characteristics of multipotency as CD24-/CD44+ cells (Figure 2.10C and Figure 2.9). Enteroids derived from both populations exhibited increased numbers of lysozyme positive cells, consistent with results in mice demonstrating that persistent WNT signaling, specifically through CHIR99021 treatment, increases Paneth cell production in enteroid culture [62]. This biased secretory lineage allocation likely explains the rare

occurrence of MUC2 positive cells in the enteroids. Importantly, co-culture with myofibroblasts did not induce growth in CD24-/CD44- and CD24+/CD44- populations, demonstrating that functional stemness under all tested growth conditions remains restricted to CD24-/CD44+ and CD24+/CD44+ populations.

The translation of scientific findings made in genetically homogeneous mouse models presents investigators with a daunting challenge when attempting to understand the vast heterogeneity underlying physiology in human populations. This report provides new methodology for epithelial dissociation, isolation, and FACS enrichment of ISCs from human small intestine through the use of a combinatorial CD-marker "signature". The ability of CD24 and CD44 to enrich for ISCs is highly consistent with observations in mice; however, in stark contrast to findings in mouse models, human Paneth cells do not appear to express high levels of CD24, highlighting the need to validate mouse genetic biomarkers on human samples [8, 21, 39, 134]. Furthermore, our results demonstrate variable levels of gene enrichment between some patients. This observation could be attributed to a wide range of factors, including phenotypic heterogeneity between samples. In the present study, tissue procurement protocols required full de-identification of samples; thus, a patient's dietary habits, state of illness, or medication history could not be correlated to gene expression variations. This highlights the need for comprehensive studies with controlled enrollment criteria, utilizing the foundation for isolation of *LGR5* and *HOPX*-enriched populations, presented here.

In addition to FACS-enrichment of 'active' *LGR5*+ ISCs, the CD24/CD44 approach facilitates differential isolation of *HOPX*+ cells, which emerging evidence suggests may represent facultative ISCs [29]. While several of the markers examined here exhibit a degree of variability between samples, it is clear that the population of cells most enriched for *LGR5* is distinct from the population most enriched for *HOPX*, unlike recent observations of *Hopx* expression in murine *Lgr5*+ cells [33]. Additionally, CD24-/CD44+ and CD24+/CD44+ populations exhibit differential behavior *in vitro*, both in response to GSK-inhibition and in basal culture conditions required for growth, further supporting the existence of phenotypically distinct ISC populations. Though overlap of *LGR5* expression with putative "reserve" ISC markers in mice remains controversial, our data suggests that more distinct populations may exist in the human small intestinal epithelium [27, 29, 33, 37]. Further work is needed to characterize the CD24+/CD44+ population and determine if it is analogous to proposed "facultative" ISCs observed in mouse models of intestinal damage and regeneration. Multiple reports have demonstrated the ability of *Lgr5*-negative facultative ISCs to drive regeneration in murine intestine following irradiation, suggesting that an analogous population may be of particular interest in human pathophysiology, especially in patients undergoing radiation treatment for tumors [36, 39]. Understanding the genetics and patient-to-patient heterogeneity of phenotypically distinct human ISC populations may provide valuable insight toward the development of novel clinical diagnostics and therapeutic interventions.

MATERIALS AND METHODS

Patients/Tissue collection and preparation. De-identified tissue from female patients ranging between 33-53 yrs of age with body mass indices of 39-60 kg/m^2 was used in this study. Tissue was obtained from laparoscopic roux-en-y gastric bypass surgery and represents jejunal segments of approximately 4 cm in length. Following resection, tissue was placed in a specimen cup on ice until a mucosectomy was performed, aided by injecting ice-cold saline between the mucosa and submucosa prior to careful dissection. Single cell dissociation was carried out on a small portion of the total mucosa (1 cm x 1 cm) for gene expression studies and a larger tissue area (4cm x 4cm) was dissociated for culture experiments. For an informative comparison, the mass of mucosa used for this preparation represents approximately 300- and 1200-times the mucosal mass of an average biopsy from endoscopy or colonoscopy at UNC (13 mg/biopsy; unpublished, Drs. Tope Keku/Robert Sandler), respectively. Following dissection, mucosa was placed in 3 mM EDTA in 1x PBS for 45 min at 4°C on a rocker to remove villi. The villus fraction was discarded (Figure 2.11A) and the remaining mucosa was then transferred into 5 mL of PBS and lightly shaken by hand (approximately 1 shake/sec for 2 min) to remove the remaining epithelium (Figure 2.11B). An equal volume of 2% Sorbitol made in 1x PBS (Sigma, St. Louis, MO) was added. To further deplete the solution of contaminating villi, the epithelial solution was passed through a 70μm filter. This procedure results in a 'crypt-enriched' epithelial preparation (Figure 2.11C). The crypts were pelleted at 150x g for 10 min at 4°C. Crypts were then digested to single cells by resuspending the pellet in 5 mL of HBSS containing 0.3 U/mL of dispase (Worthington Biochemical, Lakewood, NJ) followed by incubation at 37°C for 10 min. The crypt

solution was then manually shaken for 30 sec (3-4 shakes/sec). The solution was then checked for extent of dissociation to single cells. If cell clumps remained, shaking cycle was repeated every 5 minutes, then checked for extent of dissociation. Shaking cycles were stopped at the earliest time point at which 80-90% of crypts were dissociated to completion or up to 30 min maximum. An average of 1 x 10^7 cells was obtained from a 1 cm x 1 cm mucosal segment. Single cells were filtered using a 40μm filter to remove undissociated clumps. For FACS, 1 x 10^7 cells were placed in 500μL of ISC staining media (see *Flow cytometry* section, below) for antibody staining. Human tissue used in this study was deemed exempt from full Institutional Review Board review (approval #09-2159).

Immunostaining. A 3 cm square piece of jejunum from each case was fixed with freshly made 4% paraformaldehyde (PFA) for a 24–48h at 4°C. The tissues were then prepared for cryosectioning by immersion in 30% sucrose for at least 24h at 4°C. Tissues were embedded in Tissue-Tek optimal cutting temperature (OCT) medium (Sakura, Torrance, CA) and frozen on dry ice. 8–10μm sections were cut on a cryostat and placed on positively charged microscope slides. Prior to immunostaining, tissue sections were rinsed twice in PBS to remove OCT. Non-specific binding was blocked by applying Dako Protein Block (Dako, Carpinteria, CA, X0909) to tissue sections for 30 min at room temperature. Primary antibodies were applied in Dako Antibody Diluent (Dako, S0809) and incubated for 2h at room temperature. Dilutions were as follows: CD326/EpCAM (1:250, clone 9C4, BioLegend, San Diego, CA), CD44 (1:250, clone IM7, BioLegend), CD24 (1:100, clone ML5, Biolegend), Lysozyme (1:500, Diagnostic Biosystems,

Pleasanton, CA), Mucin2 (1:100, Santa Cruz Biotechnology, Santa Cruz, CA), Sucrase Isomaltase A-17 (1:100, Santa Cruz), ChromograninA (1:500, Immunostar, Hudson, WI). Anti-Rabbit-Cy3 (1:500 Sigma, St. Louis, MO, C2306) and anti-Rat-Cy3 (1:500 Jackson Immunoresearch, Carlsbad, CA, 112-165-003) secondary detection antibodies were diluted in Dako Antibody Diluent and applied to tissue for 30 min at room temperature. Nuclei were stained for 10 minutes with bisbenzamide (1:1,000, Sigma) diluted in PBS. Background staining was negligible as determined by nonspecific IgG staining. Images were collected by capturing ~1 μm optical sections using a Zeiss LSM 710 confocal microscope.

Flow Cytometry/FACS. Cells were stained for 90 min on ice in ISC Staining Media [Advanced DMEM/F12 (Gibco), N2 (Gibco), B27 (Gibco), Glutamax (Gibco), Penicillin/Streptomycin (Gibco), 10μM Y27632 (Selleck Chemicals, Houston, TX), 500mM N-acetyl-cysteine (Sigma), and 10% FBS (Gemini Biosciences)] with the following antibodies or isotype controls: AlexaFluor 647 (Alexa647) conjugated anti-CD24 [1:100] (clone ML5, Biolegend #311109, San Diego, CA), Pycoerythrin-Cy7 (PE-Cy7) conjugated anti-CD45 [1.6:100] (Biolegend #304015), Fluorescein isothiocyanate (FITC) conjugated anti-EpCAM [4:100](clone 9C4, Biolegend #324203), and Brilliant Violet 421 (BV421) conjugated anti-CD44 [0.6:100] (clone IM7, Biolegend #103039). For initial analysis by flow cytometry, 1 $x10^6$ cells were stained, fixed in 2% PFA for 10 min at room temperature, then rinsed (with PBS) and resupended in 500 μL of PBS for analysis using a Beckman-Coulter (Dako) CyAn ADP (Figure 2.2). For sorting experiments, cells were rinsed and resuspended in ISC Sort/Culture Media [Advanced

DMEM/F12 (Gibco), N2 (Gibco), B27 (Gibco), Glutamax (Gibco), Penicillin/Streptomycin (Gibco), 10mM HEPES (Gibco), 10μM Y27632 (Selleck Chemicals), and 500mM N-acetyl-cysteine (Sigma)]. FACS was conducted using an iCyt Reflection (Visionary BioScience) for RNA collection or FACSAria (BD Biosciences, San Jose, CA) for cell culture experiments. Dead cells and debris were first excluded based on size via bivariate plot of forward scatter (FSC) vs. side scatter (SSC) (Supplemental Figure 3A). Doublets/multimers were excluded using a bivariate plot of FSC peak vs. FSC length (Supplemental Figure 3B). Epithelial cells were FACS-enriched by sorting EpCAM (CD326) positive, CD45 negative cells (Figure 2.3C). The remaining cell events were analyzed for CD24 and CD44 expression on a bivariate plot (Figure 2.3D). Five cell populations: CD45-EpCAM+(whole), CD45-EpCAM+CD24-CD44-(negative), CD45-EpCAM+CD24+CD44-(CD24+CD44-), CD45-EpCAM+CD24-CD44+(CD24-CD44+), CD45-EpCAM+CD24+CD44+(CD24+CD44+) were collected directly into 500 μl of RNA lysis buffer (Ambion RNAqueous Micro, Grand Island, NY) for gene expression analysis. For cell culture experiments, cells were collected into 500μL of ISC Staining Media.

Primary isolation of human small intestinal myofibroblasts. Following mucosectomy for epithelial prep, remnant submucosa (~4cm x 4cm) was diced with a sterile razor blade. To eliminate blood cells and remnant epithelium, tissue chunks were washed in 20mL DMEM (Gibco), shaken for 2 minutes, and then allowed to settle before supernatant was removed and discarded. This process was repeated 8 times. The submucosal tissue was then resuspended in 5mL DMEM containing Penicillin/Streptomycin, 0.3U/mL dispase,

and 300U/mL collagenase I (Sigma) and rotated for 25 min at room temperature. 10mL of DMEM +10%FBS was added to quench the reaction and the tissue suspension was pipetted vigorously ~50 times to further mechanically dissociate myofibroblasts. The tissue suspension was centrifuged at 300g for 5min and the resulting supernatant and tissue remnants were plated separately in DMEM +10%FBS. Media was changed every 24hrs. Cultures initiated from the supernatant of the prep produced myofibroblasts, which were passaged 3 times before use in ISC culture experiments.

Intestinal epithelial stem cell culture. Cells were pelleted and resuspended in ES-qualified Matrigel (BD Biosciences) containing ISC culture growth factors (see Table 2.1). For feeder-free cultures, 10μL Matrigel droplets were plated in 96-well plates and overlaid with 100μL ISC Sort/Culture Media, with or without 2.5μM CHIR99021 (Selleck Chemical) following 15 minute of polymerization at 37°C. For feeder co-culture, 25μL Matrigel droplets were allowed to polymerize in 12-well transwell inserts (BD Biosciences) before being placed in wells containing fibroblast feeder cells and 500μL ISC Sort/Culture Media. An additional 500μL of the same media was placed in the transwell to prevent drying of the Matrigel droplet. CHIR99021 was not used in co-culture experiments. To facilitate differentiation of enteroids, Wnt3a, SB202190, and nicotinamide were withdrawn at 14 days of culture, as previously described for whole crypt cultures[83].

cDNA preparation/real-time PCR analysis. mRNA from sorted cell populations was purified by using RNAqueous Micro Kit (Ambion) according to the manufacturer's

protocols. cDNA was generated using iTaq Reverse Transcription Supermix (Bio-Rad, Hercules, CA). Real-time PCR was conducted for each sample in triplicate on 1/20,000 of the total amount of cDNA generated. Taqman probes [*18S*, HS99999901; *CD24* Hs00273561_s1; *CD44*, Hs01075861_m1; *LGR5* Hs00173664_m1; *OLFM4* HS00197437_m1; *HOPX* Hs04188695_m1; *DEFA6* Hs00427001_m1; *LYZ* Hs00426232; *MUC2* Hs00159374_m1; *CHGA* Hs00900373_m1] for each gene were obtained from Applied Biosystems (Pleasanton, CA) and used in reactions according to the manufacturer's protocol.

Statistical Analysis. qRT-PCR data was normalized for the expression of 18S. ΔΔCt values were then calculated using the CD24-CD44- (negative) cell population as the comparator. Statistical analysis compared gene expression across all cell populations by gene for each patient via one-way ANOVA followed by Bonferroni post-test for multiple comparisons between the population of interest and all other populations. Statistical analysis was performed in Graph Pad Prism (ver 4.0, LaJolla, CA).

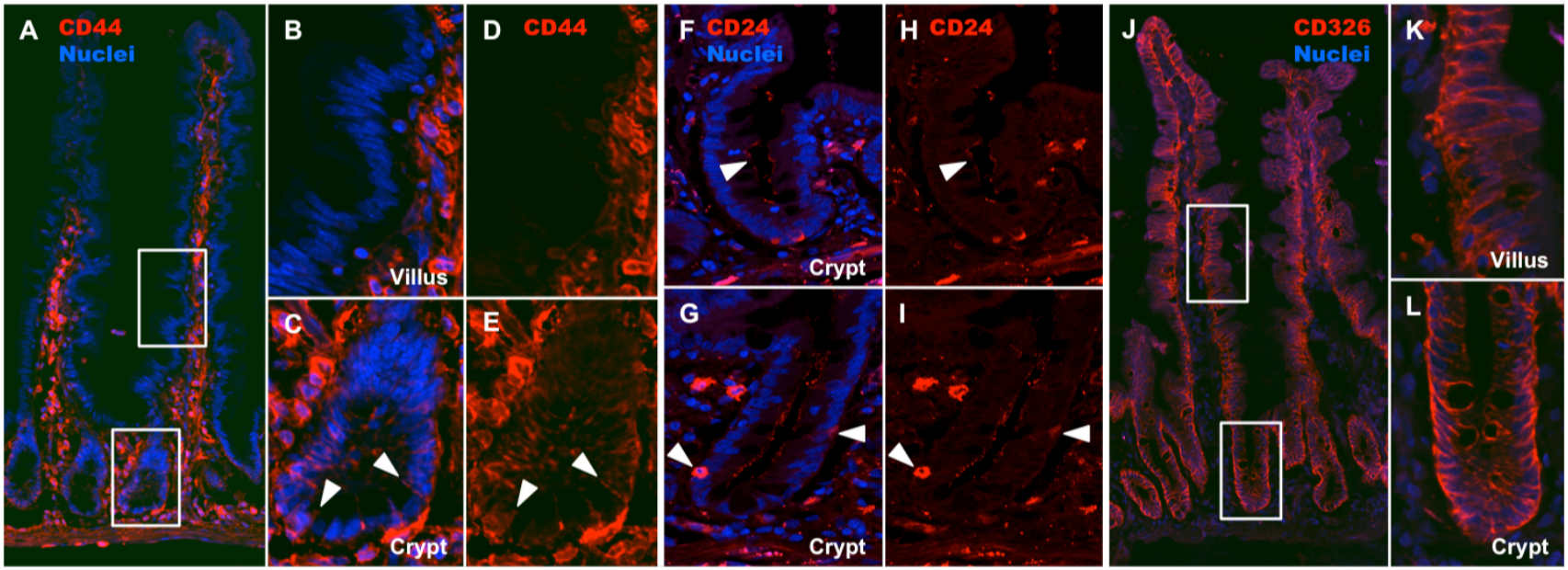


***Figure 2.1* – CD24 and CD44 are expressed in the stem cell zone of the human intestinal epithelium.** Epithelial expression of CD44 is restricted to the basolateral membranes of crypt cells (A, C, & E), with no expression in villus epithelium (B & D). CD24 is also restricted to the crypt, but is expressed on the apical membrane of epithelial cells (F & H, arrow). A subset of crypt-based epithelial cells expresses high levels of cytoplasmic CD24 (G & I, arrows). CD326 (EpCAM) is expressed throughout both the crypt and villus, but remains restricted to epithelial cells (J-L). Scale bars represent 50μm.

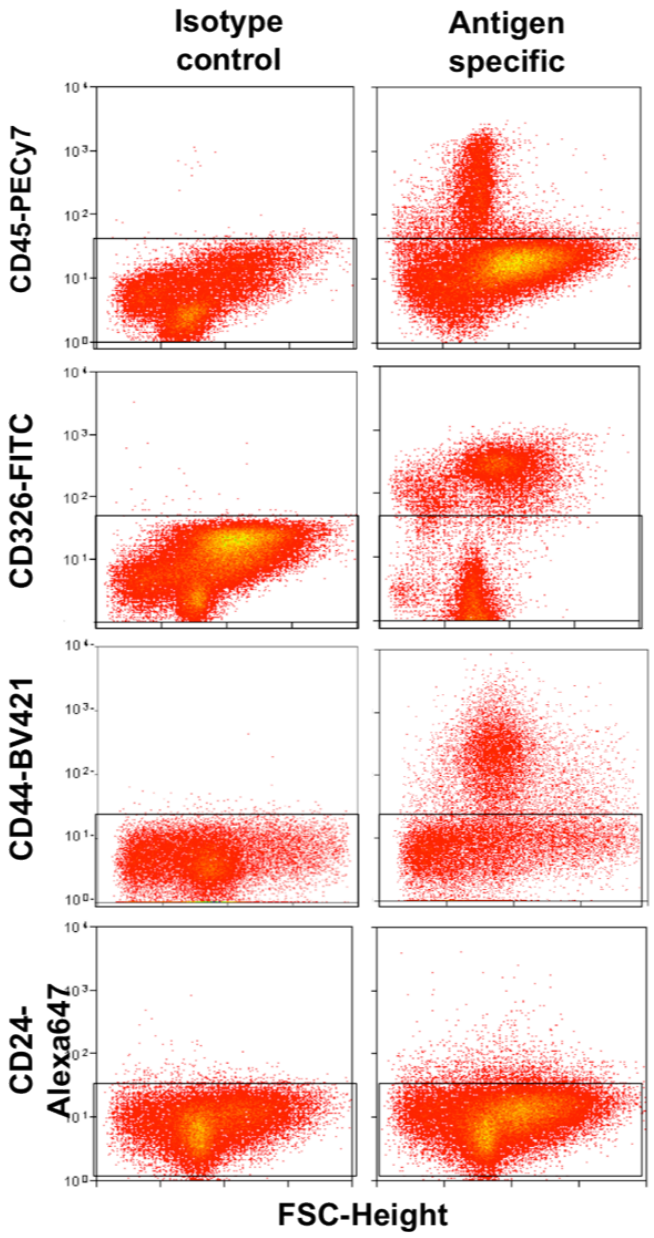
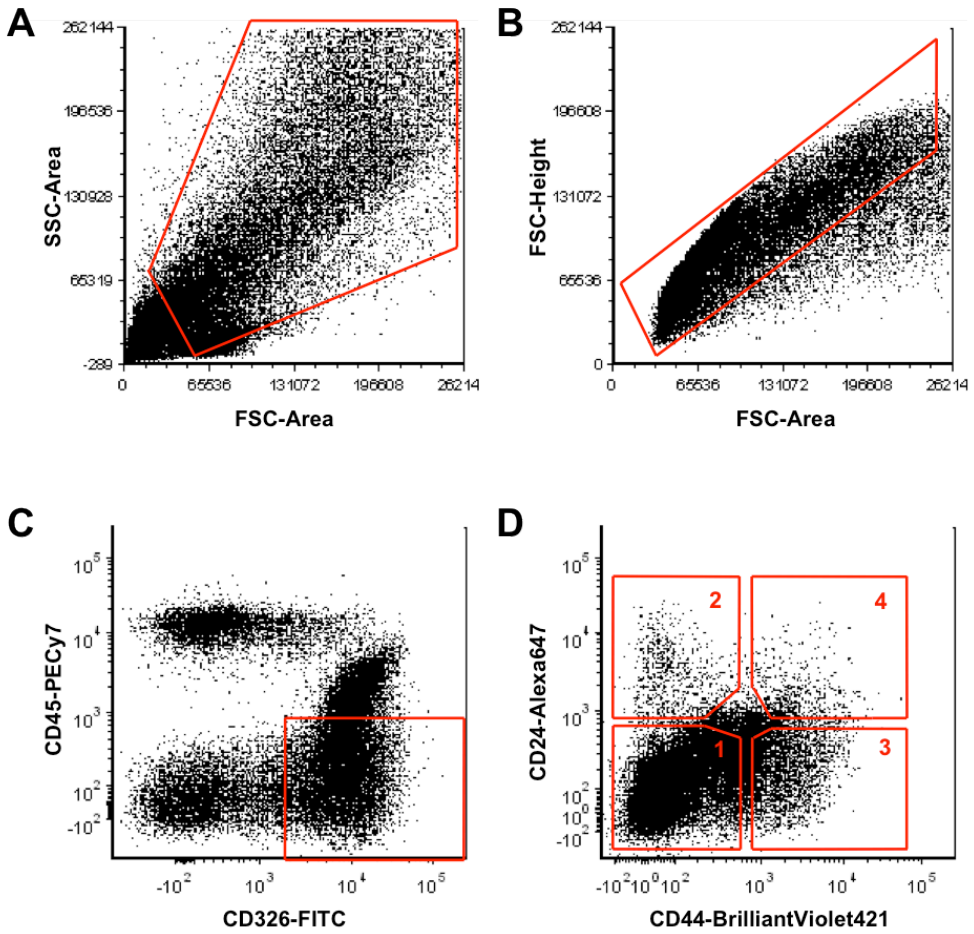


***Figure 2.2* – CD45, CD326, CD44, and CD24 antibodies label positive populations via flow cytometry.** Isotype controls and antibody staining by flow cytometry for CD45, EpCAM (CD326), CD44, and CD24 indicate positive staining of populations with negligible non-specific labeling.



E

Population	Percent collected cells
(1) CD24-/CD44-	61.4 ± 9.5
(2) CD24+/CD44-	11.5 ± 3.0
(3) CD24-/CD44+	21.2 ± 7.8
(4) CD24+/CD44+	5.9 ± 4.6

***Figure 2.3* – CD24 and CD44 differentially label human jejunal cells.** Gating strategy for FACS isolation of CD24/CD44+ epithelial populations: Removal of debris (A) Singlet gating (B) Selection of epithelial cells by CD45 exclusion and CD326 inclusion (C) Sample of gating strategy based on CD24 (y-axis) and CD44 (x-axis) (D). Composition of total jejunal epithelium by CD24/CD44 expression: table describes % of all gated cells in CD24/CD44 histogram by cell population ± SEM (n=3).

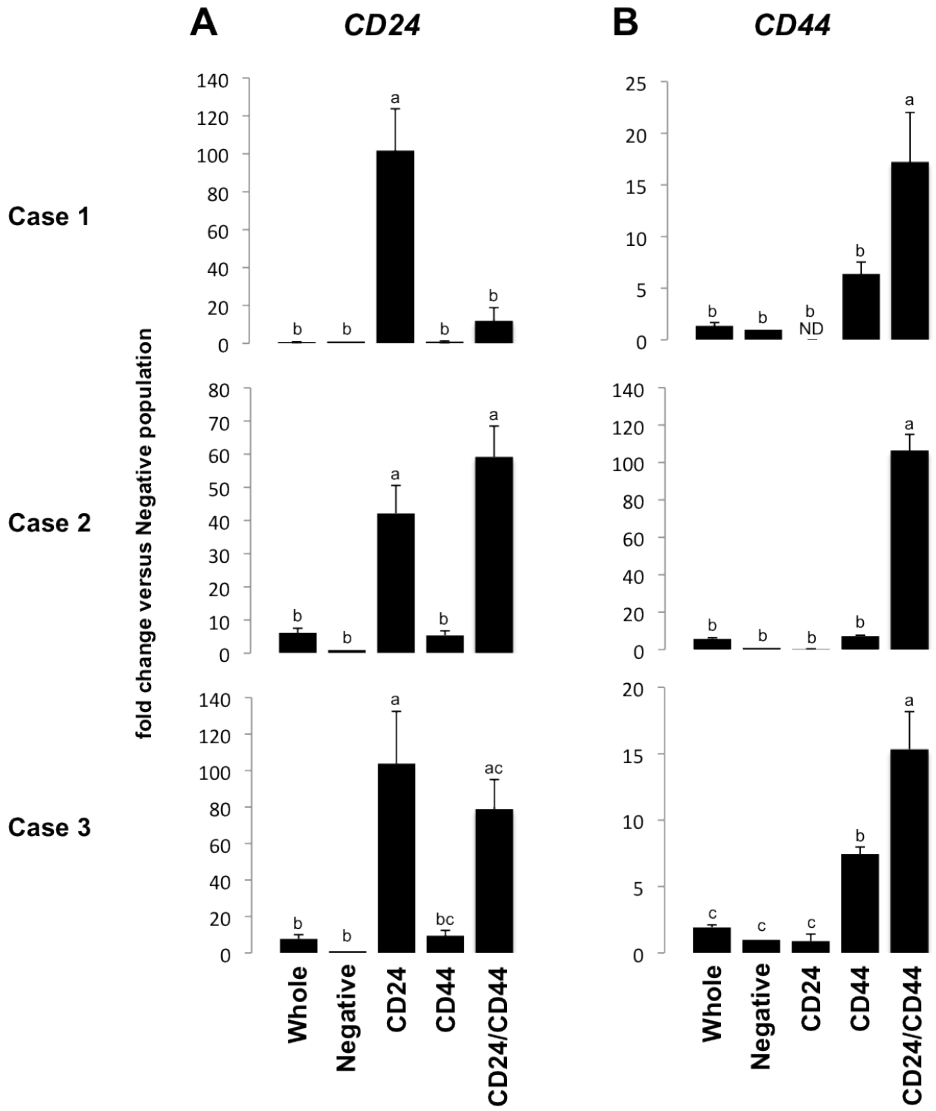


***Figure 2.4* – CD24/CD44 sorting strategy enriches for *CD24*- and *CD44*-positive populations.** qPCR for *CD24* (A) and *CD44* (B) demonstrates population-specific enrichment for each marker, validating FACS parameters. Letters a-c above each bar indicate data points that are statistically different from each other ($p < 0.05$).

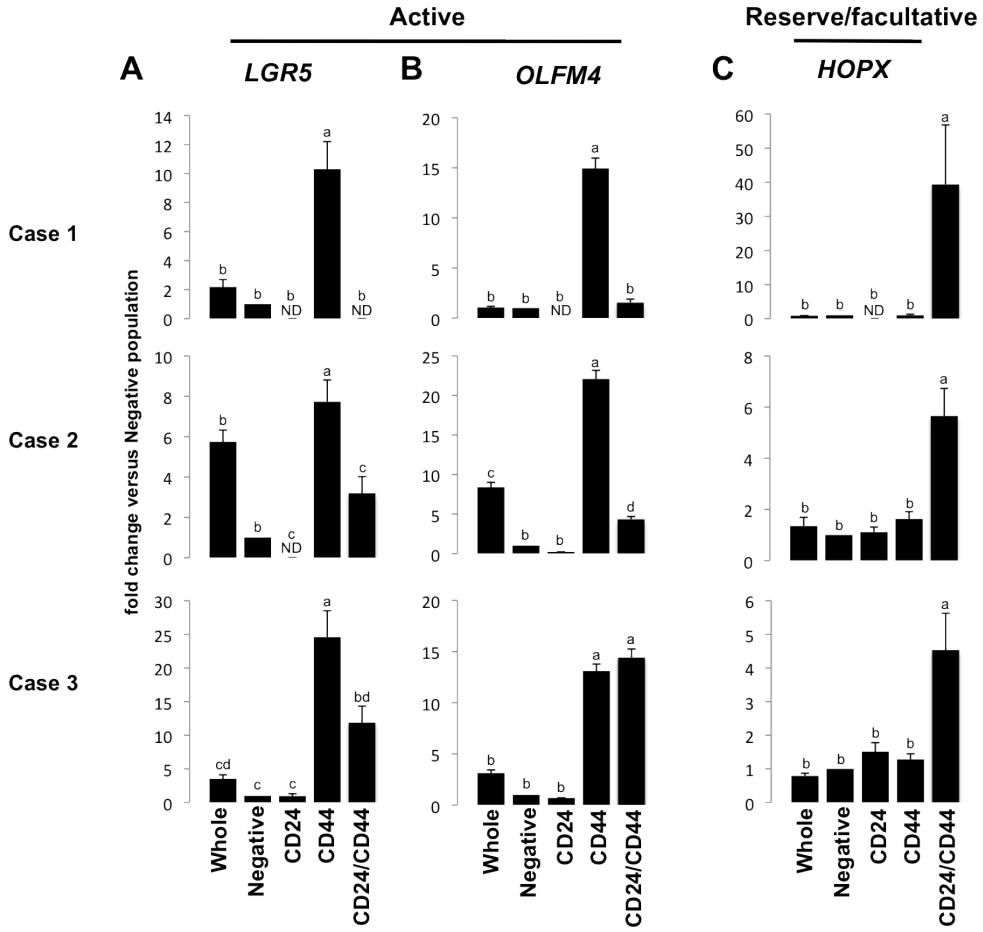


Figure 2.5 **- CD24-/CD44+ and CD24+/CD44+ intestinal epithelial cells are enriched for active and reserve/facultative markers, respectively**. CD24-/CD44+ cells are significantly enriched for active ISC markers (A) *LGR5* and (B) *OLFM4*, while CD24+/CD44+ cells exhibit enrichment for (C) *HOPX,* which is associated with reserve/facultative ISCs. Letters a-d above each bar indicate data points that are statistically different from each other ($p < 0.05$).

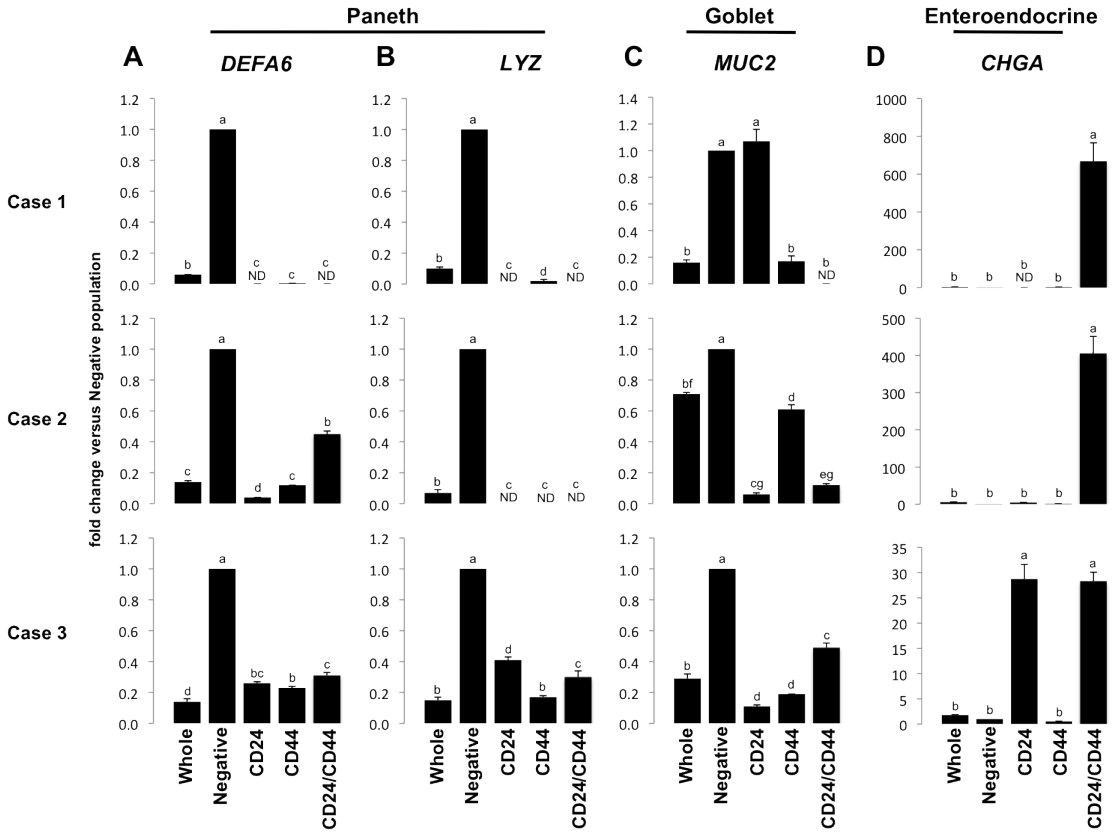


***Figure 2.6* - CD24/CD44 populations are de-enriched for Paneth and goblet cell markers.** Gene expression analysis demonstrates that CD24 and CD44 populations are significantly de-enriched for Paneth cell markers *DEFA6* (A) and *LYZ* (B) and demonstrate enriched expression of these markers in CD24/CD44 negative populations over whole epithelium. De-enrichment is also observed for goblet cell marker *MUC2* (C) in two of three cases, with no statistical difference between *MUC2* in CD24-/CD44- and CD24+/CD44- populations in Case 1, which may be attributable to patient-patient heterogeneity. High levels of enteroendocrine marker *CHGA* are observed specifically in the CD24+/CD44+ population in two of three cases, with significant enrichment in CD24+/CD44- and CD24+/CD44+ populations in Case 3 (D), consistent with reports in mice. Letters a-g above each bar indicate data points that are statistically different from each other ($p < 0.05$).

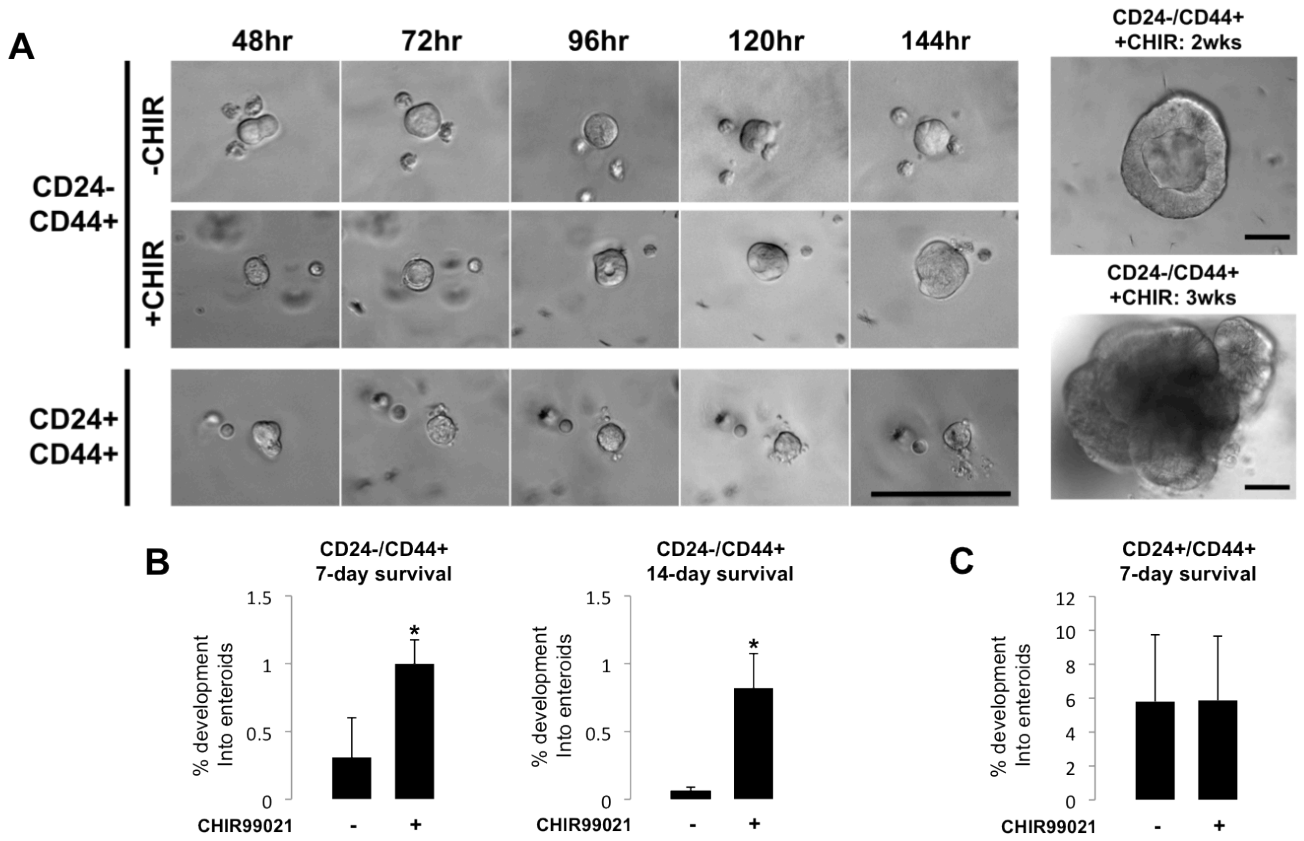


***Figure 2.7* - CD24-/CD44+ and CD24+/CD44+ populations generate enteroids *in vitro*.** By 48hrs, CD24-/CD44+ and CD24+/CD44+ cells form small enteroid structures, which increase in size over the first week of culture (A). GSK inhibition through a single dose of CHIR99021, given when cells are plated, significantly increases 7- and 14-day survival of enteroids derived from CD24-/CD44+ cells (B), but has no noticeable effect on the CD24+/CD44+ population (C). While CD24-/CD44+ cells form long-lived enteroids, the CD24+/CD44+ population does not demonstrate survival at or after 14 days in culture in feeder-free conditions (A). Student's t test was used to determine significance. Asterisks indicate a P-value of < 0.01. Scale bars represent 100μm.

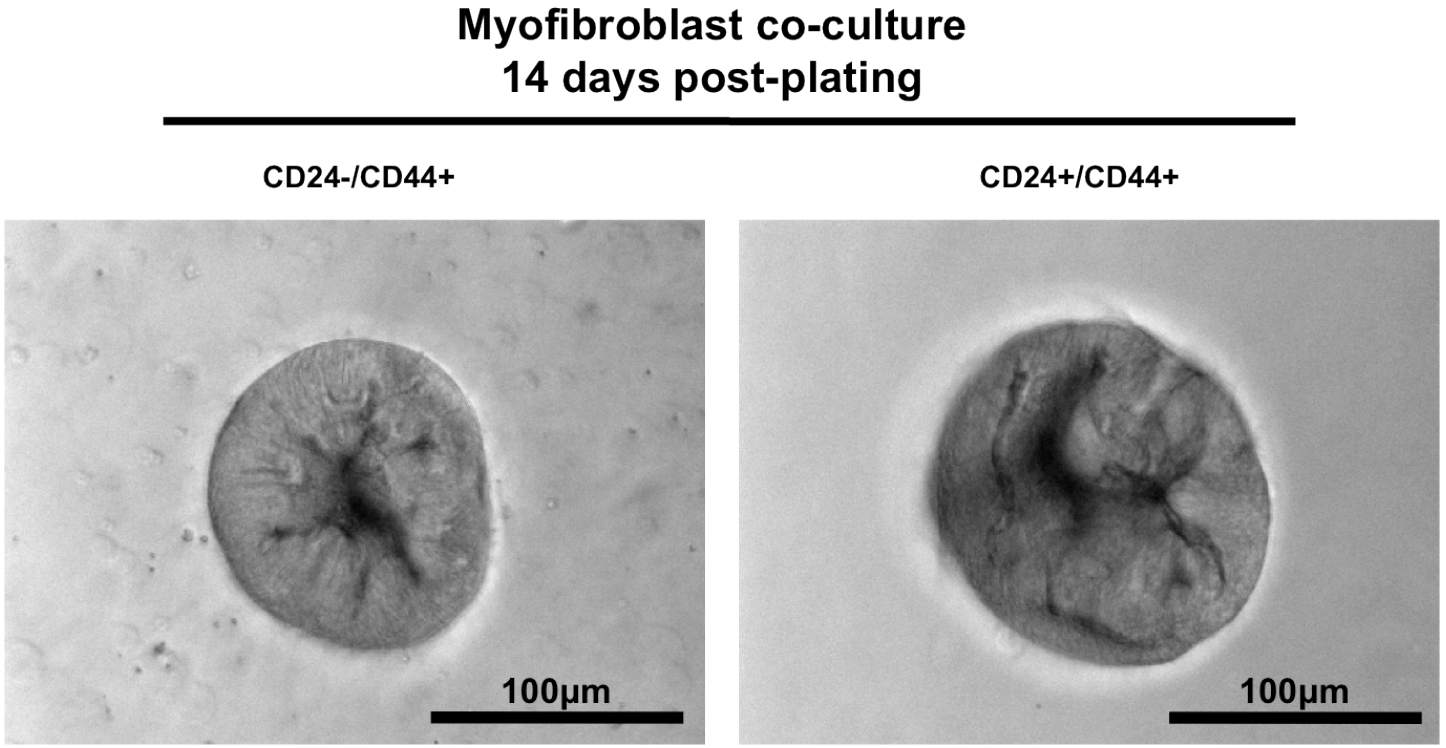


***Figure 2.8* - Co-culture with myofibroblasts induces long-lived enteroid formation in CD24-/CD44+ and CD24+/CD44+ populations.** Both *LGR5*-associated CD24-/CD44+ and *HOPX*-associated CD24+/CD44+ cells form enteroids when co-cultured on transwells with myofibroblasts isolated from human jejunal remnants. Enteroids are shown at 14 days post-plating.

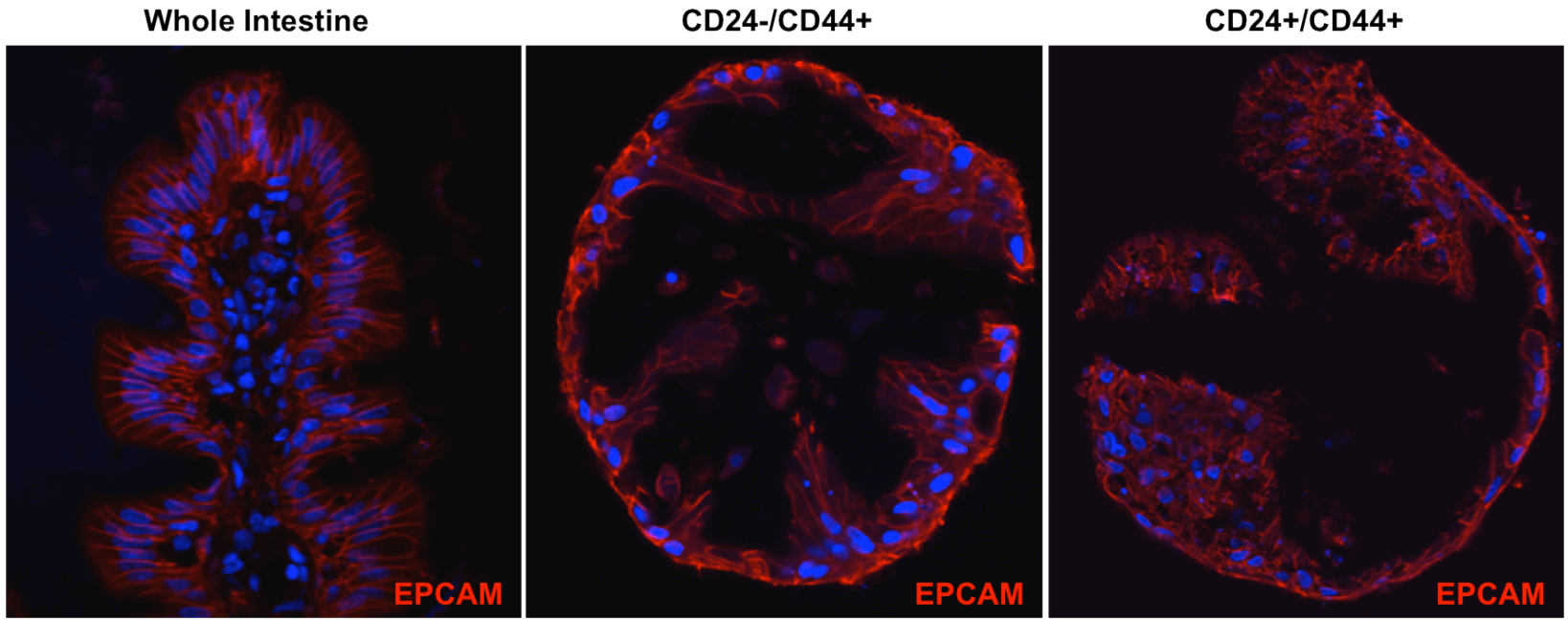


Figure 2.9 **- Human enteroids are composed of epithelial cells.** Immunofluorescence for EPCAM (CD326) demonstrates that enteroids derived from CD24-/CD44+ and CD24+/CD44+ populations are epithelial in nature. Scale bars represent 50μm.

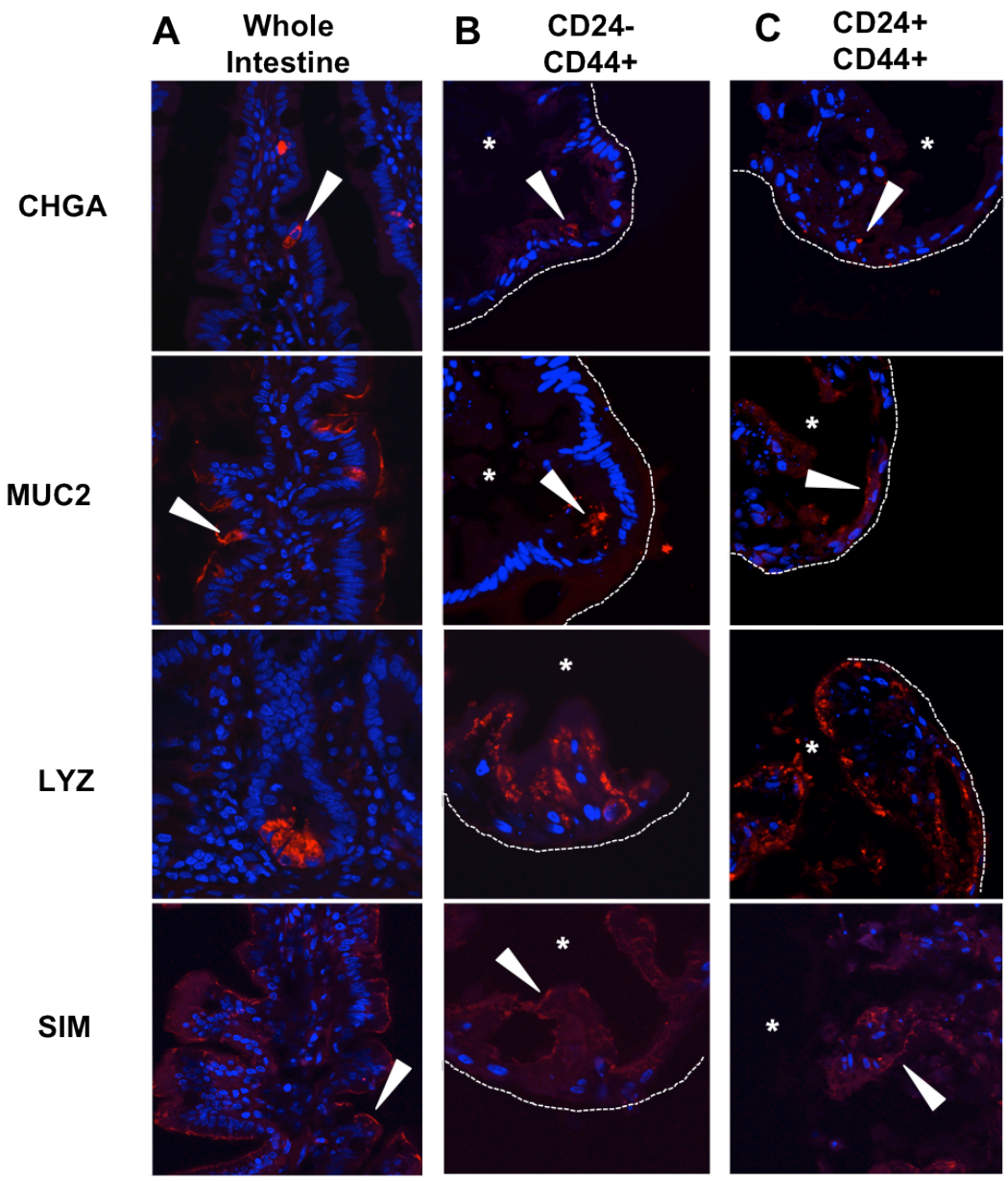


***Figure 2.10* - Enteroids derived from CD24-/CD44+ and CD24+/CD44+ populations are multipotent.** Enteroendocrine, goblet, and Paneth cells, as well as absorptive enterocytes, are detectable in whole human jejunum by expression of CHGA, MUC2, LYZ, and SIM, respectively (A). Similar expression patterns for each marker are observed by immunofluorescence in enteroids derived from the *LGR5*-associated CD24-/CD44+ population (B) and the *HOPX*-associated CD24+/CD44+ population (C). Arrows indicate positive staining, asterisks mark enteroid lumens, and dotted lines denote outer edge of enteroid structures. Scale bars represent 50μm.

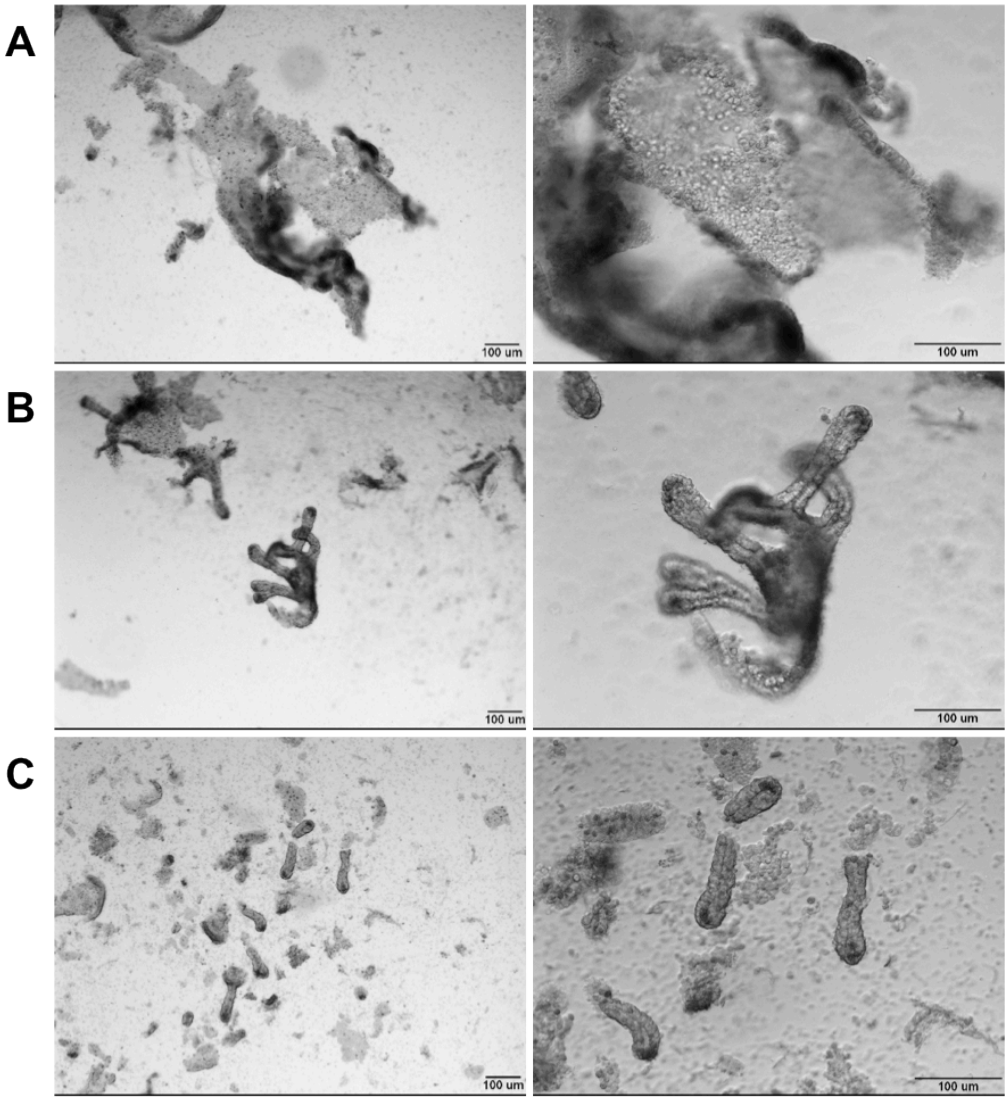


Figure 2.11 - **Epithelial preparations enrich for stem cell containing crypts.** Initial shaking steps remove large villi (A). After villus fractions are discarded, additional shaking yields both crypts and some remaining villi (B). Filtering results in single isolated crypts, with some debris and single cells from submucosal tissues (C). Scale bars represent 100μm.

		factor stock [c]	working [c] Day 0		working [c] Day 2+	working [c] Day 14+
In Matrigel	**EGF**	*1mg/mL*	*50ng/mL*	In media	*50ng/mL*	*50ng/mL*
	Noggin	*100ug/mL*	*100ng/mL*		*100ng/mL*	*100ng/mL*
	hRspo1	*250ug/mL*	*1ug/mL*		*500ng/ml*	*500ng/ml*
	hWnt3a	*40ug/mL*	*100ng/mL*		*100ng/mL*	-
	Jag1 peptide	*15mM*	*1uM*		-	-
	Gastrin	*100uM*	*10nM*		*10nM*	*10nM*
	LY2157299	*20mM*	*500nM*		-	-
	SB202190	*30mM*	*10uM*		*10uM*	-
					-	-
In media	**NAC**	*500mM*	*500uM*		-	-
	Nicotinamide	*2M*	*10mM*		*10mM*	-
	Y27632	*10mM*	*10uM*		-	-
	CHIR99021	*10mM*	*2.5uM or N/A*		-	-

***Table 2.1**:* **Culture conditions for human intestinal epithelial stem cell populations.**

CHAPTER 3

A HIGH THROUGHPUT, CLONAL PLATFORM FOR STUDYING THE INTESTINAL STEM CELL NICHE IN VITRO[3]

OVERVIEW

Rapid advances in intestinal stem cell (ISC) biology driven by biomarker identification and elegant *in vivo* lineage-tracing assays indicate that complex niche dynamics govern ISC multipotency and self-renewal. While the ISC niche is emerging as a key player in the regulation and maintenance of stemness in the intestine, current functional assays preclude statistically meaningful studies of single ISCs and ISC-niche interactions. Here we describe the development of an *in vitro* platform that facilitates high-throughput clonogenic culture and computational identification of isolated ISCs and niche cells in co-cultures at single cell and microscale (≤ 7 cells) resolution. This new *in vitro* platform has broad applicability in studies aimed at dissecting stem cell heterogeneity and stem cell-niche interactions at the single cell level.

[3] full list of contributing authors: Gracz AD, Williamson IA, Johnston MJ, Wang F, Wang Y. Attayek P, Balowski J, Liu XF, Laurenza RJ, Galanko JA, Sims CE, Li L, Allbritton NL, and Magness ST

INTRODUCTION

Throughout the life of the adult organism, somatic stem cells are responsible for maintaining tissue homeostasis following physiological "wear and tear" and damage from injury. Understanding how somatic stem cells self-renew and differentiate to produce the functional cells of their resident tissue is essential for determining the mechanisms underlying a broad range of issues related to human health and disease, including degenerative disease, cancer, and aging. Many stem cell types reside within specialized niches, where their behavior, including activity/quiescence, self-renewal, and differentiation, are likely influenced by extrinsic signaling factors. While the concept of stem cell niches was first applied to hematopoietic stem cells, the importance of niche cells has now been demonstrated for a number of mammalian tissues, including the brain, skeletal muscle, hair follicle, and intestine [46, 138]. As signals from the niche are thought to play a significant role in modulating stem cell fate decisions in physiology and disease, functional assays for studying stem cell response to experimentally modulated signaling are invaluable for understanding how stem cells behave in their native environments. However, functional stemness remains difficult to study in a statistically meaningful way *in vitro* for many adult tissue types.

The intestinal epithelium undergoes one of the most rapid rates of renewal of any mammalian tissue, making it an excellent model system for understanding stem cell

driven physiological renewal. Recent studies demonstrate that crypt-base columnar cells (CBCs), marked by high levels of *Lgr5* and low levels of *Sox9*, are capable of driving long term regeneration *in vivo* and formation of "enteroid" structures *in vitro*, both characteristics of functional ISCs [19-21, 24]. These CBC ISCs are closely associated with Paneth cells, which have been shown to function as niche cells and express soluble and insoluble factors associated with stemness, such as Wnt and Notch ligands [8, 62, 70]. Additionally, Paneth cells have been shown to increase the efficiency of enteroid formation by isolated ISCs *in vitro* [8].

Though the intestinal epithelium possesses many benefits as a model system for somatic stem cell biology, technical limitations hinder efficient and detailed functional and mechanistic studies of ISC-niche interactions. While colonic enteroids generated *in vitro* have been successfully transplanted into recipient mice, the field currently lacks a robust *in vivo* transplant assay to study single ISCs at a clonal level, a tool that has driven the understanding of stem cell niches in the hematopoietic system, mammary glands, and testis [139-141].

Recently, stem cell-niche interactions in endothelial and spermatogenic cell populations have been probed by "recapitulating" the endogenous niche environment *in vitro* through co-culturing strategies [142, 143]. Similar studies have examined ISC-Paneth cell interactions, both under normal conditions and following calorie restriction [8, 85]. However, these studies relied on the co-culture of hundreds of ISCs with hundreds of Paneth cells, and may not reflect physiologically normal conditions in the intestinal

crypts, where much smaller numbers of ISCs and Paneth cells (~15 per crypt) interact [8, 41, 144].

In the present study, we describe a novel platform to study large numbers of single ISCs simultaneously, either at the clonal level or in the presence of one or more Paneth cells. Microfabricated culture arrays modified for long-term 3-dimensional culture are used to capture and functionally assay single ISCs and clonally derived intestinal enteroids. Post-hoc computational identification of initial array contents coupled with day-to-day tracking of ISCs and ISC-Paneth cell co-cultures, through the use of unique addresses, facilitates the high-throughput analysis of ISC development into enteroid structures. To provide proof-of-principle for interrogating ISC-niche interactions, we co-culture ISCs with Paneth cells and assess enteroid formation from 15 different combinations of ISCs and Paneth cell numbers. Surprisingly, we were unable to demonstrate a statistically significant correlation between the presence of Paneth cells and the likelihood of ISCs to form enteroids. Together, our methodology provides a novel platform for high-throughput screening of primary ISCs and *in vitro* recapitulation of the ISC niche, and suggests that ISC-Paneth cell interaction may be subtler than previously thought.

RESULTS

Existing cell culture arrays are adaptable to long-term culture

Previous studies have utilized microfabricated devices consisting of arrayed reservoirs, or microwells, to conduct two-dimensional, adherent cell culture in a

microarray format [145]. We hypothesized that a similar approach could be utilized to isolate and culture single ISCs in three-dimensional extracellular matrix required for the growth of primary isolated ISCs. In the present study, we adapted a previously described platform, termed microwell arrays, designed for the isolation of viable single cells without advanced cell sorting techniques [145, 146]. The arrays are fabricated from polydimethylsiloxane (PDMS) using standard photolithography and composed of up to 12,000 microwells (Figure 3.1). As PDMS is known to be a poor substrate for some cell types, an additional element (termed microraft) is made from Petri dish grade polystyrene and embedded in each microwell of the array using dip-coating methods [145] (Figure 3.1D-F).

Since ISCs require several days to develop into enteroid structures, microraft arrays had to be amenable to media and growth factor changes [21, 24]. Additionally, separate media compartments were required in order to facilitate simultaneous examination of multiple experimental conditions on a single microraft array. To meet these requirements, optically clear polycarbonate cassettes, with dividers to create media reservoirs (termed culture chambers), were bonded to the microraft arrays (Figure 3.1H & Figure 3.2A). Cassettes were fabricated with two (~5,000 microwells per culture chamber) and four (~2,500 microwells per culture chamber) culture chambers (Figure 3.2B). Microwells were sized at 200μm^2, with 30μm spacers separating microwells from their immediate neighbors (Figure 3.2C). Microwells were designed to have a depth of 100μm to accommodate enteroid growth. Physical well addresses, stamped into

polystyrene rafts at 5 microwell intervals, were included in the array design to allow for tracking of unique enteroids across many time points (Figure 3.2C).

To facilitate high-throughput analysis, we used tile-scanning microscopy to produce high-resolution, high-magnification images of whole microraft arrays. Initial attempts revealed significant out-of-plane sagging in the z-axis, due to the elastic property of the PDMS array (Figure 3.2D&F). To prevent sagging of the array during imaging, the PDMS template was bonded to a glass slide by using a thin layer of polyacrylic acid (PAA) prior to attaching the cassette (Figure 3.1A). Microraft arrays bound to glass with PAA, which can be dissolved in PBS to release the array and attached cassette from the glass slide, were reproducibly imaged in a single Z-plane by tile-scanning without any noticeable out-of-focus sagging (n = 50) (Figure 3.2F&G).

Microraft arrays are biocompatible with long-term three-dimensional ISC culture

Next, we wished to validate that the PDMS/polystyrene microraft arrays were capable of supporting ISC growth and development into enteroids, without cytotoxic effects. To facilitate accurate localization of isolated ISCs in the microraft arrays, we crossed *Sox9EGFP* mice to *CAGdsRed* reporter mice, which express the *dsRed* fluorescent transgene ubiquitously across all cell and tissue types, and isolated $Sox9^{Low}$ ISCs from dual transgenic mice by fluorescence activated cell sorting (FACS) (Figure 3.3A) [21, 147]. Additionally, we reasoned that inclusion of a constitutive transgenic reporter in the form of *dsRed* fluorescence would allow us to identify any contaminating *GFP*-/*dsRed*+ non-ISC cell types that may have escaped exclusion from our FACS parameters

and could potentially influence functional outcomes. Isolated *Sox9*^{EGFP}low:*dsRed*+ ISCs were plated in a single culture chamber of a two-chamber microraft array, randomly seeded into microwells by centrifugation, and overlaid with Matrigel containing established growth factors mimicking the ISC niche (Figure 3.3B) [8, 84]. This plating strategy resulted in a random distribution of ISCs across individual microwells within the microraft array, with microwells containing one, multiple, or no ISCs (Figure 3.3C-E). Tile scanning of the microraft array in the *dsRed* wavelength immediately after plating and at 48hrs of culture revealed that most isolated ISCs had begun to produce primitive enteroids, indicative of biocompatibility with the arrays (Figure 3.3F-H).

Conventional Matrigel-based ISC cultures are capable of supporting enteroid growth for many weeks [24, 83]. To ask if our microwell array platform was compatible with very long-term culture of enteroids, we maintained cultures for extended periods of time. Remarkably, enteroids identified at 120hrs after plating were found in the same microwells at late culture time points, demonstrating the applicability of address-based enteroid identification in long-term culture experiments (Figure 3.3I). At 8 weeks, enteroids from isolated ISCs had grown into large structures containing many crypt buds and began to spread over the array (Figure 3.3J). Together, these observations demonstrate the applicability of the microwell array platform in the long-term culture of isolated ISCs, and its potential as a platform for high-throughput chemical screens or tissue engineering.

Post-hoc image analysis enables high-throughput identification of single ISCs

Closer examination of initial microwell contents in microwell arrays containing freshly plated ISCs revealed a number of different content combinations visible in the dsRed wavelength. In addition to wells containing isolated ISCs, we observed wells containing a combination of ISCs and debris, debris only, as well as some wells containing fluorescent artifacts, either as a result of imaging artifacts, or from glare generated by the dsRed signal of cells in neighboring microwells (Figure 3.4A). Due to the random nature of the ISC plating process in the microwell array platform, we sought to develop a semi-automated computational pipeline to achieve the following analytical goals: 1) the identification of microwells containing ISCs, 2) the exclusion of empty microwells, 3) the exclusion of microwells containing debris or fluorescence glare/imaging artifacts, and 4) the quantification of ISCs per microwell in microwells that pass exclusion criteria. We divided the pipeline for plating and identification of initial microwell contents into five steps: 1) ISC isolation and plating, 2) whole-array image acquisition by tile scan imaging and image stitching, 3) segmenting of the processed image to individual microwell images with unique addresses, 4) analysis of microwell contents using the open-source image analysis package CellProfiler 2.0, and 5) reconciliation of CellProfiler data to identify and quantify initial microwell contents (Figure 3.4B&C).

To establish parameters for cell identification and debris discrimination, we randomly selected 500 individual microwell images containing single ISCs and 500 individual microwell images containing debris or fluorescent noise from a two-chamber microwell array containing *Sox9*^{EGFP}low:*dsRed*+ ISCs. All analysis was conducted on the

dsRed wavelength, since the CAG promoter produced a significantly stronger fluorescent signal than the *Sox9*^{EGFP} BAC transgene. These images were subjected to analysis in CellProfiler 2.0, an open source software platform that facilitates computational image analysis and can be "trained" to function in specific applications through modulation of a number of filtering and analytical parameters [148]. We focused on three specific parameters to develop an analytical pipeline in CellProfiler to identify ISCs: compactness, form factor, and eccentricity. After applying standard image thresholding, CellProfiler identified objects in each microwell in our panel of 1000 images, and we found that ISCs had a distribution of values that was distinct from debris/noise for each of our parameters of interest (Figure 3.4D-F). We used these values to design a CellProfiler "pipeline" that identified objects post-thresholding, measured object size and shape, determined values for compactness, form factor, and eccentricity, and then filtered objects as ISCs or debris based on our established parameters for ISCs (Figure 3.4G). Finally, we reconciled the number of objects identified pre- and post-filtering to determine which microwells contained debris and which contained only ISCs (Figure 3.4G).

In order to validate our ISC identification pipeline using unbiased microwell images, we manually scored the contents of 2,254 randomly selected microwells. This panel of images was then subjected to analysis using our newly developed pipeline, which was able to identify initial microwell contents with a high degree of accuracy, especially for microwells containing a single ISC (99.87% accurate) (Figure 3.4H). Due

to stringency settings, this accuracy was reduced for wells containing multiple cells.

Sox9^{EGFP}negative:CD24^{high} cells represent a highly enriched Paneth cell population

Previous studies have isolated Paneth cells from whole intestinal epithelium by FACS of CD24^{High}:SSC^{High} populations, which have been shown to be enriched for Paneth cell markers, including lysozyme [8, 85]. However, CD24 is expressed throughout the intestinal crypts and is known to enrich for ISCs and TAs in murine and human tissue, opening the possibility of contaminating proliferative cells in preparations relying on CD24^{High}:SSC^{High} sorting parameters [21, 30, 84, 134, 149].

Since the current study is aimed at examining ISCs and ISC-niche interactions at the single cell level, we wished to refine isolation procedures for Paneth cells in order to meet the purity requirements of clonal and microscale co-cultures. Interestingly, our previous characterization of a *Sox9^{EGFP}* BAC transgenic mouse model revealed that, while endogenous SOX9 marks all cells in the crypt, including Paneth cells, the *EGFP* transgene is preferentially silenced in all Paneth cells, perhaps due to unknown *Sox9* regulatory differences between ISCs/TAs and Paneth cells [35]. We reasoned that we could exploit this property to isolate a highly pure population of Paneth cells from *Sox9^{EGFP}* mice. To assess relative purity of different Paneth cell isolation methods, we FACS isolated putative Paneth cells using conventional CD24^{High}:SSC^{High} parameters (Figure 3.5A&B). We also isolated the same population with the additional exclusion of all

Sox9^{EGFP}-positive cells, which should eliminate any contaminating, non-Paneth CD24-positive cells (Figure 3.5A&C). Gene expression analysis revealed a consistent 2-fold increase in lysozyme expression in Paneth cell populations isolated using *Sox9*^{EGFP} exclusion, relative to conventionally-isolated Paneth cell populations (Figure 3.5D). More importantly, we noted a nearly 5-fold decrease in *Lgr5* expression and a >10-fold decrease in *Chga* in *Sox9*^{EGFP}-excluded Paneth cells, indicating further de-enrichment of ISCs and enteroendocrine cells, respectively (Figure 3.5D). Muc2 expression, associated with both goblet and Paneth cells, was unchanged between the two sorting methods (Figure 3.5D). Over the course of our co-culture experiments, we examined 2,810 individual Paneth cells *in vitro*, and only once observed enteroid production by a putative Paneth cell isolated using *Sox9*^{EGFP}-exclusion parameters, resulting in a functional contamination rate of 0.04%, and further supporting the high purity of Paneth cell populations isolated using *Sox9*^{EGFP}neg: CD24^{High}:SSC^{High} criteria.

Paneth cells do not significantly influence ISC survival at low cell numbers in vitro

To establish the utility of the microwell array platform for ISC-niche experiments, we co-cultured isolated ISCs and *Sox9*^{EGFP}-excluded Paneth cells and assessed the impact of Paneth cell presence in microwells on the ability of ISCs to form enteroids over 120hrs *in vitro*. In order to confirm that enteroids were formed from *bona fide* ISCs and not contaminating ISCs present in Paneth cell preparations, we isolated ISCs from *Sox9*^{EGFP} or *Lgr5*^{EGFP} reporter mice and Paneth cells from *Sox9*^{EGFP}:*CAG*^{dsRed} dual-positive mice (Figure 3.6A). Under this strategy, any enteroids produced by contaminating cells from

the Paneth cell preparation would be dsRed+, and excluded from analysis (this was only observed once per 2,810 single Paneth cells, as per above).

Enteroid survival was determined by visual scoring of individual microwell addresses at 120hrs and subsequent reconciliation with initial well contents. The distribution of ISCs and Paneth cells across the microwell arrays allowed us to examine microwells with initial contents consisting of any combination of 1-5 ISCs and 0-2 PCs. In order to increase the *n* per ISC-Paneth cell combination, we conducted two technical replicate experiments for each ISC biomarker (Figure 3.6B&C). Examination of survival percentages across all combinations of ISCs and Paneth cells revealed no statistically significant or visually apparent trends, regardless of whether ISCs were isolated using $Sox9^{EGFP}$ or $Lgr5^{EGFP}$ (Figure 3.6B&C). Surprisingly, comparison between some ISC-Paneth cell combinations revealed seemingly paradoxical results. For example, in one of our $Sox9^{EGFP}$low-Paneth cell co-culture replicates, the presence of a single Paneth cell appeared to decrease percent survival compared to microwells containing a single ISC alone (Figure 3.6B). To investigate the overall effect of Paneth cells on enteroid survival, we next analyzed the percentage of enteroids formed in microwells containing either ISCs alone or ISCs and Paneth cells, regardless of the number of each cell type per microwell. These analyses also failed to produce statistically significant differences between ISC-only microwells and ISC-Paneth microwells, for both $Sox9^{EGFP}$low and $Lgr5^{EGFP}$high ISCs, despite examining a high number of microwell replicates per condition (Figure 3.6D). In contrast with previous reports, these data suggest that Paneth

cells are not sufficient to increase ISC development into enteroids *in vitro*, perhaps due to an insufficient number of Paneth cells in microwell-based cultures.

DISCUSSION

Array based culture platforms for studying ISCs

Array-based cell culture platforms are growing in use and popularity, and present an exciting, high-throughput, and cost-effective alternative to conventional cell culture. In stem cell biology, array-based systems have been applied most commonly to ES culture. These platforms often utilize printing methods adapted from gene expression microarray spotting technology to generate arrays of extracellular matrix (ECM) on glass slides, which are then seeded with ES cells and assayed for effects on proliferation and differentiation [150, 151]. Recent efforts have resulted in the development of advanced platforms for microarray-based cell culture that tether ECM molecules to biocompatible polymers, such as hydrogels, and have been applied to the culture of multiple stem cell types [152, 153]. However, these approaches are not amenable to long-term culture experiments, such as required for the development of single ISCs into enteroid structures [154].

In the present study, we adapt previously described microwell arrays to the long-term culture of primary ISCs, using established Matrigel-based culture conditions [8, 84]. To our knowledge, this study represents the first application of high-throughput culture

technology to the study of ISCs *in vitro*. While invaluable in its facilitation of *in vitro* studies on ISCs, the conventional culture system comes with limitations in terms of quantification of growth and survival. Previous reports have noted that isolated ISCs can merge to form single enteroids *in vitro*, raising the possibility that the merger of multiple ISCs into single enteroids can artificially influence survival data [8]. Additionally, the inability to follow a statistically meaningful number of individual enteroids over several days in conventional Matrigel patties precludes studying more subtle effects on enteroid survival and morphology *in vitro*. Here, we demonstrate methodology compatible with following thousands of individual ISCs over multiple days and weeks. In addition to applications aimed at dissecting ISC function at the clonal level, the microwell array platform facilitates high-throughput screening of potential ISC mitogens and morphogens for their effect on enteroid formation and development.

Enhanced isolation of Paneth cells

In order to allow for the study of ISC-Paneth cell interactions at the single cell and microscale level, the present study required isolation of highly pure Paneth cell populations. Recent studies have enriched for Paneth cells through the use of the cell surface marker CD24, but previous work in our lab and others has demonstrated that CD24 is expressed on the apical membrane of all crypt-based epithelial cells [8, 21, 30, 84, 85, 134, 149]. Although CD24 is expressed at varying levels across the intestinal crypts, and Paneth cells are further discriminated by increased SSC values, sorting on these parameters alone increases the probability of contamination with non-Paneth cell populations. In our hands, cultured cells from $CD24^{High}$:SSC^{High} populations routinely

produced enteroids in the absence of co-cultured, selected ISCs (data not shown). Here, we demonstrate that the further exclusion of the *Sox9*^{EGFP} transgene increases Paneth cell purity by both gene expression analysis and lack of enteroid-forming ability *in vitro*. This technique could be especially valuable to the direct study of Paneth cells in intestinal physiology and disease, as transgenic mice for the isolation of Paneth cells based on genetic markers are unavailable at this time.

Effects of Paneth cell co-culture on ISC enteroid-forming ability

The most striking result of our study is the observation that Paneth cells do not increase ISC survival in the microwell array platform. One key difference between our study and prior studies that demonstrate a positive correlation between Paneth cell co-culture and enteroid formation by ISCs, is that previous results have relied on the co-culture of hundreds of cells in a single Matrigel patty. In the present study, we examine the impact of no more than 2 Paneth cells on as few as 1 to as many as 5 ISCs. One possible explanation for this result is that large numbers of Paneth cells are required in order to elicit an impact on ISCs *in vitro*. *In vivo* gene expression analysis convincingly demonstrates that Paneth cells express critical ISC niche signaling components, including Wnt and Notch ligands [8, 62]. However, it is clear that niche signals, especially critical Wnt ligands, are also produced by non-epithelial sources, including subepithelial myofibroblasts [62, 70]. Additionally, the effect of the timing of Paneth cell-produced Wnt signaling on ISC growth *in vitro* is unknown. Unlike the addition of exogenous growth factors, *Wnt3a* produced by primary Paneth cells is likely secreted downstream of stochastic transcriptional bursting, resulting in waves of ligand production, as opposed to

constant secretion [155, 156]. This would explain seemingly random patterns of survival in the microwell array platform, where ISCs would be exposed to varying levels of Paneth cell-produced ligands, the production of which would be the product of stochastic, cell-autonomous mechanisms intrinsic to individual Paneth cells. In conventional cultures, the presence of hundreds of Paneth cells could result in an "averaging" effect, where production of ligands by some Paneth cells results in a steady state concentration of niche signals, which does not necessarily recapitulate physiological conditions.

Also surprising was the observation that increased numbers of ISCs per microwell at initial time points did not positively correlate with increased enteroid formation at 120hrs. These data suggest that one of two possibilities. First, is that current ISC biomarkers are more broadly expressed than initially appreciated, and mark cells with and without genotypic ISC potential. It is also possible that apparent randomness of enteroid formation by clonal ISCs reflects the relative rareness and inherent stochasticity of functional stemness. These possibilities are both supported by recent observations that potency within the ISC niche exists as a complex, shared state between multiple cell populations and that ISC-driven clonality is the byproduct of neutral drift dynamics, where ISCs are often lost to symmetric differentiation-oriented divisions [34, 37, 39, 41]. Stemness is defined as the functional ability of a cell to both self-renew and produce all differentiated lineages of its resident tissue, though the latter is a characteristic shared by some early progenitors that are only distinguished by their lack of self-renewal. Emerging data from *in vivo* studies demonstrates that self-renewal and clonal persistence is a characteristic restricted to very small numbers of putative ISCs [44]. Since the *in vitro*

culture system for ISCs is a *bona fide* functional assay of self-renewal, it is also possible that our observations mirror similar findings *in vivo*.

In summary, the present study establishes methods for the high-throughput culture and analysis of ISCs down to clonal resolution. Additionally, we demonstrate a novel method for the high purity isolation of Paneth cell populations using previously characterized *Sox9*^{EGFP} transgenic mice. We also demonstrate that Paneth cells do not support enteroid formation by ISCs at low cell numbers, and that ISC survival is not correlated with the number of ISCs present in a single microwell. Together these data suggest that Paneth cells are not sufficient to influence ISC growth at physiologically relevant numbers *in vitro*, and support emerging evidence suggesting that ISC self-renewal in response to extrinsic signaling is reliant on stochastic intrinsic regulation. Further studies examining the response of ISCs to other niche cells, such as subepithelial myofibroblasts, will be facilitated by continued use of the microwell array platform.

MATERIALS AND METHODS

Fabrication of glass mounted microwell arrays

Materials. SU-8 photoresist was purchased from MicroChem Corp. (Newton, Ma). The Sylgard 184 silicone elastomer kit was purchased from Dow Corning (Midland, MI). γ-Butyrolactone (GBL), octyltrichlorosilane, and propylene glycol monomethyl ether acetate were obtained from Sigma-Aldrich (St Luois, MO). Poly(acrylic acid) (MW ~5,000) was obtained from Polysciences, Inc. (Warrington, PA). Falcon™ Petri dishes were obtained from BD Biosciences (San Jose, CA).

Polycarbonate plates (12 inch × 12 inch × 0.5 inch) were purchased from McMaster-Carr (Los Angeles, CA) and glass slides (75 mm × 50 mm × 1 mm) were purchased from Corning Inc. (Corning, NY).

Fabrication of glass-backed PDMS microwell array. An SU-8 master mold with an array of microposts was fabricated using standard photolithography with 100 μm thick SU-8 as described previously [157]. The SU-8 master was coated by octyltrichlorosilane using vapor deposition to render the surface of the master non-sticky to PDMS [145]. Clean glass slides (75 mm × 50 mm × 1 mm) were spin coated with a 15-μm thick layer of poly(acrylic acid) (PAA) by using 50 wt% solution and a spin speed of 2000 rpm, followed by a 1 hr bake on a 100°C hotplate to remove the water. The slide coated in PAA was treated in a plasma cleaner for 10 min (Harrick Plasma, Ithaca, NY). PDMS prepolymer (10 : 1 mixture of base : curing-agent of Sylgard 184 kit) was spread on the SU-8 master mold and degassed under vacuum to remove trapped air bubbles. To control the thickness of the PDMS mold, 300-μm PDMS spacers were placed on both ends of the master mold. PDMS spacers were created by spin-coating glass with PDMS prepolymer at 200 rpm for 30 s and cured on a 120°C hotplate for 30 min. The plasma-treated PAA-coated glass slide was then placed on the master mold, flattening the PDMS prepolymer between the master mold and the glass slide (Fig. 3.1-A). The ends of the master-glass slide assembly were secured by paper clips to prevent movement during curing of the PDMS. The PDMS was cured in a 95°C oven for 1 hr (Fig. 3.1-B). The glass-backed PDMS microwell array was then obtained by separating the array from the silanized master mold (Fig. 3.1-C). The resulting microwell array has an array area of 25.4 mm ×

25.4 mm, and each microwell has a dimension of 200 μm × 200 μm × 100 μm spaced 30 μm apart. The PAA layer can later be dissolved in PBS buffer to release the PDMS microwell array from the glass backing.

Micromolding of microwell arrays by dip coating. A polystyrene solution was prepared by dissolving polystyrene Falcon™ Petri dishes in GBL at 20 wt% concentration. The polystyrene solution was spread onto the PDMS microwell array and degassed under vacuum to remove trapped air bubbles (Fig. 3.1-D). The glass-backed PDMS array was then immersed in the polystyrene solution and withdrawn vertically at a speed of 0.83 mm/min by a stepper motor controlled by a custom controller (Fig. 3.1-E). Due to surface tension difference between PDMS and polystyrene solution, polystyrene solution dewetted from the array, resulted in isolated pockets of polystyrene solution in microwells (Fig. 3.1-F). This phenomenon is called discontinuous dewetting [158]. The array was then placed in a 95°C oven overnight to evaporate the GBL solvent, and pockets of polystyrene solution shrunk into solid polystyrene microrafts (Fig. 3.1-G).

Attaching the array to the cassettes. The array was attached to a rigid 2-chamber or 4-chamber polycarbonate cassette, which was fabricated by computer numerical control (CNC) machine (Fig. 3.1-H). Prior to attachment, the cassette was cleaned by sonication in a solution of 1 wt% detergent in water for 1 hr, followed by a 1 hr sonication in a 75% ethanol solution. The cassette was rinsed thoroughly in DI water and baked for 15 min in a 95°C oven to remove any remaining solvent. Both the array and the cassette were treated in plasma cleaner for 10 min before being glued together with

PDMS (cured in a 95°C oven for 1 hr). The arrays were treated with air plasma for 5 min, sterilized with 75% ethanol, and stored in sterile 1X PBS prior to use.

Mice and genotyping

Sox9^{EGFP} transgenic mice were originally generated by the GENSAT Brain Atlas Project, and are maintained on an outbred CD-1 background [159]. *Lgr5*$^{EGFP\text{-}CreERT2}$ mice were obtained from Jackson Labs (stock number: 008875, Jackson Laboratory, Bar Harbor, ME) [19]. For microwell array experiments requiring constitutive expression of *dsRed* in isolated ISCs, heterozygous *Sox9*^{EGFP} mice were bred to homozygous *CAG*^{dsRed} mice to produce *Sox9*^{EGFP}:*CAG*^{dsRed} offspring heterozygous for both alleles. *Sox9*^{EGFP}:*CAG*^{dsRed} were phenotyped by examining tail snips taken at ~10 days postnatal for EGFP and DSRED fluorescence by epifluorescent microscopy. *Lgr5*$^{EGFP\text{-}CreERT2}$ mice were genotyped by previously described PCR protocols [19]. All experiments were conducted on adult mice between 8 and 16 weeks of age. All protocols for animal use were reviewed and approved by the University of North Carolina Institutional Animal Care and Use Committee.

Epithelial isolation and FACS

Epithelial cells were isolated from whole murine intestines, as previously described, with some modifications [160]. Briefly, intestines were opened longitudinally, rinsed in DPBS (Life Technologies, Grand Island, NY), minced, and incubated in 3mM EDTA (Sigma, St Louis, MO) in DPBS for 45min at 4°C with gentle agitation. Intestinal fragments were transferred to fresh DPBS and shaken by hand for 2 minutes to release

epithelium. Remnant submucosa was discarded, epithelium was rinsed twice in DPBS, and then dissociated to single cells by incubation in 0.3U/mL dispase (Life Technologies) in 10mL HBSS (Life Technologies) at 37°C for 8-10 min with shaking every 2 minutes. Single epithelial cells were filtered through 100, 70, and 40μm filters before being resuspended in ISC Sort/Culture Media [Advanced DMEM/F12 (Life Technologies), N2 (Life Technologies), B27 (Life Technologies), Glutamax (Life Technologies), Penicillin/Streptomycin (Life Technologies), 10mM HEPES (Life Technologies), 10μM Y27632 (Selleck Chemicals, Houston, TX), and 500mM N-acetyl-cysteine (Sigma)]. For Paneth cell isolation, cells were resuspended in ISC Sort/Culture Media with 5% FBS (Gemini Biosciences) and stained with the following antibodies: Brilliant Violet conjugated anti-CD24 (clone M1/69, Biolegend, San Diego, CA); PerCP-Cy5.5 conjugated anti-CD45 (clone 30-F11, Biolegend). All cells were stained with Sytox Blue (Life Technologies) or 7-AAD (Biolegend) and Annexin V-APC (Life Technologies) prior to FACS, for live-dead exclusion.

FACS was conducted on a MoFloXDP (Beckman Coulter,) and *Sox9*^{EGFP} and *Lgr5*^{EGFP} cells were isolated as previously described [19, 21, 160]. Paneth cells were isolated by high expression of CD24, high side-scatter (SSC), and exclusion of CD45, with or without exclusion of *Sox9*^{EGFP}. Cells were collected into ISC Sort/Culture Media for array culture and RNaqueous lysis buffer for qPCR analysis (Ambion RNaqueous Micro Kit, Life Technologies).

Plating ISCs in microraft arrays

Isolated ISCs and Paneth cells were plated in microraft arrays at ratios of ~1.5 cells: 1 microwell for ISCs and ~1 cell: 2 microwells for Paneth cells. Sorted cells were added to array reservoirs in ISC Sort/Culture Media and cells were seeded into microwells by centrifugation at 51g for 5min at 4°C. Following centrifugation, media was gently aspirated and arrays were overlaid with 600uL (2 reservoir array) or 200uL (4 reservoir array) Matrigel containing 1M JAGGED-1 peptide (AnaSpec, San Jose, CA), 50ng/ml EGF (R&D, Minneapolis, MN), 100ng/ml NOGGIN (Peprotech, Rocky Hill, NJ), and 1μg/ml R-SPONDIN1 (R&D). Arrays were centrifuged a second time at 51g for 5min at 4°C, to recapture any cells displaced by the addition of Matrigel. Matrigel was allowed to polymerize for 30min at 37°C before being overlaid with 1mL (2 reservoir array) or 600uL (4 reservoir array) ISC Sort/Culture Media. Growth factors, minus JAGGED-1 peptide and Y27632, were supplemented every two days, and media was changed every 4 days, as previously described [8, 24].

Image acquisition, stitching, and segmenting

Microwell arrays were tile-scanned in brightfield, GFP, and dsRed wavelengths using an automated stage and the Scan Slide function in the Metamorph Imaging Suite (Molecular Devices, Sunnydale, CA) immediately after plating and overlaying media. Arrays were housed in a physiological chamber mounted on a fluorescent microscope during imaging, in order to prevent cell death due to the imaging procedure. Scanned images were stitched into a single composite image using the open source image analysis

suite FIJI [161] and then segmented into address-associated individual well images using an algorithm designed in MATLAB (MathWorks, Natick, MA).

CellProfiler 2.0 analysis

Computational analysis of initial microwell contents was carried out in CellProfiler 2.0 (Broad Institute, Cambridge, MA) in two steps: (1) Segmented microwell images were loaded in sequence into CellProfiler to determine the contents of each microwell of the array. (2) Raw data generated by CellProfiler was converted into easily interpretable graphs and tables displaying the desired location, survivability, or growth of ISCs within the array culture.

Software. Our analysis pipeline (StemCellPipeline) was assembled using modules in the open source image analysis software package CellProfiler 2.0 (available for free download at http://cellprofiler.org).

Image thresholding and object identification. Autofluorescence and reflections off of microwell walls distort the processed microwell images and can reduce the accuracy of stem cell identification. Raw images loaded into CellProfiler undergo thresholding using the "MoG Global" method, which separates objects from background based on an estimate of the amount of the image occupied by objects. The image occupation of a single cell was estimated from 10 microwell images and the threshold correlation factor was increased to prevent background identification. A Laplacian of Gaussian algorithm, which is best used for objects with increasing pixel intensity towards their center (such as

dsRed+ cells), was used to distinguish between single and clumped objects.

Filtering objects based on characteristics. A panel of 500 microwell images containing only one ISC was analyzed using a pipeline containing the CellProfiler module "MeasureObjectSizeShape", which compiled data on the objects' shape and location characteristics. Distribution curves of shape-based parameters of ISCs were created and compared to distribution curves generated from a 500-microwell image panel of microwells containing debris (Figure 3.4D-F). Thresholds were set using these distribution curves to identify ISCs within total identified objects using the "FilterObjects" CellProfiler module (Figure 3.4G). This filtering allowed us to identify the micro-well images containing only intact stem cells and no debris (Figure 3.4G).

Data reconciliation. Data on identified objects and filtered cells contained in each microwell image was exported using the "ExportToSpreadsheet" CellProfiler module and processed using Microsoft Excel for Mac 2011 (Microsoft, Redmond, WA). The object delimited data exported by CellProfiler is converted to well delimited data using "Excel Module1", which displays the initial contents of each microwell, the distribution of ISCs and Paneth cell-containing wells in the array, and the addresses of microwells containing cells with no debris or fluorescent noise.

Statistical analysis of enteroid formation

Conditional logistic regression was performed with the number of enteroids alive at t = 120hr as the response and the number of ISCs and Paneth cells present in initial well contents as predictors. This was done separately for each experiment as well as

separately for each replicate within separate experiments. When analysis was conducted across both replicates in a single experiment, a term indicating individual wells was included in the model to account for well-to-well variability. All analyses were performed using SAS Version 9.3 (SAS Institute Cary NC). P-values less than 0.05 were considered statistically significant.

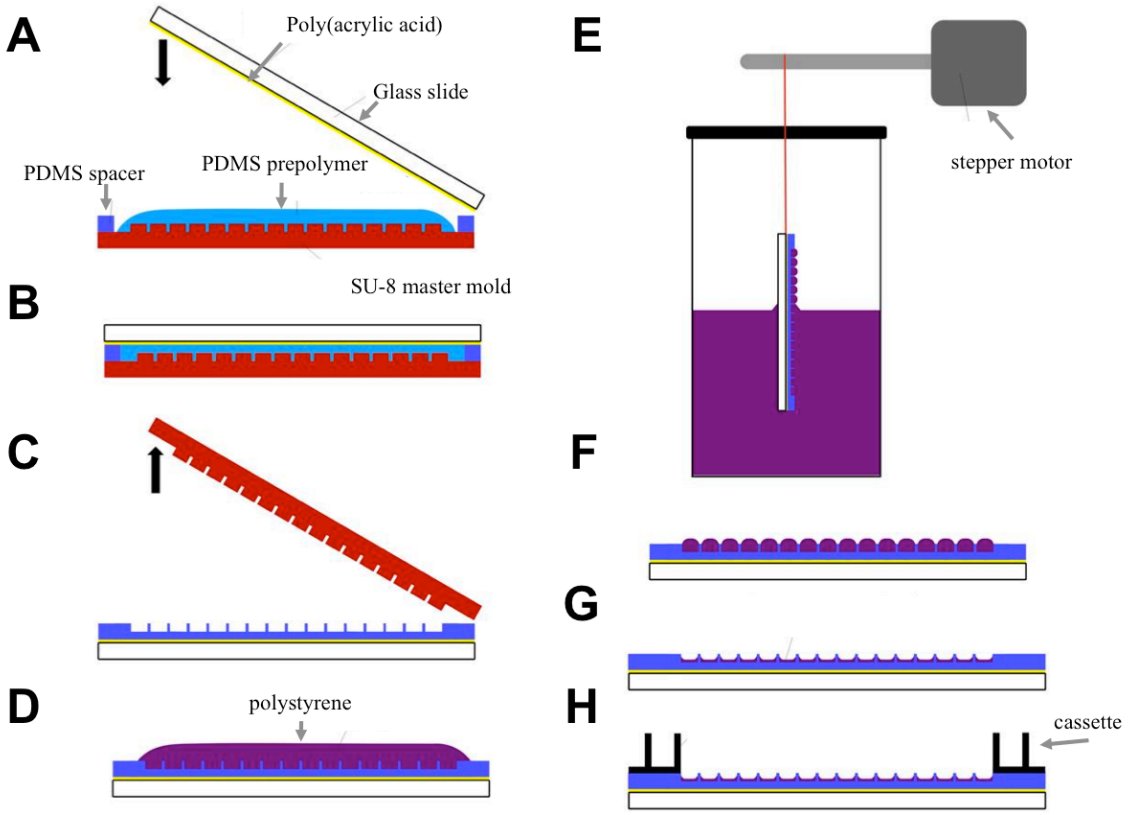


Figure 3.1 **– Fabrication of glass-mounted microwell arrays.** PDMS is liquid molded onto a PAA-coated glass slide on a SU-8 master mold (A). PDMS/SU-8 mold assemblies are cured (B), and cured, solid PDMS is gently removed from the mold (C). Polystyrene solution is added to the array and degassed (D). Dip coating with a programmed stepper motor removes excess polystyrene (D). Subsequent discontinuous dewetting generates isolated polystyrene pockets inside PDMS microwells (F). Arrays are baked at 95°C to remove solvent, resulting in solid polystyrene microrafts embedded in PDMS microwells (G). Microraft arrays are mounted to polycarbonate cassettes to complete the fabrication process (H).

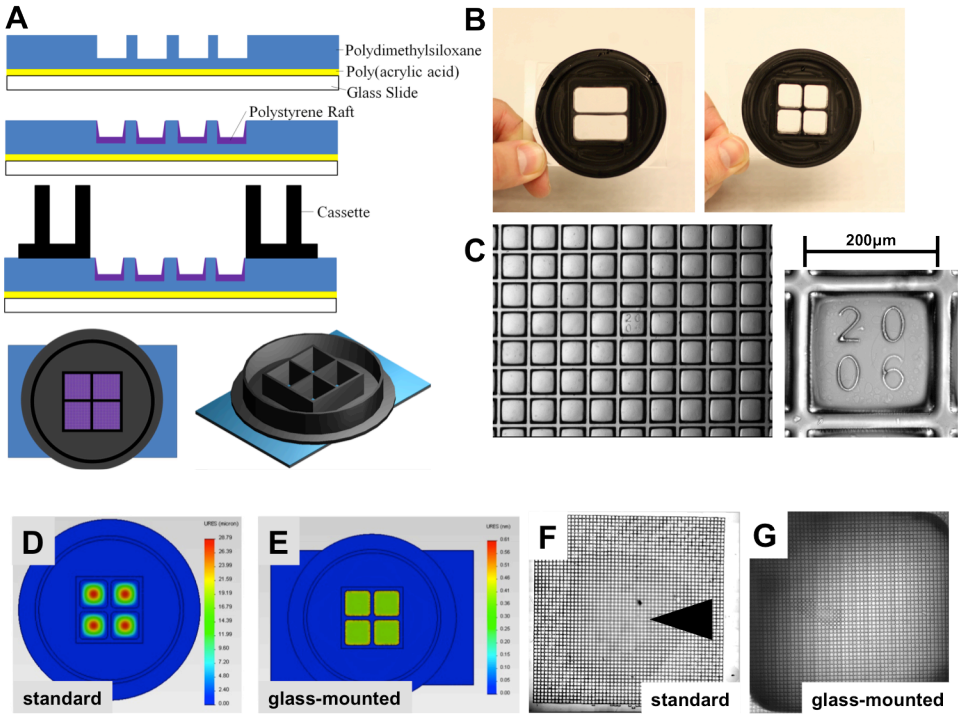


***Figure 3.2* – Modified microwell arrays meet requirements for high-throughput cell culture.** Completed microwell arrays consist of polystyrene-lined PDMS microwells mounted to a glass slide with a thin layer of PAA, and attached to a cassette containing media chambers (A). Cassettes can be scaled to divide a single microwell array into 1-4 separate media reservoirs (B). Microwells are 200µm square, arranged in a grid, and have physical addresses stamped into polystyrene rafts at 5 well intervals (C). The elastic properties of PDMS cause standard microwell arrays to sag, preventing tile-scanned imaging in a single Z-plane (D&F). Mounting arrays to glass slides prevents sagging and facilitates imaging (E&G).

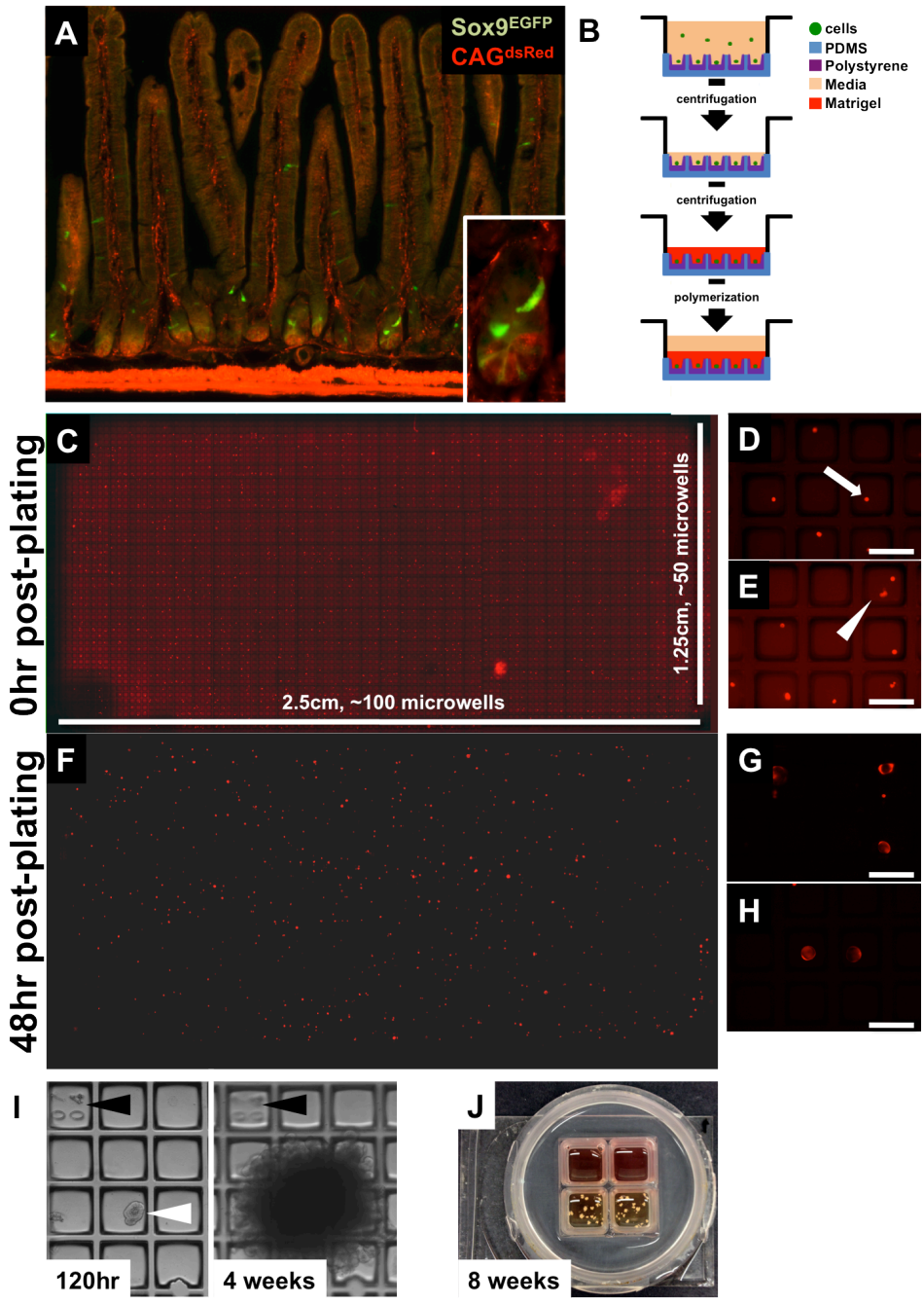


***Figure 3.3* – Microwell arrays are compatible with long-term culture of primary ISCs.** ISCs are isolated from *Sox9*^{EGFP}:*CAGdsRed* transgenic mice, which express dsRed throughout the intestinal epithelium, for ease of visualization in microwell arrays, and to identify contaminating cell types (A). Isolated cells are seeded into microwells through centrifugation in media, and then overlaid with Matrigel in order to meet the requirements of primary ISC culture (B). Tile-scan imaging of whole microwell arrays immediately following plating demonstrates the random distribution of ISCs across the array (C), with some microwells containing single ISCs (D, arrow), and others containing multiple ISCs (E, arrowhead). Imaging of the same array at 48hrs reveals widespread

enteroid formation (F), with typical cystic growth of early structures (G&H). Long-term culture experiments demonstrate that developed ISCs grow out from their original microwells over the course of 4 weeks (I), and can be sustained in the array format for 8 weeks or longer (J).

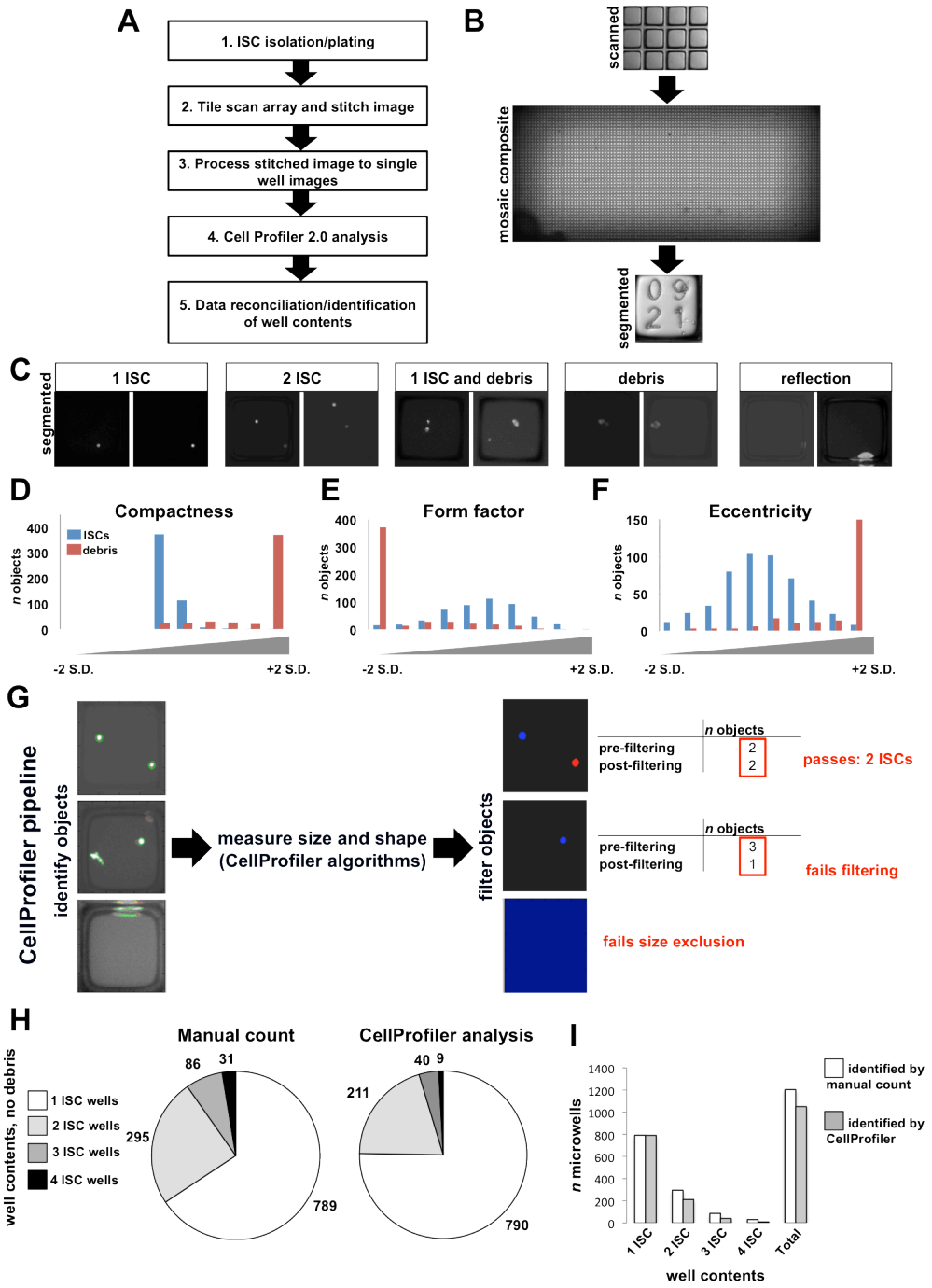


***Figure 3.4* – Software-assisted post-hoc analysis identifies initial well contents of microwell array culture.** A defined workflow facilitates post-hoc image analysis to identify well contents in microwell array cultures (A). Tile-scanned images are stitched together to form a single composite image, which is then segmented into individual microwell images, each with a unique address corresponding to its physical position in the array (B). Visual examination of individual microwells reveals the presence of debris and fluorescent glare as well as ISCs (C). To distinguish between true ISCs and debris/noise, visually scored wells were analyzed by CellProfiler to determine the distribution profiles of target events (ISC, blue) and debris (red) for the variables

compactness, form factor, and eccentricity (D-F). A novel CellProfiler pipeline that identifies objects and applies filters based on the size and shape profiles of known ISCs is used to analyze segmented microwell images (G). If the number of objects initially identified by CellProfiler is equal to those that pass filtering parameters, the microwell passes inclusion criteria (G, top microwell), otherwise it is excluded as debris containing or empty (G, middle and bottom microwells). Application of the optimized pipeline to randomly selected microwell images reveals an overall accuracy of 99.87%, 71.53%, 46.51%, and 29.03% for wells containing 1, 2, 3, and 4 ISCs, respectively (H).

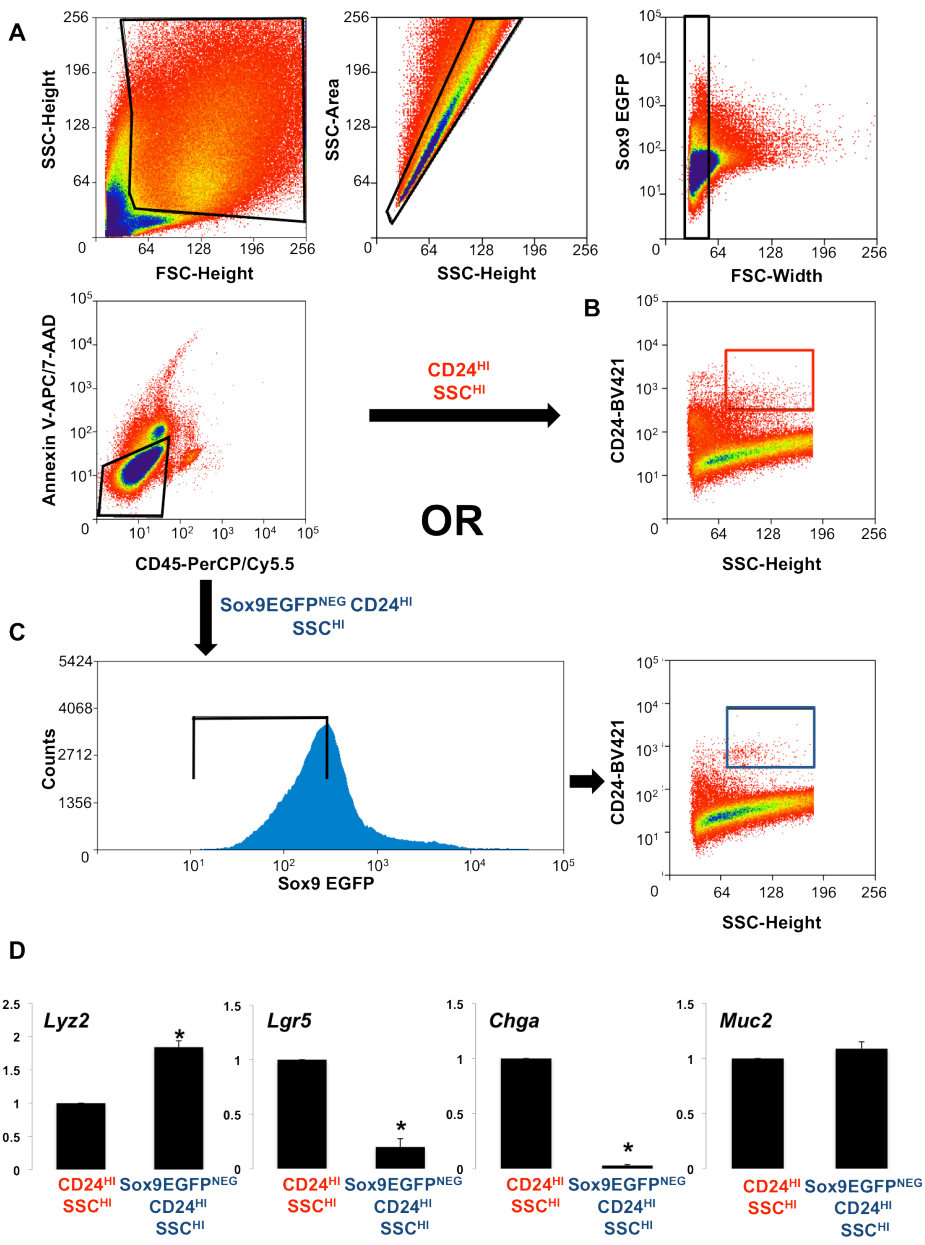


***Figure 3.5 – Sox9^{EGFP}* transgenic mice facilitate high purity FACS isolation of Paneth cells.** To increase population purity, we developed novel FACS criteria for Paneth cell sorting. Standard size, double, and live-dead exclusion criteria were applied to all FACS isolations (A). We compared putative Paneth cell populations isolated using previously described methods, which define Paneth cells as CD24^{high}:SSC^{high} (B), and our newly developed method, which applies the same parameters, but excludes all *Sox9^{EGFP}* expressing cells (C). Gene expression analysis demonstrated significant upregulation of Paneth cell marker *Lyz2*, and downregulation of ISC marker *Lgr5*, as well as EE cell marker *Chga* in Paneth cell populations isolated with *Sox9^{EGFP}* exclusion, when compared to populations isolated with CD24 and SSC alone (D).

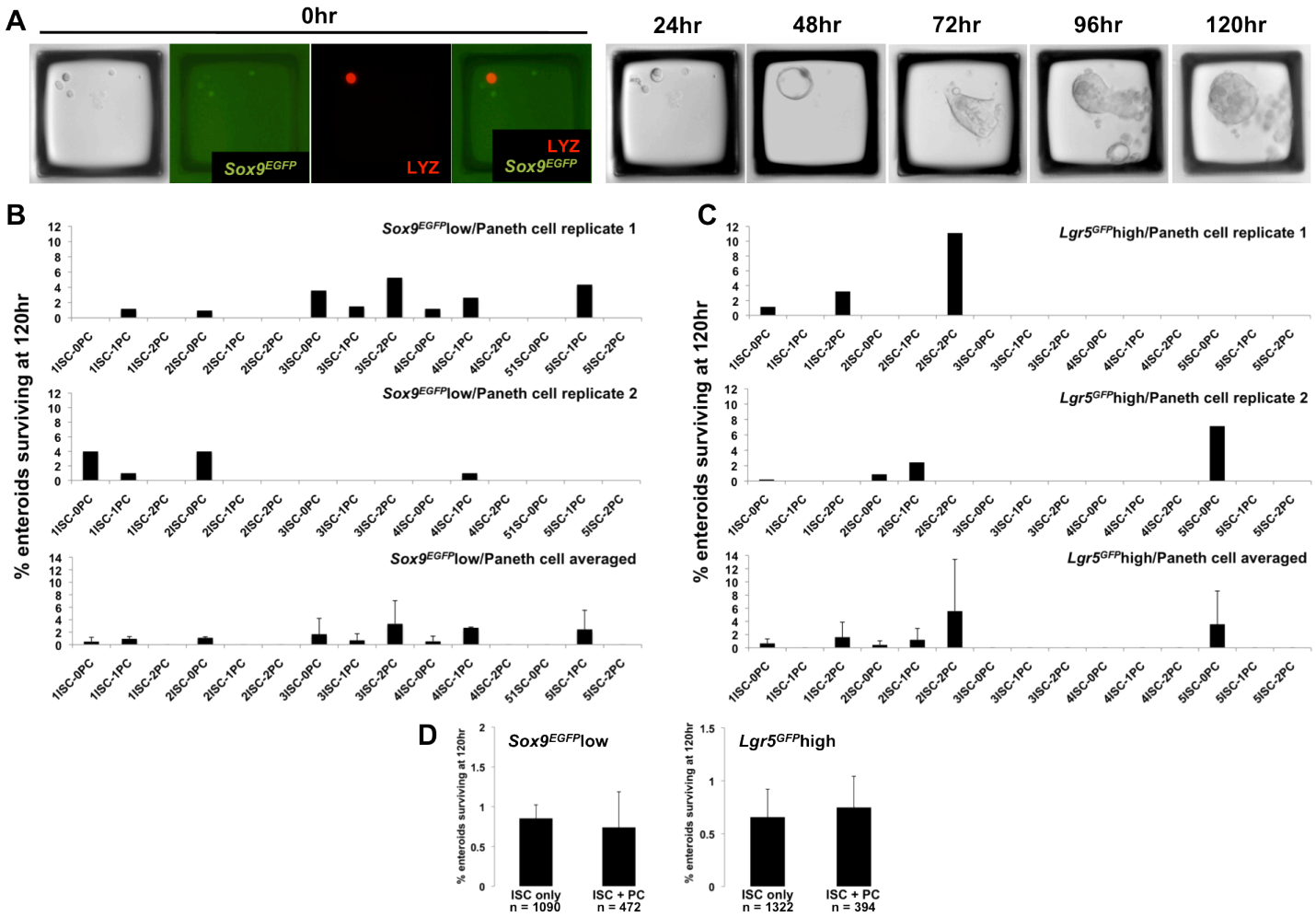


***Figure 3.6* – Paneth cells do not significantly impact ISC survival at physiologically relevant numbers *in vitro*.** Paneth cells are isolated from *Sox9*^{EGFP}:*CAGdsRed* mice and co-cultured with ISCs from *Sox9*^{EGFP} or *Lgr5*^{GFP} mice lacking the constitutive *dsRed* transgene, allowing for independent tracking of ISC and Paneth cell growth in microwell arrays (A). Quantification of enteroid formation across 15 combinations of ISCs and Paneth cells (PC) in experiments using ISCs isolated based on *Sox9* and *Lgr5*

biomarker expression demonstrates no significant increase in enteroid formation at 120hr associated with any one condition (B&C). Additionally, we were unable to identify any clear trends between Paneth cell number and enteroid development by ISCs. Averaging of replicates suggests that formation of enteroids by ISCs occurs randomly in microscale cultures. Analysis of overall enteroid formation by ISCs only and ISC-PC co-cultures, regardless of cell number per microwell, also fails to produce a significant effect associated with Paneth cell presence (D).

CHAPTER 4

SOX4 REGULATES PROLIFERATION, DIFFERENTIATION, AND 5-HYDROXYMETHYLCYTOSINE PATTERNS IN THE INTESTINAL EPITHELIUM[4]

OVERVIEW

Intestinal stem cells (ISCs) are responsible for maintaining the physiological function of the intestinal epithelium through the production of differentiated absorptive and secretory cellular lineages. In addition to producing specialized functional cells, ISCs must also drive constant proliferation in order to maintain the especially high rate of cellular turnover in the intestinal epithelium, which undergoes near total renewal every 5-7 days. Understanding the transcriptional mechanisms that control how ISCs and transit-amplifying progenitors (TAs) balance differentiation and proliferation in the intestine may provide valuable insight into common pathologies, such as inflammatory bowel disease and cancer. Here, we show that the *Sry*-box containing transcription factor, *Sox4*, is involved in the regulation of proliferation in the ISC/TA zone, the expression of ISC-associated genes, and the differentiation of enteroendocrine (EE) cells. Interestingly, we also observed a significant reduction in the expression of the methylcytosine dioxygenase, *Tet1*, in *Sox4*-deficient intestines. We find that *Tet1*, which initiates derepression of target genes via DNA demethylation, is specifically upregulated in ISC

[4] full list of contributing authors: Gracz AD, Trotier DC, McCann JV, Harrell C, and Magness ST

populations in wild-type animals. Additionally, *Sox4*-deficient intestines showed a significant reduction in levels of 5-hydroxymethylcytosine, the catalytic byproduct of TET protein activity, in the ISC/TA zone. Together, our data demonstrate that *Sox4* regulates differentiation and proliferation in the intestinal epithelium, and suggests that it may influence these processes through induction of *Tet1* and subsequent derepression of target genes through epigenetic mechanisms.

INTRODUCTION

The intestinal epithelium is a monolayer of cells arranged along a crypt-villus axis, which establishes distinct proliferative and functional compartments in the tissue. The villi, which are finger-like structures that protrude into the intestinal lumen, are made up of differentiated epithelial cells consisting mainly of absorptive enterocytes, but interspersed with secretory goblet cells (mucus-producing) and EE cells (hormone-producing) [4-6]. The crypts are glandular invaginations into the intestinal submucosa, which contain the ISC and TA populations, as well as differentiated Paneth cells (antimicrobial-producing) [3, 5, 7]. Unlike other differentiated lineages, Paneth cells reside at the base of the crypt, are intercalated between ISCs, and are thought to generate niche cues through soluble and insoluble cell-cell signaling [8]. In recent years, efforts in biomarker discovery have resulted in the identification of several genetic markers of active ISCs phenotypically associated with crypt-base columnar cells (CBCs), including *Lgr5* and *Sox9* [19-21]. While CBCs are generally appreciated as the primary ISC

population, controversy remains over the genetic signature and phenotypic nature of reserve, or facultative, ISC populations traditionally associated with the +4 position, relative to the base of the crypt. Emerging data suggests that +4 reserve ISCs are actually one or more subpopulations of short- or long-lived secretory progenitor cells, that are capable of dedifferentiating and exhibiting characteristics of functional stemness following radiation damage *in vivo* or upon exposure to extrinsic signaling cues *in vitro* [34, 36, 39]. This model of damage response, in which differentiated cells are able to adopt stem cell function following injury, is consistent with emerging evidence in other epithelial tissues, such as the lung and stomach, and may represent a common mechanism for epithelial repair [162, 163]. However, little is known about the transcriptional programs that control reacquisition of potency or terminal differentiation in intestinal progenitor cells.

Multiple signaling pathways control differentiation, lineage allocation, and proliferation in the intestinal epithelium. The Notch pathway is well characterized as a central contributor to cell fate decision between secretory and absorptive lineages (reviewed by [50]). Activation of Notch induces the expression of *Hes1*, which is associated with the production of absorptive enterocytes [78]. In the absence of Notch, progenitor cells are specified toward a secretory lineage, first through the expression of *Atoh1*, and later through lineage-specific factors such as *Ngn3* (enteroendocrine), *Spdef* (goblet/Paneth), and *Gfi1* (goblet/Paneth) [81, 164, 165]. ISCs express Notch ligands, and are dependent on Notch signaling, but appear to rely on downstream effector genes distinct from those found in differentiation programs [74, 78]. The Wnt pathway, which is critically important

for proliferation and maintenance of stemness in the intestinal epithelium, also plays a role in secretory differentiation, as increased Wnt signaling promotes Paneth cell maturation [62].

Sox factors are HMG-domain containing transcription factors that participate in the processes of progenitor proliferation and differentiation in a broad range of tissues throughout development, as well as in adulthood [97, 107]. In the intestinal epithelium, *Sox9* has been identified as a Wnt target and ISC biomarker, and distinct levels of *Sox9* expression are differentially associated with ISCs, TAs, EE cells, and Paneth cells [20, 21, 35]. SOX9 has also been shown to mark DCAMKL1-positive Tuft cells [15]. Loss of *Sox9* is associated with crypt hyperplasia, increased proliferation, and decreased Paneth cell numbers, indicating a dual role in ISC/TA proliferation and cell fate specification [108, 110]. While the expression of other *Sox* factors, including *Sox4*, *Sox6*, and *Sox17*, has been established in the intestinal epithelium, much less is known about their expression and function in different cell populations of the intestine [107]. *Sox4* belongs to the SoxC subgroup of Sox factors and plays critical roles in cardiac, hematopoietic, and pancreatic development [117, 130]. In the gastrointestinal tract, *Sox4* is associated with recurrence in colorectal cancer (CRC) and has been shown to positively regulate Wnt signaling in CRC cell lines [114, 166]. However, the role of *Sox4* in ISC proliferation and differentiation *in vivo* has not been established.

Though less appreciated in the context of ISC proliferation and differentiation, multiple epigenetic mechanisms are emerging as key determinants of potency and lineage

allocation in ESCs and iPSCs. One of the most common epigenetic marks is methylation at the 5-position of cytosine (5mC), which is associated with gene repression. Accordingly, DNA methylation patterns are known to undergo significant changes during the transition from pluripotency to lineage specification, as genes required for differentiation become active [167]. Recent work has identified the ten-eleven translocation (TET) family of proteins, consisting of TET1-3, as enzymes that oxidize 5mC to produce 5-hydroxymethylcytosine (5hmC), which is associated with increased promoter activity and thought to represent an initial step in the derepression of methylated genes [168]. *Tet1* also participates in the maintanence of *Nanog* expression in ES cells by preventing methylation of its promoter, indicative of a role in pluripotency [169]. Further studies have revealed that, in addition to its canonical role as a methylcytosine dioxygenase, *Tet1* plays a role in active repression of direct target genes, through the recruitment of polycomb repressor complex (PRC) subunits [170, 171]. *In vivo* studies into the functional role of *Tet* genes remains limited, though recent work has shown that *Tet1* participates in the maintanence of potency in adult hippocampal progenitors [172]. To date, the expression, function, and regulation of *Tet* genes remains unknown in the intestinal epithelium.

In the present study, we show that *Sox4* is expressed in ISCs, TAs, and Paneth cells. Loss of *Sox4* through conditional ablation in the intestinal epithelium results in subtle, but significant crypt hyperplasia, and increased proliferation. Additionally, we observe defective lineage allocation, with decreased numbers of enteroendocrine cells and an expansion of the Paneth cell zone in *Sox4*-deficient intestines. Gene expression analysis also reveals an upregulation of genes associated with absorptive enterocyte

differentiation and function, as well as increased expression of genes associated with ISC populations. Interestingly, we find that *Tet1* is downregulated following loss of *Sox4*, and that *Sox4*-deficient crypts exhibit significantly fewer 5hmC-positive cells compared to controls. Together, our data indicate that *Sox4* controls intestinal proliferation, is involved in enteroendocrine differentiation, and may participate in these processes through the regulation of DNA methylation patterns through *Tet1*.

RESULTS

Sox4 is expressed in ISCs and progenitor cells

Sox4 has previously been identified as a Wnt target gene, suggesting that it is expressed in the intestinal crypts. To establish the expression pattern of *Sox4*, we conducted *in situ* hybridization on intestinal tissue and found that *Sox4* mRNA was expressed at the base of the crypts, encompassing the ISC, Paneth, and supra-Paneth cell positions, consistent with its identification as a Wnt target (Figure 4.1A&B). In order to determine the precise expression of *Sox4* in crypt-based cell populations, we sorted multiple cell populations from $Sox9^{EGFP}$ transgenic mice. We have previously shown that distinct levels of $Sox9^{EGFP}$ are associated with active ISCs ($Sox9^{EGFP}$low), TAs ($Sox9^{EGFP}$sublow), EEs/reserve ISCs ($Sox9^{EGFP}$high), Paneth cells ($Sox9^{EGFP}$neg;CD24high), and differentiated villus cells ($Sox9^{EGFP}$neg;CD24neg) (Gracz, unpublished data) [21, 35, 36]. Interestingly, *Sox4* was enriched in ISCs and TAs in a similar manner as *Sox9*, but was not significantly upregulated in $Sox9^{EGFP}$high cells relative to $Sox9^{EGFP}$low ISCs (Figure 4.1C&D). Unlike *Sox9*, which is expressed throughout the crypts as well as in a subset of villus EE and tuft cells, *Sox4* expression was never

observed outside of the crypts, which could account for the lack of further upregulation of *Sox4* mRNA in *Sox9*^{EGFP}high cells [15, 35]. To further examine *Sox4* expression, we conducted immunofluorescence on intestinal tissue and identified two distinct populations of SOX4-expressing cells (Figure 4.1E&F). SOX4low cells were found at the base of the crypt, in the CBC/Paneth position (+1 to +3 cell positions from the crypt base), while SOX4high cells were most often found in the supra-Paneth cell zone, at cell positions +4 to +7 (Figure 4.1G). Together, these data demonstrate that *Sox4* is enriched in previously identified active and reserve ISC populations, as well as TA and Paneth cells. Additionally, SOX4 protein is expressed at distinct levels that are differentially correlated with the anatomical positions for active and reserve ISCs. This broad expression pattern suggests that *Sox4* may play multiple, context-dependent roles in different cell types.

SOX4high cells are not label retaining

Recent studies have established two distinct reserve ISC populations associated with the +4 position. While both are capable of driving intestinal regeneration following damage, one population is short-lived and marked by high levels of the Notch ligand *Dll1*, and the other is long-lived and identified by label retention [34, 39]. To determine if SOX4 populations are short-lived or label retaining, we implanted osmotic pumps containing the thymidine analogue 5-ethynyl-2′-deoxyuridine (EdU) subcutaneously in adult mice in order to continuously label all proliferating cells in the intestinal epithelium. After 4 weeks of continuous EdU administration, the pumps were removed and animals were sacrificed immediately and following 1 and 2 weeks of “washout”. Examination of

intestinal tissue obtained immediately following pump removal showed that the EdU had successfully labeled a majority of intestinal epithelial cells, with the expected exception of some Paneth cells, which can persist for over 4 weeks following differentiation *in vivo* and were likely produced before pump implantation (Figure 4.2A) [7]. Colocalization of EdU and SOX4 signals demonstrated that 94% of SOX4-positive cells were labeled by EdU immediately following pump removal, with the remaining 6% likely made up of Paneth cells that differentiated prior to labeling (n = 100 cells) (Figure 4.2A&D). This number significantly decreased to 11% and 9% following 1 and 2 weeks of washout, respectively (n = 100 cells) (Figure 4.2B-D).

Since our expression data demonstrated that low levels of *Sox4* mark Paneth cells, which are known to be long-lived and label retaining, we asked if label-retention in SOX4-positive populations was dependent on expression level. Interestingly, we found that label retention within SOX4-positive cells was restricted entirely to SOX4low cells, which likely represent a long-lived Paneth cell population (Figure 4.2E). In contrast, no label retaining SOX4high cells were found at 1 or 2 weeks of washout (Figure 4.2E). These data demonstrate that cells expressing high levels of SOX4 are not long-lived *in vivo* and therefore do not represent a label retaining reserve ISC population.

Conditional deletion of Sox4 in the intestinal epithelium

Constitutive deletion of *Sox4* results in embryonic lethality at E10.5, due to defects in the cardiac outflow tract that mirror common trunk arterial malformations in humans [117]. In order to study *Sox4* in adult intestine, we crossed mice carrying previously

generated floxed-*Sox4* (*Sox4*^{fl}) alleles to mice expressing *Cre* recombinase under the control of the villin promoter (*vilCre*), which is expressed throughout the intestinal epithelium starting at E13.5 [173, 174]. *Sox4*$^{fl/fl}$:*vilCre* mice were healthy, viable, and born at expected genotypic ratios. We confirmed loss of *Sox4* expression by qPCR analysis of whole jejunal tissue and found that *Sox4* was decreased greater than 2-fold in *Sox4* knockout intestines compared to *Sox4*$^{fl/fl}$ tissue from littermate controls (Figure 4.3A). Since *Sox4* was still detected in knockout intestines, we wished to determine if ablation of *Sox4* was mosaic in the intestinal epithelium, similar to *Sox9*$^{fl/fl}$:*vilCre* mice [108]. We were unable to detect any SOX4 protein in the epithelium of knockout animals by immunofluorescence, suggesting that residual *Sox4* expression observed by qPCR originates from submucosal tissues known to express *Sox4*, such as lymphocytes and peripheral nerves (Figure 4.3B&C) [115, 126].

Loss of Sox4 does not affect expression of SoxC factors

Sox factors are able to compensate for one another *in vivo*, and compensation for loss of *Sox4* by *Sox11* has been observed in the developing spinal cord [96]. To establish whether or not loss of *Sox4* in the intestinal epithelium results in compensatory upregulation of other *Sox* factors, we examined the expression levels of *Sox11* and *Sox12*, which together with *Sox4* make up the SoxC subgroup [88]. We did not observe a significant difference in the expression levels of *Sox11* and *Sox12* between knockout and control animals (Figure 4.3D). We also examined the expression of *Sox9*, which belongs to a different Sox subgroup, but is associated with a distinct proliferative phenotype, and found no difference between knockout and control tissue (Figure 4.3D). These data

indicate that loss of *Sox4* does not result in upregulation of other SoxC factors or *Sox9*.

Sox4 deficient intestines exhibit crypt hyperplasia

Initial examination of *Sox4* knockout intestine revealed generally normal epithelial architecture, with clearly defined crypt and villus units. However, we noticed an apparent increase in crypt size in knockout intestines (Figure 4.4A&B). Quantification of crypt and villus size revealed a significant increase in crypt depth in knockout intestines relative to controls, while crypt width as well as villus height and width remained unchanged (Figure 4.4C&D).

Since cell size appeared unchanged in knockout tissue, we examined proliferation in *Sox4* knockout intestines to establish hyperproliferation as the driving force behind the observed increase in crypt size. Immunofluorescence was used to detect KI67, which labels all proliferating cells, and we quantified the percent of proliferative cells per crypt unit. *Sox4* knockout crypts were composed of 65.4% ± 1.6 proliferating cells, which was significantly greater than the 51.0% ± 1.9 observed in $Sox4^{fl/fl}$ controls (Figure 4.5A-C). Since Wnt signaling is the main regulator of proliferation in the intestinal epithelium, we examined the gene expression of direct targets of the Wnt pathway in knockout tissue. We found that three direct targets of the canonical Wnt/β-catenin pathway, *Ccnd1*, *Myc*, and *Cd44* were upregulated as much as 3-fold in knockout intestine (Figure 4.5D). Together, our data demonstrate that loss of *Sox4* results in increased proliferation and crypt hyperplasia *in vivo*. These effects may be due in part to increased Wnt signaling in *Sox4* knockout intestines, and demonstrate that *Sox4* is a negative regulator of intestinal

proliferation. These results are surprising, in that studies in CRC cell lines demonstrated a positive correlation between expression of *SOX4* and Wnt target genes, suggesting that the role of *Sox4* may differ under physiological conditions and following oncogenic transformation [114].

ISC marker expression is increased in absence of Sox4

A majority of established ISC biomarkers have also been identified as targets of Wnt signaling, which induces crypt and ISC formation during development [56, 59]. To determine if the increased Wnt signaling observed in *Sox4* knockout intestines is associated with an increase in ISC gene expression, we assayed active ISC markers *Lgr5*, *Olfm4*, and *Ascl2*, as well as reserve ISC markers *Tert*, *Lrig*, and *Bmi1*. Both the active and reserve markers were significantly upregulated in knockout intestines, with the most significant upregulation found for *Ascl2*, which was expressed approximately 2-fold higher in knockout intestines compared to controls (Figure 4.6A&B). Therefore, increased Wnt signaling observed in *Sox4* knockout intestines is correlated with an increase in ISC biomarker expression.

Enteroendocrine and Paneth cell differentiation are dysregulated in Sox4 knockout intestines

Next, we wished to examine if loss of *Sox4* has an effect on differentiation in the intestinal epithelium. Secretory lineage allocation was examined by immunofluorescence against MUC2 (goblet cells), CHGA (EE cells), and LYZ (Paneth cells) (Figure 4.7A, C, &F). We found no significant difference in the number of crypt- or villus-based goblet

cells between knockout and control intestines (Figure 4.7B). Quantification of EE cell numbers revealed a significant decrease in the number of villus units containing CHGA+ cells (Figure 4.7D). Additionally, while not significant, we did observe a decrease in the number of crypts containing CHGA+ cells. Interestingly, the number of EE cells per CHGA+ crypt or villus unit was unchanged (Figure 4.7E). That is, while knockout intestines exhibited fewer crypt/villus units containing CHGA+ EE cells, the total number of EE cells in crypt/villus units containing one or more CHGA+ cells remained the same between knockouts and controls. These data suggest that *Sox4* contributes to EE differentiation in a clonal manner, but is not required for the production of all EE cells. Finally, we quantified LYZ+ cells and found an increase in Paneth cell numbers that did not reach statistical significance in knockout intestines (Figure 4.7G). However, quantification of cell position revealed the presence of LYZ+ cells outside of the expected Paneth cell zone in knockout crypts, with some crypts containing Paneth cells as high as position +12-13 from the base of the crypt (Figure 4.7H). This increase in Paneth cell number is consistent with the observed increase in Wnt signaling in knockout intestines, as Wnt is known to drive differentiation and maturation of Paneth cells [62].

To examine the possible impact of *Sox4* on upstream signaling pathways associated with lineage allocation, we assessed gene expression of several major transcriptional regulators of differentiation in the intestine by qPCR. We found a > 2-fold increase in *Notch1* expression in knockout tissue and a smaller, but significant increase in the downstream target of Notch signaling, *Hes1* (Figure 4.8A). Interestingly, *Atoh1*, which is negatively regulated by constitutive expression of *Hes1*, was not significantly

changed, suggesting that the increased *Hes1* expression was inadequate to exert a negative regulatory impact on *Atoh1* (Figure 4.8A) [78]. Consistent with our quantification data, the goblet cell associated factors, *Gfi1* and *Spdef*, were unchanged, while the EE cell associated factor *Ngn3*, was significantly downregulated in knockout intestine (Figure 4.8B). Transcription factors associated with the differentiation and function of absorptive enterocytes, *Cdx2*, *Elf3*, and *Sis*, were all significantly upregulated in knockout tissue, and may reflect the observed increase in *Notch1* and *Hes1* expression (Figure 4.8C). These data suggest that *Sox4* has a broad impact on differentiation in the intestinal epithelium, affecting secretory and absorptive lineage allocation toward multiple cell types, with the notable exception of goblet cells.

Genome-wide expression analysis of Sox4 knockout intestine

Since our observations suggested that *Sox4* exerts a broad influence on proliferation and differentiation in the intestinal epithelium, we next sought to identify possible regulatory relationships for *Sox4*, in order to establish a mechanism for these observations. To this end, we conducted genome-wide gene expression analysis by microarray on RNA isolated from $Sox4^{fl/fl}$:*vilCre* knockout intestines and $Sox4^{fl/fl}$ littermate controls. These experiments produced reliable signal values for 31,291 genes and lncRNAs that were consistently regulated across knockout replicates (Figure 4.9A). Of these, 1,546 transcripts were found to be significantly regulated with a false discovery rate (FDR) < 5.0. We examined genes that were up- or down-regulated ≥ 2.0 fold with FDR < 5.0, and found that a number of transcripts associated with enteroendocrine function were significantly downregulated in knockout tissue, consistent with our

phenotypic observations (Figure 4.9B). Of these, *Nts* and *Gcg* were the most significantly downregulated, at -90.91-fold and -37.37-fold, respectively. Additionally, microarray analysis revealed a 2.00-fold upregulation of ISC-associated *Prom1*, as well as Paneth cell-associated transcripts *Reg3g* and *Muc4* (Figure 4.8B). Interestingly, we identified the ten-eleven translocation transcript *Tet1* as being down-regulated 2-fold following loss of *Sox4* (Figure 4.8B). In ES cells, *Tet1* participates broadly in the maintanence of proliferation and initiation of differentiation through epigenetic mechanisms involving both changes in DNA methylation and recruitment of PRC2 subunits involved in gene repression [170, 171]. Furthermore, loss of *Tet1* in ES cells is associated with defective neural differentiation due to inefficient induction of *Ngn2* [175]. We reasoned that *Tet1* might play a similar role in intestinal potency and differentiation downstream of *Sox4*.

Tet1 is expressed in ISC populations and downregulated following loss of Sox4

The expression pattern and characteristics of *Tet* genes in the intestine have not previously been examined. To determine if *Tet1-3* are associated with specific cell types in the intestinal epithelium, we examined their expression in active ISCs, TAs, EE/reserve ISCs, Paneth cells, and differentiated villus cells using previously established *Sox9*^{EGFP} FACS methods [21, 36]. We found that *Tet2* and *Tet3* were expressed throughout the intestine, and not significant up- or downregulated in any specific subpopulation of epithelial cells (Figure 4.10). However, *Tet1* was enriched 8-10-fold in *Sox9*^{EGFP}low active ISCs and *Sox9*^{EGFP}high EEs/reserve ISCs, relative to *Sox9*^{EGFP}neg/CD24neg villus cells (Figure 4.10). These data are consistent with previous reports that *Tet1* is expressed in ES cells and downregulated following differentiation [168, 169].

Next, we confirmed decreased *Tet1* in *Sox4* knockout intestines by qPCR, and fold an approximately 2-fold decrease in expression (Figure 4.11). *Tet2* and *Tet3* are capable of converting 5mC to 5hmC in a manner similar to that of *Tet1* [168]. To address the possibility of compensation for decreased *Tet1* by upregulation of *Tet2* and *Tet3*, we examined their expression level in *Sox4* knockout tissue. While *Tet2* remained unchanged, we found a small but significant upregulation of *Tet3* following the loss of *Sox4* (Figure 4.11).

Altered expression pattern and decreased levels of 5hmC in Sox4 knockout intestines

Since 5hmC is the functional product of TET activity, we examined its expression pattern to determine if DNA methylation dynamics were altered following the loss of *Sox4* in the intestine. Immunofluorescence on wild-type control and $Sox4^{fl/fl}$ intestine revealed a consistent gradient in 5hmC expression in the epithelium, with low, but present levels in the crypts and increasingly stronger expression moving toward the tip of the villus (Figure 4.12A). Interestingly, several 5hmC-high cells were observed at the crypt base, anatomically and morphologically consistent with Paneth cells (Figure 4.12B). As expected, the crypts of $Sox4^{fl/fl}$:*vilCre* knockout intestines were almost completely devoid of 5hmC expression, with the notable exception of 5hmC-high Paneth cells at the crypt base (Figure 4.12C & D). Additionally, the gradient of 5hmC expression along the crypt-villus axis was perturbed in *Sox4* knockout intestines, with 5hmC levels becoming more highly expressed further up the villi than in control animals (Figure

4.12C). Quantification of the number of crypts containing 5hmC-high expressing cells revealed a significant decrease in knockout intestines, while the number of 5hmC-high cells per crypt remained approximately the same for crypts containing as least one 5hmC-high cell (Figure 4.12E & F). Together, these data demonstrate that *Tet1* is downregulated in *Sox4* knockout intestines, resulting in changes in DNA methylation patterns along the crypt-villus axis.

DISCUSSION

Our findings reveal that *Sox4* is broadly involved in proliferation and differentiation in the intestinal epithelium. *Sox4* expression has previously been localized to the crypt base, but the precise distribution of its expression pattern across ISC and TA populations was not established [131]. Here, we demonstrate that *Sox4* is specifically upregulated in $Sox9^{EGFP}$ populations previously characterized to represent active, CBC ISCs and reserve, +4 ISCs/EE cells [21, 35, 36]. Analysis of SOX4 protein expression by immunofluorescence supports these data, and further suggests that low levels of SOX4 are associated with active, CBC ISCs and Paneth cells, and high levels of SOX4 are associated with reserve, +4 ISCs. Multiple populations of +4 position ISCs have been described, with the important functional distinction of *in vivo* lifespan. Long-lived label retaining cells and short-lived *Dll1*high cells are secretory progenitor populations, and it would be consistent with our phenotypic observations for loss of *Sox4* that one or both of these populations are consistent with the +4 SOX4high population [34, 39]. Our label-retention studies demonstrate that SOX4high cells are not long-lived, which may indicate that these cells overlap with the *Dll1*high facultative ISC population and not the label

retaining population. Alternatively, increased SOX4 expression may be consistent with the initiation of terminal differentiation, and may mark cells that are no longer functionally capable of reverting to an ISC state. While further characterization studies are necessary to establish the degree of overlap between *Sox4* expression and previously characterized reserve ISC populations, the critical functional experiments required to examine potency of *Sox4*-expressing populations remain technologically impossible without a *Sox4* reporter mouse model that facilitates *in vivo* lineage tracing.

Unexpectedly, deletion of *Sox4* in the intestinal epithelium resulted in crypt hyperplasia and hyperproliferation. Previous studies on the role of *SOX4* in human CRC cell lines showed that SOX4 was able to directly bind to β-catenin and increase the output of a β-catenin reporter construct, leading the authors to conclude that *SOX4* is a positive regulator of canonical WNT signaling [114]. It is possible that *Sox4* acts in a context-dependent manner and therefore, has different functional roles in physiologically normal intestine and tumors. As precedence, studies in the nervous system have demonstrated that *Sox4* has seemingly opposing functions in neurons versus glia. In neurons, *Sox4* induces differentiated neural phenotypes, whereas it promotes an undifferentiated, progenitor-like phenotype in glial precursors [124, 127, 128]. Our studies support a model where *Sox4* negatively regulates potency in the intestinal epithelium, suppressing ISC-associated genes and downregulating Wnt signaling. Since tumorigenesis is known to rely on the dysregulation of physiologically active signaling pathways, it is possible that CRC cell lines exploit the apparently broad regulatory behavior of *Sox4* in order to activate or maintain growth pathways. As such, it will be important to examine the role of

Sox4 in the development and growth of intestinal tumors *in vivo*, to determine if dysregulation of *Sox4* plays a role in tumorigenesis.

One common theme for the role of *Sox4* across many tissues is its involvement in differentiation and the establishment of lineage-specific genetic programs. This function is true for B and T lymphocytes, as well as retinal, neural, and pancreatic progenitors [124, 130, 176]. Our data indicate a similar role in the intestinal epithelium, in that loss of *Sox4* results in decreased enteroendocrine cell numbers and gene expression, as well as an increase in the expression of genes associated with Paneth cells and absorptive enterocytes. Interestingly, loss of *Sox4* in pancreatic islets results in decreased endocrine differentiation as well, and gene expression studies identified *Ngn3* as a significantly downregulated transcript during the development of *Sox4*-deficient islets [130]. While we observed Paneth cells at significantly higher cell positions, and at slightly increased numbers in *Sox4* knockout crypts, we did not observe any differences in the gene expression of upstream transcription factors involved in Paneth cell specification, including *Spdef*, *Gfi1*, and *Sox9* [110, 164, 165]. One explanation is that the increased Wnt signaling in *Sox4* knockout intestines is sufficient to induce slightly higher numbers of Paneth cells without significantly altering the upstream factors involved in their differentiation [65]. In this model, *Sox4* would indirectly regulate Paneth cell differentiation and position through suppression of Wnt signaling. Alternatively, it is possible that *Sox4* suppresses the expression of Paneth cell specific genes, such as *Reg3g* and *Muc4* independently or downstream of *Spdef* and *Gfi1*, and that removal of this inhibition influenced more secretory progenitors to adopt a Paneth cell fate in *Sox4* knockout

intestines. Further studies are needed to establish the precise relationship between known secretory transcription factors and *Sox4*.

Our gene expression studies identified the gene encoding the TET1 enzyme as a significantly downregulated transcript in *Sox4* knockout intestines. While TET proteins are attracting great interest in the regulation of potency and differentiation in ES cells, little is known about their function *in vivo*. *Tet1* has been identified as critical for the maintanence of hippocampal progenitors in adult mice, and may play a role, along with *Tet2*, in maintaining potency in pancreatic progenitors [172, 177]. In ES cells, *Tet1* promotes both pluripotency as well as differentiation depending on context, and appears to do so through dual functions as a key mediator of DNA demethylation as well as a cofactor for PRC2-induced transcriptional repression [170, 171]. We found that *Tet1* is specifically upregulated in ISC populations in the intestine, consistent with results in ES cells and hippocampal progenitors. Interestingly, we find that *Tet2* and *Tet3*, which are functionally redundant with the methylcytosine dioxygenase activity of *Tet1*, are not specifically enriched in any one intestinal epithelial population, but expressed at equal levels across the crypt-villus axis [168]. The catalytic byproduct of TET proteins, 5hmC, is expressed as a distinct gradient across the crypt-villus axis, with the lowest levels of expression observed in the crypts, and the highest levels of expression observed at the villus tips. Compellingly, crypts become almost entirely devoid of 5hmC-positive nuclei in the absence of *Sox4*, demonstrating that functional activity of TET proteins in the intestinal crypts is dependent on *Sox4* expression. 5hmC gradients are still established in *Sox4* knockout intestines, but are initiated much higher up the crypt-villus axis. We also

found a significant increase in *Tet3* in *Sox4* knockout intestine, which may partially compensate for the decrease *Tet1*. Interestingly, combined deletion of all *Tet* genes leads to impartial embryonic lethality, suggesting that compensation is capable of rendering TET proteins dispensable at in an apparently stochastic manner [178]. Stochastic compensation and effective "rescue" of normal gene expression programs in the absence of *Tet1* could explain why the few crypt-villus units that do contain EE cells in *Sox4* knockout intestines contain the same number of EE cells as their wild type counterparts. Further investigation is needed to determine what regulatory mechanisms may bypass the need for *Tet1* expression and DNA demethylation through 5hmC in intestinal differentiation.

In summary, we show that *Sox4* is required for proper enteroendocrine differentiation and participates in the suppression of ISC-associated gene expression. A possible unifying mechanism exists in the regulation of *Tet1* by *Sox4*, which may allow expression of *Sox4* to induce broad, and context-dependent gene regulatory changes through epigenetic means. Based on our data, we propose a model in which *Sox4* participates in intestinal epithelial differentiation through the induction of *Tet1*, subsequent initiation of EE-specific genes, and simultaneous repression of gene expression programs associated with potency and stemness (Figure 4.13). Further studies will be required to discriminate direct and indirect targets of SOX4 and TET1 in the intestinal epithelium, as well as establish the methylation status and transcriptional output of these loci. Additional studies into possible distinct and redundant functions of different *Tet* genes in the intestinal differentiation will be of great interest in further understanding

fate decisions in ISCs and TAs.

MATERIALS AND METHODS

Mice and genotyping

Sox4^{fl} mice have been previously described and were the generous gift of V. Lefebvre [174]. *vilCre* mice were obtained from Jackson Labs (stock number: 004586, Jackson Laboratory, Bar Harbor, ME) [173]. *Sox9*^{EGFP} mice were obtained from the GENSAT project [21, 159]. Homozygous *Sox4*$^{fl/fl}$ mice were bred to *Sox4*$^{fl/fl}$ mice heterozygous for *vilCre* to produce *Sox4*$^{fl/fl}$*:vilCre*$^{-/-}$ (control) and *Sox4*$^{fl/fl}$*:vilCre*$^{+/-}$ (knockout) offspring. Genotyping on tail snips procured at ~10 days postnatal and lysed for isolation of genomic DNA was carried out by PCR, as previously described [173, 174]. Genotypes were observed at expected ratios and *Sox4* mutant mice appeared normal and viable with no overt physiological phenotype. *Sox4*$^{fl/fl}$ and *Sox4*$^{fl/fl}$*:vilCre*$^{+/-}$ mice were maintained on a C57Bl/6 background. *Sox9*^{EGFP} mice were maintained on a CD-1 background. All studies were conducted on the jejunum of adult mice between 8 and 20 weeks of age, and knockouts were matched to littermate controls when possible. In the absence of appropriate littermate controls, experiments were conducted on age-matched controls. All protocols for animal use were reviewed and approved by the University of North Carolina Institutional Animal Care and Use Committee.

In situ hybridization

In situ hybridization against *Sox4* was carried out as previously described [179]. cDNA templates for probes were kindly provided by M. Wegner [129]. Digoxigenin-labeled

probes were generated using the Maxiscript *in vitro* transcription kit (Life Technologies, Grand Island, NY) and digoxigenin-labeled UTP (Roche) and hybridized at 60°C for ~24hrs. Detection was carried out using alkaline phosphatase-conjugated antibodies against digoxigenin (11093274910, Roche, Indianapolis, IN) and NBT/BCIP (Sigma, St Louis, MO) for visualization of AP. Background was determined to be negligible by the lack of appreciable signal in control slides assayed using sense probes.

FACS isolation of Sox9EGFP populations

Epithelial cells were isolated from the jejunum of mice heterozygous for the *Sox9*^{EGFP} transgene, as previously described, with some modifications [160]. Briefly, intestines were opened longitudinally, rinsed in DPBS (Life Technologies), minced, and incubated in 3mM EDTA (Sigma) in DPBS for 45min at 4°C with gentle agitation. Intestinal fragments were transferred to fresh DPBS and shaken by hand for 2 minutes to release epithelium. Remnant submucosa was discarded, epithelium was rinsed twice in DPBS, and then dissociated to single cells by incubation in 0.3U/mL dispase (Life Technologies) in 10mL HBSS (Life Technologies) at 37°C for 8-10 min with shaking every 2 minutes. Single epithelial cells were filtered through 100, 70, and 40μm filters before being resuspended in ISC Sort/Culture Media [Advanced DMEM/F12 (Life Technologies), N2 (Life Technologies), B27 (Life Technologies), Glutamax (Life Technologies), Penicillin/Streptomycin (Life Technologies), 10mM HEPES (Life Technologies), 10μM Y27632 (Selleck Chemicals, Houston, TX), and 500mM N-acetyl-cysteine (Sigma)]. Single epithelial cells were stained with the following antibodies: Brilliant Violet conjugated anti-CD24 (clone M1/69, Biolegend, San Diego, CA); PerCP-

Cy5.5 conjugated anti-CD45 (clone 30-F11, Biolegend). Cells were stained with Sytox Blue (Life Technologies) or 7-AAD (Biolegend) and Annexin V-APC (Life Technologies) immediately prior to FACS, for live-dead exclusion. *Sox9*^{EGFP} subpopulations and Paneth cells were isolated as previously described using a MoFloXDP sorter (Beckman Coulter, Brea, CA) (Gracz, unpublished/Chapter 3 of this dissertation and [21, 160]). Cells were sorted directly into RNaqueous lysis buffer for RNA isolation and qPCR analysis (Ambion, RNaqueous Micro Kit, Life Technologies).

Histology and immunofluorescence

Mice were euthanized and intestines were carefully dissected and luminal contents cleared by gentle syringe flushing with PBS (Sigma). Intestines were divided up by segment (duodenum, jejunum, and ileum) and flushed with cold 4% PFA (Sigma), before being fixed overnight in 4% PFA at 4°C. Tissue was transferred to 30% sucrose (Sigma) in ddH2O for 16-24 hrs before being swiss rolled, embedded in optimal cutting temperature (OCT) media (Sakura, Torrance, CA), and frozen on dry ice. 8-10μm sections were used for immunofluorescent analyses. Prior to immunostaining, tissue sections were rinsed twice in PBS to remove OCT. Non-specific binding was blocked by applying Dako Protein Block (Dako, Carpinteria, CA, X0909) to tissue sections for 30 min at room temperature. Primary antibodies were applied in Dako Antibody Diluent (Dako, S0809) and incubated for 1-2h at room temperature. Dilutions were as follows: SOX4 (H-90, 1:100, Santa Cruz, Dallas, TX), CD326 (G8.8, 1:250, BioLegend), KI67 (TEC-3, 1:250, Dako), Lysozyme (RP028, 1:500, Diagnostic Biosystems, Pleasanton, CA), Mucin2 (H-300, 1:100, Santa Cruz), ChromograninA (1:500, Immunostar, Hudson,

WI), 5hmC (39791, 1:250, Active Motif, Carlsbad, CA). Anti-Rabbit-Cy3 (1:500 Sigma) and anti-Rabbit-Biotin (1:100 Jackson Immunoresearch, Carlsbad, CA, 112-165-003) secondary detection antibodies were diluted in Dako Antibody Diluent and applied to tissue for 30 min at room temperature. For tertiary detection of biotin-conjugated secondaries, Streptavidin-AlexaFluor549 (1:500, Jackson Immunoresearch) was applied for 30 min at room temperature. Nuclei were stained for 10 minutes with bisbenzamide (1:1,000, Sigma) diluted in PBS. Antigen retrieval (required for SOX4, CHGA, and 5hmC detection) was carried out in Reveal Decloaker Buffer (BioCare Medical, Concord, CA) at 120°C for 10 sec and 95°C for 30 sec in a pressure cooker prior to blocking. For 5hmC staining, tissue samples were incubated with 2N HCl for 10 min, followed by neutralization in 0.1M sodium borate 3 x 10 min. Background staining was negligible as determined by nonspecific IgG staining. Images were collected by capturing ~1.5 μm optical sections using a Zeiss LSM 710 confocal microscope.

EdU labeling and washout experiments

Osmotic mini-pumps (Alzet, model 1004, 0.11μl/hr, 28days) were subcutaneously implanted along the abdomen of wild-type CD-1 mice (8 weeks old) and closed by either suture or wound clip. After pump delivery (28 days), osmotic mini-pumps were surgically removed (Day 0), wound closed, and a washout period of either 1 or 2 weeks was completed prior to sacrifice and tissue collection. EdU detection was carried out as per manufacturer instructions using an EdU detection kit (C10337, Life Technologies).

Prior to surgery, pumps were filled with a concentrated EdU (Life Technologies, E10187) solution based calculations from our previous studies (EdU injections 3 x day for 5 days). Calculations to determine concentration to add to pump were based on a schedule of 75µg/injection of EdU, with 3 injections per day (totaling 225µg/day). EdU was dissolved in 50:50 DMSO:PEG-300 (Spectrum Chemicals P0108, 25322-68-2).

RNA isolation and qPCR gene expression analysis

For gene expression studies comparing knockout and control tissue, RNA was isolated from the proximal 0.5 cm of whole jejunum. Briefly, jejunal segments were thoroughly flushed with cold PBS to remove luminal contents and mesenchymal and pancreatic tissues were dissected off of the intestinal serosa. Jejunal segments were stored in RNAlater (Qiagen. Valencia, CA) overnight until RNA was extracted using the RNEasy mini kit (Qiagen) and on-column DNase treatment (Qiagen). RNA from sorted $Sox9^{EGFP}$ populations was purified using RNAqueous Micro Kit (Ambion) according to the manufacturer's protocols. cDNA was generated using iTaq Reverse Transcription Supermix (Bio-Rad, Hercules, CA). Real-time PCR was conducted in SsoFast Probes Mix with ROX (Bio-Rad) for each sample in triplicate on 1/20,000 of the total amount of cDNA generated (Taqman probes are listed in Table 4.1). qRT-PCR data was normalized for the expression of 18S. ΔΔCt values were then calculated using a single control ($Sox4^{fl/fl}$) mouse as the comparator, and significance was determined using student's t-test. Statistical analysis for $Sox9^{EGFP}$ populations compared gene expression across all cell populations by gene via one-way ANOVA followed by post-hoc t-test for

multiple comparisons using Prism (GraphPad, La Jolla, CA). $p < 0.05$ was considered significant. All knockout v. control gene expression analyses represent a biological n = 3 per group.

Gene expression microarrays and bioinformatics analysis

All tissue samples were profiled using Agilent SurePrint G3 Mouse 8x60K Gene Expression Microarrays (Agilent Technologies, Santa Clara, CA, USA) as previously described [180]. Microarray data are available at the UNC Microarray Database (https://genome.unc.edu). Statistical analysis for the microarrays was carried out with R v2.15.1. To utilize the gene expression information, the probes were retrieved and filtered by requiring the Lowess normalized intensity values in both sample and control to be > 10. In total, 31,291 gene probes had 100% good data for the three arrays used for these analyses. A one-class significance analysis of microarrays was then utilized to identify genes that up or down regulated in *Sox4* knockout mice. Genes with a false discovery rate <5 were considered significant.

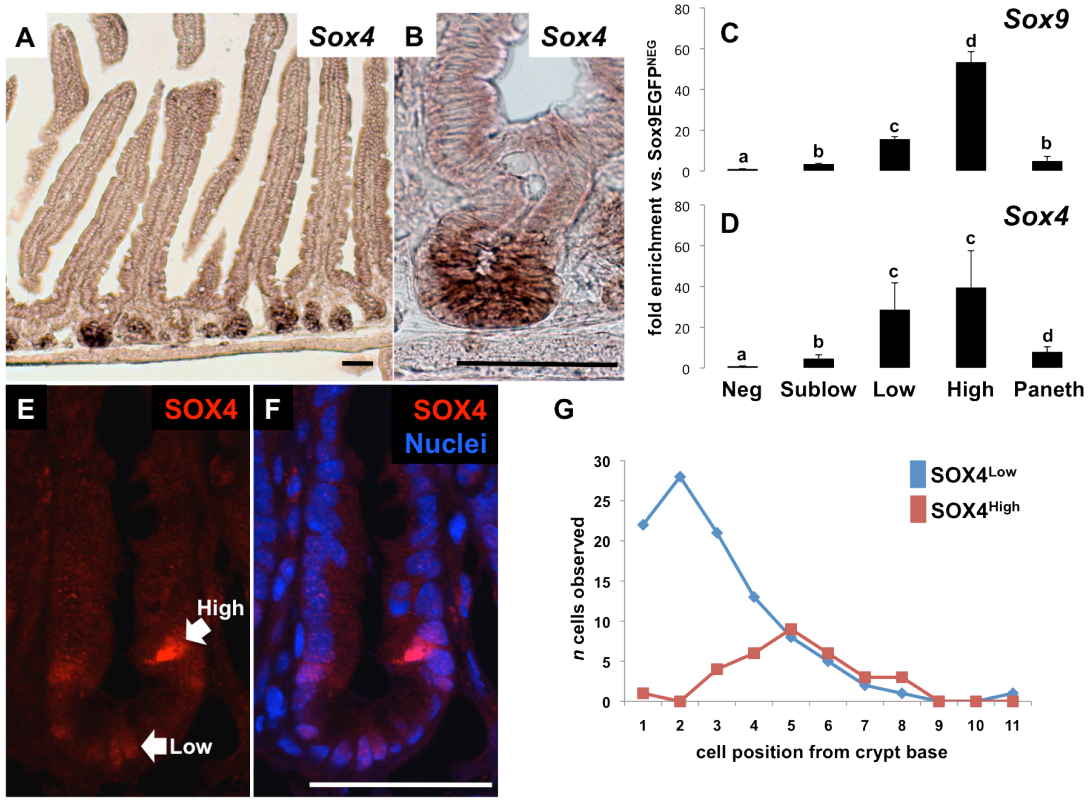


***Figure 4.1 – Sox4* in expressed in ISC and progenitor populations.** *In situ* hybridization identifies *Sox4* mRNA at the base of the intestinal crypt (A&B). qPCR on FACS isolated *Sox9*^{EGFP} epithelial populations (C) facilitates more precise detection of *Sox4* enrichment in previously established active ISCs (*Sox9*^{EGFP}low) and EEs/reserve ISCs (*Sox9*^{EGFP}high) (D). SOX4 protein is expressed at two distinct levels in the crypts (SOX4low and SOX4high), with no observable expression in the villi (E&F). SOX4low cells are found primarily in the CBC/Paneth cell zone (+1-+3), and SOX4high cells are found primarily in the supra-Paneth cell position (+4-+7) (n = 3 biological replicates, 10 crypts per mouse) (G). Scale bars represent 50μm; different letters indicate significance between groups, $p < 0.05$.

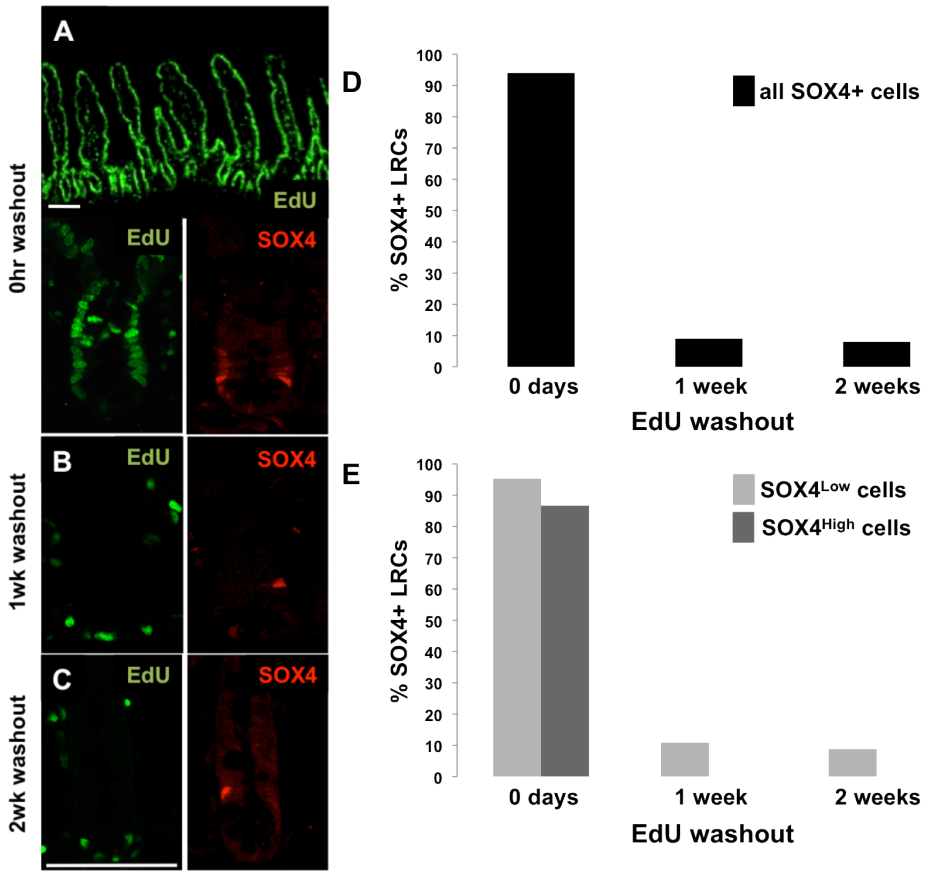


***Figure 4.2* – SOX4high cells are not label retaining.** Continuous labeling of all dividing cells using subcutaneous EdU mini-pumps results in labeling of all intestinal epithelial cells, with the exception of some long-lived Paneth cells (A). The EdU label is subsequently diminished following 1 and 2 weeks of washout, at which point only long-lived cells remain labeled (B&C). Colocalization of SOX4+ populations with EdU reveals a significant decrease in labeled cells at 1 and 2 weeks of washout (B-D). Further assessment of label retention by SOX4 expression level reveals that only SOX4low cells contribute to the label retaining SOX4 population at 1 and 2 weeks (n = 100 SOX4+ cells) (E). Scale bars represent 50μm.

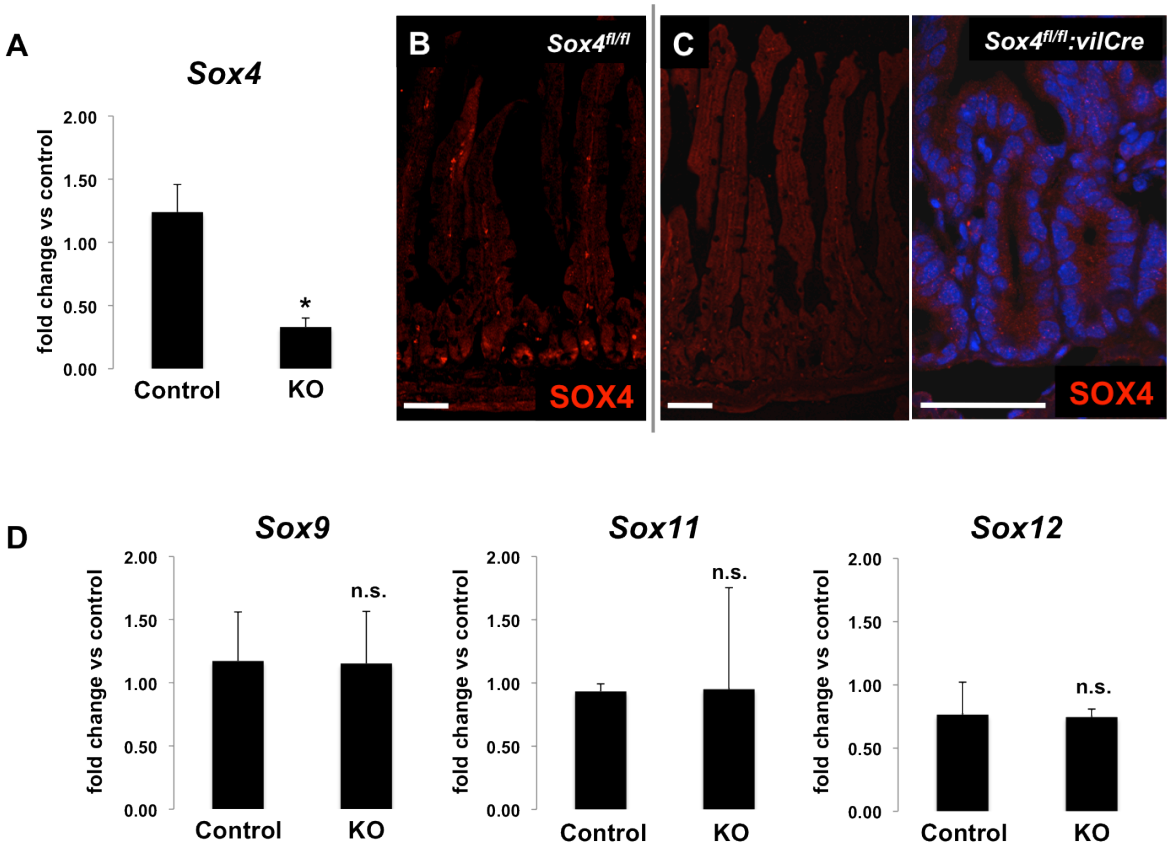


***Figure 4.3* – Conditional deletion of *Sox4* does not affect expression of other SoxC factors in the intestinal epithelium.** qPCR validation of *Sox4* ablation demonstrates a >2-fold decrease in *Sox4* expression in knockout mice (*Sox4*$^{fl/fl}$*:vilCre*) compared to littermate controls (*Sox4*$^{fl/fl}$) (A). Confirmation by immunofluorescence reveals that SOX4 is completely lost in knockout animals, suggesting that residual *Sox4* mRNA is expressed by subepithelial cells (B&C). Expression of other SoxC factors, *Sox11* and *Sox12*, as well as *Sox9*, which has known roles in intestinal epithelial proliferation and differentiation, remains unaffected (D). Scale bars represent 50μm; asterisk indicates significance p <0.001; n.s. indicates no significant difference.

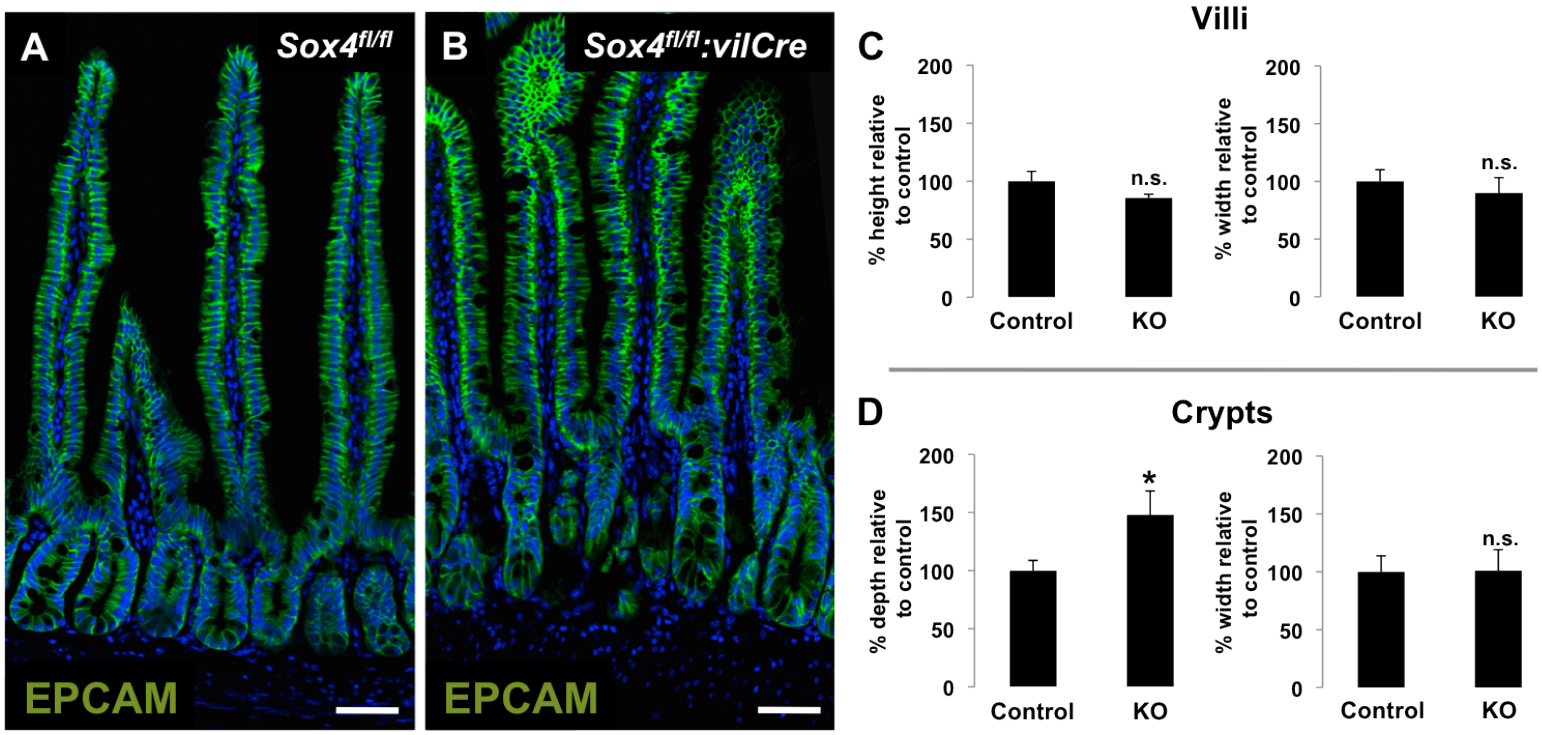


***Figure 4.4 – Sox4* knockout intestines exhibit crypt hyperplasia.** Upon histological examination, crypts in *Sox4* knockout intestine appear larger than those found in controls (A&B). While villus dimensions remain unchanged between knockout and control animals (C), the crypts of *Sox4* deficient mice are significantly deeper, suggesting changes in proliferation (D). n = 3 biological replicates per condition; 20 units per replicate. Scale bars represent 50μm; asterisk indicates significance $p < 0.01$.

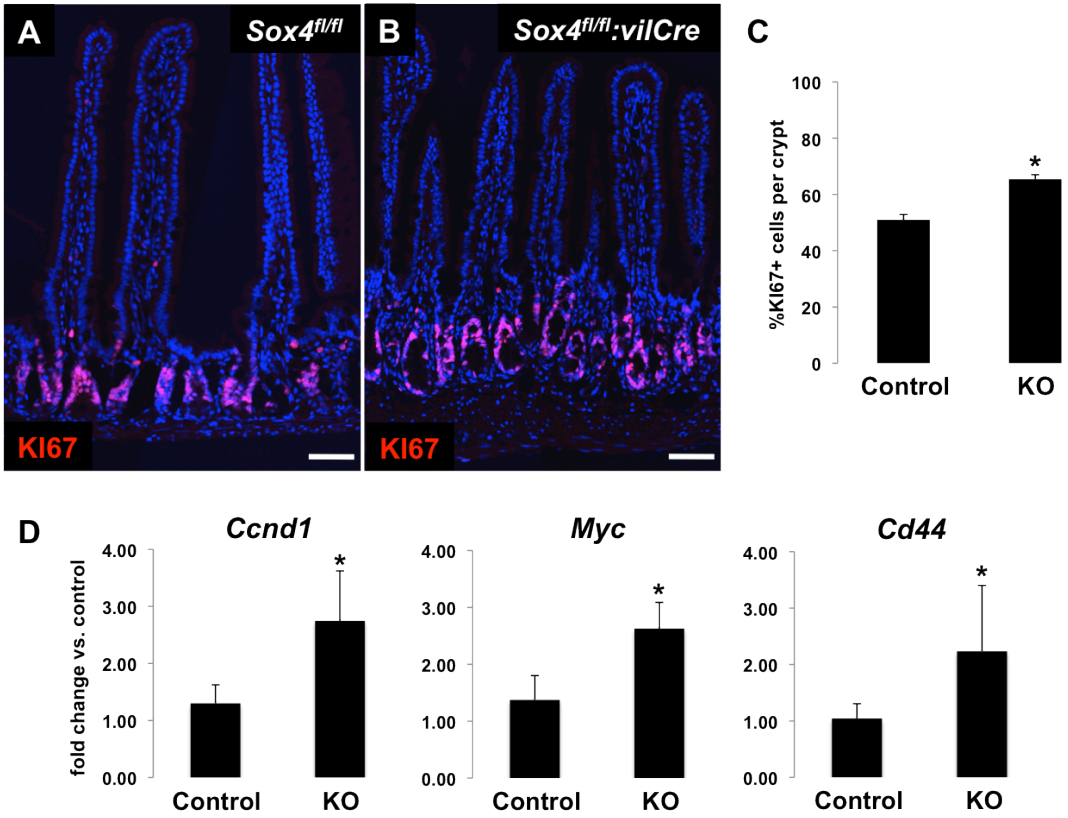


***Figure 4.5 – Sox4* negatively regulates crypt proliferation and Wnt signaling.** Detection of all proliferating cells by KI67 shows an increase in the proliferative zone of *Sox4* knockout intestines, which contain ~15% more proliferating cells than controls (n = 3 biological replicates per condition; 20 crypts per replicate) (A-C). Increased proliferation is coincident with increased expression of downstream targets of canonical Wnt signaling, *Ccnd1*, *Myc*, and *Cd44* (D). Scale bars represent 50μm; asterisk indicates significance $p < 0.01$.

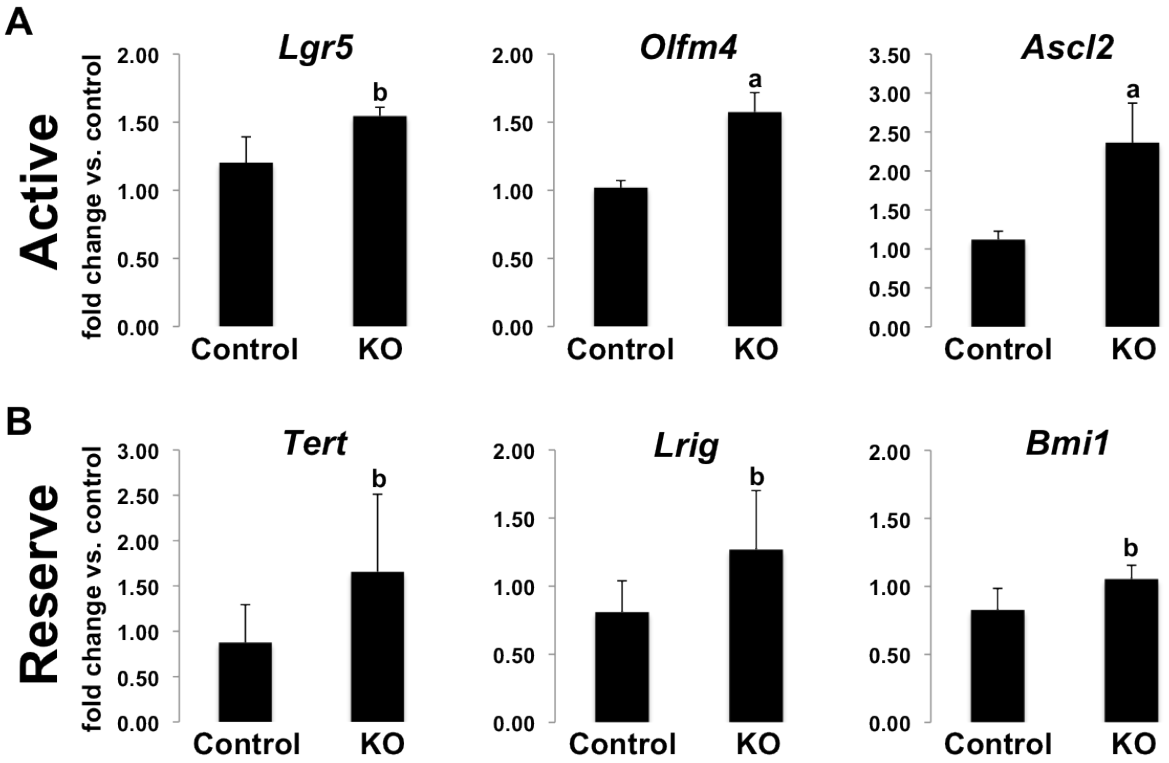


***Figure 4.6* – ISC biomarker expression is elevated in the absence of *Sox4*.** Consistent with increased proliferation and Wnt signaling, *Sox4* knockout intestines have upregulated expression of ISC-associated transcripts, including those associated with active (A) and reserve (B) ISC populations. Letters denote significance: a = $p < 0.01$; b = $p < 0.05$.

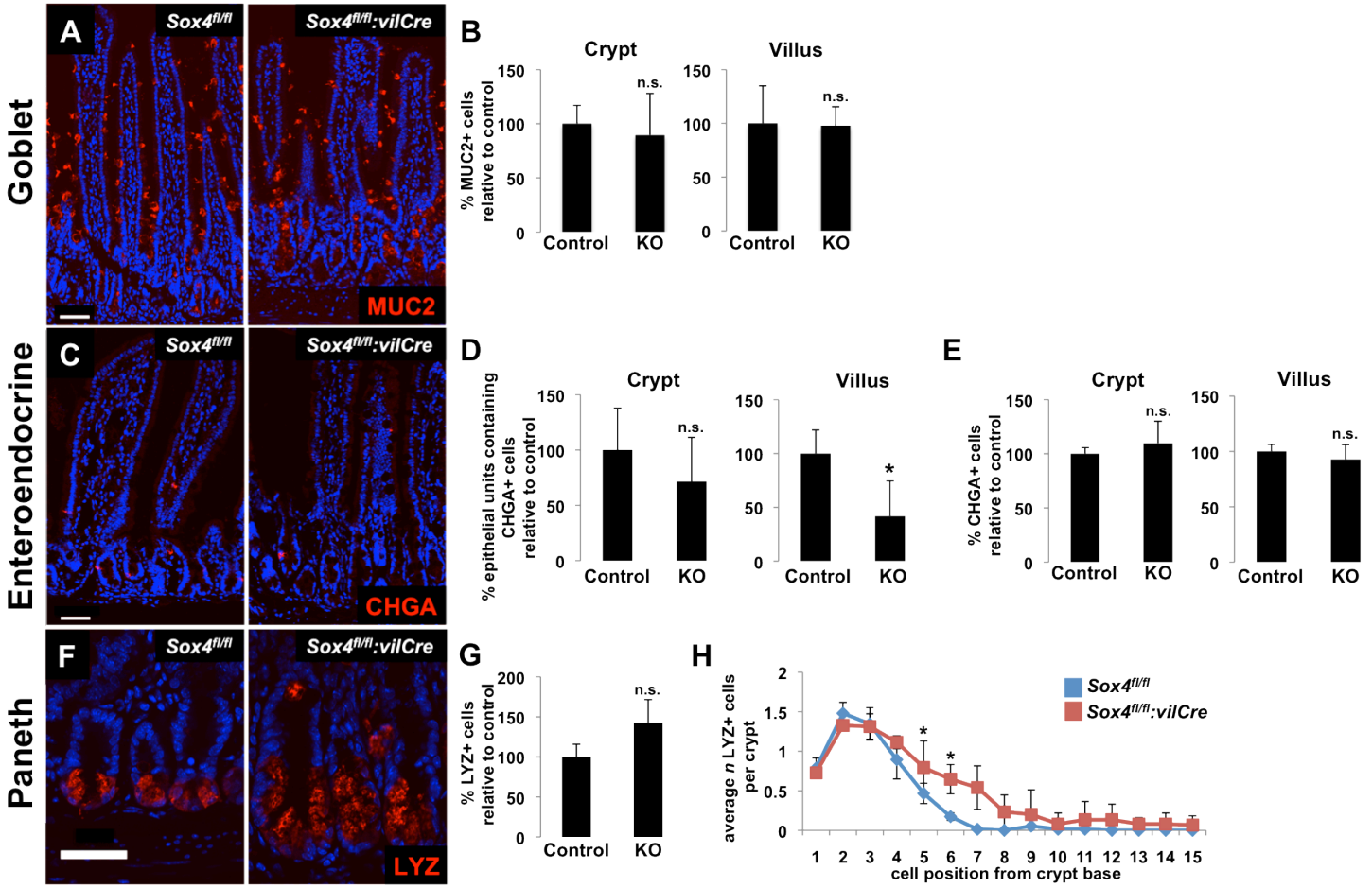


***Figure 4.7 – Sox4* regulates secretory lineage allocation in the intestinal epithelium.** Assessment of secretory lineages in *Sox4* knockout intestines demonstrates changes in differentiation patterns. Numbers of crypt- and villus-based MUC2+ goblet cells remain unchanged between knockouts and controls (n = 3 biological replicates per condition; 50 units per replicate) (A&B). Identification of EE cells by CHGA reveals an appreciable decrease in EE cells in *Sox4* knockout intestines (C) and quantification demonstrates a significant decrease in the number of villi, but not crypts, that contain CHGA+ cells (D). However, in crypt and villus units that

contain at least once CHGA+ cell, the average number of CHGA+ cells remains unchanged between knockouts and controls (n = 3 biological replicates per condition; 50 units per replicate) (E). LYZ+ Paneth cells appear at greater numbers and at higher cell positions in knockout crypts (F), and quantification shows an increase in Paneth cells per crypts in knockout intestines, but does not meet statistical significance (G). Paneth cells are, however, found at significantly increased numbers at higher cell positions in knockout crypts, especially in what is traditionally considered the supra-Paneth zone (+4-+8) (n = 4 biological replicates per condition; 25 crypts per replicate) (H). Scale bars represent 50μm; asterisk indicates significance $p < 0.05$.

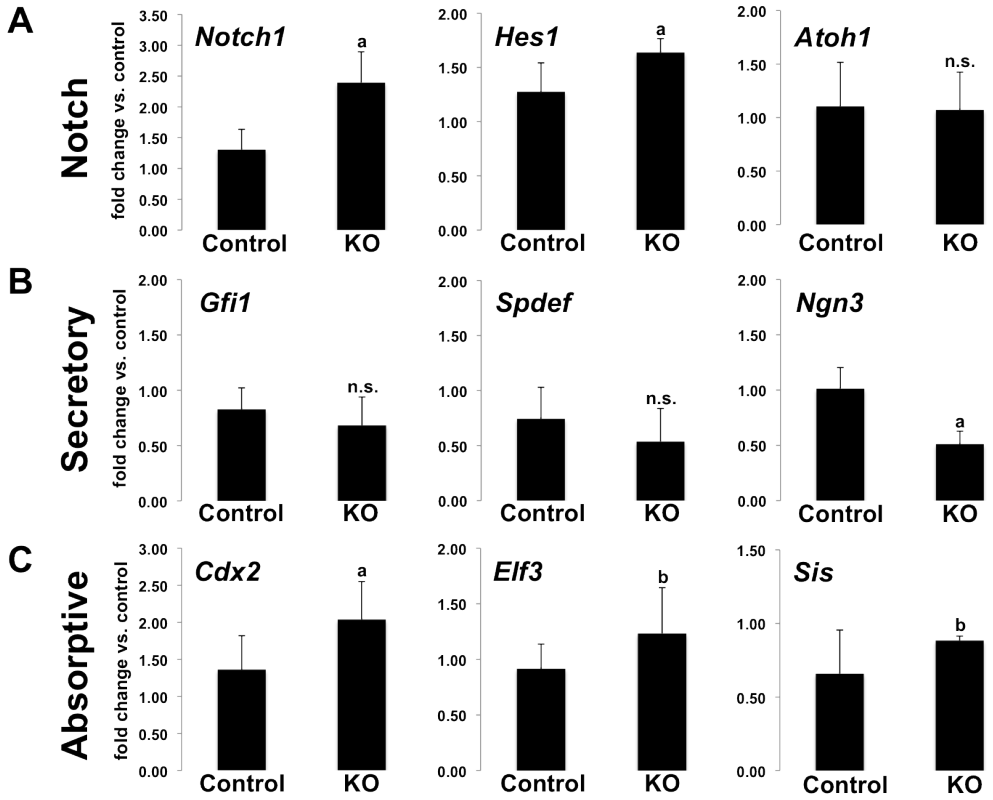


***Figure 4.8* – Expression of transcriptional regulators of differentiation is altered following the loss of *Sox4*.** Analysis of upstream pathways involved in intestinal differentiation reveals increased Notch signaling, through upregulated expression of the *Notch1* receptor and direct Notch target *Hes1*, while *Atoh1* remains unchanged (A). Consistent with histological observations, the goblet cell-associated transcription factors *Gfi1* and *Spdef* are unchanged between knockout and control intestines, but EE cell-associated *Ngn3* is significantly downregulated (~2-fold) in knockouts (B). Consistent with increased *Notch1* and *Hes1* activity, absorptive enterocyte transcription factors *Cdx2* and *Elf3* are upregulated in knockout intestines, as is the functional absorptive marker, *Sis* (C). Letters denote significance: a = $p < 0.01$; b = $p < 0.05$.

A

B

Gene symbol	Gene name	Fold Change	FDR
Nts	neurotensin	-90.91	0.00
Gcg	glucagon	-37.27	0.00
Cck	cholecystokinin	-3.53	2.91
Gip	gastric inhibitory polypeptide	-2.32	3.75
Chga	chromogranin A	-2.00	3.75
Isl1	islet 1	-2.26	3.75
Tet1	ten-eleven translocation 1	-2.03	3.75
Prom1	prominin 1	2.00	2.90
Reg3g	regenerating islet-derived 3 gamma	10.37	3.75
Muc4	mucin 4	3.43	3.75

***Figure 4.9* – RNA microarray identifies significantly regulated targets of *Sox4*.** Genome-wide gene expression analysis reveals reproducible clustering of *Sox4* knockout samples (A). Transcripts that meet analytical requirements for significant regulation show the greatest downregulation of genes associated with EE cells (*Nts*: -90.19-fold; *Gcg*: -37.27-fold), as well as upregulation of ISC- and Paneth-associated genes (B). Transcripts with a fold-change ≥ 2.00 and false discovery rate (FDR) < 5.00 were considered significantly regulated.

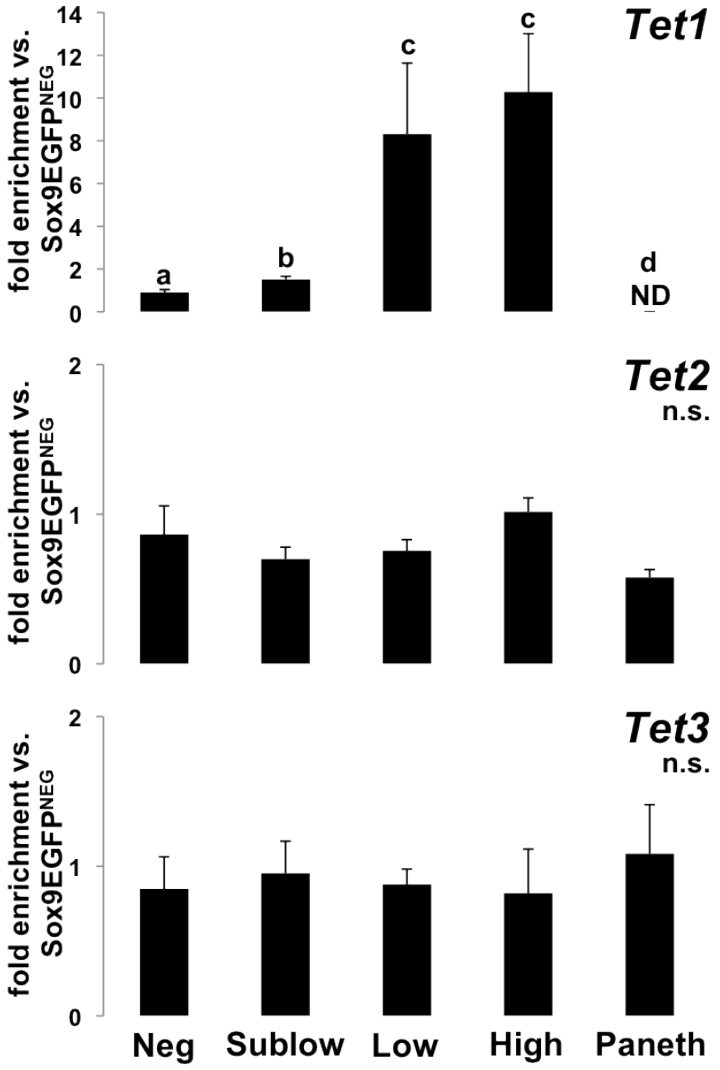


***Figure 4.10 – Tet1-3* are expressed in intestinal epithelial subpopulations.** qPCR on FACS isolated *Sox9*^{EGFP} subpopulations demonstrates that all three *Tet* genes are expressed in the intestinal epithelium. *Tet1* is specifically upregulated in active ISCs (*Sox9*^{EGFP}low) and EEs/reserve ISCs (*Sox9*^{EGFP}high). In contrast, *Tet2* and *Tet3* are detected across all epithelial populations and do not demonstrate significant upregulation in differentiated villus cells (*Sox9*^{EGFP}negative:CD24^{neg}), TAs (*Sox9*^{EGFP}sublow), active ISCs (*Sox9*^{EGFP}low), EEs/reserve ISCs (*Sox9*^{EGFP}high), or Paneth cells (*Sox9*^{EGFP}negative:CD24^{high}:SSC^{high}). Different letters indicate significance between groups, $p < 0.05$. ND = not detected.

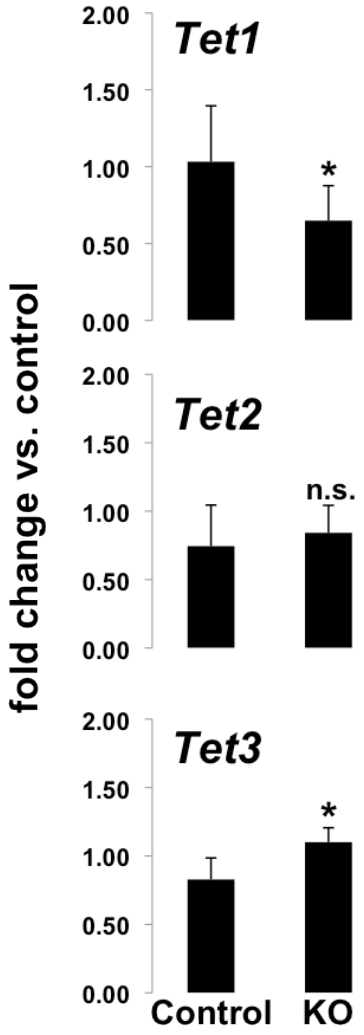


***Figure 4.11* – Dysregulation of *Tet* expression following the loss of *Sox4*.** *Tet1* is significantly downregulated in *Sox4* knockout intestines, consistent with RNA microarray data. While *Tet2* expression remains unchanged, *Tet3* is significantly upregulated, suggesting possible functional compensation for decrease *Tet1*.

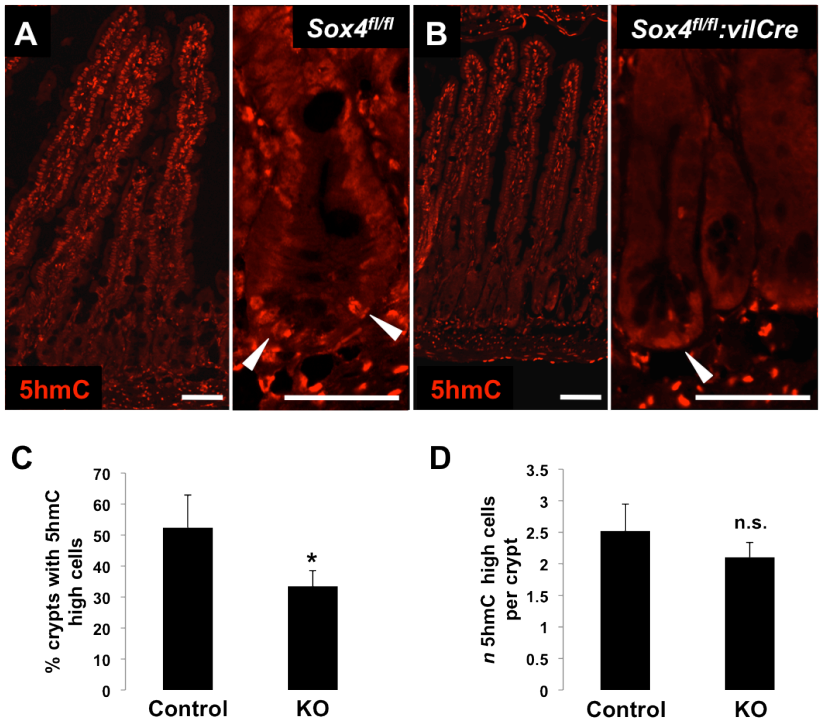


***Figure 4.12 – Sox4* knockout intestines demonstrate reduced 5hmC.** Antibody-based detection of 5hmC, the catalytic byproduct of TET activity, reveals a gradient of 5hmC intensity across the crypt-villus axis in control animals, with low 5hmC expression in the crypt and increasingly intense expression moving toward the villus tip (A). While present, this gradient is disrupted in *Sox4* knockout intestines, with knockout crypts being almost completely devoid of 5hmC-positive cells, and 5hmC intensity increasing at a slower rate up the crypt-villus axis (B). Quantification of 5hmC-high expressing cells, which are consistent with Paneth cells in both knockouts and controls as well as upper crypt cells in controls, reveals a significant decrease in this population in knockout crypts (A-C). However, in knockout crypts that contain at least one 5hmC-high cell, the average number of 5hmC-high cells per crypt unit remains unchanged between knockouts and controls (n = 3 biological replicates per condition; 50 crypts per replicate) (D). Scale bars represent 50μm; asterisk indicates significance $p < 0.05$.

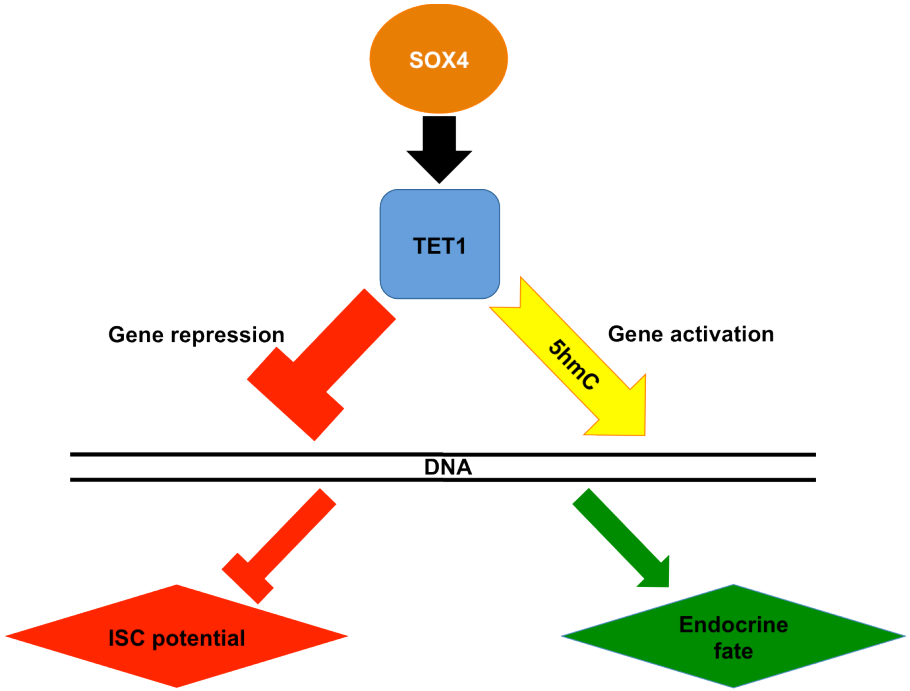


***Figure 4.13* – Proposed model for regulation of ISC potency and differentiation by *Sox4*.** Data from *Sox4* knockout studies suggests that *Sox4* regulates the exit of ISCs and early secretory progenitors from functional potency and induces genetic programs associated with secretory cell fates, especially favoring of EE cell differentiation. Because TET1 can function epigenetically as both a transcriptional activator through DNA demethylation, and a transcriptional repressor through PRC2-mediated histone methylation, it is possible that *Sox4* simultaneously regulates the loss of potency and acquisition of secretory fate through a single *Tet1*-dependent mechanism. Partial compensation by *Tet3* may influence continued Paneth and EE cell differentiation in *Sox4* knockout intestines.

Gene symbol	Probe ID
18S	Hs99999901_s1
Ascl2	Mm01268891_g1
Atoh1	Mm00476035_s1
Bmi1	Mm03053308_g1
Ccnd1	Mm00432359_m1
Cd44	Mm01277163_m1
Cdx2	Mm01212280_m1
Elf3	Mm01295975_m1
Gfi1	Mm00515855_m1
Hes1	Mm01342805_m1
Lgr5	Mm00438890_m1
Lrig	Mm00456116_m1
Myc	Mm00487804_m1
Ngn3	Mm00437606_s1
Notch1	Mm00435249_m1
Olfm4	Mm01320260_m1
Sis	Mm01210305_m1
Sox11	Mm01281943_s1
Sox12	Mm02529758_s1
Sox4	Mm00486317_s1
Sox9	Mm00448840_m1
Spdef	Mm00600221_m1
Tert	Mm00436931_m1
Tet1	Mm01169087_m1
Tet2	Mm00524395_m1
Tet3	Mm00805756_m1

Table 4.1 **– Taqman probes used in gene expression studies**

CHAPTER 5

SCOPE OF WORK, SIGNIFICANCE, AND IMPLICATIONS FOR FUTURE STUDIES

ISOLATION STRATEGIES

Isolation of human ISCs

Building on our previous work in murine ISC isolation, the data presented in Chapter 2 of this dissertation include isolation methods for human ISCs [21, 149]. Interestingly, the use of cell surface markers CD24 and CD44 was able to differentially enrich for populations with characteristics of active (CD24-/CD44+) and reserve (CD24+/CD44+) ISCs from total epithelial preparations [149]. The core contribution of these data is that they provide methodology for the isolation, genetic analysis, and *in vitro* expansion of primary human ISCs. While previous studies have described culture conditions suitable for culturing primary human intestinal crypts, our study represents the first to isolate and culture single human ISCs, as well as the first to characterize ISC populations with distinct genetic and functional characteristics in humans [83].

Interestingly, human CD24 did not enrich for *LYZ* or *DEFA* transcripts, suggesting a key difference between murine Paneth cells, which are known to express CD24, and human Paneth cells, which appear to be CD24-negative. One possibility we considered is that the qPCR probes employed in our study do not detect variants of *LYZ* and *DEFA* in human Paneth cells. We did, however, co-stain histological sections from

the same tissue samples that were subjected to FACS isolation and gene expression analysis. Confocal microscopy appears to indicate that LYZ+ human Paneth cells are indeed CD24-negative, as we never observed CD24+ staining on these cells (Figure 5.1). More comprehensive studies involving 3D reconstruction allowing for examination of the entire Paneth cell apical membrane are required to confirm these preliminary observations. Additionally, these studies would benefit from the use of additional CD24 antibodies, to assure that these data are not the result of the current antibody being unable to detect a splice variant of CD24 expressed on human Paneth cells.

Improved isolation of Paneth cells from murine intestine

In addition to the new isolation methods for human ISCs described in Chapter 2, we also presented methodology for the isolation of high purity Paneth cell populations from murine intestinal epithelium (Chapter 3). Paneth cells are implicated in a wide range of physiological and disease functions in the intestine, ranging from contribution to the ISC niche to interaction with the commensual intestinal microbiome, but studies on their molecular and genetic function are limited by the lack of Paneth cell specific reporter mice [13]. In our study, we take advantage of an earlier observation that *Sox9*^{EGFP} transgenic mice accurately report expression of *Sox9* through *EGFP* expression in all cells of the crypt, with the notable exception of Paneth cells [35]. This allowed us to further purify Paneth cells from CD24-based FACS parameters by excluding any contaminating ISCs, TAs, or enteroendocrine cells. Our results, presented in Chapter 3, demonstrate a significant de-enrichment of contaminating cell markers, including *Lgr5* (ISCs) and *Chga* (EE cells). Recent studies have demonstrated some overlap between Paneth cell markers,

Lgr5, and *Chga*, suggesting that Paneth cells may express some level of these genes *in vivo*, and possibly explaining why these genes were detected at any level in our Paneth cell population [34].

INTRINSIC REGULATION OF POTENCY IN ISCS

Transcriptional regulation of differentiation by Sox4

In order to address intrinsic regulation of potency in ISCs, we examined the role of *Sox4* in the intestinal epithelium. While multiple *Sox* factors are expressed in the intestine, information regarding their function remains largely unknown [107]. The only *Sox* factor that is well characterized in the intestinal epithelium is *Sox9*, which is involved in negatively regulating proliferation and plays a role in Paneth cell differentiation [108, 110]. Additionally, *Sox9* is expressed throughout the crypt, and distinct expression levels of *Sox9* facilitate differential isolation of ISCs, TAs, EE cells, Paneth cells, and differentiated villus-based epithelium [21, 35, 36]. We originally became interested in *Sox4* due to preliminary data demonstrating restricted expression at the crypt base, as well as multiple studies in a wide range of tissue systems that implied a role for *Sox4* in stem cell potency and differentiation (Magness, unpublished) [117, 118, 120, 125, 130, 181]. We originally hypothesized that *Sox4* and *Sox9*, as the predominantly expressed *Sox* factors in the intestinal crypt, might play opposing roles, with *Sox4* acting as a positive regulator of proliferation. This model would have been in agreement with *in vitro* studies in CRC cell lines that demonstrated positive regulation of canonical Wnt signaling by SOX4 through direct binding and enhancement of β-catenin activity [114]. We were surprised at our initial results, which reveal an opposite function. Together, our studies establish novel roles for

Sox4 as a mediator of ISC/TA proliferation, EE cell differentiation, and Paneth cell position in the intestinal epithelium. We also observe changes in gene expression programs associated with absorptive enterocytes, suggesting that *Sox4* may also regulate this lineage.

Lingering questions from the studies presented in Chapter 4 of this dissertation include those focused on understanding the precise mechanisms through which *Sox4* controls proliferation and differentiation. Under our current working model, we hypothesize that loss of *Sox4* causes differentiation of intestinal stem and progenitor cell populations to become inefficient (Figure 5.2). In this model, ISCs/TAs are still capable of differentiation through compensatory means, but the process takes place in a delayed manner, as unknown compensatory mechanisms are less efficient at inducing differentiation that *Sox4*. Precedence is found in studies that demonstrate that the loss of closely related SoxC factor *Sox11* causes delayed differentiation in retinal progenitors [176]. Under these assumptions, the proliferative defect in *Sox4*-knockout intestines would be the indirect result of accumulating progenitor cells (Figure 5.2B). This could also explain the small, but significant increase in ISC markers, as well as increased Wnt signaling. Further examination of Wnt signaling components in *Sox4*-knockout intestines hints at compensation. Our data demonstrate significant upregulation of the WNT co-receptors *Lrp6* and *Fzd5* in *Sox4*-knockout mice, but also show an upregulation of the ubiquitin ligase *Znrf3*, which negatively regulates Wnt signaling by targeting receptor complexes for ubiquitin-dependent degradation (Figure 5.3) [182]. In order to further understand the apparently complex regulatory network governed by *Sox4*, we are currently conducting

ChIP-seq experiments to identify direct targets of SOX4, which can then be correlated with our gene expression data to determine the transcriptional function of SOX4 at target loci. In the absence of ChIP data for SOX4, we were able to identify *Tet1* as a candidate target of the *Sox4* regulatory network. We propose that *Sox4* mediates ISC/TA differentiation upon upregulation by inducing *Tet1* and initiating depression of differentiation-associated genes through DNA demethylation (Figure 5.2A). Interestingly, timing of retinal differentiation by *Sox11* may also be mediated through epigenetic mechanisms, perhaps suggesting a common theme for SoxC factors in differentiation [176].

While the data presented here suggest a regulatory relationship between *Sox4* and *Tet1*, further studies are needed to confirm this relationship. Although a standard *in vitro* approach using established CRC cell lines would arguably be the most straightforward, previously published studies demonstrate a fundamentally different role for *Sox4* in CRC cell lines than we found *in vivo* [114]. Due to the possibility that *Sox4* behaves differently in normal tissue *in vivo* and in cancer cell *in vitro*, results from such studies might not accurately portray the physiological function of *Sox4*. Several recent reports have described methods for gene expression in primary intestinal cultures using viral vectors [183, 184]. We are currently optimizing these protocols in our lab, and future studies will focus on confirming our *in vivo* observations that *Sox4* positively regulates *Tet1*. We plan on inducing expression of *Sox4* to determine if this results in an increase in *Tet1* and subsequent expression of EE cell associated genes. Concomitant inhibition of *Tet1* using siRNA could be employed to test if *Sox4* induction of EE differentiation is dependent on *Tet1* expression, as hypothesized in our model (Chapter 4).

Possible roles for Sox4 in epithelial regeneration

One avenue of study, which is not addressed in this dissertation, is the possible role of *Sox4* in intestinal epithelial regeneration following damage. The most commonly employed model for intestinal regeneration is irradiation damage, which induces well-defined stages of p53-dependent and p53-independent apoptosis and proliferation in murine intestine [185-187]. Interestingly, *in vitro* studies focused on the role of *SOX4* in DNA damage response demonstrated that the SOX4 protein directly binds to and stabilizes P53 following irradiation damage, and that this interaction is required to induce p53-dependent apoptosis [188]. Our preliminary data demonstrate that apoptosis is indeed decreased in the crypts of $Sox4^{fl/fl}$:*vilCre* mice at 4.5hr following 12Gy irradiation, compared to $Sox4^{fl/fl}$littermate controls (Figure 5.4). However, the long-term implications of this reduction in initial apoptosis are unclear. Does loss of *Sox4* reduce radiosensitivity in a physiologically beneficial manner, or does the observed reduction in apoptosis cause defective regeneration by genetically compromised cells at later time points? Further investigation into the *in vivo* role of *Sox4* in intestinal regeneration is warranted.

Loss of *Sox4* may also affect intestinal regeneration through apoptosis-independent mechanisms. Our current working model of *Sox4*-mediated differentiation hypothesizes that the loss of *Sox4* results in the build-up of uncommitted progenitor cells that harbor some degree of ISC marker expression (Figure 5.2B). Additionally, our observations demonstrate lineage allocation defects in EE cells and Paneth cells (Chapter 4). Recent studies have shown that secretory progenitors with EE cell/Paneth cell

potential are capable of dedifferentiation and ISC function following irradiation damage [34, 39]. If loss of *Sox4* results in the increase of these cell numbers, it is possible that regeneration would proceed at an accelerated rate in *Sox4*-knockout intestines, as the starting pool of facultative ISCs would be greater than in control mice. Studies aimed at culturing DLL1^{High} and *Sox9*^{EGFP}high cells isolated from *Sox4*$^{fl/fl}$*:vilCre* intestinal epithelium are currently underway to determine if ISC potential is increased in these populations under physiological conditions [36, 39]. These studies would lay the groundwork for further investigation of regulation of facultative stemness by *Sox4* in the setting of irradiation injury.

Epigenetic regulation of ISC-associated genetic programs by Tet1

Regulation of DNA methylation by TET proteins is currently attracting significant interest in the fields of epigenetics and embryonic stem cell (ESC) biology. From a biochemical standpoint, the identification of TET enzymes, which oxidize 5mC to produce 5hmC, addresses important questions regarding the demethylation and subsequent derepression of methylated genes [168]. Multiple high impact studies are characterizing the impact of *Tet* function in the establishment and maintanence of pluripotency in ESCs, as well as in the initiation of differentiation. In addition to these studies, emerging work is characterizing the role of *Tet* in cancer and adult stem cell populations, including hematopoietic stem cells and neural stem cells [189]. The present study is, to our knowledge, the first to examine *Tet* genes in the intestinal epithelium.

We find that all three *Tet* genes are expressed in the intestinal epithelium (Chapter 4). However, the expression patterns of *Tet2/3*, which are expressed at apparently equal levels throughout the epithelium, differ from that of *Tet1*, which is specifically upregulated in $Sox9^{EGFP}$low and $Sox9^{EGFP}$high ISC and ISC/EE cell populations. These data suggest a differential role for *Tet1* that is unique to ISCs. Since 5hmC expression is observed throughout the intestinal epithelium, it is possible that *Tet1* establishes 5hmC marks in ISCs/TAs, and *Tet2/3* are redundantly responsible for maintaining and/or amplifying these marks in villus cells, which express higher 5hmC levels. Interestingly, *Sox4*-knockout mice, which demonstrate significant downregulation of *Tet1* still have substantial 5hmC expression in villi, despite delayed establishment of the 5hmC crypt-villus gradient. We find that *Tet3* is upregulated in the absence of *Sox4*, suggesting some degree of functional compensation for *Tet1*. Studies are underway to examine the expression of *Tet1-3* in $Sox9^{EGFP}$ subpopulations isolated from $Sox4^{fl/fl}$:*vilCre* mice, to determine if *Tet* loss and compensation occur in specific epithelial compartments.

The functional and mechanistic implications of *Tet*-mediated demethylation are perhaps so captivating because of the potential reached by broad, epigenetic transcriptional regulation. TETs are capable of activating genes associated with either potency or differentiation through a single DNA methylation-related mechanism [169, 175, 190]. Adding another level of epigenetic complexity to transcriptional regulation by TETs is their demonstrated ability to affect gene repression through recruitment of PRC2 complexes and subsequent methylation of core histones [170, 171]. The potential of TETs to

regulate gene expression at the level of both DNA and histone methylation implies central and important roles for these proteins in the regulation of stem cell behavior.

Establishing *Tet1* as a direct target of *Sox4* would provide valuable insight into the mechanism by which *Sox4* influences intestinal differentiation, and imply a broad role for *Sox4* in the initiation of epigenetic changes that influence ISC behavior. We are currently conducting parallel ChIP-seq experiments to identify direct targets of both *Sox4* and *Tet1* in the intestinal epithelium. Reconciliation of these data with our RNA microarray data from *Sox4* knockout intestines will provide valuable insight into the shared *Sox4*-*Tet1* regulatory network. If *Tet1* is confirmed as a target of *Sox4*, then future studies would benefit from identifying which direct targets of *Tet1* are regulated in response to induction by *Sox4*. Since *Tet1* is involved in a broad range of gene regulatory processes, it would be important to establish if *Sox4* is involved in all *Tet1*-mediated signaling in ISCs, or just a subset.

It would be of great interest to determine if *Tet1* is a conserved target of *Sox4* in other tissues as well. The ability of *Sox4* to act indirectly through broad epigenetic regulatory elements could help explain its diverse functions across different tissue types and in tumors. The availability of floxed *Sox4* alleles and tissue specific *Cre* driver mouse lines would facilitate relatively easy investigation in this area.

OUTSIDE INFLUENCES: THE IMPACT OF EXTRINSIC FACTORS ON ISC BEHAVIOR

Differential requirements for distinct populations of human ISCs

In Chapter 2 of this dissertation, we present data that demonstrate the differential isolation of two distinct human ISC populations. Initial gene expression studies led us to hypothesize that CD24-/CD44+ and CD24+/CD44+ populations were unique from one another given differential enrichment of *LGR5*, which is associated with active CBC ISCs, and *HOPX*, which is associated with +4 reserve ISCs [19, 29]. However, controversial genome-wide gene and protein expression studies have suggested that ISC biomarkers associated with reserve cells, including *Hopx*, *Bmi1*, *Tert*, and *Lrig* are not expressed separately from *Lgr5* in mice [33]. Therefore, we sought to establish a difference between CD24-/CD44+ and CD24+/CD44+ human intestinal epithelial populations through functional assays. We found that active-associated CD24-/CD44+ cells are functionally responsive to GSK-β inhibition, while reserve-associated CD24+/CD44+ cells are not [149]. Co-culture with subepithelial myofibroblasts in the absence of GSK-β inhibition was sufficient to induce growth and multipotency in CD24+/CD44+ cells [149]. Remarkably, the reserve-associated cells, which did not exhibit any long-term growth in the absence of myofibroblasts, survived at higher rates than the active-associated cells in co-culture conditions [149]. These data suggest different extrinsic signaling requirements for active and reserve human ISCs and provide novel insight into functional differences between these two populations.

CHIR99021: Wnt-driver or black flag operative?

While CHIR99021, which potently inhibits GSK-β, is a critical component in the culture of primary human ISCs, its exact mechanism of action remains unclear. In studies focused on the expansion of murine hematopoietic stem cells and ISCs *in vitro*, the beneficial effects of CHIR99021 have been attributed to the stabilization of β-catenin and increase in canonical Wnt signaling through the inhibition of the β-catenin destruction complex, which contains GSK-β [84, 191]. However, these studies do not address the effect of GSK inhibition on insulin signaling, which has been shown to be a critical mitogenic pathway in the intestine that drives ISC expansion and may also be an important mediator of CHIR99021's effects on ISCs *in vitro* [192]. In our studies, presented in Chapter 2 of this dissertation, we obtained tissue specimens from morbidly obese patients. After several initial attempts at culturing isolated cells from these specimens failed, we reasoned that defective insulin signaling, a common clinical sequelae of obesity, could be contributing to cell death *in vitro*. We initially tried to supplement our cultures with high doses of exogenous insulin, but noticed no change in survival (data not shown). However, CHIR99021 was capable of potently increasing survival of the CD24-/CD44+ active ISC population [149]. We found this result interesting in light of the fact that the cultures already contained high doses of recombinant WNT3A and RSPO1, both potent drivers of canonical Wnt signaling. These observations may suggest that the tissue used in our study was largely resistant to exogenous insulin, a characteristic associated with clinical obesity, and that addition of CHIR99021 helped support ISC survival and growth through bypassing and acting downstream of the insulin receptor complex. Further research in this area is also warranted by the observation that human ISC populations are

differentially responsive to CHIR99021, which may provide valuable insight into ISC population dynamics in human physiology and disease.

Paneth cells and the possibility of niche function

One of the most unexpected results presented in this dissertation is the finding that Paneth cells do not promote the formation of enteroids by single ISCs *in vitro* (Chapter 3). These data are in conflict with previous reports that the niche signaling produced by Paneth cells, particularly in the form of WNT3A, is critical for enteroid establishment and development *in vitro* [8]. However, experiments demonstrating increased ISC growth in the presence of Paneth cells have previously relied on the co-culture of very large cell numbers [8, 85]. Our study is the first, to our knowledge, to examine co-cultures of ISCs and niche cells at physiologically relevant numbers. As such, it is possible that very high numbers of Paneth cells are required to induce an effect on ISCs *in vitro*. Additionally, when we isolated Paneth cells using previously described $CD24^{high}$:SSC^{high} methods, we observed a significant amount of growth from contaminating ISCs in the Paneth cell preparation (data not shown). This may partially account for the increased survival observed in previous studies.

An additional possibility that may account for the lack of a significant increase in ISC survival in the presence of Paneth cells is that direct cell-cell contact is required for this effect. Though not definitively proven, previous studies have hinted at this requirement, as Paneth cells are known to produce Notch ligands [8]. We are currently analyzing the data presented here with the additional criteria of cell-cell contact between

ISCs and Paneth cells. These data will allow us to compare the survival percentages of ISCs and Paneth cells based on direct contact, and may reveal a significant increase in enteroid formation dependent on cell-cell contact. If so, these data would establish a mechanism for the enhancement of ISC growth by Paneth cells *in vitro*.

ISC culture methods and recapitulation of the stem cell niche in vitro

Another novel method presented in this dissertation is the establishment of a high-throughput culture platform for clonal ISC growth (Chapter 3). This is the first methodology that allows for confirmed clonal analysis of ISCs *in vitro*, as well as the first to facilitate high-throughput screening of ISCs. Future applications will include screening of proteins and small molecules to determine their effect on single ISCs. Additionally, this methodology is applicable to the *in vitro* reconstitution of the ISC niche. While the proof of principle data presented here focus on Paneth cells, we have begun preliminary studies to examine the impact of subepithelial myofibroblasts on ISCs in the microwell array platform. Myofibroblasts are known to secrete different WNT ligands than Paneth cells, though it is unknown if the functional effect of varying ligands differs *in vitro* [62]. Ongoing efforts are also focused on the retrieval of early enteroids using the raft retrieval technology integrated into the microwell array platform [145]. Single enteroid retrieval is being combined with microfludic-assisted gene expression analysis to determine the genetic regulatory programs underlying enteroid development.

Sox4 as a Wnt and Bmp response element in ISCs and progenitors

Understanding how the intrinsic regulation of differentiation by *Sox4* is affected by extrinsic cues is critical to establishing a comprehensive role for *Sox* factors in intestinal epithelial physiology. Previous studies have established that *Sox4* can respond to both Wnt and Bmp signaling pathways, though it remains unknown precisely how these pathways impact the induction of *Sox4* and downstream gene activation in the intestinal epithelium [56, 120, 193]. We are currently conducting *in vitro* studies aimed at determining the response of *Sox4* to WNT3A and BMP4 using *Sox4*-knockout and control enteroids. These data will shed light on the upstream regulatory elements involved in *Sox4* signaling and establish how the loss of *Sox4* may impact major signaling networks in the intestinal epithelium.

KEY FINDINGS

1. Cell surface markers facilitate enrichment of active and reserve human ISC populations (Chapter 2).
2. CD24 does not mark human Paneth cells (Chapter 2).
3. Differential culture requirements for active and reserve human ISCs (Chapter 2).
4. Development of a high-throughput, clonogenic platform for ISC culture (Chapter 3).
5. High purity isolation of murine Paneth cells using $Sox9^{EGFP}$ transgenic mice (Chapter 3).
6. Paneth cells do not support ISC growth at physiologically relevant numbers *in vitro* (Chapter 3).

7. *Sox4* is preferentially enriched in ISC and TA populations (Chapter 4).
8. SOX4 is expressed at different levels in CBC and +4 positions (Chapter 4).
9. SOX4high cells are not label-retaining (Chapter 4).
10. *Sox4* negatively regulates proliferation in the intestinal crypts (Chapter 4).
11. *Sox4* is involved in EE cell (positive regulation), Paneth cell (negative regulation), and absorptive enterocyte (negative regulation) differentiation, and not involved in lineage specification of goblet cells (Chapter 4).
12. *Sox4* negatively regulates Wnt signaling and ISC gene expression (Chapter 4).
13. *Sox4* positively regulates *Tet1* and contributes to DNA methylation dynamics in the intestinal crypts (Chapter 4).
14. *Tet1* is preferentially enriched in ISCs; *Tet2/3* are expressed throughout the intestinal epithelium (Chapter 4).
15. *Tet3* may be able to partially compensate for a reduction in *Tet1* expression levels (Chapter 4).

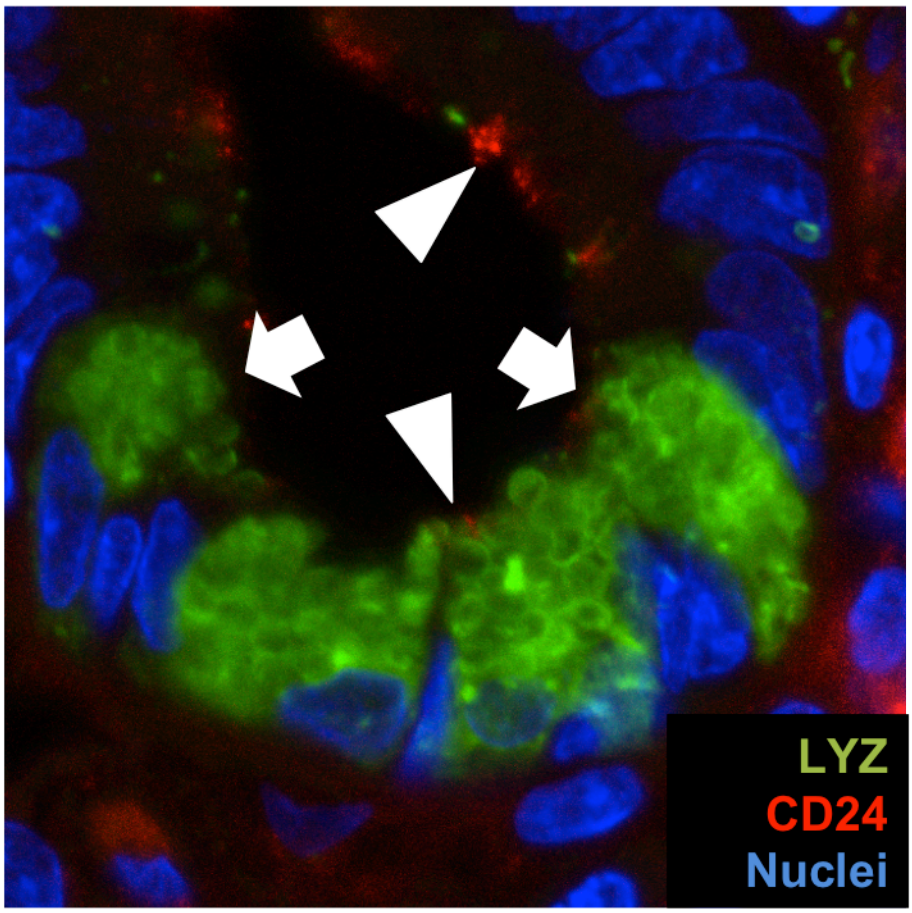


***Figure 5.1* – Human Paneth cells do not express CD24.** Immunofluorescence against CD24, which is expressed on the apical cell membrane throughout the intestinal crypts, demonstrates that human Paneth cells do not express CD24 (arrows), but are intercalated between CD24+ cells (arrowheads).

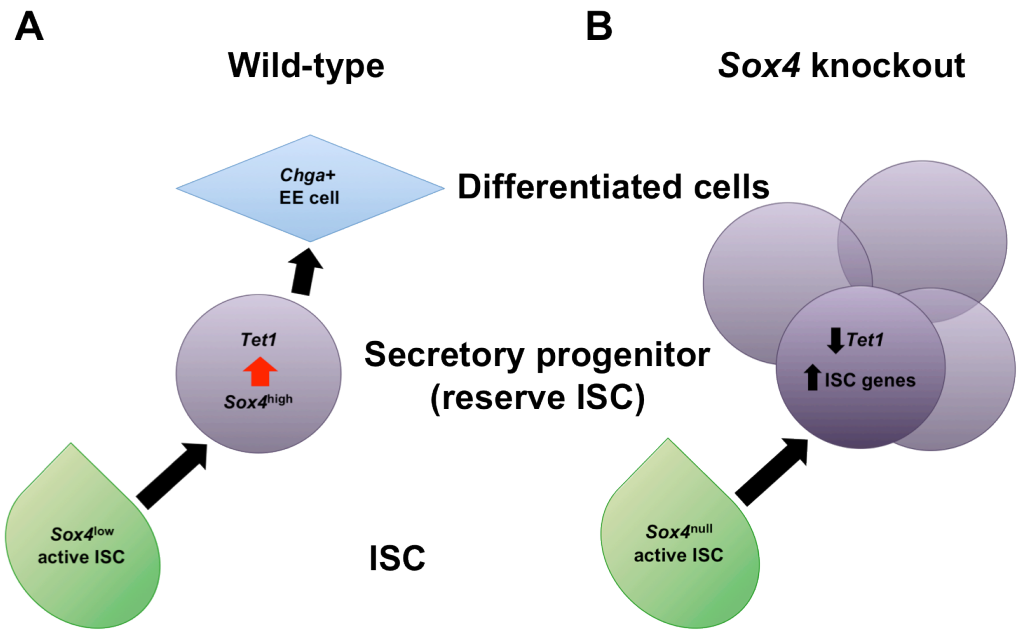


***Figure 5.2* – Working model for indirect regulation of ISC potency by *Sox4*.** Our data suggest that, under physiologic conditions, *Sox4* is expressed at low levels in ISCs and upregulated in secretory progenitor cells, which may be consistent with previously described facultative ISCs (A). Increased *Sox4* expression induces high levels of *Tet1*, which influence EE cell differentiation through the epigenetic derepression of EE-associated transcription factors, such as *Ngn3* (A). In the absence of *Sox4*, cells are eventually able to differentiate through compensatory mechanisms, but inefficient gene regulation leads to delayed differentiation and increased numbers of secretory progenitors, which also exhibit lingering expression of ISC genes and may by functionally competent following injury (B).

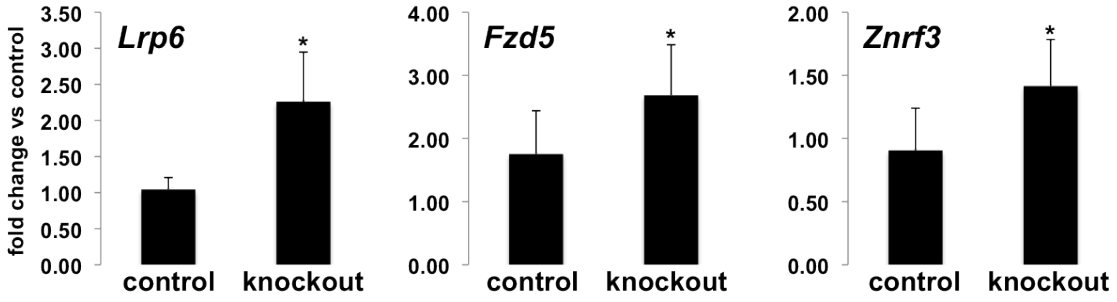


***Figure 5.3* – WNT receptor complexes and their negative regulators are upregulated in *Sox4* knockout intestines.** *Sox4* knockout intestines show significant upregulation of the WNT coreceptors *Lrp6* and *Fzd5*. *Znrf3*, a ubiquitin ligase that participates in the active degradation of *Fzd5*, is also upregulated in knockout intestines, suggesting partial compensation in the presence of increase Wnt signaling.

***Figure 5.4* – Apoptosis is impaired in *Sox4* knockout intestines following 12Gy γ-irradiation.** Examination of apoptosis by immunofluorescent detection of cleaved caspase-3 at 4.5hr post-12Gy irradiation reveals reduced apoptotic response in *Sox4* knockout intestines.

REFERENCES

1. Boron, W.F. & Boulpaep, E.L. *Medical physiology : a cellular and molecular approach*, Edn. 2nd. (Saunders/Elsevier, Philadelphia, PA; 2009).

2. Wright, N., Allison, M. *The Biology of Epithelial Cell Populations*. (Clarindon, Oxford; 1984).

3. Cheng, H. & Leblond, C.P. Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. V. Unitarian Theory of the origin of the four epithelial cell types. *Am J Anat* **141**, 537-561 (1974).

4. Cheng, H. Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. II. Mucous cells. *Am J Anat* **141**, 481-501 (1974).

5. Cheng, H. & Leblond, C.P. Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. I. Columnar cell. *Am J Anat* **141**, 461-479 (1974).

6. Cheng, H. & Leblond, C.P. Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. III. Entero-endocrine cells. *Am J Anat* **141**, 503-519 (1974).

7. Cheng, H. Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. IV. Paneth cells. *Am J Anat* **141**, 521-535 (1974).

8. Sato, T. *et al.* Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. *Nature* **469**, 415-418 (2011).

9. Potten, C.S. & Loeffler, M. Stem cells: attributes, cycles, spirals, pitfalls and uncertainties. Lessons for and from the crypt. *Development* **110**, 1001-1020 (1990).

10. Leblond, C.P. & Messier, B. Renewal of chief cells and goblet cells in the small intestine as shown by radioautography after injection of thymidine-H3 into mice. *Anat Rec* **132**, 247-259 (1958).

11. Quastler, H. & Sherman, F.G. Cell population kinetics in the intestinal epithelium of the mouse. *Exp Cell Res* **17**, 420-438 (1959).

12. Walker, B.E. & Leblond, C.P. Sites of nucleic acid synthesis in the mouse visualized by radioautography after administration of C14-labelled adenine and thymidine. *Exp Cell Res* **14**, 510-531 (1958).

13. Clevers, H.C. & Bevins, C.L. Paneth cells: maestros of the small intestinal crypts. *Annu Rev Physiol* **75**, 289-311 (2013).

14. Gerbe, F., Brulin, B., Makrini, L., Legraverend, C. & Jay, P. DCAMKL-1 expression identifies Tuft cells rather than stem cells in the adult mouse intestinal epithelium. *Gastroenterology* **137**, 2179-2180; author reply 2180-2171 (2009).

15. Gerbe, F. *et al.* Distinct ATOH1 and Neurog3 requirements define tuft cells as a new secretory cell type in the intestinal epithelium. *J Cell Biol* **192**, 767-780 (2011).

16. Potten, C.S., Booth, C. & Pritchard, D.M. The intestinal epithelial stem cell: the mucosal governor. *Int J Exp Pathol* **78**, 219-243 (1997).

17. Potten, C.S., Hume, W.J., Reid, P. & Cairns, J. The segregation of DNA in epithelial stem cells. *Cell* **15**, 899-906 (1978).

18. Li, L. & Clevers, H. Coexistence of quiescent and active adult stem cells in mammals. *Science* **327**, 542-545 (2010).

19. Barker, N. *et al.* Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature* **449**, 1003-1007 (2007).

20. Furuyama, K. *et al.* Continuous cell supply from a Sox9-expressing progenitor zone in adult liver, exocrine pancreas and intestine. *Nat Genet* **43**, 34-41 (2011).

21. Gracz, A.D., Ramalingam, S. & Magness, S.T. Sox9-Expression Marks a Subset of CD24-expressing Small Intestine Epithelial Stem Cells that Form Organoids in vitro. *Am J Physiol Gastrointest Liver Physiol* (2010).

22. van der Flier, L.G., Haegebarth, A., Stange, D.E., van de Wetering, M. & Clevers, H. OLFM4 is a robust marker for stem cells in human intestine and marks a subset of colorectal cancer cells. *Gastroenterology* **137**, 15-17 (2009).

23. van der Flier, L.G. *et al.* Transcription factor achaete scute-like 2 controls intestinal stem cell fate. *Cell* **136**, 903-912 (2009).

24. Sato, T. *et al.* Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* **459**, 262-265 (2009).

25. Stelzner, M. *et al.* A nomenclature for intestinal in vitro cultures. *Am J Physiol Gastrointest Liver Physiol* **302**, G1359-1363 (2012).

26. Sangiorgi, E. & Capecchi, M.R. Bmi1 is expressed in vivo in intestinal stem cells. *Nat Genet* **40**, 915-920 (2008).

27. Yan, K.S. *et al.* The intestinal stem cell markers Bmi1 and Lgr5 identify two functionally distinct populations. *Proc Natl Acad Sci U S A* **109**, 466-471 (2012).

28. Montgomery, R.K. *et al.* Mouse telomerase reverse transcriptase (mTert) expression marks slowly cycling intestinal stem cells. *Proc Natl Acad Sci U S A* **108**, 179-184 (2011).

29. Takeda, N. *et al.* Interconversion between intestinal stem cell populations in distinct niches. *Science* **334**, 1420-1424 (2011).

30. Powell, A.E. *et al.* The pan-ErbB negative regulator Lrig1 is an intestinal stem cell marker that functions as a tumor suppressor. *Cell* **149**, 146-158 (2012).

31. Wong, V.W. *et al.* Lrig1 controls intestinal stem-cell homeostasis by negative regulation of ErbB signalling. *Nat Cell Biol* **14**, 401-408 (2012).

32. Hua, G. *et al.* Crypt base columnar stem cells in small intestines of mice are radioresistant. *Gastroenterology* **143**, 1266-1276 (2012).

33. Munoz, J. *et al.* The Lgr5 intestinal stem cell signature: robust expression of proposed quiescent '+4' cell markers. *EMBO J* **31**, 3079-3091 (2012).

34. Buczacki, S.J. *et al.* Intestinal label-retaining cells are secretory precursors expressing Lgr5. *Nature* **495**, 65-69 (2013).

35. Formeister, E.J. *et al.* Distinct SOX9 levels differentially mark stem/progenitor populations and enteroendocrine cells of the small intestine epithelium. *Am J Physiol Gastrointest Liver Physiol* **296**, G1108-1118 (2009).

36. Van Landeghem, L. *et al.* Activation of two distinct Sox9-EGFP expressing intestinal stem cell populations during crypt regeneration after irradiation. *Am J Physiol Gastrointest Liver Physiol* (2012).

37. Tian, H. *et al.* A reserve stem cell population in small intestine renders Lgr5-positive cells dispensable. *Nature* **478**, 255-259 (2011).

38. Potten, C.S. Extreme sensitivity of some intestinal crypt cells to X and gamma irradiation. *Nature* **269**, 518-521 (1977).

39. van Es, J.H. *et al.* Dll1(+) secretory progenitor cells revert to stem cells upon crypt damage. *Nat Cell Biol* (2012).

40. Lopez-Garcia, C., Klein, A.M., Simons, B.D. & Winton, D.J. Intestinal stem cell replacement follows a pattern of neutral drift. *Science* **330**, 822-825 (2010).

41. Snippert, H.J. *et al.* Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell* **143**, 134-144 (2010).

42. Klein, A.M. & Simons, B.D. Universal patterns of stem cell fate in cycling adult tissues. *Development* **138**, 3103-3111 (2011).

43. Livet, J. *et al.* Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* **450**, 56-62 (2007).

44. Kozar, S. *et al.* Continuous Clonal Labeling Reveals Small Numbers of Functional Stem Cells in Intestinal Crypts and Adenomas. *Cell Stem Cell* (2013).

45. Morrison, S.J. & Spradling, A.C. Stem cells and niches: mechanisms that promote stem cell maintenance throughout life. *Cell* **132**, 598-611 (2008).

46. Spradling, A., Drummond-Barbosa, D. & Kai, T. Stem cells find their niche. *Nature* **414**, 98-104 (2001).

47. Clevers, H. The intestinal crypt, a prototype stem cell compartment. *Cell* **154**, 274-284 (2013).

48. Clevers, H. & Nusse, R. Wnt/beta-catenin signaling and disease. *Cell* **149**, 1192-1205 (2012).

49. Crosnier, C., Stamataki, D. & Lewis, J. Organizing cell renewal in the intestine: stem cells, signals and combinatorial control. *Nat Rev Genet* **7**, 349-359 (2006).

50. Noah, T.K. & Shroyer, N.F. Notch in the intestine: regulation of homeostasis and pathogenesis. *Annu Rev Physiol* **75**, 263-288 (2013).

51. Sugimura, R. & Li, L. Noncanonical Wnt signaling in vertebrate development, stem cells, and diseases. *Birth Defects Res C Embryo Today* **90**, 243-256 (2010).

52. Korinek, V. *et al.* Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC-/- colon carcinoma. *Science* **275**, 1784-1787 (1997).

53. Markowitz, S.D. & Bertagnolli, M.M. Molecular origins of cancer: Molecular basis of colorectal cancer. *N Engl J Med* **361**, 2449-2460 (2009).

54. Morin, P.J. *et al.* Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC. *Science* **275**, 1787-1790 (1997).

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

56. Korinek, V. *et al.* Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4. *Nat Genet* **19**, 379-383 (1998).

57. van Es, J.H. *et al.* A critical role for the Wnt effector Tcf4 in adult intestinal homeostatic self-renewal. *Mol Cell Biol* **32**, 1918-1927 (2012).

58. van de Wetering, M. *et al.* The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell* **111**, 241-250 (2002).

59. Van der Flier, L.G. *et al.* The Intestinal Wnt/TCF Signature. *Gastroenterology* **132**, 628-632 (2007).

60. Batlle, E. *et al.* Beta-catenin and TCF mediate cell positioning in the intestinal epithelium by controlling the expression of EphB/ephrinB. *Cell* **111**, 251-263 (2002).

61. Holmberg, J. *et al.* EphB receptors coordinate migration and proliferation in the intestinal stem cell niche. *Cell* **125**, 1151-1163 (2006).

62. Farin, H.F., Van Es, J.H. & Clevers, H. Redundant Sources of Wnt Regulate Intestinal Stem Cells and Promote Formation of Paneth Cells. *Gastroenterology* (2012).

63. Durand, A. *et al.* Functional intestinal stem cells after Paneth cell ablation induced by the loss of transcription factor Math1 (Atoh1). *Proc Natl Acad Sci U S A* **109**, 8965-8970 (2012).

64. Kim, T.H., Escudero, S. & Shivdasani, R.A. Intact function of Lgr5 receptor-expressing intestinal stem cells in the absence of Paneth cells. *Proc Natl Acad Sci U S A* **109**, 3932-3937 (2012).

65. van Es, J.H. *et al.* Wnt signalling induces maturation of Paneth cells in intestinal crypts. *Nat Cell Biol* **7**, 381-386 (2005).

66. Itoh, S., Itoh, F., Goumans, M.J. & Ten Dijke, P. Signaling of transforming growth factor-beta family members through Smad proteins. *Eur J Biochem* **267**, 6954-6967 (2000).

67. Howe, J.R. *et al.* Germline mutations of the gene encoding bone morphogenetic protein receptor 1A in juvenile polyposis. *Nat Genet* **28**, 184-187 (2001).

68. Howe, J.R. *et al.* Mutations in the SMAD4/DPC4 gene in juvenile polyposis. *Science* **280**, 1086-1088 (1998).

69. Haramis, A.P. *et al.* De novo crypt formation and juvenile polyposis on BMP inhibition in mouse intestine. *Science* **303**, 1684-1686 (2004).

70. He, X.C. *et al.* BMP signaling inhibits intestinal stem cell self-renewal through suppression of Wnt-beta-catenin signaling. *Nat Genet* **36**, 1117-1121 (2004).

71. Auclair, B.A., Benoit, Y.D., Rivard, N., Mishina, Y. & Perreault, N. Bone morphogenetic protein signaling is essential for terminal differentiation of the intestinal secretory cell lineage. *Gastroenterology* **133**, 887-896 (2007).

72. Itasaki, N. & Hoppler, S. Crosstalk between Wnt and bone morphogenic protein signaling: a turbulent relationship. *Dev Dyn* **239**, 16-33 (2010).

73. Schroder, N. & Gossler, A. Expression of Notch pathway components in fetal and adult mouse small intestine. *Gene Expr Patterns* **2**, 247-250 (2002).

74. Fre, S. *et al.* Notch lineages and activity in intestinal stem cells determined by a new set of knock-in mice. *PLoS One* **6**, e25785 (2011).

75. van Es, J.H. *et al.* Notch/gamma-secretase inhibition turns proliferative cells in intestinal crypts and adenomas into goblet cells. *Nature* **435**, 959-963 (2005).

76. Pellegrinet, L. *et al.* Dll1- and dll4-mediated notch signaling are required for homeostasis of intestinal stem cells. *Gastroenterology* **140**, 1230-1240 e1231-1237 (2011).

77. Riccio, O. *et al.* Loss of intestinal crypt progenitor cells owing to inactivation of both Notch1 and Notch2 is accompanied by derepression of CDK inhibitors p27Kip1 and p57Kip2. *EMBO Rep* **9**, 377-383 (2008).

78. Jensen, J. *et al.* Control of endodermal endocrine development by Hes-1. *Nat Genet* **24**, 36-44 (2000).

79. Ueo, T. *et al.* The role of Hes genes in intestinal development, homeostasis and tumor formation. *Development* **139**, 1071-1082 (2012).

80. VanDussen, K.L. *et al.* Notch signaling modulates proliferation and differentiation of intestinal crypt base columnar stem cells. *Development* **139**, 488-497 (2012).

81. Yang, Q., Bermingham, N.A., Finegold, M.J. & Zoghbi, H.Y. Requirement of Math1 for secretory cell lineage commitment in the mouse intestine. *Science* **294**, 2155-2158 (2001).

82. Jung, P. *et al.* Isolation and in vitro expansion of human colonic stem cells. *Nat Med* **17**, 1225-1227 (2011).

83. Sato, T. *et al.* Long-term Expansion of Epithelial Organoids From Human Colon, Adenoma, Adenocarcinoma, and Barrett's Epithelium. *Gastroenterology* **141**, 1762-1772 (2011).

84. Wang, F. *et al.* Isolation and characterization of intestinal stem cells based on surface marker combinations and colony-formation assay. *Gastroenterology* **145**, 383-395 e381-321 (2013).

85. Yilmaz, O.H. *et al.* mTORC1 in the Paneth cell niche couples intestinal stem-cell function to calorie intake. *Nature* **486**, 490-495 (2012).

86. Berta, P. *et al.* Genetic evidence equating SRY and the testis-determining factor. *Nature* **348**, 448-450 (1990).

87. Koopman, P., Gubbay, J., Vivian, N., Goodfellow, P. & Lovell-Badge, R. Male development of chromosomally female mice transgenic for Sry. *Nature* **351**, 117-121 (1991).

88. Bowles, J., Schepers, G. & Koopman, P. Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. *Dev Biol* **227**, 239-255 (2000).

89. Wegner, M. All purpose Sox: The many roles of Sox proteins in gene expression. *Int J Biochem Cell Biol* **42**, 381-390 (2010).

90. Schepers, G.E., Teasdale, R.D. & Koopman, P. Twenty pairs of sox: extent, homology, and nomenclature of the mouse and human sox transcription factor gene families. *Dev Cell* **3**, 167-170 (2002).

91. Pevny, L.H. & Lovell-Badge, R. Sox genes find their feet. *Curr Opin Genet Dev* **7**, 338-344 (1997).

92. Ferrari, S. *et al.* SRY, like HMG1, recognizes sharp angles in DNA. *EMBO J* **11**, 4497-4506 (1992).

93. Giese, K., Cox, J. & Grosschedl, R. The HMG domain of lymphoid enhancer factor 1 bends DNA and facilitates assembly of functional nucleoprotein structures. *Cell* **69**, 185-195 (1992).

94. Pontiggia, A. *et al.* Sex-reversing mutations affect the architecture of SRY-DNA complexes. *EMBO J* **13**, 6115-6124 (1994).

95. Smits, P. *et al.* The transcription factors L-Sox5 and Sox6 are essential for cartilage formation. *Dev Cell* **1**, 277-290 (2001).

96. Thein, D.C. *et al.* The closely related transcription factors Sox4 and Sox11 function as survival factors during spinal cord development. *J Neurochem* **115**, 131-141 (2010).

97. Sarkar, A. & Hochedlinger, K. The sox family of transcription factors: versatile regulators of stem and progenitor cell fate. *Cell Stem Cell* **12**, 15-30 (2013).

98. Avilion, A.A. *et al.* Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev* **17**, 126-140 (2003).

99. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861-872 (2007).

100. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-676 (2006).

101. Ambrosetti, D.C., Basilico, C. & Dailey, L. Synergistic activation of the fibroblast growth factor 4 enhancer by Sox2 and Oct-3 depends on protein-protein interactions facilitated by a specific spatial arrangement of factor binding sites. *Mol Cell Biol* **17**, 6321-6329 (1997).

102. Huangfu, D. *et al.* Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nat Biotechnol* **26**, 1269-1275 (2008).

103. Yuan, H., Corbi, N., Basilico, C. & Dailey, L. Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. *Genes Dev* **9**, 2635-2645 (1995).

104. Stefanovic, S. *et al.* Interplay of Oct4 with Sox2 and Sox17: a molecular switch from stem cell pluripotency to specifying a cardiac fate. *J Cell Biol* **186**, 665-673 (2009).

105. Jauch, R. *et al.* Conversion of Sox17 into a pluripotency reprogramming factor by reengineering its association with Oct4 on DNA. *Stem Cells* **29**, 940-951 (2011).

106. Aksoy, I. *et al.* Sox transcription factors require selective interactions with Oct4 and specific transactivation functions to mediate reprogramming. *Stem Cells* (2013).

107. Gracz, A.D. & Magness, S.T. Sry-box (Sox) transcription factors in gastrointestinal physiology and disease. *Am J Physiol Gastrointest Liver Physiol* **300**, G503-515 (2011).

108. Blache, P. *et al.* SOX9 is an intestine crypt transcription factor, is regulated by the Wnt pathway, and represses the CDX2 and MUC2 genes. *J Cell Biol* **166**, 37-47 (2004).

109. Jay, P., Berta, P. & Blache, P. Expression of the carcinoembryonic antigen gene is inhibited by SOX9 in human colon carcinoma cells. *Cancer Res* **65**, 2193-2198 (2005).

110. Bastide, P. *et al.* Sox9 regulates cell proliferation and is required for Paneth cell differentiation in the intestinal epithelium. *J Cell Biol* **178**, 635-648 (2007).

111. Katoh, M. Molecular cloning and characterization of human SOX17. *Int J Mol Med* **9**, 153-157 (2002).

112. Katoh, M. Expression of human SOX7 in normal tissues and tumors. *Int J Mol Med* **9**, 363-368 (2002).

113. Saitoh, T. & Katoh, M. Expression of human SOX18 in normal tissues and tumors. *Int J Mol Med* **10**, 339-344 (2002).

114. Sinner, D. *et al.* Sox17 and Sox4 differentially regulate beta-catenin/T-cell factor activity and proliferation of colon carcinoma cells. *Mol Cell Biol* **27**, 7802-7815 (2007).

115. van de Wetering, M., Oosterwegel, M., van Norren, K. & Clevers, H. Sox-4, an Sry-like HMG box protein, is a transcriptional activator in lymphocytes. *EMBO J* **12**, 3847-3854 (1993).

116. van Houte, L.P. *et al.* Solution structure of the sequence-specific HMG box of the lymphocyte transcriptional activator Sox-4. *J Biol Chem* **270**, 30516-30524 (1995).

117. Schilham, M.W. *et al.* Defects in cardiac outflow tract formation and pro-B-lymphocyte expansion in mice lacking Sox-4. *Nature* **380**, 711-714 (1996).

118. Schilham, M.W., Moerer, P., Cumano, A. & Clevers, H.C. Sox-4 facilitates thymocyte differentiation. *Eur J Immunol* **27**, 1292-1295 (1997).

119. Sun, B. *et al.* Sox4 is required for the survival of pro-B cells. *J Immunol* **190**, 2080-2089 (2013).

120. Ikushima, H. *et al.* Autocrine TGF-beta signaling maintains tumorigenicity of glioma-initiating cells through Sry-related HMG-box factors. *Cell Stem Cell* **5**, 504-514 (2009).

121. Kuwahara, M. *et al.* The transcription factor Sox4 is a downstream target of signaling by the cytokine TGF-beta and suppresses T(H)2 differentiation. *Nat Immunol* **13**, 778-786 (2012).

122. Liber, D. *et al.* Epigenetic priming of a pre-B cell-specific enhancer through binding of Sox2 and Foxd3 at the ESC stage. *Cell Stem Cell* **7**, 114-126 (2010).

123. Bernstein, B.E. *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315-326 (2006).

124. Bergsland, M., Werme, M., Malewicz, M., Perlmann, T. & Muhr, J. The establishment of neuronal properties is controlled by Sox4 and Sox11. *Genes Dev* **20**, 3475-3486 (2006).

125. Bhattaram, P. *et al.* Organogenesis relies on SoxC transcription factors for the survival of neural and mesenchymal progenitors. *Nat Commun* **1**, 9 (2010).

126. Potzner, M.R. *et al.* Sequential requirement of Sox4 and Sox11 during development of the sympathetic nervous system. *Development* **137**, 775-784 (2010).

127. Hoser, M. *et al.* Prolonged glial expression of Sox4 in the CNS leads to architectural cerebellar defects and ataxia. *J Neurosci* **27**, 5495-5505 (2007).

128. Potzner, M.R. *et al.* Prolonged Sox4 expression in oligodendrocytes interferes with normal myelination in the central nervous system. *Mol Cell Biol* **27**, 5316-5326 (2007).

129. Lioubinski, O., Muller, M., Wegner, M. & Sander, M. Expression of Sox transcription factors in the developing mouse pancreas. *Dev Dyn* **227**, 402-408 (2003).

130. Wilson, M.E. *et al.* The HMG box transcription factor Sox4 contributes to the development of the endocrine pancreas. *Diabetes* **54**, 3402-3409 (2005).

131. Itzkovitz, S. *et al.* Single-molecule transcript counting of stem-cell markers in the mouse intestine. *Nat Cell Biol* (2011).

132. Spangrude, G.J., Heimfeld, S. & Weissman, I.L. Purification and characterization of mouse hematopoietic stem cells. *Science* **241**, 58-62 (1988).

133. Magness, S.T. *et al.* A multicenter study to standardize reporting and analyses of fluorescence-activated cell-sorted murine intestinal epithelial cells. *Am J Physiol Gastrointest Liver Physiol* **305**, G542-551 (2013).

134. von Furstenberg, R.J. *et al.* Sorting mouse jejunal epithelial cells with CD24 yields a population with characteristics of intestinal stem cells. *Am J Physiol Gastrointest Liver Physiol* (2010).

135. Kemper, K. *et al.* Monoclonal antibodies against lgr5 identify human colorectal cancer stem cells. *Stem Cells* **30**, 2378-2386 (2012).

136. Ying, Q.L. *et al.* The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519-523 (2008).

137. Lahar, N. *et al.* Intestinal subepithelial myofibroblasts support in vitro and in vivo growth of human small intestinal epithelium. *PLoS One* **6**, e26898 (2011).

138. Hsu, Y.C. & Fuchs, E. A family business: stem cell progeny join the niche to regulate homeostasis. *Nat Rev Mol Cell Biol* **13**, 103-114 (2012).

139. Brinster, R.L. & Zimmermann, J.W. Spermatogenesis following male germ-cell transplantation. *Proc Natl Acad Sci U S A* **91**, 11298-11302 (1994).

140. Shackleton, M. *et al.* Generation of a functional mammary gland from a single stem cell. *Nature* **439**, 84-88 (2006).

141. Till, J.E. & Mc, C.E. A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiat Res* **14**, 213-222 (1961).

142. Kanatsu-Shinohara, M. *et al.* Reconstitution of mouse spermatogonial stem cell niches in culture. *Cell Stem Cell* **11**, 567-578 (2012).

143. Lippmann, E.S. *et al.* Derivation of blood-brain barrier endothelial cells from human pluripotent stem cells. *Nat Biotechnol* **30**, 783-791 (2012).

144. Bjerknes, M. & Cheng, H. The stem-cell zone of the small intestinal epithelium. I. Evidence from Paneth cells in the adult mouse. *Am J Anat* **160**, 51-63 (1981).

145. Wang, Y. *et al.* Micromolded arrays for separation of adherent cells. *Lab Chip* **10**, 2917-2924 (2010).

146. Gach, P.C., Wang, Y., Phillips, C., Sims, C.E. & Allbritton, N.L. Isolation and manipulation of living adherent cells by micromolded magnetic rafts. *Biomicrofluidics* **5**, 32002-3200212 (2011).

147. Vintersten, K. *et al.* Mouse in red: red fluorescent protein expression in mouse ES cells, embryos, and adult animals. *Genesis* **40**, 241-246 (2004).

148. Carpenter, A.E. *et al.* CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol* **7**, R100 (2006).

149. Gracz, A.D. *et al.* CD24 and CD44 Mark Human Intestinal Epithelial Cell Populations with Characteristics of Active and Facultative Stem Cells. *Stem Cells* (2013).

150. Anderson, D.G., Levenberg, S. & Langer, R. Nanoliter-scale synthesis of arrayed biomaterials and application to human embryonic stem cells. *Nat Biotechnol* **22**, 863-866 (2004).

151. Flaim, C.J., Chien, S. & Bhatia, S.N. An extracellular matrix microarray for probing cellular differentiation. *Nat Methods* **2**, 119-125 (2005).

152. Gobaa, S. *et al.* Artificial niche microarrays for probing single stem cell fate in high throughput. *Nat Methods* **8**, 949-955 (2011).

153. Roccio, M., Gobaa, S. & Lutolf, M.P. High-throughput clonal analysis of neural stem cells in microarrayed artificial niches. *Integr Biol (Camb)* **4**, 391-400 (2012).

154. Burdick, J.A. & Watt, F.M. High-throughput stem-cell niches. *Nat Methods* **8**, 915-916 (2011).

155. Eldar, A. & Elowitz, M.B. Functional roles for noise in genetic circuits. *Nature* **467**, 167-173 (2010).

156. Munsky, B., Neuert, G. & van Oudenaarden, A. Using gene expression noise to understand gene regulation. *Science* **336**, 183-187 (2012).

157. Pai, J.H. *et al.* Photoresist with low fluorescence for bioanalytical applications. *Anal Chem* **79**, 8774-8780 (2007).

158. Jackman, R.J., Duffy, D.C., Ostuni, E., Willmore, N.D. & Whitesides, G.M. Fabricating large arrays of microwells with arbitrary dimensions and filling them using discontinuous dewetting. *Anal Chem* **70**, 2280-2287 (1998).

159. Gong, S. *et al.* A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. *Nature* **425**, 917-925 (2003).

160. Gracz, A.D., Puthoff, B.J. & Magness, S.T. Identification, isolation, and culture of intestinal epithelial stem cells from murine intestine. *Methods Mol Biol* **879**, 89-107 (2012).

161. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-682 (2012).

162. Rawlins, E.L. *et al.* The role of Scgb1a1+ Clara cells in the long-term maintenance and repair of lung airway, but not alveolar, epithelium. *Cell Stem Cell* **4**, 525-534 (2009).

163. Stange, D.E. *et al.* Differentiated Troy Chief Cells Act as Reserve Stem Cells to Generate All Lineages of the Stomach Epithelium. *Cell* **155**, 357-368 (2013).

164. Gregorieff, A. *et al.* The ets-domain transcription factor Spdef promotes maturation of goblet and paneth cells in the intestinal epithelium. *Gastroenterology* **137**, 1333-1345 e1331-1333 (2009).

165. Shroyer, N.F., Wallis, D., Venken, K.J., Bellen, H.J. & Zoghbi, H.Y. Gfi1 functions downstream of Math1 to control intestinal secretory cell subtype allocation and differentiation. *Genes Dev* **19**, 2412-2417 (2005).

166. Andersen, C.L. *et al.* Dysregulation of the transcription factors SOX4, CBFB and SMARCC1 correlates with outcome of colorectal cancer. *Br J Cancer* **100**, 511-523 (2009).

167. Meissner, A. *et al.* Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**, 766-770 (2008).

168. Tahiliani, M. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930-935 (2009).

169. Ito, S. *et al.* Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**, 1129-1133 (2010).

170. Williams, K. *et al.* TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* **473**, 343-348 (2011).

171. Wu, H. *et al.* Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature* **473**, 389-393 (2011).

172. Zhang, R.R. *et al.* Tet1 regulates adult hippocampal neurogenesis and cognition. *Cell Stem Cell* **13**, 237-245 (2013).

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

174. Penzo-Mendez, A., Dy, P., Pallavi, B. & Lefebvre, V. Generation of mice harboring a Sox4 conditional null allele. *Genesis* **45**, 776-780 (2007).

175. Xu, Y. *et al.* Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. *Mol Cell* **42**, 451-464 (2011).

176. Usui, A. *et al.* The early retinal progenitor-expressed gene Sox11 regulates the timing of the differentiation of retinal cells. *Development* **140**, 740-750 (2013).

177. Fu, X. *et al.* MicroRNA-26a targets ten eleven translocation enzymes and is regulated during pancreatic cell differentiation. *Proc Natl Acad Sci U S A* (2013).

178. Dawlaty, M.M. *et al.* Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. *Cell Stem Cell* **9**, 166-175 (2011).

179. Gregorieff, A. & Clevers, H. In situ hybridization to identify gut stem cells. *Curr Protoc Stem Cell Biol* **Chapter 2**, Unit 2F 1 (2010).

180. Hu, Z., Troester, M. & Perou, C.M. High reproducibility using sodium hydroxide-stripped long oligonucleotide DNA microarrays. *Biotechniques* **38**, 121-124 (2005).

181. Cheung, M., Abu-Elmagd, M., Clevers, H. & Scotting, P.J. Roles of Sox4 in central nervous system development. *Brain Res Mol Brain Res* **79**, 180-191 (2000).

182. Koo, B.K. *et al.* Tumour suppressor RNF43 is a stem-cell E3 ligase that induces endocytosis of Wnt receptors. *Nature* **488**, 665-669 (2012).

183. Koo, B.K. *et al.* Controlled gene expression in primary Lgr5 organoid cultures. *Nat Methods* (2011).

184. Onuma, K. *et al.* Genetic reconstitution of tumorigenesis in primary intestinal cells. *Proc Natl Acad Sci U S A* **110**, 11127-11132 (2013).

185. Kirsch, D.G. *et al.* p53 controls radiation-induced gastrointestinal syndrome in mice independent of apoptosis. *Science* **327**, 593-596 (2010).

186. Merritt, A.J., Allen, T.D., Potten, C.S. & Hickman, J.A. Apoptosis in small intestinal epithelial from p53-null mice: evidence for a delayed, p53-independent G2/M-associated cell death after gamma-irradiation. *Oncogene* **14**, 2759-2766 (1997).

187. Merritt, A.J. *et al.* The role of p53 in spontaneous and radiation-induced apoptosis in the gastrointestinal tract of normal and p53-deficient mice. *Cancer Res* **54**, 614-617 (1994).

188. Pan, X. *et al.* Induction of SOX4 by DNA damage is critical for p53 stabilization and function. *Proc Natl Acad Sci U S A* **106**, 3788-3793 (2009).

189. Cimmino, L., Abdel-Wahab, O., Levine, R.L. & Aifantis, I. TET family proteins and their role in stem cell differentiation and transformation. *Cell Stem Cell* **9**, 193-204 (2011).

190. Blaschke, K. *et al.* Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature* **500**, 222-226 (2013).

191. Perry, J.M. *et al.* Cooperation between both Wnt/{beta}-catenin and PTEN/PI3K/Akt signaling promotes primitive hematopoietic stem cell self-renewal and expansion. *Genes Dev* **25**, 1928-1942 (2011).

192. O'Brien, L.E., Soliman, S.S., Li, X. & Bilder, D. Altered modes of stem cell division drive adaptive intestinal growth. *Cell* **147**, 603-614 (2011).

193. Kobielak, K., Stokes, N., de la Cruz, J., Polak, L. & Fuchs, E. Loss of a quiescent niche but not follicle stem cells in the absence of bone morphogenetic protein signaling. *Proc Natl Acad Sci U S A* **104**, 10063-10068 (2007).